ESTIMATING THE SPECTRUM OF A USING THE ROOTS OF THE POLYNOMIALS ASSOCIATED WITH THE QMR ITERATION

TERI BARTH * AND THOMAS MANTEUFFEL†

Abstract. In this paper we present an alternative definition for the QMR method for the solution of the linear system $A\underline{x} = \underline{b}$, as a variable metric conjugate gradient method. Based on earlier work done on the conjugate gradient method, two sets of polynomials associated with the QMR iteration, the B-orthogonal polynomials and the residual polynomials, are used to construct regions in the complex plane that approximate the spectrum of A. Although this approximation may not be as good as that resulting from a conjugate gradient iteration, it is obtained more economically because of the short recurrences for the QMR iterates.

1. Introduction.

1.1. Motivation.

In an earlier paper [MaOt92] the roots of the orthogonal and residual polynomials associated with a conjugate gradient method were examined. It was noted that, while the quasi-minimal residual (QMR) method [FrNa91] is not a conjugate gradient method, it does minimize the error at each step with respect to an inner product norm, $\|\cdot\| = \langle B\cdot,\cdot\rangle^{1/2}$. Here, B is a Hermitian positive definite matrix that depends on the initial residual, $\underline{r}_0$, as well as other parameters in the QMR algorithm, (see Section 2.3). Thus, many of the results in [MaOt92] apply to the QMR method as well. In particular, the direction vectors, $\underline{p}_i$'s, and the residual vectors, $\underline{r}_i$'s, generated by the QMR method satisfy:

$$\begin{array}{rcl} \underline{p}_i & \perp_B & \mathcal{K}_i(\underline{r}_0, A), \\ A^{-1}\underline{r}_i & \perp_B & \mathcal{K}_i(\underline{r}_0, A), \end{array}$$

where,

$$\mathcal{K}_i(\underline{r}_0, A) = sp\{\underline{r}_0, A\underline{r}_0, \ldots, A^{i-1}\underline{r}_0\}$$

is the Krylov subspace of dimension i generated by $\underline{r}_0$ and the matrix A, and $\perp_B$ represents orthogonality in the B inner product (see [AMS90]). Throughout this paper we will consider the matrix A to be preconditioned already. Since $\underline{p}_i, \underline{r}_i \in \mathcal{K}_{i+1}(\underline{r}_0, A)$ we have

$$\begin{array}{rcl} \underline{p}_i & = & \phi_i(A)\underline{r}_0, \\ \underline{r}_i & = & \rho_i(A)\underline{r}_0, \end{array}$$

for some polynomials $\phi_i(\lambda)$, $\rho_i(\lambda)$ of degree $\leq i$. We refer to these as the B-orthogonal polynomials, and the residual polynomials, respectively. In this paper we will examine the roots of these polynomials and some of their properties.

In Section 2 we develop Lanczos biorthogonalization and the QMR algorithm as an introduction to notation. It is also shown that there exists a Hermitian positive definite matrix, B, that depends upon the initial residual, $\underline{r}_0$, and that the QMR algorithm is optimal at each step in the norm induced by B. Thus, QMR is an example of a balanced projection method with a variable metric [JoMa90]. We refer

* Department of Mathematics, University of Colorado at Denver, Denver, CO 80204

† Program in Applied Mathematics, Campus Box 526, University of Colorado at Boulder, Boulder, CO 80309-0526. This work was sponsored by the National Science Foundation under grant number DMS-8704169 and Depatrtment of Energy, Applied Math Program Grant DE-FG03-94ER25217

to QMR as a variable metric conjugate gradient method. This observation is very important in that it shows that the associated polynomials have nice properties. We remark that other iterative methods based upon Lanczos biorthogonalization, such as BCG, are also variable metric conjugate gradient methods and their associated polynomials satisfy similar properties. We hope to study this in a future paper.

The paper [Fr91] introduced the concept of quasi-kernel polynomials and showed that the residual polynomials, $\rho_k(\lambda)$, associated with the QMR iteration are quasi-kernel polynomials. The roots of the residual polynomials are then used to approximate the spectrum of the iteration matrix. In our paper, we show that QMR is a variable metric method. This implies that the residual polynomials are, in fact, kernel polynomials with respect to the norm in which the Lanczos basis is orthonormal. In fact, all quasi-kernel polynomials are true kernel polynomials for some norm. As described in [Fr91], quasi-kernel polynomials are defined independent of any norm on the space of polynomials. They depend only on the selection of an arbitrary ascending basis for the polynomials. The quasi-minimal property, Equation (3.13) in [Fr91], induces a norm on the space of polynomials. The quasi-kernel polynomials of [Fr91] are true kernel polynomials in this norm. Moreover, it is possible to choose the ascending basis to generate ANY ascending sequence as quasi-kernel polynomials. Thus, the fact the QMR residual polynomials are quasi-kernel polynomials yields no information about the properties of their roots. It is only when further information about the selection of the ascending basis is supplied that such properties can be inferred. In the context of iterative methods, the ascending basis corresponds to a sequence of vectors in $\mathcal{C}^n$. In particular, for the QMR method this sequence of vectors is a result of the Lanczos biorthogonalization algorithm. In this paper, the Lanczos vectors are used to construct the matrix B, which provides the metric for the iteration and the properties of the roots of the associated polynomials as described below.

We use the roots of both the residual and the B-orthogonal polynomials to approximate the spectrum of the iteration matrix. In Section 3, we show how to compute the roots of these polynomials. The roots of both polynomials can be obtained at a lower cost than the computation of just the roots of the residual polynomials given in [Fr91].

From the results in [MaOt92], we know that the roots of the B-orthogonal polynomials are in the B-field of values of A, and the reciprocals of the roots of the residual polynomials are in the B-field of values of A^{-1}. This yields stability bounds on the roots. Section 4 shows how to use these roots to construct sets that approximate the spectrum of A using the QMR iteration. This algorithm for approximating the spectrum, can be applied to any conjugate gradient, or variable metric conjugate gradient method.

In Section 5, we implement the algorithm given in Section 4. We compare the roots of the polynomials and the resulting approximations of the spectrum obtained from the QMR iteration to that obtained via a full conjugate gradient iteration. Our numerical experiments show that the use of both sets of roots results in a much better approximation of the spectrum than that obtained by using the roots of the residual polynomials alone. The results demonstrate good approximations for definite problems, but rather poor results for indefinite problems. Also, our numerical tests indicate that good approximations of the spectrum and fast convergence of the corresponding iterative method go together.

The QMR residual polynomials are kernel polynomials, and thus optimal as residual polynomials, in the inner product norm in which the Lanczos vectors are orthonormal. This norm is related to the standard norm of the error through the matrix B which we construct in Section 2.3. Heuristically, if the Lanczos vectors are nearly orthonormal in an inner product norm of interest, then the convergence of QMR will be nearly optimal in that norm. Likewise, the roots of the orthogonal polynomials will be close to eigenvalues of an orthogonal section of A in that norm. The roots of the residual polynomials will obey a similar property. If the Lanczos vectors are not nearly orthonormal in any norm of interest, then the QMR residual polynomials will be far from optimal. This underscores the need for a better understanding of the relationship between properties of the matrix A, selection of the initial right and left Lanczos vectors, and the behavior of the Lanczos vectors.

1.2. Notation. The following notation will be employed throughout this paper. The usage should be clear from the context.

$\mathcal{R}^n, \mathcal{C}^n$	–	Vector spaces of real and complex $n-$ tuples.
$\mathcal{R}^{n\times m}, \mathcal{C}^{n\times m}$	–	Vector spaces of real and complex $n \times m$ matrices.
$\mathcal{P}_k$	–	Space of polynomials of degree at most k.
$\mathcal{K},\ \mathcal{V}, \ldots$	–	Other calligraphic letters denote subspaces of $\mathcal{R}^n$ or $\mathcal{C}^n$.
$A, B, \ldots$	–	Upper case Roman and Greek letters denote matrices.
$\underline{x}, \underline{p}, \ldots$	–	Underlined lower case Roman and Greek letters denote vectors.
$\alpha, \beta, \ldots$	–	Lower case Roman and Greek letters denote scalars.
$\langle \cdot, \cdot \rangle$	–	Euclidean inner product on $\mathcal{C}^n$.
$\langle B\cdot, \cdot \rangle$	–	$B-$inner product on $\mathcal{C}^n$.
A^*	–	Euclidean adjoint of A, that is, $A^* = \bar{A}^T$.
$A^\dagger$	–	$B-$adjoint of A.
$sp\{\underline{x}_j\}$	–	The linear span of the vectors $\underline{x}_j$.
$\Sigma(A)$	–	Spectrum of A.
$\mathcal{H}(A)$	–	Convex hull of $\Sigma(A)$.
$\mathcal{F}_B(A)$	–	$B-$field of values of A.

2. Exposition of QMR. Given an initial guess, $\underline{x}_0$, the QMR method for the solution of the $N \times N$ linear system

$$A\underline{x} = \underline{b}, \tag{2.1}$$

constructs a sequence of iterates such that

$$\underline{x}_n = \underline{x}_0 + \underline{q}^{(n)}, \qquad \underline{q}^{(n)} \in \mathcal{K}_n(\underline{r}_0, A), \tag{2.2}$$

where, $\underline{q}^{(n)}$ is chosen to satisfy a quasi-minimal residual property. This property is described in Section 2.2.

2.1. Lanczos biorthogonalization. The classical nonsymmetric Lanczos process generates a pair of bases $\{\underline{v}_j\}_{j=0}^{n-1}$, $\{\underline{w}_j\}_{j=0}^{n-1}$ for $\mathcal{K}_n(\underline{v}_0, A)$ and $\mathcal{K}_n(\underline{w}_0, A^*)$ respectively. The $\underline{v}$'s ($\underline{w}$'s) are known as right (left) Lanczos vectors. These vectors are constructed to satisfy a biorthogonality condition

$$\underline{w}_i^* \underline{v}_j = \begin{cases} \delta_i \neq 0 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \qquad i, j = 1, ..., n-1. \tag{2.3}$$

This is accomplished via a Gram Schmidt-like process using 3-term recurrences (see [Ho64]).

$$\begin{aligned} \underline{v}_n &= A\underline{v}_{n-1} - \alpha_{n-1}\underline{v}_{n-1} - \beta_{n-1}\underline{v}_{n-2}, \\ \underline{w}_n &= A^*\underline{w}_{n-1} - \bar{\alpha}_{n-1}\underline{w}_{n-1} - \bar{\beta}_{n-1}\underline{w}_{n-2}, \end{aligned}$$

where,

$$\alpha_{n-1} = \frac{\underline{w}^*_{n-1}A\underline{v}_{n-1}}{\underline{w}^*_{n-1}\underline{v}_{n-1}} \quad \text{and} \quad \beta_{n-1} = \frac{\underline{w}^*_{n-2}A\underline{v}_{n-1}}{\underline{w}^*_{n-2}\underline{v}_{n-2}} = \frac{\underline{w}^*_{n-1}\underline{v}_{n-1}}{\underline{w}^*_{n-2}\underline{v}_{n-2}}. \tag{2.4}$$

The QMR method is implemented via Lanczos biorthogonalization with a look-ahead strategy [FrGuNa91]. This look-ahead Lanczos (LAL) method was devised to avoid the breakdowns that can occur in the standard Lanczos process. For example, $\underline{w}^*_{n-1}\underline{v}_{n-1} = 0$ in (2.4) would result in a breakdown. Two types of right (left) Lanczos vectors are made, regular and inner vectors. Inner vectors are those that are constructed to avoid breakdowns. This is done by relaxing the biorthogonality condition (2.3). Grouping these vectors into blocks where the first vector in every block is a regular vector, and all other vectors are inner vectors, the biorthogonality condition resulting from the LAL process reads

$$W_i^* V_j = \begin{cases} D_i \neq [0]_{d_i \times d_j} & \text{if } i = j \\ [0]_{d_i \times d_j} & \text{otherwise} \end{cases} \qquad i,j = 1, ..., \ell - 1, \tag{2.5}$$

where V_j is the matrix whose columns are the right Lanczos vectors from the j'th block. Likewise, W_i denotes the matrix whose columns are the left Lanczos vectors from the i'th block. The number of columns of V_j is denoted by d_j, and ℓ refers to the current block of Lanczos vectors. For $i,\ j = 1, ..., \ell - 1$, the blocks are complete, the ℓ'th block is not complete unless the next pair, $\underline{v}_n,\ \underline{w}_n$, are regular vectors, thus beginning the next block.

We see that in LAL, right and left Lanczos vectors are orthogonal if they are from blocks with different indices. Like the classical method, this orthogonality can be accomplished with short recurrences. Instead of involving just the two previous vectors, the recursions in LAL involve the two previous blocks of vectors. Lanczos vectors are often scaled for stability reasons. In this paper, they will be scaled to have unit length. We begin by choosing the first right Lanczos vector to be the normalized initial residual, i.e., $\underline{v}_0 = \underline{r}_0/||\underline{r}_0||$, so that $\mathcal{K}_n(\underline{v}_0, A) = \mathcal{K}_n(\underline{r}_0, A)$. The starting left Lanczos vector, $\underline{w}_0$, is an arbitrary length one vector. These recurrences are:

Regular vectors:

$$\begin{aligned} \tilde{\underline{v}}_n &= A\underline{v}_{n-1} - V_\ell[D_\ell^{-1}W_\ell^* A\underline{v}_{n-1}] - V_{\ell-1}[D_{\ell-1}^{-1}W_{\ell-1}^* A\underline{v}_{n-1}], \\ \tilde{\underline{w}}_n &= A^*\underline{w}_{n-1} - W_\ell[(D_\ell^{-1})^* V_\ell^* A^*\underline{w}_{n-1}] - W_{\ell-1}[(D_{\ell-1}^{-1})^* V_{\ell-1}^* A^*\underline{w}_{n-1}], \end{aligned}$$

Inner vectors:

$$\begin{aligned} \tilde{\underline{v}}_n &= A\underline{v}_{n-1} - \zeta_{n-1}\underline{v}_{n-1} - \eta_{n-1}\underline{v}_{n-2} - V_{\ell-1}[D_{\ell-1}^{-1}W_{\ell-1}^* A\underline{v}_{n-1}], \\ \tilde{\underline{w}}_n &= A^*\underline{w}_{n-1} - \zeta_{n-1}\underline{w}_{n-1} - \eta_{n-1}\underline{w}_{n-2} - W_{\ell-1}[(D_{\ell-1}^{-1})^* V_{\ell-1}^* A^*\underline{w}_{n-1}], \end{aligned}$$

where $\delta_n = ||\tilde{\underline{v}}_n||$, $\xi_n = ||\tilde{\underline{w}}_n||$, $\underline{v}_n = \tilde{\underline{v}}_n/\delta_n$, $\underline{w}_n = \tilde{\underline{w}}_n/\xi_n$, V_l and W_l are the matrices whose columns are the normalized Lanczos vectors in the ℓ'th blocks, and $\zeta_{n-1},\ \eta_{n-1}$ are constants.

Notice that in order to compute $\underline{v}_n$, $\underline{w}_n$ as regular vectors and begin the $(\ell+1)$'st block, D_ℓ must be nonsingular. From this it follows that in (2.5), when $i = j$ and i is a complete block, D_i is nonsingular.

Let

$$(2.6)\quad V_n = \begin{bmatrix} | & | & & | \\ \underline{v}_0 & \underline{v}_1 & \cdots & \underline{v}_{n-1} \\ | & | & & | \end{bmatrix}, \text{ and } W_n = \begin{bmatrix} | & | & & | \\ \underline{w}_0 & \underline{w}_1 & \cdots & \underline{w}_{n-1} \\ | & | & & | \end{bmatrix},$$

denote the matrices whose columns are the normalized right and left Lanczos vectors resulting from n steps of the LAL algorithm. The recurrences for the vectors generated by this iteration can be written in matrix notation as follows:

$$(2.7)\qquad AV_n = V_{n+1}H_{n+1,n}, \qquad A^*W_n = W_{n+1}\tilde{H}_{n+1,n}$$

for the right and left Lanczos vectors respectively, where,

$$(2.8)\qquad H_{n+1,n} = \begin{bmatrix} \alpha_1 & \beta_2 & 0 & \cdots & 0 \\ \gamma_2 & \alpha_2 & \ddots & \ddots & \vdots \\ 0 & \gamma_3 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \beta_l \\ \vdots & \ddots & \ddots & \ddots & \alpha_l \\ 0 & 0 & \cdots & 0 & \gamma_{\ell+1} \end{bmatrix}_{n+1\times n},$$

is a block tridiagonal upper Hessenberg matrix. Also, from (2.5), the biorthogonality condition can be written as

$$(2.9)\qquad W_n^*V_n = \hat{D}_n = \begin{bmatrix} D_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_\ell \end{bmatrix}_{n\times n},$$

where $\hat{D}_n$ is a block diagonal matrix with nonsingular blocks, D_k, for $k = 1, ..., \ell - 1$.

If only regular vectors are made in LAL, the blocks of Lanczos vectors are all 1×1 and like classical Lanczos we have 3-term recurrences, $H_{n,n}$ is tridiagonal and $\hat{D}_n$ is diagonal. For a full exposition of LAL see [FrGuNa91].

2.2. The quasi-minimal residual property. Since $\{\underline{v}_j\}_{j=0}^{n-1}$ forms a basis for $\mathcal{K}_n(\underline{r}_0, A)$ we have

$$(2.10)\qquad \mathcal{K}_n(\underline{r}_0, A) = sp\{\underline{v}_0, \cdots, \underline{v}_{n-1}\} = \{V_n\underline{z},\ \underline{z} \in \mathcal{C}^n\}.$$

Thus, our iterates (2.2) can be written

$$(2.11)\qquad \underline{x}_n = \underline{x}_0 + V_n\underline{z},\ \underline{z} \in \mathcal{C}^n,$$

and the corresponding residual vectors satisfy

$$\underline{r}_n = \underline{b} - A\underline{x}_n = \underline{b} - A(\underline{x}_0 + V_n\underline{z}) = \underline{r}_0 - AV_n\underline{z}.$$

Let $\underline{s}_{n+1} = [\delta_0, 0, \cdots, 0]^T$ have length $n+1$, where $\delta_0 = \|\underline{r}_0\|$. Using (2.7) and recalling that the first column of V_{n+1} is $\underline{r}_0/\delta_0$, we see that

$$\underline{r}_n = \underline{r}_0 - V_{n+1} H_{n+1,n} \underline{z} = V_{n+1}(\underline{s}_{n+1} - H_{n+1,n}\underline{z}). \tag{2.12}$$

Minimizing $\|\underline{r}_n\|_2$ at each step is rather expensive, so instead QMR chooses $\underline{z}^{(n)}$ to satisfy

$$\min_{\underline{z} \in \mathcal{C}^n} \|\underline{s}_{n+1} - H_{n+1,n}\underline{z}\|_2. \tag{2.13}$$

This is called the quasi-minimal residual property, [FrNa91].

To allow for the modification of the scaling of the problem, in [FrNa91], they introduce an $(n+1)\times(n+1)$ diagonal weight matrix, $\Omega_{n+1} = \mathrm{diag}(\omega_1, \ldots, \omega_{n+1})$, where the ω_j's are all greater than zero. Except where noted, to simplify the development in this paper, we have omitted this matrix, which is the same as setting $\Omega_{n+1} = I_{n+1}$, the $(n+1)\times(n+1)$ identity matrix. All results apply to $\Omega_{n+1} \neq I_{n+1}$.

2.3. The B-norm of the error. In the absence of an incurable breakdown of the look-ahead Lanczos process [FrGuNa91], we know that the matrices V_d, W_d, and $H_{d,d}$ exist, and $\hat{D}_d = W_d^* V_d$ is nonsingular. Here, $d = d(\underline{r}_0, A)$, is the dimension of the maximal Krylov subspace generated by $\underline{r}_0$ and A. So, (2.12) can be rewritten as

$$\underline{r}_n = \begin{bmatrix} | & & | \\ \underline{v}_0 & \cdots & \underline{v}_{d-1} \\ | & & | \end{bmatrix}_{N\times d} \left(\begin{pmatrix} \underline{s}_{n+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_d - H_{d,d} \begin{pmatrix} \underline{z} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_d \right).$$

Multiplying on the left by $\hat{D}_d^{-1} W_d^*$ yields

$$\hat{D}_d^{-1} W_d^* \underline{r}_n = \begin{pmatrix} \underline{s}_{n+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_d - H_{d,d} \begin{pmatrix} \underline{z} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_d . \tag{2.14}$$

Now by substituting $\underline{r}_n = A\underline{e}_n$ into (2.14) we see that QMR minimizes

$$\|\hat{D}_d^{-1} W_d^* A \underline{e}_n\|_2. \tag{2.15}$$

By definition, this is the same as

$$\langle \hat{D}_d^{-1} W_d^* A\underline{e}_n, \hat{D}_d^{-1} W_d^* A\underline{e}_n \rangle^{1/2} = \langle (\hat{D}_d^{-1} W_d^* A)^* (\hat{D}_d^{-1} W_d^* A)\underline{e}_n, \underline{e}_n \rangle^{1/2},$$

which is just

$$\langle B\underline{e}_n, \underline{e}_n \rangle^{1/2} = \|\underline{e}_n\|_B, \tag{2.16}$$

where,

$$B = (\hat{D}_d^{-1} W_d^* A)^* (\hat{D}_d^{-1} W_d^* A) = A^* W_d (\hat{D}_d^{-1})^* \hat{D}_d^{-1} W_d^* A. \tag{2.17}$$

If $d = N$, the matrix B will be Hermitian positive definite. If $d < N$, then $\mathcal{K}_d(\underline{r}_0, A)$ is an invariant subspace of both A and A^{-1}, since we can write $A^{-1}\underline{r}_0 = p_{d-1}(A)\underline{r}_0$, where $p_{d-1}(\lambda)$ is a polynomial of degree $d-1$. Thus, we may recast the problem by restricting A to $\mathcal{K}_d(\underline{r}_0, A)$. In this context, B is a Hermitian positive definite operator on $\mathcal{K}_d(\underline{r}_0, A)$. If we had allowed for a scaling matrix, Ω_d, then (2.17) would read

$$B = A^* W_d (\hat{D}_d^{-1})^* \Omega_d^* \Omega_d \hat{D}_d^{-1} W_d^* A. \tag{2.18}$$

THEOREM 2.1. *The iterates generated by the QMR method are uniquely defined by the following properties:*

$$\begin{array}{rclc} \underline{x}_n & = & \underline{x}_0 + \underline{q}^{(n)} & \underline{q}^{(n)} \in \mathcal{K}_n(\underline{r}_0, A), \\ \underline{e}_n & \perp_B & \mathcal{K}_n(\underline{r}_0, A), & \end{array} \tag{2.19}$$

that is, $\underline{q}^{(n)}$ is chosen to minimize $||\underline{e}_n||_B$ over $\mathcal{K}_n(\underline{r}_0, A)$, for B given by (2.17).

Proof. The proof follows from the above discussion. □

Remark: Substituting $\underline{e}_n = A^{-1}\underline{r}_n$ into (2.16) shows the residuals are optimal in the norm

$$\hat{B} = W_d (\hat{D}_d^{-1})^* \hat{D}_d^{-1} W_d^*. \tag{2.20}$$

Viewed in this way, we see that QMR is a balanced projection method, (cf. [JoMa90]). Since B is Hermitian positive definite, this definition also looks like the definition for the iterates produced by a conjugate gradient method, [AMS90]. To see the difference between these two methods, we first review some terminology.

Assume B is Hermitian positive definite and recall that the B-adjoint of A is the unique matrix, $A^\dagger$, satisfying

$$\langle BA\underline{x}, \underline{y}\rangle = \langle B\underline{x}, A^\dagger \underline{y}\rangle, \quad \forall \underline{x}, \underline{y} \in \mathcal{C}^N.$$

This yields, $A^\dagger = B^{-1}A^*B$. The matrix A is B-normal if it satisfies the following equivalent conditions:

i) $AA^\dagger = A^\dagger A$,

ii) A, $A^\dagger$ have the same complete set of B-orthogonal eigenvectors,

iii) $A^\dagger$ can be written as a polynomial of some degree η in the matrix A.

We say that A is B-normal(s) if it is B-normal, and s is the smallest degree for which $A^\dagger = \mathcal{P}_\eta(A)$.

In [FaMa84] it was shown that a conjugate gradient method that minimizes in a fixed B-norm can be carried out with a single s-term recursion for any choice of $\underline{x}_0$ and $\underline{b}$ if and only if A is B-normal(s-2) or the dimension of the minimal polynomial of A is less than or equal to s. We have shown above that the QMR method is optimal in the B norm, where B is defined as in (2.17). This does not contradict the above result in that here B depends upon $\underline{r}_0$, and thus on $\underline{x}_0$ and $\underline{b}$. To see this, recall that $\hat{D}_d = W_d^* V_d$, where the first column of V_d is $\underline{r}_0/\delta_0$. Since QMR minimizes the error at each step in a norm that depends on the initial residual, as well as other parameters, it could be referred to as a variable metric conjugate gradient method.

The above observation was confirmed by numerical experiments. The B obtained after d steps of LAL was then used in a full conjugate gradient algorithm (Orthodir), along with the same choice of $\underline{x}_0$ and $\underline{b}$. For our test problems, the iterates from QMR and the conjugate gradient iterations were identical.

We now develop some identities that will be useful in Section 3. Consider (2.7) for $n = d = d(\underline{r}_0, A)$; from this it follows that

$$AV_d = V_d H_{d,d}, \quad A^* W_d = W_d \tilde{H}_{d,d}. \tag{2.21}$$

Multiplying the first equation from the left by W_d^* and the second equation from the left by V_d^* yields

$$W_d^* A V_d = \hat{D}_d H_{d,d}, \quad V_d^* A^* W_d = \hat{D}_d^* \tilde{H}_{d,d}.$$

Since the first equation is the adjoint of the second equation, we have that

$$(\hat{D}_d H_{d,d})^* = H_{d,d}^* \hat{D}_d^* = \hat{D}_d^* \tilde{H}_{d,d},$$

and finally,

$$\tilde{H}_{d,d}(\hat{D}_d^{-1})^* = (\hat{D}_d^{-1})^* H_{d,d}^*. \tag{2.22}$$

By using (2.18),(2.21), and (2.22), B can be rewritten equivalently in two forms:

$$\begin{aligned} B &= W_d(\hat{D}_d^{-1})^* H_{d,d}^* \Omega_d^* \Omega_d H_{d,d} \hat{D}_d^{-1} W_d^*, \\ B &= W_d(\hat{D}_d^{-1})^* H_{d,d}^* \Omega_d^* \Omega_d \hat{D}_d^{-1} W_d^* A. \end{aligned} \tag{2.23}$$

3. The polynomials associated with the QMR iteration. The solution of the least squares problem (2.13) is done using QR decomposition of $H_{n+1,n}$,

$$H_{n+1,n} = Q_{n+1}^* \begin{bmatrix} R_n \\ \underline{0}^* \end{bmatrix}, \tag{3.1}$$

where Q_{n+1} is an $(n+1) \times (n+1)$ unitary matrix, constructed as a product of Givens rotations, and R_n is an $n \times n$ nonsingular upper triangular matrix [FrNa91]. Substituting (3.1) into (2.13) and simplifying leaves us with solving the following least squares problem. At the n'th step, choose $\underline{z}^{(n)}$ to satisfy

$$\min_{\underline{z} \in \mathcal{C}^n} \|Q_{n+1} \underline{s}_{n+1} - \begin{bmatrix} R_n \\ \underline{0}^* \end{bmatrix} \underline{z}\|_2. \tag{3.2}$$

The solution to this problem is given by

$$\begin{aligned} \underline{z}^{(n)} &= R_n^{-1} \underline{t}_n, \quad \text{where} \\ \underline{t}_n &= \begin{bmatrix} \tau_0 \\ \vdots \\ \tau_{n-1} \end{bmatrix}, \quad \hat{\underline{t}}_n = \begin{bmatrix} \underline{t}_n \\ \tilde{\tau}_n \end{bmatrix} = Q_{n+1} \underline{s}_{n+1}. \end{aligned} \tag{3.3}$$

Since Q_{n+1} is the product of Givens rotation matrices, call them $G_k, k = 1, ..., n$, the vector $\hat{\underline{t}}_n$ is updated from the previous step by applying only the last Givens rotation matrix,

$$\hat{\underline{t}}_n = G_n \begin{bmatrix} \tau_0 \\ \vdots \\ \tau_{n-2} \\ \tilde{\tau}_{n-1} \\ 0 \end{bmatrix} = \begin{bmatrix} \tau_0 \\ \vdots \\ \tau_{n-2} \\ \tau_{n-1} \\ \tilde{\tau}_n \end{bmatrix}.$$

Notice that $\tau_0, \ldots, \tau_{n-2}$ do not change in this update from step n to step $n+1$. Using (3.3), (2.11) and (2.13), we rewrite the QMR iterates as

$$\underline{x}_n = \underline{x}_0 + V_n (R_n)^{-1} \underline{t}_n. \tag{3.4}$$

Let

$$P_n = \begin{bmatrix} | & & | \\ \underline{p}_0 & \cdots & \underline{p}_{n-1} \\ | & & | \end{bmatrix} = V_n (R_n)^{-1}, \tag{3.5}$$

denote our matrix of direction vectors. With this, (3.4) becomes

$$\underline{x}_n = \underline{x}_0 + P_n \underline{t}_n.$$

Since $\underline{t}_n = [\tau_0, \ldots, \tau_{n-2}, \tau_{n-1}]^T$, where $\tau_0, \ldots, \tau_{n-2}$ are the same as on the previous step, this can be written equivalently as

$$\underline{x}_n = \underline{x}_{n-1} + \underline{p}_{n-1} \tau_{n-1}. \tag{3.6}$$

As a consequence of Theorem 2.1 we know that

$$\tau_{n-1} = \frac{\langle B \underline{e}_{n-1}, \underline{p}_{n-1} \rangle}{\langle B \underline{p}_{n-1}, \underline{p}_{n-1} \rangle}. \tag{3.7}$$

Furthermore, it follows from (2.10),(2.12), and (3.5) that

$$\underline{p}_n, \underline{r}_n \in \mathcal{K}_{n+1}(\underline{r}_0, A), \tag{3.8}$$

and the following properties are satisfied: (see [AMS90])

$$\begin{array}{rcll} \langle B \underline{p}_i, \underline{p}_j \rangle & = & 0 & i \neq j, \\ \langle B \underline{e}_i, \underline{r}_j \rangle & = & 0 & i > j. \end{array} \tag{3.9}$$

Thus, from (3.8), $\underline{p}_n$, and $\underline{r}_n$ can be expressed as

$$\begin{array}{rcl} \underline{p}_n & = & \phi_n(A) \underline{r}_0, \\ \underline{r}_n & = & \rho_n(A) \underline{r}_0, \end{array} \tag{3.10}$$

where the polynomials, $\phi_n(\lambda)$ and $\rho_n(\lambda)$, are called the B-orthogonal polynomials and the residual polynomials respectively. As mentioned in Section 2.3, the residuals are optimal in the norm given by (2.20). Therefore, the residual polynomials can be referred to as kernel polynomials with respect to that norm. We note here that this is also the norm for which the right Lanczos vectors are orthonormal.

3.1. The roots of the B-orthogonal polynomials. Using (3.8) and (3.9) we see that $\underline{p}_n$ is the unique (up to scale) vector such that

$$\begin{array}{rcl} \underline{p}_n & \in & \mathcal{K}_{n+1}(\underline{r}_0, A), \\ \underline{p}_n & \perp_B & \mathcal{K}_n(\underline{r}_0, A). \end{array}$$

The following theorem is a result from [MaOt92].

THEOREM 3.1. *Suppose that* $\underline{p}_i = \phi_i(A)\underline{r}_0 \neq \underline{0}$ *with* $\phi_i(\lambda) = \prod_{j=1}^{b_i}(\lambda - \lambda_j)^{m_j}$ *and* $d_i \equiv \sum_{j=1}^{b_i} m_j = i$. *Suppose*

$$\underline{p}_i \perp_B \mathcal{K}_i(\underline{r}_0, A). \tag{3.11}$$

Let S_i *and* T_i *be any two* $N \times i$ *matrices whose columns span* $\mathcal{K}_i(\underline{r}_0, A)$. *Then the roots of*

$$\phi_i(\lambda) = 0 \tag{3.12}$$

are the eigenvalues of the generalized eigenvalue problem

$$T_i^* BAS_i \underline{w} = \lambda T_i^* BS_i \underline{w}. \tag{3.13}$$

If $m_j > 1$, *then* $(T_i^* BS_i)^{-1}(T_i^* BAS_i)$ *has a Jordan block of size* m_j *associated with the eigenvalue* λ_j.

To find the roots of the B-orthogonal polynomials, $\phi_n(\lambda)$, we begin by choosing $S_n = T_n = P_n$, which is an $N \times n$ matrix whose columns span $\mathcal{K}_n(\underline{r}_0, A)$, so that (3.13) in the above theorem becomes

$$(P_n^* BAP_n)\underline{w} = \lambda(P_n^* BP_n)\underline{w}. \tag{3.14}$$

Using (3.5) and (2.7) we see that

$$AP_n = AV_n R_n^{-1} = V_{n+1} H_{n+1,n} R_n^{-1} = P_{n+1} R_{n+1} H_{n+1,n} R_n^{-1}.$$

Substituting this into (3.14) gives

$$(P_n^* BP_{n+1} R_{n+1} H_{n+1,n} R_n^{-1})\underline{w} = \lambda(P_n^* BP_n)\underline{w}.$$

The direction vectors are B-orthogonal, so $P_n^* BP_n = \Delta_n$, a diagonal matrix, and we rewrite this as

$$[\Delta_n : \underline{0}] R_{n+1} H_{n+1,n} R_n^{-1} \underline{w} = \lambda \Delta_n \underline{w}.$$

By multiplying both sides on the left by Δ_n^{-1} we have

$$[I_n : \underline{0}] R_{n+1} H_{n+1,n} R_n^{-1} \underline{w} = \lambda \underline{w},$$

which is the same as

$$[R_n : \tilde{u}_n] H_{n+1,n} R_n^{-1} \underline{w} = \lambda \underline{w},$$

where $\underline{\tilde{u}}_n$ is the off diagonal part of the last column of R_{n+1}, that is, $\underline{\tilde{u}}_n = [u_{1,n+1}, ..., u_{n,n+1}]^T$. Letting $\underline{w} = R_n \underline{\tilde{w}}$, and multiplying by R_n^{-1} on the left yields

$$[I_n : \underline{\eta}^{(n)}] H_{n+1,n} \underline{\tilde{w}} = \lambda \underline{\tilde{w}}, \tag{3.15}$$

where,

$$\underline{\eta}^{(n)} = R_n^{-1} \underline{\tilde{u}}_n. \tag{3.16}$$

Consequently, (3.13) can be rewritten as:

$$[H_{n,n} + \gamma_{n+1}(\underline{\eta}^{(n)}\underline{\epsilon}_n^*)]\underline{\tilde{w}} = \lambda\underline{\tilde{w}}, \tag{3.17}$$

where $\underline{\epsilon}_n = [0,\ldots,0,1]^T$ has length n, γ_{n+1} is the $((n+1)\times n)$'th component of the matrix $H_{n+1,n}$, and $\underline{\eta}^{(n)}$ can be found by backsubstitution of

$$R_n\underline{\eta}^{(n)} = \underline{\tilde{u}}_n, \tag{3.18}$$

or, recursively as follows:

since for every k, R_k is an upper triangular matrix,

$$R_k^{-1} = \begin{bmatrix} R_{k-1} & \underline{\tilde{u}}_{k-1} \\ \underline{0}^* & u_{k,k} \end{bmatrix}^{-1} = \begin{bmatrix} R_{k-1}^{-1} & -\underline{\eta}^{(k-1)}/u_{k,k} \\ \underline{0}^* & 1/u_{k,k} \end{bmatrix}. \tag{3.19}$$

Thus, from (3.16) and (3.19) we have

$$\underline{\eta}^{(n)} = \begin{bmatrix} R_{n-1}^{-1} & -\underline{\eta}^{(n-1)}/u_{n,n} \\ \underline{0}^* & 1/u_{n,n} \end{bmatrix}\underline{\tilde{u}}_n. \tag{3.20}$$

Expanding (3.20) gives

$$\underline{\eta}^{(n)} = \begin{bmatrix} R_{n-1}^{-1} \\ \underline{0}^* \end{bmatrix}\begin{pmatrix} u_{1,n+1} \\ \vdots \\ u_{n-1,n+1} \end{pmatrix} + \frac{u_{n,n+1}}{u_{n,n}}\begin{pmatrix} -\underline{\eta}^{(n-1)} \\ 1 \end{pmatrix}.$$

The result after applying (3.19) $n-2$ times in (3.20) and expanding is

$$\underline{\eta}^{(n)} = \frac{u_{1,n+1}}{u_{1,1}}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_n + \frac{u_{2,n+1}}{u_{2,2}}\begin{pmatrix} -\underline{\eta}^{(1)} \\ 1 \\ \underline{0} \end{pmatrix}_n + \cdots + \frac{u_{n,n+1}}{u_{n,n}}\begin{pmatrix} -\underline{\eta}^{(n-1)} \\ 1 \end{pmatrix}_n. \tag{3.21}$$

Since $H_{n+1,n+1}$ is a block tridiagonal upper Hessenberg matrix, the $(n+1)$'st column of this matrix, $\underline{h}_{n+1} = [0,...,0,*,...,*]^T$, has nonzero entries starting from the s'th component. Here, s refers to the starting index of the previous block of Lanczos vectors. Therefore the $(n+1)$'st column of R_{n+1} is obtained from $\underline{h}_{n+1}$ by applying only a few Given's rotation matrices,

$$G_{n+1}\cdots G_s G_{s-1}\underline{h}_{n+1} = \begin{pmatrix} \underline{\tilde{u}}_n \\ u_{n+1,n+1} \\ 0 \end{pmatrix}_{n+2}.$$

This process produces one row of nonzero fill in, thus $\underline{\tilde{u}}_n$ has nonzero entries starting from the $(s-1)$'st component. For details, see [FrGuNa91].

From this, we see that (3.21) can be updated with recurrences involving only a few previous $\underline{\eta}^{(k)}$'s,

$$\underline{\eta}^{(n)} = \frac{u_{s-1,n+1}}{u_{s-1,s-1}}\begin{pmatrix} -\underline{\eta}^{(s-2)} \\ 1 \\ \underline{0} \end{pmatrix}_n + \cdots + \frac{u_{n,n+1}}{u_{n,n}}\begin{pmatrix} -\underline{\eta}^{(n-1)} \\ 1 \end{pmatrix}_n. \tag{3.22}$$

If LAL builds all 1×1 blocks, the recursions for the $\underline{\eta}^{(k)}$'s involve only 4-terms.

We summarize the above results in the following theorem.

THEOREM 3.2. *The roots of the B-orthogonal polynomials, $\phi_n(\lambda) = 0$, are the eigenvalues of*

$$H_{n,n} + \gamma_{n+1}(\underline{\eta}^{(n)}\underline{e}_n^*), \tag{3.23}$$

the block tridiagonal, upper Hessenberg matrix, $H_{n,n}$, plus a rank one matrix.

Proof. See above discussion. □

3.2. The roots of the residual polynomials. We can write $\underline{e}_n = A^{-1}\underline{r}_n$, and thus equation (3.8) and (3.9) show that the residual vectors, $\underline{r}_n$, are characterized by

$$\begin{array}{rcl} \underline{r}_n & \in & \mathcal{K}_{n+1}(\underline{r}_0, A), \\ A^{-1}\underline{r}_n & \perp_B & \mathcal{K}_n(\underline{r}_0, A). \end{array}$$

From [MaOt92] we have the following theorem:

THEOREM 3.3. *Suppose that $\underline{r}_i = \rho_i(A)\underline{r}_0 \neq 0$ with $\rho_i(\lambda) = \prod_{j=1}^{b_i}(1 - \frac{\lambda}{\lambda_j})^{m_j}$ and $d_i \equiv \sum_{j=1}^{b_i} m_j \leq i$. Suppose*

$$A^{-1}\underline{r}_i \perp_B \mathcal{K}_i(\underline{r}_0, A). \tag{3.24}$$

Let S_i and T_i be any two $(N \times i)$ matrices whose columns span $\mathcal{K}_i(\underline{r}_0, A)$. If λ is a root of

$$\rho_i(\lambda) = 0 \tag{3.25}$$

then $\frac{1}{\lambda}$ is an eigenvalue of the generalized eigenvalue problem

$$T_i^* B A^{-1} S_i \underline{w} = \frac{1}{\lambda} T_i^* B S_i \underline{w}. \tag{3.26}$$

If $m_j > 1$, then $(T_i^* B S_i)^{-1}(T_i^* B A^{-1} S_i)$ has a Jordan block of size m_j associated with eigenvalue $\frac{1}{\lambda_j}$. If $d_i < i$, then $(T_i^* B S_i)^{-1}(T_i^* B A^{-1} S_i)$ has a nondiagonal Jordan block of size i - d_i associated with a zero eigenvalue.

To compute the roots of the residual polynomials, we begin with the choice of the two $N \times n$ matrices, S_n and T_n, whose columns span $\mathcal{K}_n(\underline{r}_0, A)$. One possibility is to set $S_n = T_n = V_n$. With this substitution into (3.26), along with further substitutions of (2.23), and (2.9), we can obtain

$$(H_{n+1,n})^* H_{n+1,n} \underline{w} = \lambda H_{n,n}^* \underline{w}, \tag{3.27}$$

which is the same generalized eigenvalue problem given in [Fr91], page 148. There, an equivalent formulation of (3.27) is given by

$$R_n \underline{w} = \lambda \tilde{Q}_n \underline{w}, \tag{3.28}$$

where R_n is the $n \times n$ upper triangular matrix given in (3.1) and

$$\tilde{Q}_n = [I_n : \underline{0}]_{n,n+1} Q_{n+1} \begin{bmatrix} I_n \\ \underline{0}^* \end{bmatrix}_{n+1,n}.$$

See [Fr91] for details.

Since R_n is always nonsingular, the roots of the residual polynomials can be found as the reciprocals of the eigenvalues of $R_n^{-1}\tilde{Q}_n$. If $\rho_{n-1}(\lambda) = \rho_n(\lambda)$, then $R_n^{-1}\tilde{Q}_n$ will have a zero eigenvalue. The matrix $R_n^{-1}\tilde{Q}_n$ is dense. In general, solving (3.28) via the QZ algorithm requires approximately $30n^3$ operations.

If $\rho_{n-1}(\lambda) \neq \rho_n(\lambda)$, it is possible to restructure the problem so that we are finding the eigenvalues of a matrix that is block tridiagonal, upper Hessenberg except for the last column which may be dense. Since the matrix is already upper Hessenberg, this reduces the work required in computing the roots of the residual polynomials to approximately $10n^3$ operations. This alternative formulation of the eigenvalue problem is summarized in the following theorem. Details can be obtained from [BaMa94b].

THEOREM 3.4. *The roots of the residual polynomials, $\rho_n(\lambda) = 0$, are eigenvalues of the generalized eigenvalue problem*

$$\lambda \begin{bmatrix} I_{n-1} & \vdots & \underline{z}^{(n)} \\ \underline{0}^* & \vdots & \end{bmatrix} \underline{\tilde{w}} = [H_{n,n-1} : \underline{s}_n]\underline{\tilde{w}}, \tag{3.29}$$

where $\underline{s}_n = [\delta_0, 0, ..., 0]_n^T$, and $\underline{z}^{(n)} = [z_0, z_1, ..., z_{n-1}]_n^T = R_n^{-1}\underline{t}_n$. Since $R_n\underline{z}^{(n)} = \underline{t}_n$, and R_n is a nonsingular, upper triangular matrix, it follows that if $|z_{n-1}| \equiv 0$, then $|\tau_{n-1}| \equiv 0$. It follows from (3.6) that $\underline{r}_n = \underline{r}_{n-1}$, and the roots at step n are the same as those at step $n-1$. If $|z_{n-1}|$ is not zero, then the roots of the residual polynomials are eigenvalues of

$$[H_{n,n-1} : \underline{\xi}_n], \tag{3.30}$$

where $\underline{\xi}_n$ is given by

$$\underline{\xi}_n = \begin{pmatrix} \frac{\delta_n}{z_{n-1}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_n - H_{n,n-1} \begin{pmatrix} \frac{z_0}{z_{n-1}} \\ \vdots \\ \frac{z_{n-2}}{z_{n-1}} \end{pmatrix}_{n-1}. \tag{3.31}$$

Proof. See [BaMa94b] □

We note here that by testing $|z_{n-1}|$ it is possible to detect when $\rho_{n-1}(\lambda)$ is close to $\rho_n(\lambda)$. When this is the case, the roots of the residual polynomials will not have changed much from the previous step. We can either use the roots from the previous step as a "close" approximation, or we can compute them via (3.29). When $\rho_{n-1}(\lambda)$ is not close to $\rho_n(\lambda)$, then we solve for the roots by computing the eigenvalues of the upper Hessenberg matrix given in (3.30).

4. Approximating the spectrum of A. In this section we will apply the results from [MaOt92] to form a set that approximates the spectrum of A using information that is available during the QMR iteration. The algorithm that accomplishes this is based on two results that were presented in [MaOt92] as corollaries to Theorems' 3.1 and 3.3. These results are:

$$\text{If} \quad \lambda \quad \text{is a root of an orthogonal polynomial, then } \lambda \in \mathcal{F}_B(A). \tag{4.1}$$

$$\text{If} \quad \lambda \quad \text{is a root of a residual polynomial, then } \frac{1}{\lambda} \in \mathcal{F}_B(A^{-1}), \tag{4.2}$$

where, the B-field of values of A is defined as

$$\mathcal{F}_B(A) = \{\lambda : \lambda = \frac{\langle BA\underline{x}, \underline{x}\rangle}{\langle B\underline{x}, \underline{x}\rangle} \text{ for some } \underline{x} \neq \underline{0} \in \mathcal{C}^n\}. \tag{4.3}$$

This set satisfies the following properties: (cf. [Ho64], [HoJo91])

i) $\mathcal{F}_B(A)$ is a convex set in $\mathcal{C}$.
ii) $\Sigma(A) \subseteq \mathcal{H}(A) \subseteq \mathcal{F}_B(A)$.
iii) $\mathcal{H}(A) \equiv \mathcal{F}_B(A)$ if A is B-normal.

We make the following definition:

DEFINITION: The reciprocal B-field of values of A is the set

$$\mathcal{F}_B^{-1}(A^{-1}) = \{\lambda : \frac{1}{\lambda} \in \mathcal{F}_B(A^{-1})\}. \tag{4.4}$$

Using this definition, and (4.2), it follows that if λ is a root of a residual polynomial, then $\lambda \in \mathcal{F}_B^{-1}(A^{-1})$. Next, notice that if $\lambda_k \in \Sigma(A) \subseteq \mathcal{F}_B(A)$, then $\frac{1}{\lambda_k} \in \Sigma(A^{-1}) \subseteq \mathcal{F}_B(A^{-1})$. From (4.4) we have that $\lambda_k \in \mathcal{F}_B^{-1}(A^{-1})$. Therefore, $\Sigma(A) \subseteq \mathcal{F}_B(A)$ and $\Sigma(A) \subseteq \mathcal{F}_B^{-1}(A^{-1})$ which is expressed as:

$$\Sigma(A) \subseteq \mathcal{G} \equiv \mathcal{F}_B(A) \cap \mathcal{F}_B^{-1}(A^{-1}).$$

Let $\mathcal{H}_j$ and $\hat{\mathcal{H}}_j$ denote the convex polygons formed by the roots of the orthogonal polynomials and the reciprocals of the roots of the residual polynomials respectively, obtained at the j'th step of the QMR algorithm. That is,

$$\begin{aligned} \mathcal{H}_j &\equiv \mathcal{H}(\{\lambda_i^j\}_{i=1}^j) \subseteq \mathcal{F}_B(A), \\ \hat{\mathcal{H}}_j &\equiv \mathcal{H}(\{\frac{1}{\hat{\lambda}_i^j}\}_{i=1}^j) \subseteq \mathcal{F}_B(A^{-1}), \end{aligned}$$

and (4.4) shows that

$$\hat{\mathcal{H}}_j^{-1} \equiv \{z : \frac{1}{z} \in \mathcal{H}(\{\frac{1}{\hat{\lambda}_i^j}\}_{i=1}^j)\} \subseteq \mathcal{F}_B^{-1}(A^{-1}).$$

Observe that $\mathcal{H}_j \cap \hat{\mathcal{H}}_j^{-1} \subseteq \mathcal{G}$, which can be used as an approximation to $\Sigma(A)$.

This results in the following algorithm:

1) Given the linear system, (2.1), and an initial guess, $\underline{x}_0$, begin the QMR iteration.
2) At step j,
i) Compute the roots of the orthogonal polynomials and the residual polynomials.
ii) Form the convex polygons, $\mathcal{H}_j$ and $\hat{\mathcal{H}}_j$.
iii) Form the set $\mathcal{G}_j = \mathcal{H}_j \cap \hat{\mathcal{H}}_j^{-1}$, and use this set as an approximation of $\Sigma(A)$.

5. Numerical results and concluding remarks. Extensive tests were performed on a set of problems and are reported in [BaMa94b]. Here we present two representative examples. We are interested in comparing the approximation of $\Sigma(A)$ obtained using the QMR iteration to that obtained using a full conjugate gradient iteration. Throughout these examples the conjugate gradient (CG) algorithm used is Orthodir with $B = A^*A$. For each example, the same random starting and right-hand side vectors were used for both the CG and QMR iterations.

5.1. Annulus segment. This set of test results is on a matrix whose eigenvalues lie in a segment of an annulus centered at the origin in the complex plane. Figure 5.1 shows the plot of the eigenvalues. $\mathcal{H}(A)$ is the region enclosed by the solid line. $\mathcal{H}^{-1}(A^{-1})$ is the area to the right of the dashed line.

Figures 5.3 and 5.4 show side-by-side comparisons of the roots of the polynomials associated with the CG and QMR iterations for this example. In our experiments, the approximation of $\Sigma(A)$, $\mathcal{G}_j = \mathcal{H}_j \cap \hat{\mathcal{H}}_j^{-1}$, is only guaranteed to be in $\mathcal{F}_B(A) \cap \mathcal{F}_B^{-1}(A^{-1})$ since the matrix considered is not B-normal. However, early in the iteration the roots lie close to the spectrum, some even converging to the outermost members of $\Sigma(A)$. By the 40'th iteration, except for an occassional root associated with the QMR iteration, the roots are all in $\mathcal{H}(A)$, many of which are converging to members of the spectrum. The approximation of $\Sigma(A)$ obtained via the QMR iteration is comparable to that obtained using a full CG iteration. In fact, many of the roots obtained with QMR are qualitatively close to those obtained from the CG iteration. In [BaMa94b] we obseved that annulus matrices whose eigenvalues wrap further around the origin, into the left half plane, presented a greater challenge for both the CG and QMR iterations. The roots obtained via the CG and QMR iterations were still comparable, however the approximating set, $\mathcal{G}_j$, grew at a slower rate for both iterations. The result is that more steps are needed to obtain a good approximation to the spectrum. Here, we note that the convergence of the corresponding iterative methods also took longer. In general, indefinite problems are hard problems to solve for any polynomial iterative method.

5.2. Elliptic partial differential equation. The matrix in this test arises from central difference approximations to the partial differential operator

$$\mathcal{L}u = -\Delta u + \beta u_x + \gamma u$$

on a square domain $[0,1] \times [0,1]$ with Dirichlet boundary conditions on a 10×10 mesh. The graph in Figure 5.2 shows the spectrum along with the sets $\mathcal{H}(A)$ and $\mathcal{H}^{-1}(A^{-1})$. The values, $\beta = 22.5$ and $\gamma = -410$ were used in this example. Here, $\mathcal{H}^{-1}(A^{-1})$ is the area to the outside of the dashed line, thus we see that $\mathcal{H}(A) \cap \mathcal{H}^{-1}(A^{-1})$ is the union of two disjoint sets.

Comparisons of the roots of the polynomials associated with the QMR and CG iterations are given in Figures 5.5 and 5.6. This indefinite example is a more difficult problem for both the CG and QMR iterations. Although the approximating set $\mathcal{G}_j$ for CG, does not contain any of the spectrum until about the 36'th iteration, the roots of the polynomials encircle the spectrum. Notice that the set $\mathcal{G}_j$ grows very slowly. The figures also show that this situation gets worse using the QMR iteration. The roots still surround the spectrum in the early iterations, but the roots of the residual polynomials keep more distance from the spectrum than those obtained using CG.

Fig. 5.1. *Annulus segment*

Fig. 5.2. *PDE matrix*

ROOTS OF THE POLYNOMIALS ASSOCIATED WITH THE CG AND QMR ITERATIONS
ANNULUS MATRIX - Fig. 5.1
$*$'s denote roots of the B-orthogonal polynomials, their hull, a solid line,
$+$'s denote roots of the residual polynomials, their reciprocal hull, a dashed line.
o's denote the eigenvalues of the matrix.

FIG. 5.3. *Iteration 20, CG on left, QMR on right.*

FIG. 5.4. *Iteration 40, CG on left, QMR on right.*

ROOTS OF THE POLYNOMIALS ASSOCIATED WITH THE CG AND QMR ITERATIONS
PDE MATRIX - Fig. 5.2
$*$'s denote roots of the B-orthogonal polynomials, their hull, a solid line,
$+$'s denote roots of the residual polynomials, their reciprocal hull, a dashed line.
o's denote the eigenvalues of the matrix.

FIG. 5.5. *Iteration 36, CG on left, QMR on right*

FIG. 5.6. *Iteration 81, CG on left, QMR on right*

5.3. Concluding remarks. In this paper we have shown that the QMR method can be viewed as a variable metric conjugate gradient method, or a balanced projection method, [JoMa90]. At each step, the QMR method minimizes the B-norm of the error, where B is a Hermitian positive definite matrix that depends on the initial residual. From this, we have demonstrated that there are B-orthogonal polynomials, and residual polynomials associated with the QMR iteration. Since minimizing $||\underline{e}_n||_B$ is equivalent to minimizing $||\underline{r}_n||_{\hat{B}}$, the residual polynomials are kernel polynomials in the $\hat{B}$ norm. Methods were described for finding the roots of these polynomials. Earlier work proved that the roots of the B-orthogonal polynomials are contained in $\mathcal{F}_B(A)$, and the reciprocals of the roots of the residual polynomials are contained in $\mathcal{F}_B(A^{-1})$. Thus, the same techniques for approximating the spectrum that were used in [MaOt92] with the conjugate gradient iteration can be utilized with the QMR iteration. Because of the short recurrences for the QMR iterates, the work and storage requirements needed to obtain this approximation via the QMR iteration are much less than that required using a full CG iteration.

In the numerical examples, we have compared the roots of the polynomials and the resulting approximations of the spectrum obtained from the QMR and the CG iterations. We showed an example of an annulus matrix, where the approximations realized using the QMR iteration were comparable to those obtained using a full CG iteration. In some cases, we see only small differences in the plots of the roots obtained by the two methods. However, as with CG, these experiments also show that there are examples where a large number of steps are required in order to realize a good approximation via the techniques described in this paper. We found this to be the case with many of the indefinite problems we tested, such as the PDE example. Here, we noted that while both the CG and QMR iterations struggled, QMR was much worse than CG. The examples which yielded good approximation results also demonstrated rapid convergence of the corresponding iterative method. Our test results also illustrate that better approximations of the spectrum are obtained by the inclusion/exclusion regions given by the set, $\mathcal{G}_j$, than can be obtained from the roots alone. For more complete numerical results, see [BaMa94b].

In [BaMa94a] we showed that bounds on the relative error for a conjugate gradient iteration as well as any variable metric CG method are given by

$$\frac{||\underline{e}_n||}{||\underline{e}_0||} \leq \mathcal{C}(B)^{1/2} \min_{p_n \in \mathcal{P}_n^\circ} \frac{||p_n(A)\underline{e}_0||}{||\underline{e}_0||},$$

where, $\mathcal{C}(B)^{1/2}$ is the square root of the condition number of B. For the CG iteration in this paper, $B = A^*A$, while for QMR, B depends upon $\underline{r}_0$. The problem is that for QMR, B can be arbitrarily bad. In near breakdown, $\mathcal{C}(B)$ becomes unbounded, and the sets, $\mathcal{F}_B(A)$, and $\mathcal{F}_B(A^{-1})$ get arbitrarily large, thus we loose stability bounds on the roots.

For matrices that are not B-normal, theory only guarantees that the roots of the B-orthogonal polynomials will be contained in $\mathcal{F}_B(A)$, and the roots of the residual polynomials will lie in $\mathcal{F}_B^{-1}(A^{-1})$. From the examples, we noticed that even in the early iterations, the roots of these polynomials lie relatively close to the spectrum of the matrix. In [BaMa94b] it was demonstrated that there can be a big discrepancy in the size of the sets, $\mathcal{F}_B(A)$ and $\mathcal{F}_B^{-1}(A^{-1})$, between QMR and CG. However, the larger size of these sets did not necessarily imply that the roots would be more widespread. These examples suggest that we may be able to confine the roots of these polynomials to regions containing the spectrum that may be much smaller than the sets, $\mathcal{F}_B(A)$ and $\mathcal{F}_B^{-1}(A^{-1})$, even in the early iterations. Perhaps the roots contain more information about the spectrum than what our current theory has revealed. We hope that further research into these ideas will lead to more accurate, and efficient ways of obtaining approximations of the spectrum.

REFERENCES

[AMS90] S. F. Ashby, T. A. Manteuffel and P. E. Saylor, *A Taxonomy for Conjugate Gradient Methods*, SIAM J. Numer. Anal., 26 (1990), pp. 1542-1568.

[BaMa94a] T. Barth and T. A. Manteuffel, *Variable Metric Conjugate Gradient Methods*, Procedings for the conference in Matrix Analysis and Parallel Computing, Advances in Numerical Methods for Large Sparse Sets of Linear Equations, No. 10, March 1994, Keio University, Japan.

[BaMa94b] T. Barth and T. A. Manteuffel, *Estimating the Spectrum of A Using the Roots of the Polynomials Associated with the QMR Iteration*, Los Alamos National Laboratory Report, Oct (1994).

[FaMa84] V. Faber and T. A. Manteuffel, *Necessary and Sufficient Conditions for the Existence of a Conjugate Gradient Method*, SIAM J. Numer. Analysis, Vol. 21, No. 2, pp. 352–362 (1984).

[Fr91] R. W. Freund, *Quasi-kernel polynomials and their use in non-Hermitian matrix iterations*, J. Comput. Appl. Math. 43, pp. 135-158 (1992).

[FrGuNa91] R. W. Freund, M. H. Gutknecht, and N. M. Nachtigal, *An Implementation of the Look-Ahead Lanczos Algorithm for Non-Hermitian Matrices*, SIAM J. Sci. Comput., Vol. 14, No. 1, pp. 137-158, January 1993.

[FrNa91] R. W. Freund and N. M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numerische Mathematik 60 (1991), pp. 315-339.

[HoJo91] R. A. Horn and C. R. Johnson, " Topics in Matrix Analysis,", Cambridge University Press, Cambridge, 1991.

[Ho64] A. S. Householder, " The Theory of Matrices in Numerical Analysis," pp. 37–57, Blaisdell Press, New York, 1964.

[JoMa90] W. D. Joubert, and T. A. Manteuffel, *Iterative Methods for Nonsymmetric Linear Systems.*, "Iterative Methods for Large Linear Systems,", Academic Press, San Diego, 1990, pp. 149-171.

[MaOt92] T. A. Manteuffel, and J. S. Otto, *On The Roots of the Orthogonal Polynomials and Residual Polynomials Associated with a Conjugate Gradient Method*, to appear, Journal of Numerical Linear Algebra, 1993.