Notes on the convergence of the restarted GMRES

Eugene Vecharynski and Julien Langou

January 17, 2010

Abstract

The current paper contains a summary of our recent theoretical results on the convergence of restarted GMRES. We prove that the cycle-convergence of restarted GMRES applied to a system of linear equations with a normal coefficient matrix is sublinear (at best linear). In the general case, we design examples for which restarted GMRES cycle-convergence follows any admissible curve, moreover the matrix spectrum can be chosen arbitrarily. The latter result can be viewed as an extension of the result for full GMRES of Greenbaum, Pták, and Strakoš [10] to the case of restarted GMRES.

1 Introduction

The *generalized minimal residual method* (GMRES) was originally introduced by Saad and Schultz [14] in 1986, and has become a popular method for solving non-Hermitian systems of linear equations

$$Ax = b, \quad A \in \mathbb{C}^{n \times n}, \quad b \in \mathbb{C}^n. \tag{1}$$

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) 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.

$$r_i = \min_{u \in \mathcal{K}_i(A, r_0)} \|r_0 - Au\|, \tag{2}$$

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

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

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).

As usual, in a linear setting, a notion of minimality is associated with some orthogonality condition. In our case, the minimization (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 [14].

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, m is called the restart parameter. 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 k-th GMRES(m) cycle (as opposed to the residual vectors r_i (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

$$r_k = \min_{u \in \mathcal{K}_m(A, r_{k-1})} \|r_{k-1} - Au\|, \tag{3}$$

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

$$\mathcal{K}_m(A, r_{k-1}) = \text{span}\{r_{k-1}, Ar_{k-1}, \dots, A^{m-1}r_{k-1}\}. \tag{4}$$

The price paid for the reduction of the computational work in GMRES(m) is the loss of global optimality (2). Although (3) implies a monotonic decrease of the norms of the residual vectors r_k, GMRES(m) can stagnate [14, 22]. This is in contrast with full GMRES which is guaranteed to converge to the exact solution of (1) in n steps (assuming exact arithmetic and nonsingular A). However, a proper choice of a preconditioner or/and a restart parameter, e.g. [7, 8, 13], can significantly accelerate the convergence of GMRES(m), thus making the method practically attractive.

While a lot of efforts have been devoted to the characterization of the convergence of full GMRES, e.g. [18, 4, 10, 11, 12, 16, 17], our understanding of the behavior of GMRES(m) is far from complete, leaving us with more questions than answers, e.g. [7]. The results presented in this paper contribute to the topic.

In Section 3 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 proof of this fact is based mainly on the apparatus developed in Section 2. Note that for certain classes of normal matrices, e.g. (normal) positive definite matrices, our result of *at most* linear cycle-convergence can be combined with available estimates establishing *at least* linear cycle-convergence, e.g. [5, 6, 9], which results in quite a sharp characterization of GMRES(m) behavior.

In Section 4 we show that the cycle-convergence can become superlinear as the coefficient matrix departs from normality.

Section 5 is concerned with the general case. We are able to show that given any admissible cycle-convergence curve at q initial GMRES(m) cycles, $q < n/m$, $1 \le m \le n-1$, and an arbitrary set Λ of n nonzero complex numbers, it is possible to construct a matrix A and a right-hand side vector b, such that GMRES(m) applied to (1) generates the prescribed cycle-convergence curve and the spectrum of A is Λ. This can be viewed as an extension of the well known result [10] to the case of *restarted* GMRES. Although Greenbaum, Pták, and Strakoš laid the path, there are several specific difficulties appearing in the analysis of the *restarted* method which make the extension nontrivial.

Due to the space limitations we had to skip several proofs. In particular, we only present an *outline of the proof* of Theorem 3 in Section 5 which establishes the possibility of any admissible cycle-convergence regardless of the matrix spectrum at a number of initial GMRES(m) cycles. The references that we provide below, however, directly point to the sources with full versions of the missing proofs.

The results presented in this paper were originally introduced in two separate manuscripts: [19] and [20]. The former is to appear in 2008 SISC Copper Mountain Special Issue; the latter is undergoing revision in Numerical Linear Algebra with Applications. The preprints are available online (see the corresponding references). Note that while [19] was in revision, we learned of the independent but related work of Baker, Jessup and Kolev [1] who present one of the main results of this paper (Theorem 1). 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 Section 2 and Section 3 of the current paper we will assume (unless otherwise explicitly stated) A to be nonsingular and normal, i.e. A allows the decomposition

$$A = V\Lambda V^H, \tag{5}$$

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 $(1 \leq m \leq n-1)$, let us denote the k-th cycle of GMRES(m) applied to the system (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 $\mathcal{K}_m\left(A, r_{k-1}\right)$ given in (4). For each $\mathcal{K}_m\left(A, r_{k-1}\right)$ we define a matrix $K\left(A, r_{k-1}\right) \in \mathbb{C}^{n \times (m+1)}$, such that

$$K\left(A, r_{k-1}\right) = \left[r_{k-1} \quad A r_{k-1} \quad \ldots \quad A^m r_{k-1}\right], \quad k = 1, 2, \ldots, q, \tag{6}$$

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

The matrix (6) is called the Krylov matrix. We will say that $K\left(A, r_{k-1}\right)$ corresponds to the cycle GMRES(A, m, r_{k-1}). Note that the columns of $K\left(A, r_{k-1}\right)$ span the next, $(m+1)$–dimensional, Krylov subspace $\mathcal{K}_{m+1}(A, r_{k-1})$. By the assumption that $r_k \neq 0$,

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

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

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

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

Lemma 1 *Given $A \in \mathbb{C}^{n \times n}$ (not necessarily normal) and the full rank Krylov matrix $K\left(A, r_{k-1}\right) \in \mathbb{C}^{n \times (m+1)}$, corresponding to the cycle GMRES(A, m, r_{k-1}) for any $k = 1, 2, \ldots, q$. Then*

$$\left(K^{\dagger}\left(A, r_{k-1}\right)\right)^H e_1 = \frac{1}{\|r_k\|^2} r_k, \tag{7}$$

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

Proof. See Ipsen [12, Theorem 2.1], as well as [3, 15]. □

Another important ingredient, first described in [12] and intensively used in [21, 22], is the so-called spectral factorization of the Krylov matrix $K\left(A, r_{k-1}\right)$. This factorization is made of three components, each one encapsulating separately the information on eigenvalues of A, its eigenvectors and the previous residual vector r_{k-1}.

Lemma 2 *Let $A \in \mathbb{C}^{n \times n}$ satisfy (5). Then the Krylov matrix $K\left(A, r_{k-1}\right)$, for any $k = 1, 2, \ldots q$, can be factorized as*

$$K\left(A, r_{k-1}\right) = V D_{k-1} Z, \tag{8}$$

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

$$Z = \left[e \quad \Lambda e \quad \ldots \quad \Lambda^m e\right], \tag{9}$$

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

Proof. See [19, Lemma 3]. □

It is clear that the statement of Lemma 2 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) let us consider the system

$$A^H x = b \tag{10}$$

with the matrix A replaced by its conjugate transpose. Clearly, according to (5),

$$A^H = V\overline{\Lambda}V^H. \tag{11}$$

It turns out that m steps of GMRES applied to the systems (1) and (10) 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 3 *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) and (10) respectively, $1 \le m \le n-1$. Then*

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

provided that the initial approximate solutions of (1) and (10) induce the same initial residual vector r_0.

Moreover, if $p_m(z)$ and $q_m(z)$ are the (GMRES) polynomials which minimize $\|p(A)r_0\|$ and $\|p(A^H)r_0\|$, respectively, 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. See [19, Lemma 4]. □

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 3 implies that the cycles GMRES(A, m, r_{k-1}) and GMRES(A^H, m, r_{k-1}) result in residual vectors r_k and $\hat{r}_k$, respectively, that have the same norm.

Theorem 1 (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) with a nonsingular normal matrix $A \in \mathbb{C}^{n\times n}$, $1 \le m \le n-1$. Then*

$$\frac{\|r_k\|}{\|r_{k-1}\|} \le \frac{\|r_{k+1}\|}{\|r_k\|}, \quad k = 1, \dots, q-1. \tag{12}$$

Proof. Left multiplication of both parts of (7) 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 (8) in Lemma 2, we factorize the Krylov matrix $K(A, r_{k-1})$ in the equality above:

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

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 2, $D_{k-1}\overline{d}_k = \overline{D}_k d_{k-1}$, thus

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

From (8) and (11) 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

$$e_1 = \frac{1}{\|r_k\|^2} K^H\left(A^H, r_k\right) r_{k-1}. \tag{13}$$

Considering the residual vector r_{k-1} as a solution of the underdetermined system (13), we can represent the latter as

$$r_{k-1} = \|r_k\|^2 \left(K^H\left(A^H, r_k\right)\right)^\dagger e_1 + w_k, \tag{14}$$

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 1), 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 3, hence the corresponding Krylov matrix $K\left(A^H, r_k\right)$ is of the full rank, and thus the pseudoinverse in (14) 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 $\left(K^H\left(A^H, r_k\right)\right)^\dagger = \left(K^\dagger\left(A^H, r_k\right)\right)^H$), we get

$$\begin{aligned}\|r_{k-1}\|^2 &= \|r_k\|^4 \| \left(K^\dagger\left(A^H, r_k\right)\right)^H e_1\|^2 + \|w_k\|^2, \qquad \text{and then by (7),}\\ &= \frac{\|r_k\|^4}{\|\hat{r}_{k+1}\|^2} + \|w_k\|^2\\ &\geq \frac{\|r_k\|^4}{\|\hat{r}_{k+1}\|^2},\end{aligned}$$

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

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

so that

$$\frac{\|r_k\|}{\|r_{k-1}\|} \leq \frac{\|\hat{r}_{k+1}\|}{\|r_k\|}. \tag{15}$$

By Lemma 3, 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, the theorem 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). Below we state several corollaries which follow from Theorem 1.

Corollary 1 (The cycle-convergence of GMRES(m)) *Suppose that $\|r_0\|$ and $\|r_1\|$ are given. Then, under assumptions of the Theorem 1, norms of the residual vectors r_k at the end of each GMRES(m) cycle obey the following inequality*

$$\|r_{k+1}\| \geq \|r_1\| \left(\frac{\|r_1\|}{\|r_0\|}\right)^k, \quad k = 1, \ldots, q-1. \tag{16}$$

Proof. Directly follows from (12). □

The inequality (16) shows that we are able to provide a lower bound for the residual norm at any cycle $k > 1$ after performing only one cycle of GMRES(m), applied to the system (1) with a nonsingular normal matrix A.

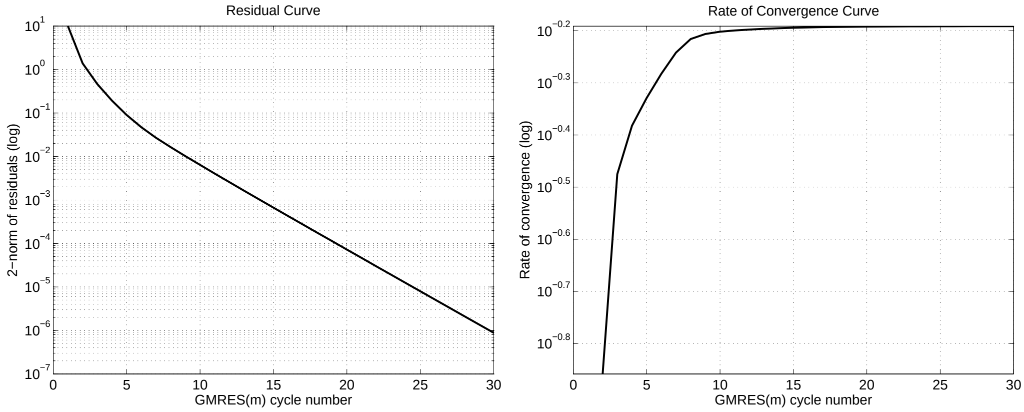

Figure 1: Cycle-convergence of GMRES(5) applied to a 100–by −100 normal matrix.

From the proof of the Theorem 1 it is clear that, for a fixed k, the equality in (12) holds if and only if the vector w_k (14) 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 (13) becomes an n–by–n nonsingular matrix, hence with a zero null space, and thus the inequality (12) is indeed an equality when $m = n-1$.

It turns out that the cycle-convergence of GMRES($n-1$), applied to the system (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 2 (The cycle-convergence of GMRES($n-1$)) *Suppose that $\|r_0\|$ and $\|r_1\|$ are given. Then, under assumptions of the Theorem 1, norms of the residual vectors r_k at the end of each GMRES($n-1$) cycle obey the following formula*

$$\|r_{k+1}\| = \|r_1\| \left(\frac{\|r_1\|}{\|r_0\|}\right)^k, \quad k = 1, \ldots, q-1. \tag{17}$$

Proof. See [19, Corollary 6]. □

Another observation in the proof of the Theorem 1 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 3 (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) with a nonsingular Hermitian or skew-Hermitian matrix $A \in \mathbb{C}^{n\times n}$. Then*

$$r_{k+1} = \alpha_k r_{k-1}, \quad \alpha_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2} \in (0, 1], \quad k = 1, 2, \ldots, q-1. \tag{18}$$

Proof. See [19, Corollary 7]. □

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 illustration purpose and one can observe the claim. Indeed, for nonnormal matrices, it has been observed that the cycle-convergence of restarted GMRES can be superlinear [23].

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

$$A = V\Lambda V^{-1}, \quad A^H = V^{-H}\overline{\Lambda}V^H. \tag{19}$$

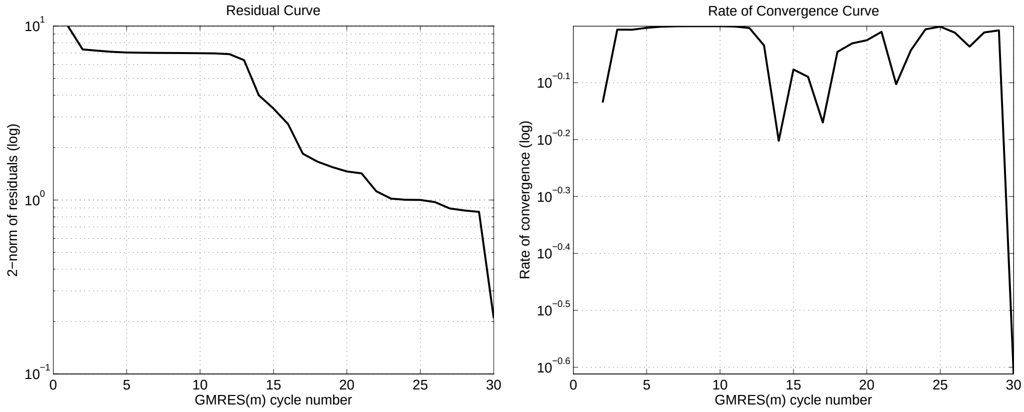

Figure 2: Cycle-convergence of GMRES(5) applied to a 100–by–100 diagonalizable (nonnormal) matrix.

The analysis performed in Theorem 1 can be generalized for the case of a diagonalizable matrix ([21]), resulting in the inequality (15). However, as we depart from normality, Lemma 3 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 (19) 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 gives a chance for breaking the sublinear convergence of GMRES(m), provided that the subspace $A\mathcal{K}_m(A, r_k)$ results in a better approximation (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 4 *Let $\{r_k\}_{k=0}^{q}$ be a sequence of nonzero residual vectors produced by GMRES(m) applied to the system (1) with a nonsingular diagonalizable matrix $A \in \mathbb{C}^{n\times n}$ as in (19), $1 \le m \le n-1$. Then*

$$\frac{\|r_k\|}{\|r_{k-1}\|} \le \frac{\alpha\left(\|r_{k+1}\| + \beta_k\right)}{\|r_k\|}, \quad k = 1, \ldots, q-1, \tag{20}$$

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. See [19, Lemma 8]. □

5 Any admissible cycle-convergence curve is possible for restarted GMRES

In this section we consider the behavior of the restarted GMRES in the general case, i.e. when the method is applied to the system (1) with a nonsingular non-Hermitian matrix A.

For the general case, the following result is available for *full* GMRES:

Theorem 2 (Greenbaum, Pták, and Strakoš, 1996, [10]) *Given a nonincreasing positive sequence $f(0) \ge f(1) \ge \cdots \ge f(n-1) > 0$, there exists an n-by-n matrix A and a vector r_0 with $\|r_0\| = f(0)$ such that $f(k) = \|r_k\|$, $k = 1, \ldots, n-1$, where r_k is the residual at step k of the GMRES algorithm applied to the linear system $Ax = b$, with initial residual $r_0 = b - Ax_0$. Moreover, the matrix A can be chosen to have any desired (nonzero) eigenvalues.*

The theorem claims that any nonincresing convergence curve is possible for *full* GMRES and eigenvalues alone do not determine the convergence. Our work extends Theorem 2 to the case of the *restarted* GMRES as follows

Theorem 3 *Given a matrix order n, a restart parameter m ($1 \le m \le n$), and a positive sequence $\{f(k)\}_{k=0}^{q}$, such that $f(0) > f(1) > \cdots > f(s) > 0$ and $f(s) = f(s+1) = \ldots = f(q)$, where $0 < q < n/m$, $0 \le s \le q$. There exists an n-by-n matrix A and a vector r_0 with $\|r_0\| = f(0)$ such that $\|r_k\| = f(k)$, $k = 1, \ldots, q$, where r_k is the residual at cycle k of restarted GMRES with restart parameter m applied to the linear system $Ax = b$, with initial residual $r_0 = b - Ax_0$. Moreover, the matrix A can be chosen to have any desired (nonzero) eigenvalues.*

Full GMRES has a nonincreasing convergence (for any $i \ge 0$, $f(i) \ge f(i+1)$) and it computes the exact solution in at most n steps ($f(n) = 0$). Note that assumptions on $\{f(i)\}_{i=1}^{n-1}$ in Theorem 2 do not cover the class of convergence curves corresponding to the convergence to the exact solution *before* step n. This can be shown, however, to follow directly (though implicitly) from Theorem 2. In this sense it is remarkable that Greenbaum, Pták, and Strakoš are able to characterize any *admissible* convergence for *full* GMRES.

At the same time we would like to note that the cycle-convergence of *restarted* GMRES can have two *admissible* scenarios: either for any i, $f(i) > f(i+1)$, in other words, the cycle-convergence is decreasing; or there exists s such that $f(i) > f(i+1)$ for any $i < s$, and then for any $i > s$, $f(i) = f(s)$, in other words, if restarted GMRES stagnates at cycle $s+1$, it stagnates forever. Thus assumptions on $\{f(k)\}_{k=0}^{q}$ in Theorem 3 reflect any *admissible* cycle-convergence behavior of *restarted* GMRES at the first q cycles, except for the case when the convergence to the exact solution happens *before* the cycle q. This case turns out to be an immediate implication of Theorem 3.

The detailed proof of Theorem 3, as well as further discussions and generalizations, can be found in [20]. In what follows we briefly give an outline of ideas behind our construction.

Let n be a matrix order and m a restart parameter ($m < n$), $\Lambda = \{\lambda_1, \lambda_2, \ldots \lambda_n\} \subset \mathbb{C} \setminus \{0\}$ be a set of n nonzero complex numbers, and $\{f(k)\}_{k=0}^{q}$ be a positive sequence, such that $f(0) > f(1) > \cdots > f(s) > 0$ and $f(s) = f(s+1) = \ldots = f(q)$, where $0 < q < n/m$, $0 \le s \le q$. In [20] we construct a matrix $A \in \mathbb{C}^{n\times n}$ and an initial residual vector $r_0 = b - Ax_0 \in \mathbb{C}^n$ such that GMRES(m) applied to the system (1) with the initial approximate solution x_0, produces a sequence $\{x_k\}_{k=1}^{q}$ of approximate solutions with corresponding residual vectors $\{r_k\}_{k=0}^{q}$ having the prescribed norms: $\|r_k\| = f(k)$. Moreover the spectrum of A is Λ.

For the purpose of clarity, we restrict our attention only to the case of the strictly decreasing cycle-convergence, i.e. $f(0) > f(1) > \cdots > f(q) > 0$. The (remaining) case of stagnation, i.e. $f(0) > f(1) > \cdots > f(s) > 0$ and $f(s) = f(s+1) = \ldots = f(q)$, $0 \le s < q$ is then handled by a slight change in the proof for the considered case of the strictly decreasing cycle-convergence (see [20]).

The general approach is similar to the approach of Greenbaum, Pták, and Strakoš [10]: we fix an initial residual vector, construct an appropriate basis of $\mathbb{C}^n$ and use this basis to define a linear operator $\mathcal{A}$. This operator is represented by the matrix A in the canonical basis. It has the prescribed spectrum Λ and provides the desired cycle-convergence at the first q cycles of GMRES(m). However, the presence of restarts somewhat complicates the construction: the choice of the basis vectors, as well as the structure of the resulting operator $\mathcal{A}$, become less transparent. Below we briefly describe our *three-step construction*.

At the *first step* we construct q sets of vectors $\mathcal{W}_m^{(k)} = \{w_1^{(k)}, \ldots, w_m^{(k)}\}$, $k = 1, \ldots, q$, each set $\mathcal{W}_m^{(k)}$ is the orthonormal basis of the Krylov residual subspace $A\mathcal{K}_m(A, r_{k-1})$ generated at the k-th GMRES(m) cycle such that

$$\text{span}\ \mathcal{W}_j^{(k)} = A\mathcal{K}_j(A, r_{k-1}), \quad j = 1, \ldots, m. \tag{21}$$

(With this definition, $\mathcal{W}_m^{(k)}$ is defined up to multiplication by a complex number of unit modulus.)

The orthonormal basis $\mathcal{W}_m^{(k)}$ needs to be chosen in order to generate residual vectors r_k with the prescribed (strictly decreasing) norms $f(k)$ at the end of each cycle subject to the additional requirement that the set of $mq+1 (\le n)$ vectors

$$\overline{\mathcal{S}} = \{r_0, w_1^{(1)}, \ldots, w_{m-1}^{(1)}, r_1, w_1^{(2)}, \ldots, w_{m-1}^{(2)}, \ldots, r_{q-1}, w_1^{(q)}, \ldots, w_{m-1}^{(q)}, r_q\} \tag{22}$$

is linearly independent.

Once we have the set $\overline{\mathcal{S}}$, we will complete it to have a basis for $\mathbb{C}^n$. When the number of vectors in $\overline{\mathcal{S}}$ is less than n, a basis $\mathcal{S}$ of $\mathbb{C}^n$ is obtained by completion of $\overline{\mathcal{S}}$ with a set $\widehat{\mathcal{S}}$ of $n - mq - 1$ vectors, i.e. $\mathcal{S} = \{\overline{\mathcal{S}}, \widehat{\mathcal{S}}\}$. This will provide a representation of $\mathbb{C}^n$ as the direct sum

$$\mathbb{C}^n = \text{span}\ \mathcal{S} = \text{span}\{r_0, \mathcal{W}_{m-1}^{(1)}\} \oplus \cdots \oplus \text{span}\{r_{q-1}, \mathcal{W}_{m-1}^{(q)}\} \oplus \text{span}\{r_q, \widehat{\mathcal{S}}\}. \tag{23}$$

The latter translates in terms of Krylov subspaces into

$$\mathbb{C}^n = \text{span } \mathcal{S} = \mathcal{K}_m (A, r_0) \oplus \cdots \oplus \mathcal{K}_m (A, r_{q-1}) \oplus \text{span}\{r_q, \widehat{\mathcal{S}}\}.$$

At the *second step* of our construction, we define a linear operator $\mathcal{A} : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ with spectrum Λ which generates the Krylov residual subspaces in (21) at each GMRES(m) cycle, by its action on the basis vectors $\mathcal{S}$, such that the desired matrix A is the operator $\mathcal{A}$'s representation in the canonical basis. The *third step* accomplishes the construction by a similarity transformation. Most of the paper [20] mainly justifies that the described above pattern indeed allows to prove Theorem 3.

6 Conclusion

In this paper we established several results which address the cycle-convergence behavior of restarted GMRES. We prove that the cycle-convergence of the method applied to a system of linear equations with a normal coefficient matrix is sublinear. In the general case, we show that any cycle-convergence curve is possible for a certain number of initial cycles independently of the spectrum of the coefficient matrix. While our results are theoretical, we foresee several pratical applications. For example, for a normal matrix, based on the fact that the convergence is sublinear, one can predict a lower bound for the minimum computing time required to complete the calculation. In practice, this can enable to allocate more resources to the application if possible or abort it early otherwise.

References

[1] A. H. Baker, E. R. Jessup, and Tz. V. Kolev. A simple strategy for varying the restart parameter in GMRES(m). *Journal of Computational and Applied Mathematics*, 230(2):751–761, 2009.

[2] A. H. Baker, E. R. Jessup, and T. Manteuffel. A Technique for Accelerating the Convergence of Restarted GMRES. *SIAM Journal on Matrix Analysis and Applications*, 26(4):962–984, 2005.

[3] S. Chandrasekaran and I. C. F. Ipsen. On the sensitivity of solution components in linear systems of equations. *SIAM Journal on Matrix Analysis and Applications*, 16(1):93–112, 1995.

[4] M. Eiermann. Fields of values and iterative methods. *Numerical Linear Algebra with Applications*, 180:167–197, 1993.

[5] S. C. Eisenstat, H. C. Elman, and M. H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM Journal on Numerical Analysis*, 20:345–357, 1983.

[6] H. C. Elman. *Iterative methods for large sparse nonsymmetric systems of linear equations.* PhD thesis, Yale University: New Haven, CT, 1982.

[7] M. Embree. The tortoise and the hare restart GMRES. *SIAM Review*, 45(2):259–266, 2003.

[8] J. Erhel, K. Burrage, and B. Pohl. Restarted GMRES preconditioned by deflation. *Journal of Computational and Applied Mathematics*, 69(2):303–318, 1996.

[9] A. Greenbaum. *Iterative Methods for Solving Linear Systems.* SIAM, 1997.

[10] A. Greenbaum, V. Pták, and Z. Strakoš. Any nonincreasing convergence curve is possible for GMRES. *SIAM Journal on Matrix Analysis and Applications*, 17(3):465–469, 1996.

[11] A. Greenbaum and Z. Strakoš. *Matrices that generate the same Krylov residual spaces, in: Recent Advances in Iterative Methods*, pages 95–118. Springer, 1994.

[12] I. C. F. Ipsen. Expressions and bounds for the GMRES residual. *BIT*, 40(3):524–535, 2000.

[13] W. Joubert. On the convergence behavior of the restarted GMRES algorithm for solving nonsymmetric linear systems. *Numerical Linear Algebra with Applications*, 1:427–447, 1994.

[14] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856–869, 1986.

[15] G. W. Stewart. Collinearity and least squares regression. *Statistical Science*, 2(1):68–84, 1987.

[16] K.-C. Toh. GMRES vs. Ideal GMRES. *SIAM Journal on Matrix Analysis and Applications*, 18(1):30–36, 1997.

[17] L. N. Trefethen. *Approximation theory and numerical linear algebra, in Algorithms for Approximation II.* Chapman and Hall, London, U.K., 1990.

[18] H. A. van der Vorst and C. Vuik. The superlinear convergence behaviour of GMRES. *Journal of Computational and Applied Mathematics*, 48(3):327–341, 1993.

[19] E. Vecharynski and J. Langou. The cycle-convergence of restarted GMRES for normal matrices is sublinear. Research Report 270, University of Colorado Denver, Center for Computational Mathematics, 2008. Available online at http://arxiv.org/abs/0806.3260.

[20] E. Vecharynski and J. Langou. Any decreasing cycle-convergence curve is possible for restarted GMRES. Research Report 279, University of Colorado Denver, Center for Computational Mathematics, 2009. Available online at http://arxiv.org/abs/0907.3573.

[21] I. Zavorin. Spectral factorization of the Krylov matrix and convergence of GMRES. Technical report, University of Maryland Computer Science Department, http://hdl.handle.net/1903/1168, 2002.

[22] I. Zavorin, D. P. O'Leary, and H. Elman. Complete stagnation of GMRES. *Linear Algebra and its Applications*, 367:165–183, 2003.

[23] B. Zhong and R. B. Morgan. Complementary cycles of restarted GMRES. *Numerical Linear Algebra with Applications*, 15(6):559–571, 2008.