SIAM J. Sci. Comput.
Vol. 32, No. 1, pp. 186–196

THE CYCLE-CONVERGENCE OF RESTARTED GMRES FOR NORMAL MATRICES IS SUBLINEAR*

E. VECHARYNSKI† AND J. LANGOU†

Abstract. We prove that the cycle-convergence of the restarted GMRES applied to a system of linear equations with a normal coefficient matrix is sublinear.

Key words. restarted generalized minimal residual method, convergence, normal, diagonalizable, sublinear, superlinear

AMS subject classification. 65F10

DOI. 10.1137/080727403

1. Introduction.

The *generalized minimal residual method* (GMRES) was originally introduced by Saad and Schultz [13] in 1986, and has become a popular method for solving non-Hermitian systems of linear equations,

$$Ax = b, \quad A \in \mathbb{C}^{n\times n}, \quad b \in \mathbb{C}^n. \tag{1.1}$$

GMRES is classified as a Krylov subspace (projection) iterative method. At every new iteration i, GMRES constructs an approximation x_i to the exact solution of (1.1) such that the 2-norm of the corresponding residual vector $r_i = b - Ax_i$ is minimized over the affine space $r_0 + A\mathcal{K}_i(A, r_0)$, i.e.,

$$r_i = \min_{u\in\mathcal{K}_i(A,r_0)} \|r_0 - Au\|, \tag{1.2}$$

where $\mathcal{K}_i(A, r_0)$ is the i-dimensional Krylov subspace

$$\mathcal{K}_i(A, r_0) = \operatorname{span}\{r_0, Ar_0, \ldots, A^{i-1}r_0\}$$

induced by the matrix A and the initial residual vector $r_0 = b - Ax_0$ with x_0 being an initial approximate solution of (1.1).

As usual, in a linear setting, a notion of minimality is associated with some orthogonality condition. In our case, the minimization (1.2) is equivalent to forcing the new residual vector r_i to be orthogonal to the subspace $A\mathcal{K}_i(A, r_0)$ (also known as the Krylov residual subspace). In practice, for a large problem size, the latter orthogonality condition results in a costly procedure of orthogonalization against the expanding Krylov residual subspace. Orthogonalization together with storage requirement makes the GMRES method complexity and storage prohibitive for practical application. A straightforward treatment for this complication is the so-called restarted GMRES [13].

The *restarted GMRES*, or GMRES(m), is based on restarting GMRES after every m iterations. At each restart, we use the latest approximate solution as the initial approximation for the next GMRES run. Within this framework a single run of m GMRES iterations is called a GMRES(m) cycle, and m is called the restart parameter.

*Received by the editors June 15, 2008; accepted for publication (in revised form) April 1, 2009; published electronically February 5, 2010.
http://www.siam.org/journals/sisc/32-1/72740.html

†Department of Mathematical and Statistical Sciences, University of Colorado Denver, Denver, CO 80217 (yaugen.vecharynski@ucdenver.edu, julien.langou@ucdenver.edu).

Consequently, restarted GMRES can be regarded as a sequence of GMRES(m) cycles. When the convergence happens without any restart occurring, the algorithm is known as *full GMRES*.

In the context of restarted GMRES, our interest will shift towards the residual vectors r_k at the end of every kth GMRES(m) cycle (as opposed to the residual vectors r_i (1.2) obtained at each iteration of the algorithm).

DEFINITION 1 (cycle-convergence). *We define the cycle-convergence of restarted GMRES(m) to be the convergence of the residual norms $\|r_k\|$, where, for each k, r_k is the residual at the end of the kth GMRES(m) cycle.*

We note that each r_k satisfies the local minimality condition

$$r_k = \min_{u \in \mathcal{K}_m(A, r_{k-1})} \|r_{k-1} - Au\|, \tag{1.3}$$

where $\mathcal{K}_m(A, r_{k-1})$ is the m-dimensional Krylov subspace produced at the kth GMRES(m) cycle,

$$\mathcal{K}_m(A, r_{k-1}) = \operatorname{span}\{r_{k-1}, Ar_{k-1}, \ldots, A^{m-1}r_{k-1}\}. \tag{1.4}$$

The price paid for the reduction of the computational work in GMRES(m) is the loss of global optimality (1.2). Although (1.3) implies a monotonic decrease of the norms of the residual vectors r_k, GMRES(m) can stagnate [13, 18]. This is in contrast with full GMRES which is guaranteed to converge to the exact solution of (1.1) in n steps (assuming exact arithmetic and nonsingular A). However, a proper choice of a preconditioner and/or a restart parameter, e.g., [6, 7, 12], can significantly accelerate the convergence of GMRES(m), thus making the method attractive for real applications.

While a great deal of effort has been devoted to the characterization of the convergence of full GMRES, e.g., [4, 5, 8, 9, 11, 15, 16], our understanding of the behavior of GMRES(m) is far from complete, leaving us with more questions than answers, e.g., [6]. In this manuscript, we prove that the cycle-convergence of restarted GMRES for normal matrices is sublinear. This statement means that, for normal matrices, the reduction in the norm of the residual vector at the current GMRES(m) cycle cannot be better than the reduction at the previous cycle.

The current manuscript was inspired by ideas introduced in the technical report [17] by Zavorin. In that work, Zavorin is interested in the worst-case behavior of (full) GMRES. The worst-case behavior at a given step m is the maximum residual norm obtained for a given matrix over all the right-hand sides. Zavorin shows that, at every step of GMRES, a diagonalizable matrix A and its conjugate transpose A^H yield the same worst-case. He also derives a necessary condition (the so-called cross-equality) for the worst-case right-hand side vector. We inherit the mathematical tools for our analysis from [17], as well as [11, 18], and give their brief description—slightly adapted to the case of the restarted GMRES and a normal matrix A—in section 2. The main result of the sublinear cycle-convergence is proved in section 3. In section 4, the behavior of GMRES(m) in the nonnormal case is discussed.

While this manuscript was in revision, we learned of the independent but related work of Baker, Jessup, and Kolev [1], who present the main result of this paper (Theorem 5). While the result is the same, both proofs and contexts of application are substantially different. We encourage readers of our manuscript to read [1] as well.

Throughout the manuscript we will assume (unless otherwise explicitly stated) A

to be nonsingular and normal, i.e., A allows the decomposition

$$A = V\Lambda V^H, \tag{1.5}$$

where $\Lambda \in \mathbb{C}^{n\times n}$ is a diagonal matrix with the diagonal elements being the nonzero eigenvalues of A, and $V \in \mathbb{C}^{n\times n}$ is a unitary matrix of the corresponding eigenvectors. $\|\cdot\|$ denotes the 2-norm throughout.

2. Krylov matrix, its pseudoinverse, and spectral factorization. For a given restart parameter m ($1 \le m \le n-1$), let us denote the kth cycle of GMRES(m) applied to the system (1.1) with the initial residual vector r_{k-1} as GMRES(A, m, r_{k-1}). We assume that the residual vector r_k, produced at the end of GMRES(A, m, r_{k-1}), is nonzero.

A run of GMRES(A, m, r_{k-1}) generates the Krylov subspace $\mathcal{K}_m(A, r_{k-1})$ given in (1.4). For each $\mathcal{K}_m(A, r_{k-1})$ we define a matrix $K(A, r_{k-1}) \in \mathbb{C}^{n\times(m+1)}$ such that

$$K(A, r_{k-1}) = [r_{k-1} \quad Ar_{k-1} \quad \cdots \quad A^m r_{k-1}], \quad k = 1, 2, \ldots, q, \tag{2.1}$$

where q is the total number of GMRES(m) cycles.

The matrix (2.1) is called the Krylov matrix. We will say that $K(A, r_{k-1})$ corresponds to the cycle GMRES(A, m, r_{k-1}). Note that the columns of $K(A, r_{k-1})$ span the next $(m+1)$-dimensional Krylov subspace $\mathcal{K}_{m+1}(A, r_{k-1})$. By the assumption that $r_k \ne 0$,

$$\operatorname{rank}(K(A, r_{k-1})) = m + 1.$$

This latter equality allows us to introduce the Moore–Penrose pseudoinverse of the matrix $K(A, r_{k-1})$,

$$K^\dagger(A, r_{k-1}) = \left(K^H(A, r_{k-1})\, K(A, r_{k-1})\right)^{-1} K^H(A, r_{k-1}) \in \mathbb{C}^{(m+1)\times n},$$

which is well-defined and unique. The following lemma shows that the first column of $\left(K^\dagger(A, r_{k-1})\right)^H$ is the next residual vector r_k up to a scaling factor.

LEMMA 2. *Given $A \in \mathbb{C}^{n\times n}$ (not necessarily normal), for any $k = 1, 2, \ldots, q$, we have*

$$\left(K^\dagger(A, r_{k-1})\right)^H e_1 = \frac{1}{\|r_k\|^2} r_k, \tag{2.2}$$

where $e_1 = [1 \quad 0 \quad \cdots \quad 0]^T \in \mathbb{R}^{m+1}$.

Proof. See Ipsen [11, Theorem 2.1], as well as [3, 14]. ∎

Another important ingredient, first described in [11] and intensively used in [17, 18], is the so-called spectral factorization of the Krylov matrix $K(A, r_{k-1})$. This factorization is made of three components that encapsulate separately the information on eigenvalues of A, its eigenvectors, and the previous residual vector r_{k-1}.

LEMMA 3. *Let $A \in \mathbb{C}^{n\times n}$ satisfy (1.5). Then the Krylov matrix $K(A, r_{k-1})$, for any $k = 1, 2, \ldots q$, can be factorized as*

$$K(A, r_{k-1}) = VD_{k-1}Z, \tag{2.3}$$

where $d_{k-1} = V^H r_{k-1} \in \mathbb{C}^n$, $D_{k-1} = \operatorname{diag}(d_{k-1}) \in \mathbb{C}^{n\times n}$, and $Z \in \mathbb{C}^{n\times(m+1)}$ is the Vandermonde matrix computed from the eigenvalues of A,

$$Z = [e \quad \Lambda e \quad \cdots \quad \Lambda^m e], \tag{2.4}$$

$e = [1 \quad 1 \quad \cdots \quad 1]^T \in \mathbb{R}^n$.

Proof. Starting from (1.5) and the definition of the Krylov matrix (2.1),

$$\begin{aligned} K\left(A, r_{k-1}\right) &= \left[r_{k-1} \quad Ar_{k-1} \quad \cdots \quad A^m r_{k-1}\right] \\ &= \left[VV^H r_{k-1} \quad V\Lambda V^H r_{k-1} \quad \cdots \quad V\Lambda^m V^H r_{k-1}\right] \\ &= V\left[d_{k-1} \quad \Lambda d_{k-1} \quad \cdots \quad \Lambda^m d_{k-1}\right] \\ &= V\left[D_{k-1}e \quad \Lambda D_{k-1}e \quad \cdots \quad \Lambda^m D_{k-1}e\right] \\ &= VD_{k-1}\left[e \quad \Lambda e \quad \cdots \quad \Lambda^m e\right] = VD_{k-1}Z. \qquad \square \end{aligned}$$

It is clear that the statement of Lemma 3 can be easily generalized to the case of a diagonalizable (nonnormal) matrix A providing that we define $d_{k-1} = V^{-1}r_{k-1}$ in the lemma.

3. The sublinear cycle-convergence of GMRES(m). Along with (1.1) let us consider the system

$$A^H x = b \tag{3.1}$$

with the matrix A replaced by its conjugate transpose. Clearly, according to (1.5),

$$A^H = V\overline{\Lambda}V^H. \tag{3.2}$$

It turns out that m steps of GMRES applied to the systems (1.1) and (3.1) produce residual vectors of equal norms at each step—provided that the initial residual vector is identical. This observation is crucial in concluding the sublinear cycle-convergence of GMRES(m) and is formalized in the following lemma.

LEMMA 4. *Assume that $A \in \mathbb{C}^{n\times n}$ is a nonsingular normal matrix. Let r_m and $\hat{r}_m$ be the nonzero residual vectors obtained by applying m steps of GMRES to the systems* (1.1) *and* (3.1)*; $1 \le m \le n-1$. Then*

$$\|r_m\| = \|\hat{r}_m\|,$$

provided that the initial approximate solutions of (1.1) *and* (3.1) *induce the same initial residual vector r_0.*

Moreover, if $p_m(z)$ and $q_m(z)$ are the (GMRES) polynomials which minimize, respectively, $\|p(A)r_0\|$ and $\|p(A^H)r_0\|$ over $p(z) \in \mathcal{P}_m$, where $\mathcal{P}_m$ is the set of all polynomials of degree at most m defined on the complex plane such that $p(0) = 1$, then

$$\overline{p}_m(z) = q_m(z),$$

where $\overline{p}(z)$ denotes the polynomial obtained from $p(z) \in \mathcal{P}_m$ by the complex conjugation of its coefficients.

Proof. Consider a polynomial $p(z) \in \mathcal{P}_m$. Let r_0 be a nonzero initial residual vector for the systems (1.1) and (3.1) simultaneously. Since the matrix A is normal, so is $p(A)$; thus $p(A)$ commutes with its conjugate transpose $p^H(A)$. We have

$$\begin{aligned} \|p(A)r_0\|^2 &= \langle p(A)r_0, p(A)r_0\rangle = \langle r_0, p^H(A)p(A)r_0\rangle \\ &= \langle r_0, p(A)p^H(A)r_0\rangle = \langle p^H(A)r_0, p^H(A)r_0\rangle \\ &= \langle \left(Vp(\Lambda)V^H\right)^H r_0, \left(Vp(\Lambda)V^H\right)^H r_0\rangle = \langle V\overline{p}(\overline{\Lambda})V^H r_0, V\overline{p}(\overline{\Lambda})V^H r_0\rangle \\ &= \langle \overline{p}(V\overline{\Lambda}V^H)r_0, \overline{p}(V\overline{\Lambda}V^H)r_0\rangle = \|\overline{p}(V\overline{\Lambda}V^H)r_0\|^2, \end{aligned}$$

where $\overline{p}(z)$ is the polynomial obtained from $p(z)$ by conjugating its coefficients. By (3.2) we conclude that

$$\|p(A)r_0\| = \|\overline{p}(A^H)r_0\|.$$

Note that the last statement is true for any polynomial p, for any r_0, and for any normal A.

Now, let us look at $\|r_m\|$ and $\|\hat{r}_m\|$. On the one hand,

$$\begin{aligned}\|r_m\| &= \min_{p\in\mathcal{P}_m} \|p(A)r_0\| = \|p_m(A)r_0\| = \|\overline{p}_m(A^H)r_0\| \\ &\geq \min_{p\in\mathcal{P}_m} \|\overline{p}(A^H)r_0\| = \min_{p\in\mathcal{P}_m} \|p(A^H)r_0\| = \|\hat{r}_m\|.\end{aligned}$$

On the other hand,

$$\begin{aligned}\|\hat{r}_m\| &= \min_{p\in\mathcal{P}_m} \|p(A^H)r_0\| = \|q_m(A^H)r_0\| = \|\overline{q}_m(A)r_0\| \\ &\geq \min_{p\in\mathcal{P}_m} \|\overline{p}(A)r_0\| = \min_{p\in\mathcal{P}_m} \|p(A)r_0\| = \|r_m\|.\end{aligned}$$

Combining both results, we conclude that

$$\|r_m\| = \|\hat{r}_m\|,$$

which proves the first part of the lemma.

To prove the second part of the lemma, consider the following equalities:

$$\begin{aligned}\|q_m(A^H)r_0\| &= \min_{p\in\mathcal{P}_m} \|p(A^H)r_0\| = \|\hat{r}_m\| = \|r_m\| = \min_{p\in\mathcal{P}_m} \|p(A)r_0\| = \|p_m(A)r_0\| \\ &= \|\overline{p}_m(A^H)r_0\|.\end{aligned}$$

By uniqueness of the GMRES polynomial [10, Theorem 2], we conclude that $\overline{p}_m(z) = q_m(z)$. ∎

The previous lemma is a general result for full GMRES which states that, given a nonsingular normal matrix A and an initial residual vector r_0, GMRES applied to A with r_0 produces the same convergence curve as GMRES applied to A^H with r_0. In the framework of restarted GMRES, Lemma 4 implies that the cycles GMRES(A, m, r_{k-1}) and GMRES(A^H, m, r_{k-1}) result in, respectively, residual vectors r_k and $\hat{r}_k$ that have the same norm.

THEOREM 5 (the sublinear cycle-convergence of GMRES(m)). *Let* $\{r_k\}_{k=0}^{q}$ *be a sequence of nonzero residual vectors produced by GMRES(m) applied to the system* (1.1) *with a nonsingular normal matrix* $A \in \mathbb{C}^{n\times n}$, $1 \leq m \leq n-1$. *Then*

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq \frac{\|r_{k+1}\|}{\|r_k\|}, \quad k = 1, \ldots, q-1. \tag{3.3}$$

Proof. Left multiplication of both parts of (2.2) by $K^H(A, r_{k-1})$ leads to

$$e_1 = \frac{1}{\|r_k\|^2} K^H(A, r_{k-1})\, r_k.$$

By (2.3) in Lemma 3, we factorize the Krylov matrix $K(A, r_{k-1})$ in the previous equality:

$$e_1 = \frac{1}{\|r_k\|^2} (VD_{k-1}Z)^H r_k = \frac{1}{\|r_k\|^2} Z^H \overline{D}_{k-1} V^H r_k$$
$$= \frac{1}{\|r_k\|^2} Z^H \overline{D}_{k-1} d_k.$$

Applying complex conjugation to this equality (and observing that e_1 is real), we get

$$e_1 = \frac{1}{\|r_k\|^2} Z^T D_{k-1} \overline{d}_k.$$

According to the definition of D_{k-1} in Lemma 3, $D_{k-1}\overline{d}_k = \overline{D}_k d_{k-1}$; thus

$$e_1 = \frac{1}{\|r_k\|^2} Z^T \overline{D}_k d_{k-1} = \frac{1}{\|r_k\|^2} \left(Z^T \overline{D}_k V^H\right) r_{k-1} = \frac{1}{\|r_k\|^2} \left(VD_k\overline{Z}\right)^H r_{k-1}.$$

From (2.3) and (3.2) we notice that

$$K\left(A^H, r_k\right) = K\left(V\overline{\Lambda}V^H, r_k\right) = VD_k\overline{Z},$$

and so therefore

$$e_1 = \frac{1}{\|r_k\|^2} K^H\left(A^H, r_k\right) r_{k-1}. \tag{3.4}$$

Considering the residual vector r_{k-1} as a solution of the underdetermined system (3.4), we can represent the latter as

$$r_{k-1} = \|r_k\|^2 \left(K^H\left(A^H, r_k\right)\right)^\dagger e_1 + w_k, \tag{3.5}$$

where $w_k \in \text{null}\left(K^H\left(A^H, r_k\right)\right)$. Note that since r_{k+1} is nonzero (assumption in Theorem 5), the residual vector $\hat{r}_{k+1}$ at the end of the cycle GMRES(A^H, m, r_k) is nonzero as well by Lemma 4; hence the corresponding Krylov matrix $K\left(A^H, r_k\right)$ is of the full rank, and thus the pseudoinverse in (3.5) is well defined. Moreover, since

$$w_k \perp \left(K^H\left(A^H, r_k\right)\right)^\dagger e_1,$$

by the Pythagorean theorem we obtain

$$\|r_{k-1}\|^2 = \|r_k\|^4 \| \left(K^H\left(A^H, r_k\right)\right)^\dagger e_1\|^2 + \|w_k\|^2.$$

Now since $(K^H(A^H, r_k))^\dagger = (K^\dagger(A^H, r_k)^H)$, we get

$$\|r_{k-1}\|^2 = \|r_k\|^4 \| \left(K^\dagger\left(A^H, r_k\right)\right)^H e_1\|^2 + \|w_k\|^2,$$

and then by (2.2),

$$= \frac{\|r_k\|^4}{\|\hat{r}_{k+1}\|^2} + \|w_k\|^2$$
$$\geq \frac{\|r_k\|^4}{\|\hat{r}_{k+1}\|^2},$$

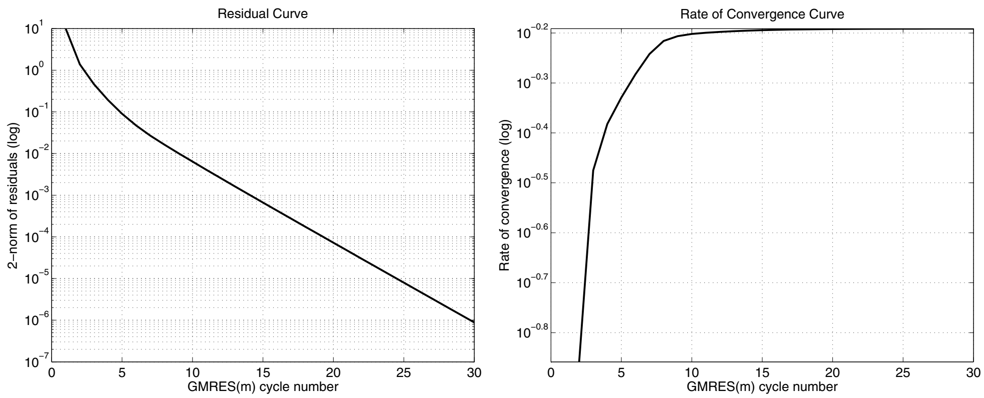

FIG. 1. *Cycle-convergence of GMRES(5) applied to a* 100*-by-*100 *normal matrix.*

where $\hat{r}_{k+1}$ is the residual vector at the end of the cycle GMRES(A^H, m, r_k). Finally,

$$\frac{\|r_k\|^2}{\|r_{k-1}\|^2} \le \frac{\|r_k\|^2\|\hat{r}_{k+1}\|^2}{\|r_k\|^4} = \frac{\|\hat{r}_{k+1}\|^2}{\|r_k\|^2},$$

so that

$$\frac{\|r_k\|}{\|r_{k-1}\|} \le \frac{\|\hat{r}_{k+1}\|}{\|r_k\|}. \tag{3.6}$$

By Lemma 4, the norm of the residual vector $\hat{r}_{k+1}$ at the end of the cycle GMRES(A^H, m, r_k) is equal to the norm of the residual vector r_{k+1} at the end of the cycle GMRES(A, m, r_k), which completes the proof of the theorem. □

Geometrically, Theorem 5 suggests that any residual curve of a restarted GMRES, applied to a system with a nonsingular normal matrix, is nonincreasing and concave up (Figure 1).

From the proof of Theorem 5 it is clear that, for a fixed k, the equality in (3.3) holds if and only if the vector w_k (3.5) from the null space of the corresponding matrix $K^H\left(A^H, r_k\right)$ is zero. In particular, when the restart parameter is chosen to be one less than the problem size, i.e., $m = n-1$, the matrix $K^H\left(A^H, r_k\right)$ in (3.4) becomes an n-by-n nonsingular matrix, hence with a zero null space, and thus the inequality (3.3) is indeed an equality when $m = n-1$.

It turns out that the cycle-convergence of GMRES($n-1$), applied to the system (1.1) with a nonsingular normal matrix A, can be completely determined by norms of the two initial residual vectors r_0 and r_1.

COROLLARY 6 (the cycle-convergence of GMRES($n-1$)). *Suppose that $\|r_0\|$ and $\|r_1\|$ are given. Then, under the assumptions of Theorem* 5*, norms of the residual vectors r_k at the end of each GMRES($n-1$) cycle obey the following formula:*

$$\|r_{k+1}\| = \|r_1\|\left(\frac{\|r_1\|}{\|r_0\|}\right)^k, \quad k = 1, \dots, q-1. \tag{3.7}$$

Proof. The representation (3.5) of the residual vector r_{k-1}, for $m = n-1$, turns into

$$r_{k-1} = \|r_k\|^2\left(K^H\left(A^H, r_k\right)\right)^{-1} e_1, \tag{3.8}$$

implying, by the proof of the Theorem 5, that the equality in (3.3) holds at each GMRES($n-1$) cycle. Thus,

$$\|r_{k+1}\| = \|r_k\| \frac{\|r_k\|}{\|r_{k-1}\|}, \quad k = 1, \ldots, q-1.$$

We show (3.7) by induction in k. Using the formula above, it is easy to verify (3.7) for $\|r_2\|$ and $\|r_3\|$ ($k = 1, 2$). Let us assume that for some k, $3 \le k \le q-1$, $\|r_{k-1}\|$ and $\|r_k\|$ can also be computed by (3.7). Then

$$\|r_{k+1}\| = \|r_k\| \frac{\|r_k\|}{\|r_{k-1}\|} = \|r_1\| \left(\frac{\|r_1\|}{\|r_0\|}\right)^{k-1} \frac{\|r_1\| \left(\frac{\|r_1\|}{\|r_0\|}\right)^{k-1}}{\|r_1\| \left(\frac{\|r_1\|}{\|r_0\|}\right)^{k-2}}$$
$$= \|r_1\| \left(\frac{\|r_1\|}{\|r_0\|}\right)^{k-1} \left(\frac{\|r_1\|}{\|r_0\|}\right) = \|r_1\| \left(\frac{\|r_1\|}{\|r_0\|}\right)^{k}.$$

Thus, (3.7) holds for all $k = 1, \ldots, q-1$. ∎

Another observation in the proof of Theorem 5 leads to a result from Baker, Jessup, and Manteuffel [2]. In this paper, the authors prove that, when GMRES($n-1$) is applied to a system with Hermitian or skew-Hermitian matrix, the residual vectors at the end of each restart cycle alternate direction in a cyclic fashion [2, Theorem 2]. In the following corollary we (slightly) refine this result by providing the exact expression for the constants α_k in [2, Theorem 2].

COROLLARY 7 (the alternating residuals). *Let $\{r_k\}_{k=0}^{q}$ be a sequence of nonzero residual vectors produced by GMRES($n-1$) applied to the system* (1.1) *with a nonsingular Hermitian or skew-Hermitian matrix $A \in \mathbb{C}^{n\times n}$. Then*

$$r_{k+1} = \alpha_k r_{k-1}, \quad \alpha_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2} \in (0, 1], \quad k = 1, 2, \ldots, q-1. \tag{3.9}$$

Proof. For the case of a Hermitian matrix A, i.e., $A^H = A$, the proof follows directly from (3.8) and (2.2).

Let A be skew-Hermitian, i.e., $A^H = -A$. Then, by (3.8) and (2.2),

$$r_{k-1} = \left(K^H\left(A^H, r_k\right)\right)^{-1} e_1 = \left(K^H\left(-A, r_k\right)\right)^{-1} e_1 = \frac{\|r_k\|^2}{\|\hat{r}_{k+1}\|^2}\hat{r}_{k+1},$$

where $\hat{r}_{k+1}$ is the residual vector produced at the end of the cycle GMRES($-A$, $n-1$, r_k).

According to (1.3), the residual vectors r_{k+1} and $\hat{r}_{k+1}$ at the end of the cycles GMRES(A, $n-1$, r_k) and GMRES($-A$, $n-1$, r_k) are obtained by orthogonalizing r_k against the Krylov residual subspaces $A\mathcal{K}_{n-1}(A, r_k)$ and $(-A)\mathcal{K}_{n-1}(-A, r_k)$, respectively. But $(-A)\mathcal{K}_{n-1}(-A, r_k) = A\mathcal{K}_{n-1}(A, r_k)$, and hence $\hat{r}_{k+1} = r_{k+1}$. ∎

4. Note on the departure from normality. In general, for systems with nonnormal matrices, the cycle-convergence behavior of the restarted GMRES is not sublinear. In Figure 2, we consider a nonnormal diagonalizable matrix for the purpose of illustration and one can observe the claim. Indeed, for nonnormal matrices, it has been observed that the cycle-convergence of restarted GMRES can be superlinear [19].

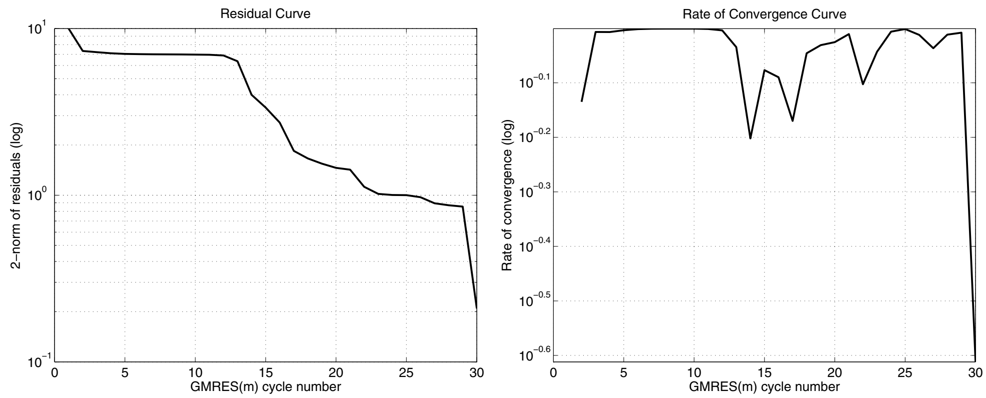

FIG. 2. *Cycle-convergence of GMRES(*5*) applied to a* 100*-by-*100 *diagonalizable (nonnormal) matrix.*

In this concluding section we restrict our attention to the case of a diagonalizable matrix A,

$$A = V\Lambda V^{-1}, \quad A^H = V^{-H}\overline{\Lambda}V^H. \tag{4.1}$$

The analysis performed in Theorem 5 can be generalized for the case of a diagonalizable matrix [17], resulting in the inequality (3.6). However, as we depart from normality, Lemma 4 fails to hold and the norm of the residual vector $\hat{r}_{k+1}$ at the end of the cycle GMRES(A^H, m, r_k) is no longer equal to the norm of the vector r_{k+1} at the end of GMRES(A, m, r_k). Moreover, since the eigenvectors of A can be significantly changed by transpose-conjugation, as (4.1) suggests, the matrices A and A^H can have almost nothing in common, so that the norms of $\hat{r}_{k+1}$ and r_{k+1} are, possibly, far from being equal. This creates an opportunity to break the sublinear convergence of GMRES(m), provided that the subspace $A\mathcal{K}_m(A, r_k)$ results in a better approximation (1.3) of the vector r_k than the subspace $A^H\mathcal{K}_m\left(A^H, r_k\right)$.

It is natural to expect that the convergence of the restarted GMRES for "almost normal" matrices will be "almost sublinear." We quantify this statement in the following lemma.

LEMMA 8. *Let* $\{r_k\}_{k=0}^{q}$ *be a sequence of nonzero residual vectors produced by GMRES(m) applied to the system* (1.1) *with a nonsingular diagonalizable matrix* $A \in \mathbb{C}^{n\times n}$ *as in* (4.1), $1 \le m \le n-1$. *Then*

$$\frac{\|r_k\|}{\|r_{k-1}\|} \le \frac{\alpha\left(\|r_{k+1}\| + \beta_k\right)}{\|r_k\|}, \quad k = 1, \dots, q-1, \tag{4.2}$$

where $\alpha = \frac{1}{\sigma^2_{min}(V)}$, $\beta_k = \|p_k(A)(I - VV^H)r_k\|$, $p_k(z)$ *is the polynomial constructed at the cycle GMRES(A, m,* r_k*), and where q is the total number of GMRES(m) cycles. Note that as* $V^HV \longrightarrow I$, $0 < \alpha \longrightarrow 1$ *and* $0 < \beta_k \longrightarrow 0$.

Proof. Consider the norm of the residual vector $\hat{r}_{k+1}$ at the end of the cycle GMRES(A^H, m, r_k). Then we have

$$\|\hat{r}_{k+1}\| = \min_{\hat{p}\in\mathcal{P}_m} \|\hat{p}(A^H)r_k\| \le \|p(A^H)r_k\|,$$

where $p(z) \in \mathcal{P}_m$ is any polynomial of degree at most m such that $p(0) = 1$. Then, using (4.1),

$$
\begin{aligned}
\|\hat{r}_{k+1}\| &\le \|p(A^H)r_k\| \\
&= \|V^{-H}p(\overline{\Lambda})V^H r_k\| \\
&= \|V^{-H}p(\overline{\Lambda})(V^{-1}V)V^H r_k\| \\
&= \|V^{-H}p(\overline{\Lambda})V^{-1}(VV^H)r_k\| \\
&= \|V^{-H}p(\overline{\Lambda})V^{-1}(I-(I-VV^H))r_k\| \\
&= \|V^{-H}p(\overline{\Lambda})\left(V^{-1}r_k - V^{-1}(I-VV^H)r_k\right)\| \\
&\le \|V^{-H}\|\|p(\overline{\Lambda})\left(V^{-1}r_k - V^{-1}(I-VV^H)r_k\right)\|.
\end{aligned}
$$

Note that

$$\|p(\overline{\Lambda})\left(V^{-1}r_k - V^{-1}(I-VV^H)r_k\right)\| = \|\overline{p}(\Lambda)\left(V^{-1}r_k - V^{-1}(I-VV^H)r_k\right)\|.$$

Thus,

$$
\begin{aligned}
\|\hat{r}_{k+1}\| &\le \|V^{-H}\|\|\overline{p}(\Lambda)\left(V^{-1}r_k - V^{-1}(I-VV^H)r_k\right)\| \\
&= \|V^{-H}\|\|(V^{-1}V)\overline{p}(\Lambda)\left(V^{-1}r_k - V^{-1}(I-VV^H)r_k\right)\| \\
&\le \|V^{-H}\|\|V^{-1}\|\|V\overline{p}(\Lambda)V^{-1}r_k - V\overline{p}(\Lambda)V^{-1}(I-VV^H)r_k\| \\
&= \frac{1}{\sigma^2_{min}(V)}\|\overline{p}(V\Lambda V^{-1})r_k - \overline{p}(V\Lambda V^{-1})(I-VV^H)r_k\| \\
&\le \frac{1}{\sigma^2_{min}(V)}\left(\|\overline{p}(A)r_k\| + \|\overline{p}(A)(I-VV^H)r_k\|\right),
\end{aligned}
$$

where σ_{min} is the smallest singular value of V.

Since the last inequality holds for any polynomial $\overline{p}(z) \in \mathcal{P}_m$, it will also hold for $\overline{p}(z) = p_k(z)$, where $p_k(z)$ is the polynomial constructed at the cycle GMRES(A, m, r_k). Hence,

$$\|\hat{r}_{k+1}\| \le \frac{1}{\sigma^2_{min}(V)}\left(\|r_{k+1}\| + \|p_k(A)(I-VV^H)r_k\|\right).$$

Setting $\alpha = \frac{1}{\sigma^2_{min}(V)}$, $\beta_k = \|p_k(A)(I-VV^H)r_k\|$, and observing that $\alpha \longrightarrow 1$, $\beta_k \longrightarrow 0$ as $V^HV \longrightarrow I$, from (3.6) we obtain (4.2). ∎

REFERENCES

[1] A. H. Baker, E. R. Jessup, and T. Kolev, *A simple strategy for varying the restart parameter in GMRES(m)*, J. Comput. Appl. Math., 230 (2009), pp. 751–761.

[2] A. H. Baker, E. R. Jessup, and T. Manteuffel, *A technique for accelerating the convergence of restarted GMRES*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 962–984.

[3] S. Chandrasekaran and I. C. F. Ipsen, *On the sensitivity of solution components in linear systems of equations*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 93–112.

[4] H. A. van der Vorst and C. Vuik, *The superlinear convergence behaviour of GMRES*, J. Comput. Appl. Math., 48 (1993), pp. 327–341.
[5] M. Eiermann, *Fields of values and iterative methods*, Linear Algebra Appl., 180 (1993), pp. 167–197.
[6] M. Embree, *The tortoise and the hare restart GMRES*, SIAM Rev., 45 (2003), pp. 259–266.
[7] J. Erhel, K. Burrage, and B. Pohl, *Restarted GMRES preconditioned by deflation*, J. Comput. Appl. Math., 69 (1996), pp. 303–318.
[8] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469.
[9] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent Advances in Iterative Methods, Springer, New York, 1994, pp. 95–118.
[10] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.
[11] I. C. F. Ipsen, *Expressions and bounds for the GMRES residual*, BIT, 40 (2000), pp. 524–535.
[12] W. Joubert, *On the convergence behavior of the restarted GMRES algorithm for solving nonsymmetric linear systems*, Numer. Linear Algebra Appl., 1 (1994), pp. 427–447.
[13] 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.
[14] G. W. Stewart, *Collinearity and least squares regression*, Statist. Sci., 2 (1987), pp. 68–109.
[15] K.-C. Toh, *GMRES vs. ideal GMRES*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 30–36.
[16] L. N. Trefethen, *Approximation theory and numerical linear algebra*, in Algorithms for Approximation II, Chapman and Hall, London, UK, 1990, pp. 336–360.
[17] I. Zavorin, *Spectral Factorization of the Krylov Matrix and Convergence of GMRES*, Technical report CS-TR-4309, Computer Science Department, University of Maryland, College Park, MD, 2002.
[18] I. Zavorin, D. P. O'Leary, and H. Elman, *Complete stagnation of GMRES*, Linear Algebra Appl., 367 (2003), pp. 165–183.
[19] B. Zhong and R. B. Morgan, *Complementary cycles of restarted GMRES*, Numer. Linear Algebra Appl., 15 (2008), pp. 559–571.