SIAM J. MATRIX ANAL. APPL.
Vol. 17, No. 2, pp. 223–247, April 1996

RELATIONS BETWEEN GALERKIN AND NORM-MINIMIZING ITERATIVE METHODS FOR SOLVING LINEAR SYSTEMS*

JANE CULLUM† AND ANNE GREENBAUM‡

Abstract. Several iterative methods for solving linear systems $Ax = b$ first construct a basis for a Krylov subspace and then use the basis vectors, together with the Hessenberg (or tridiagonal) matrix generated during that construction, to obtain an approximate solution to the linear system. To determine the approximate solution, it is necessary to solve either a linear system with the Hessenberg matrix as coefficient matrix or an extended Hessenberg least squares problem. In the first case, referred to as a *Galerkin* method, the residual is orthogonal to the Krylov subspace, whereas in the second case, referred to as a *norm-minimizing* method, the residual (or a related quantity) is minimized over the Krylov subspace. Examples of such pairs include the full orthogonalization method (FOM) (Arnoldi) and generalized minimal residual (GMRES) algorithms, the biconjugate gradient (BCG) and quasi-minimal residual (QMR) algorithms, and their symmetric equivalents, the Lanczos and minimal residual (MINRES) algorithms. A relationship between the solution of the linear system and that of the least squares problem is used to relate the residual norms in Galerkin processes to the norms of the quantities minimized in the corresponding norm-minimizing processes. It is shown that when the norm-minimizing process is converging rapidly, the residual norms in the corresponding Galerkin process exhibit similar behavior, whereas when the norm-minimizing process is converging very slowly, the residual norms in the corresponding Galerkin process are significantly larger. This is a generalization of the relationship established between Arnoldi and GMRES residual norms in P. N. Brown, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Statist. Comput., 12, 1991, pp. 58–78. For MINRES and Lanczos, and for two nonsymmetric bidiagonalization procedures, we extend the arguments to incorporate the effects of finite precision arithmetic.

Key words. GMRES, Arnoldi, biconjugate gradients, QMR, iterative methods

AMS subject classifications. 65F10, 65F15

1. Introduction.

The Arnoldi algorithm [1] (also known as the full orthogonalization method or FOM [22]) and the generalized minimal residual (GMRES) algorithm [23] are two recently developed Krylov methods for solving nonsymmetric linear systems

$$Ax = b. \tag{1}$$

In Brown [3], a theoretical comparison of the two methods is presented. Brown exhibits connections between the singularity of the Hessenberg matrices generated in the Arnoldi algorithm and the stagnation of the corresponding iterates in the GMRES algorithm. From this he infers a relationship between the stagnation (the plateaus) observed in GMRES and near-singularity of these Hessenberg matrices. He also obtains relationships between the norms of the residuals generated by the Arnoldi algorithm and the norms of the residuals generated by the GMRES algorithm which, when combined with a relationship in [23], can be used to infer that if the iterates in both methods are well defined, then if one of the methods performs very well on a

* Received by the editors April 2, 1993; accepted for publication (in revised form) by R. Freund April 11, 1995.

† IBM T. J. Watson Research Center, Yorktown Heights, NY 10598. Part of this research was supported by National Science Foundation grant GER-9450081 while this author was visiting the Department of Computer Science, University of Maryland, College Park, MD.

‡ Courant Institute of Mathematical Sciences, 251 Mercer Street, New York, NY 10012 (greenbau@nyu.edu). Part of this research was performed while this author was visiting IBM T. J. Watson Research Center and was supported by National Science Foundation grant 25968-5375 and DOE contract DEFG0288ER25053.

particular problem the other one will also, and if one method performs very poorly, then so will the other one.

In this paper we obtain a slightly different formulation of these GMRES/Arnoldi residual norm relationships from which it is easier to see this interdependence. This formulation explains the general correspondence between peaks in the plot of residual norms in the Arnoldi algorithm and plateaus in the GMRES algorithm, as observed in [4] for a different pair of iterative methods. Moreover, we demonstrate, by numerical example, that plateaus (stagnation) need not be associated with near-singularity of the Hessenberg matrices. In fact the proof given in [3] connecting stagnation of the GMRES iterates to singularity of the Hessenberg matrices is not applicable to the nearly singular case.

The biconjugate gradient (BCG) algorithm [7] and the quasi-minimal residual (QMR) algorithm [8] are another pair of Krylov methods for solving linear systems. In [8], relationships are established between these two algorithms that are very similar to the relationships obtained in [3] for the Arnoldi and GMRES residuals. In this case, however, the appropriate quantities to compare are the BCG residuals and what we refer to as the QMR *quasi-residuals*, the vectors whose norms are actually minimized at each step of the QMR algorithm. From these relationships one can infer that if the norm of the QMR quasi-residual is greatly reduced at a step, then the norm of the BCG residual at that step is approximately equal to the norm of the QMR quasi-residual, whereas if the QMR quasi-residual norm remains nearly constant, the BCG residual norm is significantly greater than the QMR quasi-residual norm. The norms of the actual QMR residuals may be somewhat larger than those of the quasi-residuals (by as much as a factor $\sqrt{k+1}$ at step k [8]) or they may be somewhat smaller, but it is, in some sense, a happy accident if the actual residual norms turn out to be much smaller than the quasi-residual norms. No attempt is made to produce the sort of cancellation that is needed to make this happen and in practice it seldom occurs. Thus the relationship between BCG residuals and QMR quasi-residuals can be thought of as an approximate relationship between BCG residuals and QMR residuals. Roughly speaking, then, for a given problem these algorithms will either both perform well or both perform poorly.

The above statements assume exact arithmetic. For *real symmetric* problems, similar relationships are shown to hold in finite precision arithmetic as well. For real symmetric problems, and in exact arithmetic, both of these pairs of algorithms reduce to the MINRES [20] and Lanczos [17] algorithms. In this case, the Lanczos vectors are orthogonal and the quasi-residual norms are the same as the residual norms, so the relationship between BCG residual norms and QMR quasi-residual norms becomes a relationship between Lanczos residual norms and MINRES residual norms. In finite precision computations, the computed Lanczos vectors may be far from orthogonal, yet numerical experiments reported in [4] suggest that for certain types of problems these relationships hold to a close approximation, even after orthogonality of the Lanczos vectors is lost. The report [4] describes a series of numerical experiments using a pair of bidiagonalization procedures, denoted by SQMR and BLanczos, for solving nonsymmetric systems of equations. These procedures replace the original nonsymmetric problem by a larger symmetric problem, and then use MINRES and Lanczos on the associated symmetric iteration matrices.

We prove that for certain classes of real symmetric problems, the relationships between the residuals demonstrated for GMRES/Arnoldi in exact arithmetic hold approximately for the MINRES and Lanczos residuals even in finite precision. We then

prove that if a nonsymmetric problem (1) is well conditioned, then the real symmetric problems generated in the nonsymmetric bidiagonalization algorithm considered in [4] belong to this class. We therefore obtain the result that the residual relationship, as demonstrated numerically in [4], is theoretically valid in finite precision arithmetic for that pair of bidiagonalization algorithms. These proofs rely on the fact that the tridiagonal matrices generated by a finite precision Lanczos computation are the same as those that would be generated by the exact Lanczos algorithm applied to a certain larger matrix with nearby eigenvalues, and that components of the computed Lanczos vectors can be related to corresponding groups of components in the exact Lanczos vectors associated with this larger matrix [12].

The exact arithmetic results obtained in this paper are easy consequences of relations between linear system solutions and extended least squares solutions that have been established in a number of places. See, for example, [3, 8, 9, 15, 20, 28]. It is indeed surprising that the precise relation between Galerkin and norm-minimizing methods was not recognized and explicitly stated much earlier!

In §2, the theorem relating the solution of a linear system and an extended least squares problem is established. In §3 the Arnoldi and GMRES algorithms are described and the theorem of §2 is used to establish the relationship between Arnoldi and GMRES residual norms. In §4 the BCG and QMR algorithms are described and the analogous relationship between the BCG residual norms and the QMR quasi-residual norms is derived. In both sections, numerical examples are given to demonstrate the relationships.

In §5 we first define the real symmetric Lanczos and MINRES algorithms and then define the two nonsymmetric bidiagonalization algorithms considered in [4]. In §6 we focus on the real symmetric case in finite precision arithmetic and establish the analogous relationship between the norms of computed Lanczos and MINRES residuals for a class of real symmetric problems. We then prove in §7 that if a nonsymmetric problem (1) is well conditioned, then the real symmetric problems generated by the two bidiagonalization algorithms described in §5 belong to this class, so that in fact the residual relationship holds for these nonsymmetric bidiagonalization algorithms in finite precision arithmetic.

2. Relation between linear systems and least squares. Let H_k, $k = 1, 2, \ldots$, denote a family of upper Hessenberg matrices, where H_k is k by k and H_{k-1} is the $k-1$ by $k-1$ principal submatrix of H_k. For each k, define the $k+1$ by k matrix $H_k^{(e)}$ by

$$H_k^{(e)} = \begin{pmatrix} H_k \\ h_{k+1,k} e_k^T \end{pmatrix},$$

where $e_k^T = (0, \ldots, 0, 1)$.

The matrix $H_k^{(e)}$ can be factored in the form $Q_k^* R_k^{(e)}$, where Q_k is a $k+1$ by $k+1$ unitary matrix and $R_k^{(e)}$ is a $k+1$ by k matrix whose top k by k block, denoted R_k, is upper triangular and whose last row consists of zeros. This factorization can be performed using plane rotations:

$$(F_k \cdots F_1) H_k^{(e)} = R_k^{(e)}, \quad \text{where } F_i = \begin{pmatrix} I_{i-1} & & & \\ & c_i & -s_i & \\ & s_i & c_i & \\ & & & I_{k-i} \end{pmatrix}.$$

Note that the first $k-1$ sines and cosines, s_i, c_i, $i = 1, \ldots, k-1$, in the Givens rotations used to factor $H_k^{(e)}$ are the same as those used to factor $H_{k-1}^{(e)}$.

Let $\beta > 0$ and let e_1 denote the first unit vector, either a k-vector or a $(k+1)$-vector, depending on the context. Assume that H_k is nonsingular, and let $\check{y}_k$ denote the solution of the linear system $H_k y = \beta e_1$. Let y_k denote the solution of the least squares problem $\min_y \|H_k^{(e)} y - \beta e_1\|$. Finally, let

$$\check{\nu}_k = H_k^{(e)} \check{y}_k - \beta e_1, \qquad \nu_k = H_k^{(e)} y_k - \beta e_1.$$

The following result is established in a slightly different form in [9] and is implicit in a number of other works [3, 8, 15, 20, 28].

THEOREM 1. *Using the above notation, the norms of ν_k and $\check{\nu}_k$ are related to the sines and cosines of the Givens rotations by*

$$\|\nu_k\| = \beta |s_1 s_2 \cdots s_k| \quad \textit{and} \quad \|\check{\nu}_k\| = \beta \frac{1}{|c_k|} |s_1 s_2 \cdots s_k|. \tag{2}$$

It follows that

$$\|\check{\nu}_k\| = \frac{\|\nu_k\|}{\sqrt{1 - (\|\nu_k\| / \|\nu_{k-1}\|)^2}}, \tag{3}$$

or, equivalently,

$$\left(\frac{\|\nu_k\|}{\|\check{\nu}_k\|}\right)^2 + \left(\frac{\|\nu_k\|}{\|\nu_{k-1}\|}\right)^2 = 1. \tag{4}$$

Proof. Let $Q_k = F_k \cdots F_1$ be the $k+1$ by $k+1$ unitary matrix reducing $H_k^{(e)}$ to $R_k^{(e)}$. The least squares problem can be written in the form

$$\min_y \|H_k^{(e)} y - \beta e_1\| = \min_y \|Q_k (H_k^{(e)} y - \beta e_1)\| = \min_y \|R_k^{(e)} y - \beta Q_k e_1\|.$$

The solution y_k is determined by solving the upper triangular linear system with coefficient matrix R_k and right-hand side equal to the first k entries of $\beta Q_k e_1$. The remainder $R_k^{(e)} y_k - \beta Q_k e_1$ is therefore zero except for the last entry, which is just the last entry of $-\beta Q_k e_1 = -\beta (F_k \cdots F_1) e_1$, which is easily seen to be $-\beta s_1 \cdots s_k$. This establishes the first equality in (2).

For the linear system solution $\check{y}_k = H_k^{-1} \beta e_1$, we have

$$\check{\nu}_k = H_k^{(e)} H_k^{-1} \beta e_1 - \beta e_1,$$

which is zero except for the last entry, which is $\beta h_{k+1,k}$ times the $(k,1)$ entry of H_k^{-1}. Now H_k can be factored in the form $\tilde{Q}_k^* \tilde{R}_k$, where $\tilde{Q}_k = \tilde{F}_{k-1} \cdots \tilde{F}_1$, and $\tilde{F}_i$ is the k by k principal submatrix of F_i. The matrix $H_k^{(e)}$, after applying the first $k-1$ plane rotations, has the form

$$(F_{k-1}\cdots F_1)\,H_k^{(e)} = \begin{pmatrix} x & x & \dots & x \\ & x & \dots & x \\ & & \ddots & \vdots \\ & & & r \\ & & & h \end{pmatrix},$$

where r is the (k,k) entry of the upper triangular matrix $\tilde{R}_k$ and $h = h_{k+1,k}$. The kth rotation is chosen to annihilate the nonzero entry in the last row:

$$c_k = \frac{r}{\sqrt{r^2+h^2}}, \quad s_k = -\frac{h}{\sqrt{r^2+h^2}}.$$

Note that r and hence c_k is nonzero since H_k is nonsingular. Now we have $H_k^{-1} = \tilde{R}_k^{-1}\tilde{Q}_k$, and the $(k,1)$ entry of this is $1/r$ times the $(k,1)$ entry of $\tilde{Q}_k = \tilde{F}_{k-1}\cdots\tilde{F}_1$, and this is just $s_1\cdots s_{k-1}$. It follows that the nonzero entry of $\check{\nu}_k$ is $\beta(h_{k+1,k}/r)s_1\cdots s_{k-1}$. Finally, using the fact that $|s_k/c_k| = |h/r| = |h_{k+1,k}/r|$, we obtain the second equality in (2).

From (2), it is clear that

$$\frac{\|\nu_k\|}{\|\nu_{k-1}\|} = |s_k|, \quad \frac{\|\nu_k\|}{\|\check{\nu}_k\|} = |c_k|.$$

The results (3) and (4) follow from the fact that $|c_k|^2 + |s_k|^2 = 1$. □

3. The Arnoldi and GMRES algorithms. Consider a system of linear equations $Ax = b$, where A is an N by N nonsingular matrix and b is a given N-vector. For ease of notation we will assume that the matrix A and the vectors involved in the solution algorithms are real, but our results here and in other sections are easily modified for complex matrices. Given an initial guess x_0 for the solution, the Arnoldi and GMRES algorithms construct approximate solutions x_k, $k = 1, 2, \dots$, of the form

$$x_k = x_0 + t_k, \quad t_k \in K_k(A, r_0) \equiv \operatorname{span}\{r_0, Ar_0, \dots, A^{k-1}r_0\}, \tag{5}$$

where $r_0 \equiv b - Ax_0$ is the initial residual and $K_k(A, r_0)$ is referred to as the kth Krylov space. The residual $r_k \equiv b - Ax_k$ is given by

$$r_k = r_0 - At_k.$$

The two methods differ in how the approximate solutions are chosen from the space (5). For the Arnoldi method (see [23]), the kth residual vector, denoted r_k^A, satisfies

$$r_k^A \perp K_k(A, r^0), \tag{6}$$

while the kth GMRES residual vector, denoted r_k^G, satisfies

$$r_k^G \perp AK_k(A, r_0). \tag{7}$$

From (7) it follows that of all vectors x_k of the form (5), the GMRES approximation x_k^G has the residual of smallest Euclidean norm:

$$\|r_k^G\| = \min_{x_k \in\, x_0 + K_k(A,r_0)} \|b - Ax_k\|. \tag{8}$$

The properties (5) and (6) completely characterize the Arnoldi iterates, while the properties (5) and either (7) or (8) completely define the GMRES iterates.

In order to generate these approximate solutions, both algorithms construct an orthonormal basis for the Krylov space $K_k(A, r_0)$. This can be accomplished by using the modified Gram–Schmidt procedure, for example:

Modified Gram–Schmidt procedure:

1. Compute $r_0 = b - Ax_0$ and set $u_1 = r_0/\|r_0\|$.
2. For $j = 1, \ldots, k$ do:
   $u_{j+1} := Au_j$
   for $i = 1, \ldots, j$ do:

$$h_{ij} := u_i^T u_{j+1}, \quad u_{j+1} := u_{j+1} - h_{ij} u_i$$

$$h_{j+1,j} := \|u_{j+1}\|, \quad u_{j+1} := u_{j+1}/h_{j+1,j}.$$

Other methods have also been proposed for computing these basis vectors [27], but we will not be concerned with the particular implementation used. Note that it is necessary to save all of the basis vectors and at each step to orthogonalize the new vector against each of the previous basis vectors.

Let U_k denote the N by k matrix whose columns are the orthonormal basis vectors $u_1, \ldots, u_k$, and let H_k denote the k by k upper Hessenberg matrix whose nonzero entries are the scalars h_{ij}. The above recurrence can be written in matrix form as

$$AU_k = U_k H_k + h_{k+1,k} u_{k+1} e_k^T, \tag{9}$$

where u_{k+1} is the $(k+1)$st normalized basis vector and e_k is the kth unit k-vector $(0, \ldots, 0, 1)^T$. If H_k is nonsingular, then it can be seen from expression (9) and the definition (5), (6) of the Arnoldi iterates that the kth Arnoldi iterate x_k^A is of the form

$$x_k^A = x_0 + U_k y_k^A, \tag{10}$$

where y_k^A satisfies

$$H_k y_k^A = \beta e_1, \quad \beta = \|r_0\|. \tag{11}$$

If H_k is singular, then it is shown in [3] that the kth Arnoldi iterate does not exist and, moreover, that the kth GMRES iterate does not improve.

Defining the $k+1$ by k matrix $H_k^{(e)}$ to be

$$H_k^{(e)} = \begin{pmatrix} H_k \\ 0 \cdots 0 \; h_{k+1,k} \end{pmatrix},$$

equation (9) can be written in the form

$$AU_k = U_{k+1} H_k^{(e)}. \tag{12}$$

Using this equation and the characterization (5), (8) of the GMRES iterates, it is shown in [23] that the kth GMRES iterate x_k^G is of the form

$$x_k^G = x_0 + U_k y_k^G, \tag{13}$$

where y_k^G satisfies the least squares problem

$$\|\beta e_1 - H_k^{(e)} y_k^G\| = \min_y \|\beta e_1 - H_k^{(e)} y\|. \tag{14}$$

The following theorem is an immediate consequence of Theorem 1. Let s_i and c_i, $i = 1, \ldots, k$, be the sines and cosines of the Givens rotations used to factor $H_k^{(e)}$. Relations between GMRES and Arnoldi residuals and the sines and cosines of Givens rotations were established in [23] and [3], but the direct relation between GMRES and Arnoldi residuals was never stated explicitly.

THEOREM 2. *In exact arithmetic, if $c_k \neq 0$ at iteration k, then the Arnoldi and GMRES residuals are related by*

$$\|r_k^A\| = \frac{\|r_k^G\|}{\sqrt{1 - (\|r_k^G\|/\|r_{k-1}^G\|)^2}}. \tag{15}$$

Proof. From (10), (13), and (12), it follows that the Arnoldi and GMRES residuals can each be written in the form

$$\begin{aligned} r_k^{A,G} &= r_0 - AU_k y_k^{A,G} \\ &= r_0 - U_{k+1} H_k^{(e)} y_k^{A,G} \\ &= U_{k+1}(\beta e_1 - H_k^{(e)} y_k^{A,G}). \end{aligned} \tag{16}$$

Since the columns of U_{k+1} are orthonormal, it follows that

$$\|r_k^{A,G}\| = \|\beta e_1 - H_k^{(e)} y_k^{A,G}\|, \tag{17}$$

and the desired relation (15) now follows from Theorem 1 and the definitions (11) and (14) of y_k^A and y_k^G. □

Note that for the Arnoldi method, relation (17) follows from (16), even if the columns of U_{k+1} are not orthonormal. It requires only that $\|u_{k+1}\| = 1$, since the quantity $\beta e_1 - H_k^{(e)} y_k^A$ has only the last component nonzero.

Theorem 2 shows that if the GMRES residual norm is reduced by a significant factor at step k, then the Arnoldi residual norm will be approximately equal to the GMRES residual norm at step k since the denominator in the right-hand side of (15) will be close to 1. If the GMRES residual norm remains almost constant, however, then the denominator in the right-hand side of (15) is close to 0 and the Arnoldi residual norm will be much larger. Table 1 shows the relation between the GMRES residual norm reduction and the ratio of Arnoldi to GMRES residual norm. Note that the GMRES residual norm must be *very* flat before the Arnoldi residual norm is orders of magnitude larger than the GMRES residual norm.

To illustrate these results, we consider a real matrix A of the following form:

$$A = \Sigma V^T \Lambda V \Sigma^{-1}, \tag{18}$$

where Σ is a diagonal matrix with positive entries, V is an orthogonal matrix, and Λ is a real block diagonal matrix, consisting of at most two by two blocks, each corresponding to a complex conjugate pair of eigenvalues of A. We note that since any of the 2×2 blocks in Λ can be diagonalized by a 2×2 unitary transformation, Σ

TABLE 1
Relation between GMRES residual norm reduction and ratio of Arnoldi to GMRES residual norm.

$\|r_k^G\|/\|r_{k-1}^G\|$	$\|r_k^A\|/\|r_k^G\|$
.5	1.2
.9	2.3
.99	7.1
.9999	70.7
.999999	707

specifies the singular values of an eigenvector matrix of A. Every real diagonalizable matrix B is unitarily similar to a matrix of the form (18), since if $B = X\Lambda X^{-1}$ for some real matrix X and $X = U\Sigma V^T$ is a singular value decomposition of X, then we have $B = UAU^T$. Since the iterative methods we consider are invariant under unitary similarity transformations—the residual norms at each step of the algorithm for solving $Ax = b$ are the same as those at each step of the algorithm for solving $UAU^T y = Ub$, $x = U^T y$—it follows that all possible residual norm plots corresponding to diagonalizable real matrices can be obtained by considering matrices of the form (18).

For our example, we set $N = 111$. We chose Σ to have one *small* singular value (.8), two *large* singular values (10 and 10.3), and the remaining singular values ranging from 2.6 upward with a uniform spacing of .02 between successive singular values. The matrix Λ was defined by specifying three randomly generated complex eigenvalues of magnitudes .02, .1, and 10 and a real eigenvalue of magnitude 1, generating the remainder of the spectrum randomly as complex numbers in the box $1 \leq x \leq 3$, $2 \leq y \leq 4$, and then defining a 2×2 real block in Λ for each corresponding complex conjugate pair of eigenvalues. The V matrix was set equal to the permutation matrix which for $1 \leq j \leq N-1$ maps each coordinate vector e_j into e_{j+1} and maps e_N into e_1. The solution was set equal to the vector whose components are all 1, and the initial guess was the zero vector. The convergence tolerance was 10^{-13}, as measured by the ratio of the norm of the residual at iteration k to the norm of the initial residual.

Figure 1 shows a plot of the logarithms of the Arnoldi and GMRES residual norms versus iteration number for this example problem. The solid line is the GMRES convergence curve and the dashed line is the Arnoldi curve. The norm of the starting residual was 92.7.

Observe that for the specified convergence tolerance these algorithms converged simultaneously in 95 iterations. From iterations 63 to 95, the GMRES convergence is basically fast and, as the picture indicates, on that portion of the curve and in fact on similar steep portions of the GMRES curve, the Arnoldi norms converge in a similar fashion.

Also observe the matching of the peaks in the Arnoldi residual norm plot with the plateaus in the GMRES residual norm plot. The double peak corresponding approximately to iterations 22 to 36 coincides with the rough recognition of the members of the conjugate pair of size 10^{-1} as eigenvalues in the spectra of the associated Hessenberg matrices in (9). The second double peak from approximately iterations 42 to 62 corresponds to the identification of the members of the conjugate pair of size 2×10^{-2}. In test problems with smaller eigenvalues, these double peaks are more clearly visible and may be either overlapping or split apart. We note that in [25, 26] connections between the appearance of certain eigenvalues in the spectra of the GMRES/Arnoldi Hessenberg matrices and subsequent speedups in the convergence of GMRES were

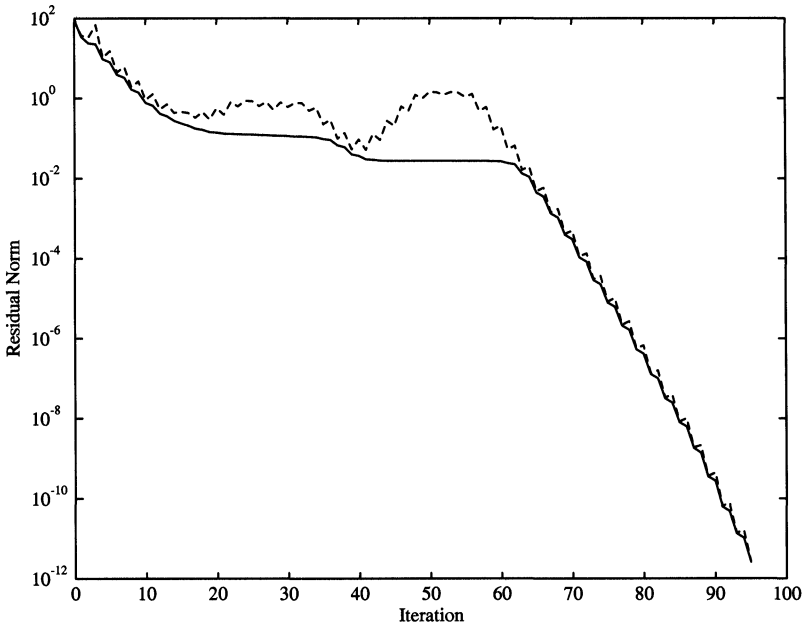

FIG. 1. *GMRES (solid) and Arnoldi (dashed) residual norms.*

observed.

On the scale of Fig. 1, it is difficult to see the precise correlation between the degree of flatness of the GMRES curve and the height of the Arnoldi curve above the GMRES curve. However, for example, at iterations 24 and 50 within the double peaks in the residual norm curve for the Arnoldi iterations, the norm of the GMRES residual was reduced, respectively, by factors of .989529 and .999825, and the corresponding ratios of the Arnoldi to GMRES residual norms were approximately 6.93 and 53.4. These values are as predicted by Theorem 2.

In this example, across the first double peak/plateau the condition numbers of the corresponding Hessenberg matrices vary from 261 to 2351 and are less than the condition number 4625.8 of the original iteration matrix A. Across the second double peak/plateau, however, this is not the case. Over iterations 42 to 62 the condition numbers of the Hessenberg matrices range from 965 to 106,146. After iteration 62, these condition numbers settle down to the condition number of A.

4. The BCG and QMR algorithms. The BCG [7] and QMR [8] algorithms also construct approximate solutions $x_k, k = 1, 2, \ldots$, of the form (5). They differ from the Arnoldi and GMRES algorithms in that, instead of constructing an orthonormal basis for the Krylov space $K_k(A, r_0)$, they construct sequences of biorthogonal vectors spanning the spaces $K_k(A, r_0)$ and $K_k(A^T, \hat{r}_0)$, where $\hat{r}_0$ is an arbitrary vector, often chosen equal to r_0. This can be accomplished using the nonsymmetric Lanczos algorithm and requires two simple three-term recurrences:

Nonsymmetric Lanczos algorithm:

1. Set $v_1 = r_0/\|r_0\|$ and $w_1 = \hat{r}_0/\|\hat{r}_0\|$. Set $\rho_1 = 1$ and $\xi_1 = 1$ and $v_0 = w_0 = 0$.

2. For $j = 1, \ldots, k$ do

$$\alpha_j = w_j^T A v_j / w_j^T v_j,$$
$$\beta_1 = 0, \quad \beta_j = \xi_j w_j^T v_j / w_{j-1}^T v_{j-1}, \text{ if } j > 1,$$
$$\nu_{j+1} = A v_j - \alpha_j v_j - \beta_j v_{j-1},$$
$$\rho_{j+1} = \|\nu_{j+1}\|, \quad v_{j+1} = \nu_{j+1}/\rho_{j+1},$$
$$\omega_{j+1} = A^T w_j - \alpha_j w_j - (\beta_j \rho_j / \xi_j) w_{j-1},$$
$$\xi_{j+1} = \|\omega_{j+1}\|, \quad w_{j+1} = \omega_{j+1}/\xi_{j+1}.$$

Here we have used the nonsymmetric Lanczos formulation that scales by setting the norms of the Lanczos vectors to unity. Note that each step of the nonsymmetric Lanczos algorithm requires matrix vector multiplications by A and A^T but does not require saving and orthogonalizing against all previous basis vectors, as is required by the Arnoldi/GMRES methods.

Unfortunately, the nonsymmetric Lanczos algorithm can break down. If $w_j^T v_j = 0$ for some j, the coefficients in the above algorithm are undefined. If $w_j = 0$ or $v_j = 0$, then this means an invariant subspace for either A^T or A has been found, but if $w_j^T v_j = 0$ when neither w_j nor v_j is zero, then this is referred to as a *serious* breakdown. While an exact breakdown is unlikely, near breakdowns can cause numerical instabilities. To avoid such problems, various look-ahead strategies have been proposed; see, e.g., [2, 8, 10, 21]. The relationship we establish between BCG and QMR residuals holds in exact arithmetic provided the same look-ahead steps are used in the underlying Lanczos recurrence for each algorithm.

Let V_k denote the N by k matrix whose columns are the basis vectors $v_1, \ldots, v_k$ generated for the space $K_k(A, r_0)$ and let W_k denote the N by k matrix whose columns are the basis vectors $w_1, \ldots, w_k$ generated for the space $K_k(A^T, \hat{r}_0)$. Let the tridiagonal matrix T_k be defined by

$$T_k = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \rho_2 & \alpha_2 & \beta_3 & & \\ & \rho_3 & \ddots & \ddots & \\ & & \ddots & \alpha_{k-1} & \beta_k \\ & & & \rho_k & \alpha_k \end{pmatrix}.$$

If no look-ahead steps are performed, then the α's, β's, and ρ's are numbers. If look-ahead steps have been performed then T_k can still be written in this form but now the entries are matrices of size determined by the number of look-ahead steps necessary before a regular Lanczos vector can be produced. For details see, for example, [8].

The above recurrences can be written in matrix form as

$$AV_k = V_k T_k + \rho_{k+1} v_{k+1} e_k^T, \tag{19}$$
$$A^T W_k = W_k \Gamma_k^{-1} T_k \Gamma_k + \xi_{k+1} w_{k+1} e_k^T,$$

where

$$\Gamma_k = \text{diag}(\gamma_1, \ldots, \gamma_k), \quad \gamma_1 = 1, \quad \gamma_j = \gamma_{j-1} \rho_j / \xi_j, \quad j > 1.$$

The kth BCG iterate x_k^B is chosen so that the residual r_k^B satisfies

$$r_k^B \perp K_k(A^T, \hat{r}_0). \tag{20}$$

This is somewhat analogous to the condition (6) defining the Arnoldi iterates and, like (6), this condition may be impossible to satisfy with an iterate of the form (5). Using expression (19), condition (20) can be written in the form

$$x_k^B = x_0 + V_k y_k^B, \tag{21}$$

where y_k^B satisfies

$$T_k y_k^B = \beta e_1, \quad \beta = \|r_0\|. \tag{22}$$

It is shown in [8] that this equation has a solution if and only if the (block) tridiagonal matrix T_k is nonsingular. For the remainder of this discussion we will assume that the matrices $T_1, \ldots, T_k$ are nonsingular. Here again look-ahead strategies can be used to deal with near-singularity of the tridiagonal matrices.

The QMR algorithm is derived in much the same way as the GMRES algorithm described in the previous section. Define the $k+1$ by k matrix $T_k^{(e)}$ by

$$T_k^{(e)} = \begin{pmatrix} T_k \\ 0 \ldots 0 \ \rho_{k+1} \end{pmatrix}. \tag{23}$$

Equation (19) can be written in the form

$$AV_k = V_{k+1} T_k^{(e)}. \tag{24}$$

The kth QMR iterate x_k^Q is of the form

$$x_k^Q = x_0 + V_k y_k^Q,$$

so the kth QMR residual r_k^Q is of the form

$$r_k^Q = r_0 - AV_k y_k^Q = r_0 - V_{k+1} T_k^{(e)} y_k^Q = V_{k+1}(\beta e_1 - T_k^{(e)} y_k^Q), \tag{25}$$

where $\beta = \|r_0\|$ and e_1 is the first unit $(k+1)$-vector. Ideally, one would like to choose y_k^Q to minimize $\|r_k^Q\|$, but since the columns of V_{k+1} are not orthogonal, this would not be practical. Instead, the QMR iterate is defined by taking y_k^Q to minimize the quantity in the parentheses in (25). That is, y_k^Q satisfies the least squares problem

$$\|\beta e_1 - T_k^{(e)} y_k^Q\| = \min_y \|\beta e_1 - T_k^{(e)} y\|. \tag{26}$$

We refer to the vector $\beta e_1 - T_k^{(e)} y_k^Q$ as the QMR *quasi-residual* and denote it z_k^Q. The actual QMR residual is

$$r_k^Q = V_{k+1} z_k^Q. \tag{27}$$

In [8] a more general definition of the QMR iterate is given, allowing for an arbitrary diagonal scaling of the least squares problem (26). It is not clear how this diagonal scaling should be chosen, however, and here we consider only the scaling inherent in (26) with the right and left Lanczos vectors each having norm one.

Since the columns of V_{k+1} are not orthonormal, the norms of the true residuals are not the same as those of the quasi-residuals. One can give upper and lower bounds

on the ratios of these norms, however. Since the columns of V_{k+1} each have norm one, it is shown in [8] that

$$\|r_k^Q\| \leq \sqrt{k+1}\, \|z_k^Q\|. \tag{28}$$

A lower bound on $\|r_k^Q\|$ is given by

$$\|r_k^Q\| \geq \sigma_{\min}(V_{k+1})\|z_k^Q\|, \tag{29}$$

where $\sigma_{\min}(V_{k+1})$ denotes the smallest singular value of V_{k+1}. While it is possible that $\sigma_{\min}(V_{k+1})$ could be very small (especially in finite precision arithmetic, where this is usually the case!), it is unlikely, in such cases, that the inequality (29) will be a near equality, since the approximate solution x_k^Q is chosen to satisfy (26) without regard to the matrix V_{k+1}.

The following theorem is an immediate consequence of Theorem 1 from §2. Let $\hat{s}_i$ and $\hat{c}_i$, $i = 1, \ldots, k$, be the sines and cosines of the Givens rotations used to factor $T_k^{(e)}$. Relations between BCG residuals and QMR quasi-residuals and the sines and cosines of the Givens rotations were established in [8], but the direct relation between BCG residuals and QMR quasi-residuals was never explicitly stated.

THEOREM 3. *In exact arithmetic, if $\hat{c}_k \neq 0$ at iteration k, then the BCG residual and the QMR quasi-residual are related by*

$$\|r_k^B\| = \frac{\|z_k^Q\|}{\sqrt{1 - (\|z_k^Q\|/\|z_{k-1}^Q\|)^2}}. \tag{30}$$

Proof. From (21) and (24), it follows that the BCG residual can be written in the form

$$\begin{aligned} r_k^B &= r_0 - AV_k y_k^B \\ &= r_0 - V_{k+1} T_k^{(e)} y_k^B \\ &= V_{k+1}(\beta e_1 - T_k^{(e)} y_k^B). \end{aligned} \tag{31}$$

From the definition (22) of y_k^B it follows that the quantity in parentheses in (31) has a nonzero entry only in the $(k+1)$st component, and since $\|v_{k+1}\| = 1$, we have

$$\|r_k^B\| = \|\beta e_1 - T_k^{(e)} y_k^B\|. \tag{32}$$

Using relation (32) and the definition (26) of the QMR quasi-residual, the desired result now follows from Theorem 1. □

Note that while the choice of the starting vector $\hat{r}_0$ affects the tridiagonal matrix that is generated and hence affects the sines and cosines of the Givens transformations, it does not affect the relationship (30). This relationship holds provided only that the same vector $\hat{r}_0$ is used for both the BCG and QMR computations.

Figure 2 shows a plot of the logarithms of the norms of the BCG residuals (dashed line), the QMR residuals (dotted line), and the QMR quasi-residuals (solid line) versus iteration number for the same example described in the previous section. Observe that for the convergence tolerance used, both algorithms converged in 101 iterations if the norm of the true residuals is used to measure convergence. Note that on the log plot it can be seen that the QMR residual norm and the quasi-residual norm are of the same order of magnitude, although they are not identical.

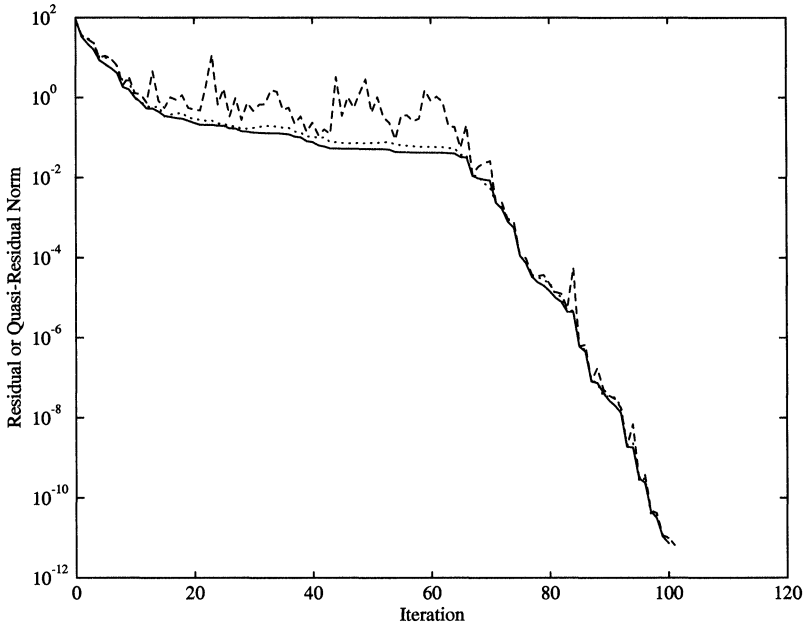

FIG. 2. *QMR quasi-residual norm (solid), BCG residual norm (dashed), and QMR residual norm (dotted).*

On the scale of Fig. 2 it is again difficult to see the precise correlation between the degree of flatness of the QMR quasi-residual norm curve and the height of the corresponding BCG curve above the QMR curve. However, for example, at iterations 33 and 59 in the plot of the residual norms of the BCG iterates, the QMR quasi-residual norms were reduced, respectively, by factors of .99647 and .99962. The corresponding ratios of the BCG norms to the QMR quasi-residual norms were 11.9 and 36.4. These values fit well with the predictions from Theorem 3. The corresponding ratios of the BCG residual norms to the QMR residual norms were 7.95 and 26.04.

The BCG peaks covering approximately iterations 20 to 40 correspond to the appearance in the spectra of the associated tridiagonal matrices in (22) of the conjugate pair of eigenvalues of magnitude 10^{-1}. The next and more recognizable two peaks, from approximately iterations 43 to 64, correspond to the identification of the members of the conjugate pair of magnitude 2×10^{-2}.

In this example, the condition numbers of the tridiagonal matrices converged more or less monotonically to 62,609, a factor of almost 15 times greater than the condition number of A. On iterations 20 to 40 the condition numbers ranged from 473 to 34,310, and from iterations 43 to 64 they ranged from 33,500 to 62,609.

5. Bidiagonalization/SQMR/BLanczos and symmetric Lanczos.

5.1. Real symmetric Lanczos algorithm. The Lanczos algorithm for constructing an orthonormal basis for the Krylov space $K_k(A, r_0)$, where A is a real symmetric matrix, can be written as follows.

Real symmetric Lanczos algorithm:

1. Set $v_1 = r_0/\|r_0\|$. Set $\rho_1 = 1$ and $v_0 = 0$.

2. For $j = 1, \ldots, k$ do

$$\alpha_j = v_j^T(Av_j - \rho_j v_{j-1}),$$
$$\nu_{j+1} = Av_j - \alpha_j v_j - \rho_j v_{j-1},$$
$$\rho_{j+1} = \|\nu_{j+1}\|, \quad v_{j+1} = \frac{\nu_{j+1}}{\rho_{j+1}}.$$

If T_k denotes the symmetric tridiagonal matrix

$$T_k = \begin{pmatrix} \alpha_1 & \rho_2 & & \\ \rho_2 & \ddots & \ddots & \\ & \ddots & \alpha_{k-1} & \rho_k \\ & & \rho_k & \alpha_k \end{pmatrix}$$

and $T_k^{(e)}$ the extended matrix (23), then formulas (19) and (24) express this recurrence in matrix form. The MINRES and Lanczos algorithm iterates are defined using equations (26) and (22), respectively.

For real symmetric problems, assuming exact arithmetic, the Lanczos vectors are orthonormal and the norm of the quasi-residual and the actual residual defined in (27) are the same. In this case relation (30) becomes

$$\|r_k^L\| = \frac{\|r_k^M\|}{\sqrt{1 - (\|r_k^M\|/\|r_{k-1}^M\|)^2}}. \tag{33}$$

5.2. Bidiagonalization of nonsymmetric systems. Any nonsymmetric system (1) can be solved by solving a larger symmetrized version of the problem. The use of bidiagonalization to symmetrize a nonsymmetric problem was suggested in [17] and was subsequently used to compute singular values of A [11] and to solve (1) and associated least squares problems [19] and [20]. Simple bidiagonalization replaces (1) by the following $2N \times 2N$ real symmetric but indefinite system

$$B\bar{x} = \bar{b}, \text{ where } B \equiv \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}, \ \bar{x} \equiv \begin{pmatrix} y \\ x \end{pmatrix}, \ \bar{b} \equiv \begin{pmatrix} b \\ 0 \end{pmatrix}, \tag{34}$$

whose solution contains the desired solution. We have the following lemma relating the eigenvalues of B to the singular values of A and the eigenvectors of B to concatenations of left and right singular vectors of A.

LEMMA 1 (see [5]). *Let A be any real nonsymmetric $N \times N$ matrix with singular value decomposition $A = X\Sigma Y^T$, where $\Sigma = \text{diag}\ \{\sigma_1, \sigma_2, \ldots, \sigma_N\}$ and $Y^TY = X^TX = I$. Then*

$$BZ = Z \begin{pmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{pmatrix}, \text{ where } Z = \frac{1}{\sqrt{2}} \begin{pmatrix} X & X \\ Y & -Y \end{pmatrix}. \tag{35}$$

Both bidiagonalization procedures SQMR and BLanczos map (1) into (34) and then use the real symmetric Lanczos recursion to map (34) into simple tridiagonal problems. Specifically, if we apply the real symmetric Lanczos recursions to B with starting vector $w_1 = \bar{r}_0/\|\bar{r}_0\|$, where $\bar{r}_0 = -B\bar{x}_0 + \bar{b}$ with $\bar{x}_0 = (0, x_0^T)^T$, then in exact arithmetic we obtain the matrix recursion

$$BW_k = W_k T_k + \rho_{k+1} w_{k+1} e_k^T, \tag{36}$$

where

$$T_k = \begin{pmatrix} 0 & \rho_2 & & & \\ \rho_2 & 0 & \rho_3 & & \\ & \rho_3 & 0 & \ddots & \\ & & \ddots & \ddots & \rho_k \\ & & & \rho_k & 0 \end{pmatrix}. \tag{37}$$

Because of the special structure of B and of w_1, all of the $\alpha_j \equiv 0$. Theoretically, the $W_k \equiv \{w_1, \ldots, w_k\}$ are orthonormal, and $T_k = W_k^T B W_k$ is an orthogonal projection of B onto the Krylov subspace $K_k(B, w_1)$.

If $\rho_j \neq 0$, $2 \leq j \leq 2k$, then each T_{2j} is nonsingular and its eigenvalues occur in $\pm$ pairs. Each T_{2j-1} is singular. For details see [5]. Therefore, the Lanczos iterates are defined only on even-numbered steps. We denote the corresponding Lanczos and MINRES iterates by $\bar{x}_k^L$ and $\bar{x}_k^M$ and the corresponding residuals by $\bar{r}_k^L$ and $\bar{r}_k^M$. In Lemmas 2 and 3 we use these quantities to define the BLanczos and SQMR iterates (and residuals) for (1). In exact arithmetic, the kth Lanczos iterate is obtained by solving

$$T_{2k}\bar{z} = \rho e_1, \quad \rho = \|\bar{r}_0\| \tag{38}$$

and forming

$$\bar{x}_k^G = \bar{x}_0 + W_{2k}\bar{z}, \text{ where } \bar{x}_0 = \begin{pmatrix} 0 \\ x_0 \end{pmatrix}. \tag{39}$$

In exact arithmetic, the kth MINRES iterate $\bar{x}_k^M$ is obtained by solving the least squares problem

$$\min_{\bar{z}} \|T_{2k}^{(e)}\bar{z} - \rho e_1\|, \text{ where } T_{2k}^{(e)} = \begin{pmatrix} T_{2k} \\ \rho_{2k+1} e_{2k}^T \end{pmatrix} \tag{40}$$

and forming

$$\bar{x}_k^M = \bar{x}_0 + W_{2k}\bar{z}. \tag{41}$$

Lemmas 2 and 3 extract the corresponding BLanczos and SQMR iterates from the above relationships.

LEMMA 2 (see [4]). *If we apply the Lanczos method to* (34), *then all odd-numbered components of* $\bar{z}$ *in* (38) *are zero. Furthermore, if we specify the* k*th BLanczos iterate* x_k^{BL} *to consist of the last* N *components of* $\bar{x}_k^L$, *then* $r_k^{BL} \equiv b - Ax_k^{BL}$ *consists of the first* N *components of* $\bar{r}_k^L$, *and* $x_k^{BL} = x_0 + V_k\bar{z}^*$, *where* $*$ *denotes the even-numbered components of* $\bar{z}$, $V_k \equiv \{v_1, \ldots, v_k\}$ *and each* v_j *consists of the last* N *components of* w_{2j}. *In addition, in exact arithmetic,*

$$x_k^{BL} = x_{k-1}^{BL} + \bar{z}(2k)\ v_k \quad \textit{and} \quad \|r_k^{BL}\| = |\rho_{2k+1}\bar{z}(2k)|, \tag{42}$$

where $\bar{z}(2k) = (-1)^{k+1} \prod_{j=1}^{k} \rho_{2j-1} / \prod_{j=1}^{k} \rho_{2j}$.

Now consider the SQMR iterates x_k^{SQ}. The least squares problem in (40) can be solved by successively applying Givens transformations $F_j(c_j, s_j)$ to $T_{2k}^{(e)}$ to obtain

$$(F_{2k} \cdots F_1)\, T_{2k}^{(e)} = R_{2k}^{(e)} = \begin{pmatrix} \bar{R}_{2k} \\ 0 \end{pmatrix}, \tag{43}$$

where $\bar{R}_{2k}$ is $2k \times 2k$ and upper triangular. For each j, $c_{2j-1} = 0$ and $s_{2j-1} = 1$, and we therefore use c_j, s_j to denote the cosine and sine which define F_{2j}. If we set $\delta_j = \bar{R}_{2k}(j,j)$ and $\bar{P}_{2k} = W_{2k}\bar{R}_{2k}^{-1}$, then

$$p_k = \left[v_k - \rho_{2k-1}\rho_{2k}\delta_{2k-2}^{-1}p_{k-1}\right]\delta_{2k}^{-1}, \tag{44}$$

where p_k is the N-vector consisting of the last N components of the $2k$th column of $\bar{P}_{2k}$.

LEMMA 3. *If we apply MINRES to* (34), *then all odd-numbered components of* $\bar{z}$ *in* (40) *are zero. Furthermore, if we specify the* k*th SQMR iterate* x_k^{SQ} *to consist of the last* N *components of* $\bar{x}_k^M$, *then* $r_k^{SQ} \equiv b - Ax_k^{SQ}$ *consists of the first* N *components of* $\bar{r}_k^M$, *and* $x_k^{SQ} = x_0 + P_k\bar{z}^*$, *where* $*$ *denotes the even-numbered components of* $\bar{z}$, $P_k \equiv \{p_1, \ldots, p_k\}$ *and* p_j *consists of the last* N *components of* $\bar{p}_{2j}$. *In addition, in exact arithmetic,*

$$\begin{aligned} x_k^{SQ} &= x_{k-1}^{SQ} - c_k\|r_{k-1}^{SQ}\|p_k, \\ \|r_{k-1}^{SQ}\| &= \min_{\bar{z}} \|T_{2k-2}^e\bar{z} - \rho e_1\| = \left|\prod_{j=1}^{k-1} s_j\right| \cdot \|r_0\|, \end{aligned} \tag{45}$$

where c_j, s_j *define the* $2j$*th Givens transformations which were used in the factorization of* $T_{2k}^{(e)}$.

Proofs of Lemmas 2 and 3 are in [4]. If we were defining only SQMR then there is no apparent reason not to consider the $T_{2k-1}^{(e)}$. However from Brown [3] we know, at least in exact arithmetic, that since T_{2k-1} is singular, $T_{2k-1}^{(e)}$ and $T_{2k}^{(e)}$ would yield the same SQMR iterate. In the tests presented in [4] there was no reorthogonalization of any Lanczos vectors.

In §6 we consider the real symmetric Lanczos procedures, Lanczos and MINRES, in finite precision arithmetic and demonstrate that a relationship analogous to (33) exists for a certain class of symmetric problems. In §7 we then show that if a nonsymmetric problem (1) is well conditioned, then the real symmetric problems generated by the bidiagonalization algorithms defined in this section are in that class. Then, using Lemmas 2 and 3, we obtain a relationship analogous to (33) which is valid for these bidiagonalization methods in finite precision arithmetic.

6. Finite precision arithmetic, Lanczos/MINRES. Quantities generated by the Lanczos and the MINRES algorithms will be denoted with superscripts L and M, respectively. Finite precision quantities will be denoted with tildes. In finite precision computations, the matrix equations (19) and (24) are replaced by

$$A\tilde{V}_k = \tilde{V}_k\tilde{T}_k + \tilde{\rho}_{k+1}\tilde{v}_{k+1}e_k^T + F_k = \tilde{V}_{k+1}\tilde{T}_k^{(e)} + F_k. \tag{46}$$

For standard implementations of the real symmetric Lanczos algorithm it is shown in [18] that the Frobenius norm of the roundoff matrix F_k satisfies

$$\|F_k\|_F \leq c\sqrt{k}\,\epsilon\,\|A\|_F, \tag{47}$$

where ϵ is the machine precision and c is a moderate size constant. We will not use this bound explicitly but will simply express error estimates in terms of $\|F_k\|$.

Suppose the computed Lanczos and MINRES approximations satisfy

$$\tilde{x}_k^L = x_0 + \tilde{V}_k\tilde{y}_k^L + g_k^L, \quad \tilde{x}_k^M = x_0 + \tilde{V}_k\tilde{y}_k^M + g_k^M, \tag{48}$$

where $\tilde{y}_k^L$ is the exact solution to the tridiagonal system

$$\tilde{T}_k y = \beta e_1 \tag{49}$$

and $\tilde{y}_k^M$ is the exact solution to the least squares problem

$$\min_y \|\tilde{T}_k^{(e)} y - \beta e_1\|. \tag{50}$$

If the tridiagonal matrix T_k is not too badly conditioned, then the error due to the inexact solution of the linear system or least squares problem will be small. The Lanczos and MINRES residuals for the computed quantities satisfy

$$\begin{aligned} \tilde{r}_k^{L,M} &= r_0 - A\tilde{V}_k \tilde{y}_k^{L,M} - Ag_k^{L,M} \\ &= r_0 - \tilde{V}_{k+1}\tilde{T}_k^{(e)}\tilde{y}_k^{L,M} - F_k\tilde{y}_k^{L,M} - Ag_k^{L,M} \\ &= \tilde{V}_{k+1}(\beta e_1 - \tilde{T}_k^{(e)}\tilde{y}_k^{L,M}) - F_k\tilde{y}_k^{L,M} - Ag_k^{L,M}. \end{aligned} \tag{51}$$

In finite precision computations, the columns of Lanczos vectors $\tilde{V}_{k+1}$ in (51) frequently lose orthogonality. Yet numerical experiments in [4] suggest that relation (33) holds to a close approximation, even after orthogonality of the Lanczos vectors is lost. We now show why this is to be expected, assuming that the terms $-F_k\tilde{y}_k^{L,M} - Ag_k^{L,M}$ in (51) are small compared to $\|\tilde{r}_k^L\|$ and $\|\tilde{r}_k^M\|$.

It is shown in [12] that for any given K, the tridiagonal matrices generated by K steps of a finite precision Lanczos recurrence with a real symmetric matrix A are the same as those that would be generated by the exact Lanczos algorithm applied to a larger real symmetric matrix $\bar{A}$ whose eigenvalues all lie within tiny intervals about the eigenvalues of A. The size of the intervals depends on the machine precision and on an upper bound K for the number of steps that will be run. It is further shown that components of the computed Lanczos vectors associated with a particular eigenvector of A are related to the components of the corresponding exact Lanczos vectors associated with the eigenvectors of $\bar{A}$ whose eigenvalues lie in the interval about this eigenvalue of A, in the following way:

$$(\tilde{v}_k(i))^2 \doteq \sum_\ell (\bar{v}_k(i_\ell))^2. \tag{52}$$

Here $\tilde{v}_k(i)$ represents the component of the computed Lanczos vector $\tilde{v}_k$ in the direction of the ith eigenvector of A. If $\bar{v}_k$ is the corresponding exact Lanczos vector obtained from the recurrence with $\bar{A}$, then $\bar{v}_k(i_\ell)$, $\ell = 1, \ldots$ denotes the components of $\bar{v}_k$ in the directions of the eigenvectors of $\bar{A}$ whose eigenvalues lie in the tiny interval about the ith eigenvalue of A.

Let $\bar{r}_k^L$ and $\bar{r}_k^M$ denote the corresponding exact arithmetic residual vectors obtained by applying the Lanczos and the MINRES algorithms to a linear system $\bar{A}\bar{x} = \bar{b}$, with an initial guess $\bar{x}_0$ such that $\bar{r}_0 = \bar{b} - \bar{A}\bar{x}_0$ has the same norm as r_0 and is parallel to the first Lanczos vector $\bar{v}_1$. The following lemmas relate the norms of the computed residual vectors $\tilde{r}_k^L$ and $\tilde{r}_k^M$ to the norms of the exact residual vectors $\bar{r}_k^L$ and $\bar{r}_k^M$.

LEMMA 4. *Let $\tilde{r}_k^L$ satisfy* (51) *and let $\bar{r}_k^L$ be the residual vector obtained after k steps of the exact Lanczos algorithm applied to the linear system $\bar{A}\bar{x} = \bar{b}$, as described above. Then*

$$\|\tilde{r}_k^L\| = \|\bar{r}_k^L\| + h_k^L, \quad \text{where } |h_k^L| \le \|F_k\| \cdot \|\tilde{y}_k^L\| + \|A\| \cdot \|g_k^L\|. \tag{53}$$

Proof. Since the exact Lanczos recurrence generates the same tridiagonal matrix $\tilde{T}_k$ and the same parameter $\tilde{\rho}_{k+1}$ (since this is an element of $\tilde{T}_{k+1}$) as the finite precision computation, the exact residual vector $\bar{r}_k^L$ satisfies

$$\bar{r}_k^L = -\tilde{\rho}_{k+1}\bar{v}_{k+1}\beta(e_k^T\tilde{T}_k^{-1}e_1).$$

It follows from (51) that

$$\tilde{r}_k^L = -\tilde{\rho}_{k+1}\tilde{v}_{k+1}\beta(e_k^T\tilde{T}_k^{-1}e_1) - F_k\tilde{y}_k^L - Ag_k^L,$$

since the quantity in parentheses in (51) has only its last component nonzero. Since $\|\bar{v}_{k+1}\| = \|\tilde{v}_{k+1}\| = 1$, the desired result (53) follows. □

LEMMA 5. *Let $\tilde{r}_k^M$ satisfy* (51) *and let $\tilde{V}_{k+1}$ satisfy*

$$A\tilde{V}_{k+1} = \tilde{V}_{k+2}\tilde{T}_{k+1}^{(e)} + F_{k+1}. \tag{54}$$

Then $A\tilde{r}_k^M$ is given by

$$A\tilde{r}_k^M = \tilde{v}_{k+1}\gamma_{k+1} + \tilde{v}_{k+2}\gamma_{k+2} + F_{k+1}\tilde{z}_k^M - AF_k\tilde{y}_k^M - A^2g_k^M, \tag{55}$$

where $\tilde{z}_k^M = \beta e_1 - \tilde{T}_k^{(e)}\tilde{y}_k^M$ and the coefficients γ_{k+1} and γ_{k+2} depend only on the elements of the extended tridiagonal matrix $\tilde{T}_{k+1}^{(e)}$.

Proof. Note that since $\tilde{y}_k^M$ minimizes $\|\beta e_1 - \tilde{T}_k^{(e)}y\|$, the remainder $\tilde{z}_k^M = \beta e_1 - \tilde{T}_k^{(e)}\tilde{y}_k^M$ is orthogonal to the columns of $\tilde{T}_k^{(e)}$. Note also that the extended tridiagonal matrix $\tilde{T}_{k+1}^{(e)}$ can be written in the form

$$\tilde{T}_{k+1}^{(e)} = \begin{pmatrix} \tilde{T}_k^{(e)^T} & \\ 0\ldots0\ \tilde{\rho}_{k+1} & \tilde{\alpha}_{k+1} \\ 0\ldots0\quad 0 & \tilde{\rho}_{k+2} \end{pmatrix},$$

where the elements $\tilde{\alpha}_{k+1}$ and $\tilde{\rho}_{k+1}$, $\tilde{\rho}_{k+2}$ are those generated by the symmetric Lanczos recurrence. Multiplying (54) by $\tilde{z}_k^M$ on the right gives

$$A\tilde{V}_{k+1}\tilde{z}_k^M = \tilde{v}_{k+1}\gamma_{k+1} + \tilde{v}_{k+2}\gamma_{k+2} + F_{k+1}\tilde{z}_k^M, \tag{56}$$

where the coefficients

$$\gamma_{k+1} = \tilde{\rho}_{k+1}\tilde{z}_k^M(k) + \tilde{\alpha}_{k+1}\tilde{z}_k^M(k+1), \quad \gamma_{k+2} = \tilde{\rho}_{k+2}\tilde{z}_k^M(k+1)$$

are functions of the elements of the extended tridiagonal matrix $\tilde{T}_{k+1}^{(e)}$. Using (51) to substitute for $\tilde{V}_{k+1}\tilde{z}_k^M$ in (56) gives

$$A(\tilde{r}_k^M + F_k\tilde{y}_k^M + Ag_k^M) = \tilde{v}_{k+1}\gamma_{k+1} + \tilde{v}_{k+2}\gamma_{k+2} + F_{k+1}\tilde{z}_k^M,$$

from which the result (55) follows. □

LEMMA 6. *The exact arithmetic residual vector $\bar{r}_k^M$ satisfies*

$$\bar{A}\bar{r}_k^M = \bar{v}_{k+1}\gamma_{k+1} + \bar{v}_{k+2}\gamma_{k+2}, \tag{57}$$

where γ_{k+1} and γ_{k+2} are the same coefficients as in (55).

Proof. Since the exact Lanczos vectors satisfy

$$\bar{A}\bar{V}_{k+1} = \bar{V}_{k+2}\tilde{T}_{k+1}^{(e)},$$

where $\tilde{T}_{k+1}^{(e)}$ is the same tridiagonal matrix as in (54), the result follows by the same arguments as used in Lemma 5. □

We wish to show that $\|\tilde{r}_k^M\| \approx \|\bar{r}_k^M\|$. To see this, it is necessary to translate into bases in which A and $\bar{A}$ are diagonal. That is, suppose $A = W\Lambda W^T$ and $\bar{A} = \bar{W}\bar{\Lambda}\bar{W}^T$, where Λ and $\bar{\Lambda}$ are diagonal and W and $\bar{W}$ are orthogonal matrices. Let $\tilde{v}(i)$ denote the ith component of a vector $W^T\tilde{v}$ associated with the finite precision computation for Λ and let $\bar{v}(i_\ell)$ denote the i_ℓth component of a vector $\bar{W}^T\bar{v}$ associated with the exact calculation for $\bar{\Lambda}$. The index i_ℓ will range over all eigenvalues $\bar{\lambda}_{i_\ell}$ of $\bar{\Lambda}$ that lie in the interval about eigenvalue λ_i of Λ. There is no loss in generality in making this transformation. The arguments used need only the fact that the error term in (54) is small and the size of that term is independent of the orthogonal matrix W.

Since the bound proved in [12] on the size of the intervals containing the eigenvalues of $\bar{\Lambda}$ appears to be a large overestimate of their actual size [13], it is not very enlightening to include this bound in our estimates. Instead we will simply assume

$$\max_\ell |\lambda_i - \bar{\lambda}_{i_\ell}| \le \xi \quad \forall i. \tag{58}$$

The bound in [12] on the difference between the left- and right-hand sides in (52) is also an overestimate. Therefore, we will simply assume

$$\left| (\tilde{v}_j(i))^2 - \sum_\ell (\bar{v}_j(i_\ell))^2 \right| \le \delta \quad \forall i, j. \tag{59}$$

The following lemma establishes one more relation between the components $W^T\tilde{v}_k$ of the computed Lanczos vectors and the components $\bar{W}^T\bar{v}_k$ of the exact Lanczos vectors.

LEMMA 7. *The following relation holds between components of the computed Lanczos vectors for Λ and those of the exact Lanczos vectors for $\bar{\Lambda}$:*

$$\left| \tilde{v}_{k+1}(i)\tilde{v}_k(i) - \sum_\ell \bar{v}_{k+1}(i_\ell)\bar{v}_k(i_\ell) \right| \le \frac{1}{\tilde{\rho}_{k+1}} \left[\delta \cdot \sum_{j=1}^{k} |\lambda_i - \tilde{\alpha}_j| \right.$$

$$\left. + \xi \cdot \sum_{j=1}^{k} \sum_\ell (\bar{v}_j(i_\ell))^2 + f \cdot \sum_{j=1}^{k} |\tilde{v}_j(i)| \right], \tag{60}$$

where $f = \max_{i,j} |F_k(i,j)|$ and ξ and δ are as defined in (58) *and* (59).

Proof. Writing the three-term Lanczos recurrences for the relevant components we have

$$\tilde{\rho}_{k+1}\tilde{v}_{k+1}(i) = (\lambda_i - \tilde{\alpha}_k)\tilde{v}_k(i) - \tilde{\rho}_k\tilde{v}_{k-1}(i) - F_k(i,k),$$

$$\tilde{\rho}_{k+1}\bar{v}_{k+1}(i_\ell) = (\bar{\lambda}_{i_\ell} - \tilde{\alpha}_k)\bar{v}_k(i_\ell) - \tilde{\rho}_k\bar{v}_{k-1}(i_\ell).$$

Multiplying the first of these equations by $\tilde{v}_k(i)$ and multiplying the second by $\bar{v}_k(i_\ell)$ and summing over ℓ gives

$$\tilde{\rho}_{k+1}\tilde{v}_{k+1}(i)\tilde{v}_k(i) = (\lambda_i - \tilde{\alpha}_k)(\tilde{v}_k(i))^2 - \tilde{\rho}_k\tilde{v}_k(i)\tilde{v}_{k-1}(i) - F_k(i,k)\tilde{v}_k(i), \tag{61}$$

$$\tilde{\rho}_{k+1}\sum_\ell \bar{v}_{k+1}(i_\ell)\bar{v}_k(i_\ell) = (\lambda_i - \tilde{\alpha}_k)\sum_\ell(\bar{v}_k(i_\ell))^2 - \tilde{\rho}_k\sum_\ell \bar{v}_k(i_\ell)\bar{v}_{k-1}(i_\ell)$$

$$+\sum_\ell(\bar{\lambda}_{i_\ell} - \lambda_i)(\bar{v}_k(i_\ell))^2. \tag{62}$$

Subtract (62) from (61) to obtain

$$\tilde{\rho}_{k+1}d_{k+1,i} = (\lambda_i - \tilde{\alpha}_k)\delta_{k,i} - \tilde{\rho}_k d_{k,i} - F_k(i,k)\tilde{v}_k(i) - \xi_{k,i}\sum_\ell(\bar{v}_k(i_\ell))^2, \tag{63}$$

where we have defined

$$d_{j,i} \equiv \tilde{v}_j(i)\tilde{v}_{j-1}(i) - \sum_\ell \bar{v}_j(i_\ell)\bar{v}_{j-1}(i_\ell),$$

$$\delta_{j,i} \equiv (\tilde{v}_j(i))^2 - \sum_\ell(\bar{v}_j(i_\ell))^2,$$

$$\xi_{j,i} \ : \ \sum_\ell(\bar{\lambda}_{i_\ell} - \lambda_i)(\bar{v}_j(i_\ell))^2 = \xi_{j,i}\sum_\ell(\bar{v}_j(i_\ell))^2.$$

Clearly, $|\delta_{j,i}| \le \delta$ and $|\xi_{j,i}| \le \xi$ for all i and j. Applying formula (63) recursively gives

$$\tilde{\rho}_{k+1}d_{k+1,i} = \sum_{j=1}^{k}(-1)^{k-j}\left[\delta_{j,i}(\lambda_i - \tilde{\alpha}_j) - F_k(i,j)\tilde{v}_j(i) - \xi_{j,i}\sum_\ell(\bar{v}_j(i_\ell))^2\right].$$

Dividing by $\tilde{\rho}_{k+1}$, taking absolute values on each side, and bounding the quantities $|\delta_{j,i}|$, $|F_k(i,j)|$, and $|\xi_{j,i}|$ on the right-hand side gives the desired result (60). □

LEMMA 8. *The residual vectors $\tilde{r}_k^M$ and $\bar{r}_k^M$ are related by*

$$\|\tilde{r}_k^M\| = \|\bar{r}_k^M\| + h_k^M,$$

where

$$|h_k^M| \le \|\tilde{r}_k^M\|\,\frac{1}{\lambda_{\min}^2}\left[\frac{1}{2}N\cdot\delta + \frac{1}{2}d + |\lambda_{\min}|\cdot\frac{\|\zeta\|}{\|\tilde{r}_k^M\|} + |\lambda_{\min}|\cdot\xi\right] + O(\Delta^2), \tag{64}$$

where $\lambda_{\min}$ is the eigenvalue of A of smallest absolute value, δ and ξ are defined by (59) *and* (58), *ζ is given by*

$$\zeta = F_{k+1}\tilde{z}_k^M - AF_k\tilde{y}_k^M - A^2g_k^M,$$

and d satisfies

$$d = \sum_{i=1}^{N}|d_{k+2,i}|,$$

$$d \le \frac{1}{\tilde{\rho}_{k+2}}\left[\delta\cdot 2(k+1)N|\lambda_{\max}| + \xi\cdot(k+1) + f\cdot(k+1)\sqrt{N}\right] \tag{65}$$

with $\lambda_{\max}$ the eigenvalue of A of largest absolute value. The term $O(\Delta^2)$ denotes higher-order terms in δ, ξ, $\|\zeta\|$, and d.

Proof. Equation (55) can be written in component form as

$$\lambda_i \tilde{r}_k^M(i) = \gamma_{k+1}\tilde{v}_{k+1}(i) + \gamma_{k+2}\tilde{v}_{k+2}(i) + \zeta(i), \tag{66}$$

and equation (57) becomes

$$\bar{\lambda}_{i_\ell} \bar{r}_k^M(i_\ell) = \gamma_{k+1}\bar{v}_{k+1}(i_\ell) + \gamma_{k+2}\bar{v}_{k+2}(i_\ell). \tag{67}$$

Squaring both sides in (66) gives

$$\begin{aligned}\lambda_i^2(\tilde{r}_k^M(i))^2 &= \gamma_{k+1}^2(\tilde{v}_{k+1}(i))^2 + \gamma_{k+2}^2(\tilde{v}_{k+2}(i))^2 + 2\gamma_{k+1}\gamma_{k+2}\tilde{v}_{k+1}(i)\tilde{v}_{k+2}(i) \\ &\quad + 2\lambda_i\tilde{r}_k^M(i)\zeta(i) - (\zeta(i))^2,\end{aligned} \tag{68}$$

and squaring both sides and summing over ℓ in (67) gives

$$\begin{aligned}\lambda_i^2 \sum_\ell (\bar{r}_k^M(i_\ell))^2 &= \gamma_{k+1}^2 \sum_\ell (\bar{v}_{k+1}(i_\ell))^2 + \gamma_{k+2}^2 \sum_\ell (\bar{v}_{k+2}(i_\ell))^2 \\ &\quad + 2\gamma_{k+1}\gamma_{k+2} \sum_\ell \bar{v}_{k+1}(i_\ell)\bar{v}_{k+2}(i_\ell) + \sum_\ell (\lambda_i - \bar{\lambda}_{i_\ell})(\lambda_i + \bar{\lambda}_{i_\ell})(\bar{r}_k^M(i_\ell))^2.\end{aligned} \tag{69}$$

Finally, subtracting (69) from (68) gives

$$\begin{aligned}\lambda_i^2 \left[(\tilde{r}_k^M(i))^2 - \sum_\ell (\bar{r}_k^M(i_\ell))^2 \right] &= \gamma_{k+1}^2\delta_{k+1,i} + \gamma_{k+2}^2\delta_{k+2,i} + 2\gamma_{k+1}\gamma_{k+2}d_{k+2,i} \\ &\quad + 2\lambda_i\tilde{r}_k^M(i)\zeta(i) - (\zeta(i))^2 - (\lambda_i - \eta_{k,i})(\lambda_i + \eta_{k,i}) \sum_\ell (\bar{r}_k^M(i_\ell))^2,\end{aligned} \tag{70}$$

where $\eta_{k,i}$ satisfies

$$\sum_\ell (\lambda_i - \bar{\lambda}_{i_\ell})(\lambda_i + \bar{\lambda}_{i_\ell})(\bar{r}_k^M(i_\ell))^2 = (\lambda_i - \eta_{k,i})(\lambda_i + \eta_{k,i}) \sum_\ell (\bar{r}_k^M(i_\ell))^2.$$

By the mean value theorem we have $|\lambda_i - \eta_{k,i}| \le \xi$. Divide each side of (70) by λ_i^2, sum over i, and use the bounds on $|\delta_{j,i}|$, $|d_{j,i}|$, and $|\lambda_i - \eta_{j,i}|$ to obtain

$$\begin{aligned}|\, \|\tilde{r}_k^M\|^2 - \|\bar{r}_k^M\|^2 \,| &\le \frac{1}{\lambda_{\min}^2} [\, (\gamma_{k+1}^2 + \gamma_{k+2}^2)n\delta + 2|\gamma_{k+1}\gamma_{k+2}|d \\ &\quad + 2|\lambda_{\min}|\, \|\tilde{r}_k^M\|\, \|\zeta\| + 2|\lambda_{\min}|\, \xi\|\bar{r}_k^M\|^2 \,] + O(\Delta^2).\end{aligned} \tag{71}$$

Use the fact that $\gamma_{k+1}^2 + \gamma_{k+2}^2 = \|\bar{r}_k^M\|^2$ and $2|\gamma_{k+1}\gamma_{k+2}| \le \|\bar{r}_k^M\|^2$ and divide each side in (71) by $\|\tilde{r}_k^M\| + \|\bar{r}_k^M\|$ to obtain the desired result (64). The bound (65) is obtained by summing over i in expression (60). □

THEOREM 4. *The computed Lanczos and MINRES residuals are related by*

$$\begin{aligned}\|\tilde{r}_k^L\| &= \frac{\|\tilde{r}_k^M\|}{\sqrt{1 - (\|\tilde{r}_k^M\|/\|\tilde{r}_{k-1}^M\|)^2}} - \frac{h_k^M - h_{k-1}^M(\|\tilde{r}_k^M\|/\|\tilde{r}_{k-1}^M\|)^3}{[1 - (\|\tilde{r}_k^M\|/\|\tilde{r}_{k-1}^M\|)^2]^{3/2}} \\ &\quad + h_k^L + O(\Delta^2).\end{aligned} \tag{72}$$

Proof. The exact arithmetic residual vectors associated with $\bar{A}$ satisfy

$$\|\bar{r}_k^L\| = \frac{\|\bar{r}_k^M\|}{\sqrt{1 - (\|\bar{r}_k^M\|/\|\bar{r}_{k-1}^M\|)^2}},$$

so from Lemmas 4 and 8 we have

$$\|\tilde{r}_k^L\| - h_k^L = \frac{\|\tilde{r}_k^M\| - h_k^M}{\sqrt{1 - ((\|\tilde{r}_k^M\| - h_k^M)/(\|\tilde{r}_{k-1}^M\| - h_{k-1}^M))^2}}.$$

Manipulating this expression gives the result (72). □

Note that Theorem 4 implies that relation (33) holds to a close approximation in finite precision arithmetic, provided the roundoff terms h_k^L, h_k^M, and h_{k-1}^M are much smaller than $\|\tilde{r}_k^M\|$ and provided the reduction factor $\|\tilde{r}_k^M\|/\|\tilde{r}_{k-1}^M\|$ is not too close to one.

7. Finite precision arithmetic, BLanczos/SQMR. If the tridiagonal matrices generated by the Lanczos algorithm are well conditioned, then one can expect the roundoff terms h_k^L and h_k^M in Theorem 4 to be small since the roundoff term F_k in (46) is tiny and g_k^L and g_k^M in (48) will be small if the tridiagonal systems are solved accurately. In this section we show that the even-order tridiagonal matrices generated by the BLanczos and SQMR algorithms described in §5 are essentially as well conditioned as the original matrix A in (1).

For each tridiagonal matrix T_k generated by the BLanczos and SQMR algorithms, let $\lambda_i^{(k)}$, $1 \le i \le k$ denote the eigenvalues of T_k. The proof that the even-order T_{2k} are as well conditioned as A uses the interlacing property of eigenvalues of tridiagonal matrices. (See, for example, [24, p. 46] or, later, [16].) This property says that if T_k is any principal submatrix of a symmetric tridiagonal matrix T, then between each pair of eigenvalues of T_k is at least one eigenvalue of T. Also, there is an eigenvalue of T that is less than the smallest eigenvalue of T_k and an eigenvalue of T that is greater than the largest eigenvalue of T_k. We also need the following lemma.

LEMMA 9 (see [4]). *Each unreduced, even-ordered tridiagonal matrix T_{2k} defined by* (37) *is nonsingular and has eigenvalues that occur in $\pm$ pairs. Each odd-order T_{2k+1} is singular and has a simple zero eigenvalue.*

Using these properties and results from [12] relating the tridiagonal matrices generated by a finite precision computation to those that would be generated by an exact calculation for a certain larger matrix with nearby eigenvalues, we are able to prove the following theorem.

THEOREM 5. *Let A be any real nonsymmetric matrix with singular values $0 < \sigma_N \le \sigma_{N-1} \le \cdots \le \sigma_1$. Let $\tilde{T}_{2k}$, $k = 1, 2, \ldots, K$ be the even-ordered tridiagonal matrices generated by applying either BLanczos or SQMR to* (1) *in finite precision arithmetic. Then for all $1 \le j \le K$, the eigenvalues of $\tilde{T}_{2j}$ lie in the intervals $[-\sigma_1 - \xi, -\sigma_N + \xi] \cup [\sigma_N - \xi, \sigma_1 + \xi]$, where ξ is a bound on the distance between the eigenvalues of B in* (34) *and those of a corresponding exact arithmetic matrix $\bar{B}$, as described in the previous section.*

Proof. Since bidiagonalization is an application of the real symmetric Lanczos procedure, the results in [12] are applicable. Therefore, there exists a matrix $\bar{B}$ whose eigenvalues lie in tiny intervals about the eigenvalues of B such that the exact Lanczos algorithm applied to $\bar{B}$ generates tridiagonal matrices $\bar{T}_{2j}$ identical to the tridiagonal

matrices $\tilde{T}_{2j}$, $j = 1, \ldots, K$ generated by the finite precision bidiagonalization process applied to B in (34). Let ξ be a bound on the width of these intervals. Since the Lanczos computation for $\bar{B}$ is exact, there exists some $M \geq 2K$ such that the eigenvalues of $\bar{T} \equiv \bar{T}_M$ are the eigenvalues of $\bar{B}$. It follows from the interlacing theorem that between each pair of eigenvalues of $\tilde{T}_{2j}$ is an eigenvalue of $\bar{T}$ and hence of $\bar{B}$. Additionally, all eigenvalues of $\tilde{T}_{2j}$ must lie between the smallest and largest eigenvalues of $\bar{B}$, $[-\sigma_1 - \xi, \sigma_1 + \xi]$. From Lemma 9 the eigenvalues of T_{2j} occur in $\pm$ pairs. Therefore, if for some j, $\tilde{T}_{2j}$ had an eigenvalue in the interval $(-\sigma_N + \xi, \sigma_N - \xi)$, then it would necessarily have a pair of eigenvalues in this interval and hence $\bar{B}$ would have to have an eigenvalue in this interval. This is a contradiction, and therefore the eigenvalues of each $\tilde{T}_{2j}$ must be contained in the intervals given in the theorem. □

From Theorem 5 it follows that if the original matrix A in (1) is well conditioned, then the error term in Theorem 4 will be small. Therefore, using Lemmas 2 and 3 we get that the BLanczos and the SQMR residual norms will satisfy an approximate relationship of the form (72).

8. Conclusions. In exact arithmetic we have derived a precise relation between the sizes of the Arnoldi and GMRES residuals at any iteration k and between the sizes of the BCG residual and the QMR quasi-residual at any iteration k. This relation implies roughly that if the Galerkin iterates are well defined and if one member of either pair of algorithms converges very well, then the other member of the pair will also converge very well, and if one member performs very poorly then the other member will also perform poorly. While the residual (or a related quantity) in the norm-minimizing method cannot grow, as it can in a Galerkin method, it is no more useful to have a near constant residual norm than it is to have a growing one. If one prefers to see a (weakly) monotonically decreasing convergence curve, one can always plot the norm of the smallest residual obtained so far.

While those proofs assumed exact arithmetic, the relation between GMRES and Arnoldi residual norms can be expected to hold to a close approximation in finite precision arithmetic as well, since orthogonality of the Arnoldi vectors is maintained, or can be maintained, with a sufficiently careful implementation of the algorithm [6].

For the QMR and BCG algorithms, precise details of the implementation and use of look-ahead steps will determine whether or not the relation continues to hold in finite precision arithmetic. If the BCG iterate produced at some step has a very large norm, then future iterates updated from this one may never approach the true solution [14]. This situation can be avoided through the use of look-ahead procedures or by storing certain intermediate quantities and using these to generate the BCG approximations. For example, the BCG iterates can be generated from the QMR iterates [8]. There seems to be little if any reason, however, to choose the Galerkin variant over the norm-minimizing variant when each can be generated with essentially the same amount of work and storage. A possible exception may be the case of very ill conditioned symmetric problems, where it was observed in [20] that the SYMMLQ implementation of the Lanczos algorithm (the symmetric equivalent of BCG, but implemented in such a way that large intermediate iterates are not used to generate future iterates) sometimes attained a higher level of accuracy than the MINRES algorithm (the symmetric equivalent of QMR).

For real symmetric problems with well-conditioned tridiagonal matrices we have shown that, despite the loss of orthogonality of the Lanczos vectors, the relation between the Lanczos and MINRES residuals holds to a close approximation in finite precision arithmetic. We then used this result to prove that if a nonsymmetric

problem (1) is well conditioned, then the residuals generated by the nonsymmetric bidiagonalization algorithms, BLanczos and SQMR, also satisfy these relationships to a close approximation in finite precision arithmetic.

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] C. Brezinski, M. Redivo Zaglia, and H. Sadok, *A breakdown-free Lanczos type algorithm for solving linear systems*, Numer. Math., 63 (1992), pp. 29–38.

[3] P. N. Brown, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 58–78.

[4] J. K. Cullum, *Peaks and plateaus in Lanczos methods for solving nonsymmetric systems of equations $Ax = b$*, IBM research report RC 18084, T. J. Watson Research Center, Yorktown Heights, NY, 1992.

[5] J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*, Vol. 1, Theory, Progress in Scientific Computing Vol. 3, S. Abarbanel et al. eds., Birkhäuser, Basel, Switzerland, 1985.

[6] J. Drkošová, A. Greenbaum, M. Rozložník, and Z. Strakoš, *Numerical stability of GMRES*, BIT, 35 (1995), pp. 309–330.

[7] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Numerical Analysis Dundee 1975, G. A. Watson, ed., Lecture Notes in Mathematics 506, Springer-Verlag, Berlin, 1976, pp. 73–89.

[8] R. W. Freund and N. M. Nachtigal, *QMR: A quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.

[9] R. W. Freund, *A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems*, SIAM J. Sci. Comput., 14 (1993), pp. 470–482.

[10] 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., 14 (1993), pp. 137–158.

[11] G. H. Golub and W. Kahan, *Calculating the singular values and pseudo-inverse of a matrix*, SIAM J. Numer. Anal., 2 (1965), pp. 205–224.

[12] A. Greenbaum, *Behavior of slightly perturbed Lanczos and conjugate gradient recurrences*, Lin. Algebra Appl., 113 (1989), pp. 7–63.

[13] A. Greenbaum and Z. Strakos, *Predicting the behavior of finite precision Lanczos and conjugate gradient computations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 121–137.

[14] A. Greenbaum, *Accuracy of computed solutions from conjugate-gradient-like methods*, in PCG '94: Advances in Numerical Methods for Large Sparse Sets of Linear Equations, M. Natori and T. Nodera, eds., Keio University, Yokahama, Japan, 1994, pp. 126–138.

[15] M. H. Gutknecht, *Changing the norm in conjugate gradient type algorithms*, SIAM J. Numer. Anal., 30 (1993), pp. 40–56.

[16] R. O. Hill, Jr. and B. N. Parlett, *Refined interlacing properties*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 239–247.

[17] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Natl. Bur. Stand., 45 (1950), pp. 255–282.

[18] C. C. Paige, *Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix*, J. Inst. Math. Appl., 18 (1976), pp. 341–349.

[19] ———, *Bidiagonalization of matrices and solution of linear equations*, SIAM J. Numer. Anal., 11 (1974), pp. 197–209.

[20] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[21] B. N. Parlett, D. R. Taylor, and Z. A. Liu, *A look-ahead Lanczos algorithm for unsymmetric matrices*, Math. Comp., 44 (1985), pp. 105–124.

[22] Y. Saad, *Krylov subspace methods for solving unsymmetric linear systems*, Math. Comp., 37 (1981), pp. 105–126.

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

[24] G. Szego, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, Volume XXIII, New York, 1959.

[25] H. A. van der Vorst and C. Vuik, *The superlinear convergence behavior of GMRES*, J. Comp. Appl. Math., 48 (1993), pp. 327–341.

[26] A. VAN DER SLUIS AND H. A. VAN DER VORST, *The convergence behavior of Ritz values in the presence of close eigenvalues*, Lin. Algebra Appl. 88/89 (1987), pp. 651–694.

[27] H. F. WALKER, *Implementation of the GMRES method using Householder transformations*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 152–163.

[28] R. WEISS, *Convergence behavior of generalized conjugate gradient methods*, Ph.D. thesis, University of Karlsruhe, Karlsruhe, Germany, 1990.