Any admissible cycle-convergence behavior is possible for restarted GMRES at its initial cycles

Eugene Vecharynski and Julien Langou*, †

Department of Mathematical and Statistical Sciences, University of Colorado Denver, 1250, 14th Street, Denver, CO 80202, U.S.A.

SUMMARY

We show that any admissible cycle-convergence behavior is possible for restarted GMRES at a number of initial cycles, moreover the spectrum of the coefficient matrix alone does not determine this cycle-convergence. The latter can be viewed as an extension of the result of Greenbaum, Pták and Strakoš (*SIAM Journal on Matrix Analysis and Applications* 1996; **17**(3):465–469) to the case of restarted GMRES. Copyright © 2010 John Wiley & Sons, Ltd.

1. INTRODUCTION

We consider the *generalized minimal residual method* (GMRES) and its restarted variant (GMRES(m)) [1] for the solution of a nonsingular non-Hermitian system of linear equations

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

In order to better distinguish between the two methods we will further refer to them as *full* GMRES and *restarted* GMRES, respectively.

For a few classes of matrices, convergence estimates are available for *restarted* GMRES and *full* GMRES. For example, for real positive-definite matrices (that is, for matrices A for which $H = (A + A^T)/2$ is symmetric positive definite, or, equivalently, for matrices A for which $x^T A x > 0$ for any nonzero $x \in \mathbb{R}^n$), Elman's bound [1–4] can be stated as follows:

$$\|r_k\|^2 \leqslant (1-\rho)^k \|r_0\|^2 \quad \text{where } 0 < \rho \equiv (\lambda_{\min}(H)/\|A\|)^2 \leqslant 1.$$

The latter guarantees linear convergence of GMRES(m) for a positive-definite matrix. Improvements and generalizations of this bound can be found in [5–7]. Note that by r_k here we understand a residual vector *at the end of the kth GMRES(m) cycle*. Thus, the above estimate actually characterizes the *cycle-convergence* of restarted GMRES(m).

Definition 1
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.

---

*Correspondence to: Julien Langou, Department of Mathematical and Statistical Sciences, University of Colorado Denver, 1250, 14th Street, Denver, CO 80202, U.S.A.
†E-mail: julien.langou@ucdenver.edu

For normal matrices the convergence of both full and restarted GMRES is well studied. In particular, the convergence of *full* GMRES for normal matrices is known to be governed solely by the spectrum of A [8, 9]. The cycle-convergence of *restarted* GMRES for normal matrices, however, is sublinear [10, 11]. The current paper is concerned with the general case.

In the general case, for full GMRES, the following theorem shows that we cannot prove convergence results based only on the spectrum of the coefficient matrix alone.

Theorem 1 (*Greenbaum*, *Pták and Strakoš, 1996*, [*12*])
Given a nonincreasing positive sequence $f(0) \geqslant f(1) \geqslant \cdots \geqslant f(n-1) > 0$, there exists an $n \times n$ 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.

This result states that, in general, eigenvalues alone do not determine the convergence of *full* GMRES. A complete description of the set of all pairs $\{A, b\}$ for which *full* GMRES applied to (1) generates the prescribed convergence curve while the matrix A has any (nonzero) eigenvalues is given in [13].

Assuming that the coefficient matrix A is diagonalizable, some characterizations of the convergence of full GMRES rely on the condition number of the eigenbasis [9]. Other characterizations of the convergence of full GMRES rely on pseudospectra [14]. More commonly, the field of values is used [1–7]. A discussion on how descriptive some of these bounds are is given by Embree [15].

The main result of this paper is stated as the following

Theorem 2
Given a matrix order n, a restart parameter m ($0 < m < n$) and a positive sequence $\{f(k)\}_{k=0}^{q}$, such that $f(0) > f(1) > \cdots > f(s) > 0$ and $f(s) = f(s+1) = \cdots = f(q)$, where $0 < q < n/m$, $0 \leqslant s \leqslant q$. There exists an $n \times n$ 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.

Section 2 contains a proof of Theorem 2. One can view Theorem 2 as an extension of Theorem 1 to the case of *restarted* GMRES. The proof we provide is constructive and directly inspired by the article of Greenbaum *et al.* [12]. Although Greenbaum, Pták and Strakoš laid the path, there are several specific difficulties ahead in the analysis of the *restarted* method.

Full GMRES has a nonincreasing convergence (for any $i \geqslant 0$, $f(i) \geqslant f(i+1)$) and it computes the exact solution in at most n steps ($f(n) = 0$). Note that assumptions on $\{f(k)\}_{k=1}^{n-1}$ in Theorem 1 do not cover the class of convergence sequences corresponding to the convergence to the exact solution *before* step n. One can see, however, that these assumptions are sufficient to conclude that the theorem holds in the 'missing' case as well. In this sense it is remarkable that Greenbaum, Pták and Strakoš are able to prove that any *admissible* convergence behavior is possible for *full* GMRES at its n steps. 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 (strictly) 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 2 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 *within* these q cycles. It turns out that the assumptions are sufficient to guarantee that Theorem 2 also holds in the above-mentioned case of 'early' convergence. In Section 3 we exactly point out how the assumptions of Theorems 1 and 2 allow to conclude that any *admissible* convergence behavior is possible for *full* and *restarted* GMRES (at its q initial cycles).

As mentioned above, the maximum number of iterations of *full* GMRES is at most n, and the method delivers the exact solution in a finite number of steps. *Restarted* GMRES, however, may never provide the exact solution. It will (hopefully) decrease the residual norm at each cycle, that

is, provide a more and more accurate approximation to the exact solution. With n^2 parameters in A and n parameters in b we are not able to control the convergence for an infinite amount of cycles. For this reason, we consider only the first $q < n/m$ initial GMRES(m) cycles. We note that, in practice, $n \gg m$ so q is relatively large.

In Section 3, we generalize the result given by Theorem 2 for the case of variable restart parameters. The sequence of restart parameters m_k needs to be known *a priori*. We show that GMRES(m_k) can result in any admissible cycle-convergence behavior at the q initial cycles, regardless of the spectrum of the coefficient matrix, where q is such that $\sum_{i=1}^{q} m_k < n$. We note that our construction can be a reasonable tool for generating examples/counter-examples for different strategies for varying the restart parameter.

The cycle-convergence of restarted GMRES for normal matrices is sublinear [10, 11]. However, for general matrices, through Theorem 2, one can expect any admissible cycle-convergence behavior at least at a number of initial steps. In particular, it is possible to construct matrices for which the cycle-convergence of GMRES(m) is fast (e.g. superlinear). This relates to the observations of Zhong and Morgan [16] who report superlinear cycle-convergence for their particular cases of nonnormal matrices, as well as to [11], where it is shown that the cycle-convergence can become superlinear as the coefficient matrix departs from normality.

In a pedagogical paper, Embree [17] presents a 3×3 linear system of equations and attempts to solve it with GMRES(1) and GMRES(2). While GMRES(1) converges to the exact solution in 3 cycles, GMRES(2) (almost) stagnates [18]. Similar observations were also made before in [18]. Our main result, basically, reaffirms these observations in the sense that the increase in the restart parameter (and thus, in the computational complexity at each cycle) does not necessarily imply a faster convergence.

In order to improve the convergence of restarted GMRES, several techniques [19–23] have been proposed which consist of augmenting (or enriching) the Krylov space with eigenvectors or, alternatively, deflating some of the eigenvalues from the spectrum of the original matrix A. The eigenvalues targeted are the closest to zero. These techniques have proved to be effective and the convergence of restarted GMRES is, in practice, greatly improved. Theorem 2 states that, in the general case, eigenvalues alone do not determine the convergence of restarted GMRES at a reasonable, possibly large, number of initial cycles, therefore it is hard to provide a theoretical justification for the choice of removing the eigenvalues closest to zero. A beginning of theoretical understanding has been provided by Zítko [24].

We have generated two Matlab functions that correspond to Theorems 1 and 2 (the case of the strictly decreasing cycle-convergence). Given a matrix size, a restart parameter, a convergence curve and a spectrum, we construct the appropriate matrix and right-hand side. See: http://www.math.ucdenver.edu/~langou/research/.

The main message that we would like our readers to retain from this paper is that in the context of GMRES(m), for a certain number of initial cycles, any cycle-convergence behavior is possible independently of the spectrum of the coefficient matrix.

2. CONSTRUCTIVE PROOF OF THEOREM 2

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 $f(s) = f(s+1) = \cdots = f(q)$, where $0 < q < n/m$, $0 \leqslant s \leqslant q$.

In this Section, 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 first restrict our attention to the case of the strictly decreasing cycle-convergence, and in Section 2.2 we prove Theorem 2 under the assumption that $f(0) > f(1) > \cdots > f(q) > 0$ (i.e. we assume that $s = q$). Next, in Section 2.3 we complete the

proof of Theorem 2 by handling the (remaining) case of stagnation, i.e. $f(0) > f(1) > \cdots > f(s) > 0$ and $f(s) = f(s+1) = \cdots = f(q)$, $0 \leqslant s < q$. This is done by a slight change in the proof for the previously considered case of the strictly decreasing cycle-convergence.

2.1. Outline of the proof

The general approach described in this paper is similar to the approach of Greenbaum *et al.* [12]: we fix an initial residual vector, construct an appropriate basis of $\mathbb{C}^n$ and use this basis to define a linear operator $\mathscr{A}$. This operator is represented by the matrix A in the canonical basis. It has the prescribed spectrum 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 $\mathscr{A}$, become less transparent. Below we briefly describe our *three-step construction for the case of the strictly decreasing cycle-convergence* and then suggest its easy modification to prove the general case which includes stagnation.

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

$$\text{span } \mathscr{W}_j^{(k)} = A \mathscr{K}_j(A, r_{k-1}), \quad j = 1, \ldots, m. \tag{2}$$

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

The orthonormal bases $\mathscr{W}_m^{(k)}$ need to be chosen to ensure generation of 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 $mq + 1 (\leqslant n)$ vectors

$$\overline{\mathscr{S}} = \{r_0, w_1^{(1)}, \ldots, w_{m-1}^{(1)}, r_1, w_1^{(2)}, \ldots, w_{m-1}^{(2)}, \ldots, r_{q-1}, w_1^{(q)}, \ldots, w_{m-1}^{(q)}, r_q\} \tag{3}$$

is linearly independent.

Once we have the set $\overline{\mathscr{S}}$ in (3), we will complete it to have a basis $\mathscr{S}$ for $\mathbb{C}^n$. When the number of vectors in $\overline{\mathscr{S}}$ is less than n, the basis $\mathscr{S}$ is obtained by expanding $\overline{\mathscr{S}}$ with a set $\widehat{\mathscr{S}}$ of $n - mq - 1$ vectors, i.e. $\mathscr{S} = \{\overline{\mathscr{S}}, \widehat{\mathscr{S}}\}$ (otherwise $\mathscr{S} = \overline{\mathscr{S}}$, $\widehat{\mathscr{S}} = \emptyset$). This will provide a representation of $\mathbb{C}^n$ as the direct sum

$$\mathbb{C}^n = \text{span } \mathscr{S} = \text{span}\{r_0, \mathscr{W}_{m-1}^{(1)}\} \oplus \cdots \oplus \text{span}\{r_{q-1}, \mathscr{W}_{m-1}^{(q)}\} \oplus \text{span}\{r_q, \widehat{\mathscr{S}}\}. \tag{4}$$

The latter translates in terms of Krylov subspaces into

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

At the *second step* of our construction, we define a linear operator $\mathscr{A} : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ with spectrum Λ, which generates the Krylov residual subspaces in Equation (2) at each GMRES(m) cycle by its action on the basis vectors $\mathscr{S}$, such that the desired matrix A is the operator $\mathscr{A}$'s representation in the canonical basis. The *third step* accomplishes the construction by a similarity transformation.

In the following subsection we show that this three-step approach indeed allows us to prove Theorem 2 in the case of a strictly decreasing positive sequence $\{f(k)\}_{k=0}^q$. In order to deal with the particular case of stagnation, i.e. $f(0) > f(1) > \cdots > f(s) > 0$ and $f(s) = f(s+1) = \cdots = f(q)$, we keep the same framework but set $q = s + 1$ and redefine the vector r_q (r_q is the last vector in Equation (3)). More details are provided in Section 2.3.

2.2. Proof of Theorem 2 for the case of a strictly decreasing cycle-convergence

Throughout this subsection we let the positive sequence $\{f(k)\}_{k=0}^q$ only to be *strictly decreasing*. We also assume here that $q = \max\{z \in \mathbb{Z} : z < n/m\}$. This means that for the given n and m we perform our construction along the largest number of initial cycles where we are able to come up with A (having a prescribed spectrum) and r_0 which provide the desired cycle-convergence.

Although our proof is formally valid for any $0 < q < n/m$, the assumption emphasizes the extend to which we can take control over the process. Note that any case with $q < \max\{z \in \mathbb{Z} : z < n/m\}$ can be extended to the one assumed above by properly defining a number of additional elements in $\{f(k)\}_{k=0}^{q}$.

2.2.1. *Step 1*: *Construction of a sequence of Krylov subspaces which provide the prescribed cycle-convergence.*

At the kth GMRES(m) cycle, the residual vector r_k satisfies the following minimality condition:

$$\|r_k\| = \min_{u \in A\mathcal{K}_m(A, r_{k-1})} \|r_{k-1} - u\|. \tag{5}$$

We assume that each set $\mathscr{W}_m^{(k)}$ is an orthonormal basis of a corresponding Krylov residual subspace $A\mathcal{K}_m(A, r_{k-1})$, therefore the condition (5) implies

$$r_k = r_{k-1} - \sum_{j=1}^{m} \langle r_{k-1}, w_j^{(k)} \rangle w_j^{(k)}, \quad k = 1, \ldots, q. \tag{6}$$

At this stage, in order to simplify the forthcoming justification of the linear independence of the set $\overline{\mathscr{S}}$ in Equation (3), we impose a stricter requirement on the residual change inside the cycle. We will require that the residual vector r_{k-1} remains constant during the first $m-1$ inner steps of GMRES and is reduced only at the last, mth, step. Thus, the equality in (6) can be written as

$$r_k = r_{k-1} - \langle r_{k-1}, w_m^{(k)} \rangle w_m^{(k)}, \quad k = 1, \ldots, q. \tag{7}$$

This implies that the vectors $w_j^{(k)}$, $j = 1, \ldots, m-1$, are orthogonal to the residual vector r_{k-1}, i.e.

$$\langle r_{k-1}, w_j^{(k)} \rangle = 0, \quad j = 1, \ldots, m-1, \quad k = 1, \ldots, q. \tag{8}$$

From Equation (7), using the fact that $r_k \perp w_m^{(k)}$ and the Pythagorean theorem, we obtain

$$|\langle r_{k-1}, w_m^{(k)} \rangle| = \sqrt{\|r_{k-1}\|^2 - \|r_k\|^2}, \quad k = 1, \ldots, q.$$

By defining angles $\psi_k \equiv \angle(r_{k-1}, w_m^{(k)})$ and cosines of the corresponding angles $\cos \psi_k \equiv |\langle r_{k-1}, w_m^{(k)} \rangle| / \|r_{k-1}\|$, we can equivalently rewrite the identity above in the following form:

$$\cos \psi_k = \frac{\sqrt{f(k-1)^2 - f(k)^2}}{f(k-1)} \in (0, 1), \quad k = 1, \ldots, q, \tag{9}$$

where $f(k-1)$ and $f(k)$ are the prescribed values for the norm of the residual vectors r_{k-1} and r_k, respectively. Thus, if we are given r_{k-1}, one way to ensure the desired cycle-convergence at cycle k of GMRES(m) is to choose the unit vectors $w_j^{(k)}$ such that (7)–(9) holds.

In the following lemma, we show constructively that the described approach (7)–(9) leads to an appropriate set $\overline{\mathscr{S}}$.

Lemma 1
Given a strictly decreasing positive sequence $\{f(k)\}_{k=0}^{q}$ and an initial vector r_0, $\|r_0\| = f(0)$, there exist vectors r_k, $\|r_k\| = f(k)$ and orthonormal sets $\mathscr{W}_m^{(k)}$ such that Equations (7), (8) and (9) hold, and the set $\overline{\mathscr{S}}$ in Equation (3) is linearly independent, $k = 1, \ldots, q$.

Proof
The proof is by induction.

Let $k = 1$. Given the initial vector r_0, $\|r_0\| = f(0)$, we pick $\mathscr{W}_{m-1}^{(1)} = \{w_1^{(1)}, \ldots, w_{m-1}^{(1)}\}$ an orthonormal set in $r_0^{\perp}$ in order to satisfy Equation (8). The set $\{r_0, \mathscr{W}_{m-1}^{(1)}\}$ is linearly independent.

In order to choose the unit vector $w_m^{(1)}$ orthogonal to the previously constructed vectors $\mathscr{W}_{m-1}^{(1)}$ and to satisfy Equation (9), we introduce a unit vector $y^{(1)} \in \{r_0, \mathscr{W}_{m-1}^{(1)}\}^{\perp}$, so that

$$w_m^{(1)} = \frac{r_0}{f(0)} \cos \psi_1 + y^{(1)} \sin \psi_1.$$

We find the vector r_1 by satisfying Equation (7). Equation (9) guarantees that $\|r_1\| = f(1)$, as desired. Finally, we append the constructed vector r_1 to $\{r_0, \mathscr{W}_{m-1}^{(1)}\}$ and get the set $\{r_0, \mathscr{W}_{m-1}^{(1)}, r_1\}$, which is linearly independent, since, by construction, r_1 is not in span$\{r_0, \mathscr{W}_{m-1}^{(1)}\}$.

The induction assumption is that we have constructed $k-1$ vectors $r_1, \dots, r_{k-1}$ with the prescribed norms $f(1), \dots, f(k-1)$ and orthonormal sets $\mathscr{W}_m^{(1)}, \dots, \mathscr{W}_m^{(k-1)}$, such that the equalities (7), (8) and (9) hold, and the set

$$\{r_0, \mathscr{W}_{m-1}^{(1)}, \dots, r_{k-2}, \mathscr{W}_{m-1}^{(k-1)}, r_{k-1}\} \tag{10}$$

is linearly independent. We want to show that we can construct the next vector r_k, $\|r_k\| = f(k)$, and the orthonormal set $\mathscr{W}_m^{(k)}$, satisfying Equation (7), (8) and (9), such that

$$\{r_0, \mathscr{W}_{m-1}^{(1)}, \dots, r_{k-2}, \mathscr{W}_{m-1}^{(k-1)}, r_{k-1}, \mathscr{W}_{m-1}^{(k)}, r_k\} \tag{11}$$

is linearly independent, $k \leqslant q$.

We start by constructing orthonormal vectors $\mathscr{W}_{m-1}^{(k)} = \{w_1^{(k)}, \dots, w_{m-1}^{(k)}\}$, satisfying Equation (8), with the additional requirement that the set $\mathscr{W}_{m-1}^{(k)}$ is not in the span of the previously constructed vectors given in the set (10). From these considerations we choose $\mathscr{W}_{m-1}^{(k)}$ as an orthonormal set in the orthogonal complement of (10), i.e.

$$w_j^{(k)} \in \{r_0, \mathscr{W}_{m-1}^{(1)}, \dots, r_{k-2}, \mathscr{W}_{m-1}^{(k-1)}, r_{k-1}\}^{\perp}, \quad j = 1, \dots, m-1.$$

Appending $\mathscr{W}_{m-1}^{(k)}$ to the set (10) will give a linearly independent set.

To finish the proof, we need to construct the vector $w_m^{(k)}$, satisfying Equation (9) and orthogonal to $\mathscr{W}_{m-1}^{(k)}$. For this reason we introduce a unit vector $y^{(k)}$,

$$y^{(k)} \in \{r_0, \mathscr{W}_{m-1}^{(1)}, \dots, r_{k-2}, \mathscr{W}_{m-1}^{(k-1)}, r_{k-1}, \mathscr{W}_{m-1}^{(k)}\}^{\perp},$$

so that

$$w_m^{(k)} = \frac{r_{k-1}}{f(k-1)} \cos \psi_k + y^{(k)} \sin \psi_k.$$

We define the vector r_k by (7). Equation (9) guarantees $\|r_k\| = f(k)$. The set (11) is linearly independent, since, by construction, the vector r_k is not in span $\{r_0, \mathscr{W}_{m-1}^{(1)}, \dots, r_{k-2}, \mathscr{W}_{m-1}^{(k-1)}, r_{k-1}, \mathscr{W}_{m-1}^{(k)}\}$. $\qquad \square$

2.2.2. Step 2: *Definition of a linear operator with any prescribed spectrum.* So far we have shown that, given a strictly decreasing positive sequence $\{f(k)\}_{k=0}^q$ and an initial residual vector r_0, $\|r_0\| = f(0)$, it is possible to construct vectors r_k, $\|r_k\| = f(k)$, and orthonormal vectors $\mathscr{W}_m^{(k)}$, $k = 1, \dots, q$, satisfying Equations (7), (8) and (9), such that the set $\overline{\mathscr{S}}$ of $mq + 1$ vectors in Equation (3) is linearly independent.

In order to define a unique linear operator, we need to have a valid basis of $\mathbb{C}^n$ at hand. Thus, we expand the set $\overline{\mathscr{S}}$ by linearly independent vectors $\widehat{\mathscr{S}} = \{\widehat{s}_1, \dots, \widehat{s}_t\}$, $t = n - mq - 1$ ($< m$, since we assumed that $q = \max\{z \in \mathbb{Z} : z < n/m\}$):

$$\mathscr{S} = \{r_0, \mathscr{W}_{m-1}^{(1)}, \dots, r_{q-1}, \mathscr{W}_{m-1}^{(q)}, r_q, \widehat{s}_1, \dots, \widehat{s}_t\}, \tag{12}$$

so that $\mathscr{S}$ is a basis of $\mathbb{C}^n$.

Before we define a linear operator $\mathscr{A}$, let us consider the set $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ of nonzero numbers in the complex plane that will define $\mathscr{A}$'s spectrum. We split Λ into $q+1$ disjoint subsets

$$\Lambda = \{\Lambda_1, \Lambda_2, \ldots, \Lambda_q, \Lambda_{q+1}\}, \tag{13}$$

such that each Λ_k, $k=1, \ldots, q$, contains m elements of Λ, and the remaining $n-mq$ elements are included into Λ_{q+1}.

For each set Λ_k, $k=1, \ldots, q$, we define a monic polynomial $p_k(x)$, such that the roots of this polynomial are exactly the elements of the corresponding Λ_k

$$p_k(x) = x^m - \sum_{j=0}^{m-1} \alpha_j^{(k)} x^j, \quad k=1, \ldots, q, \tag{14}$$

with $\alpha_j^{(k)}$'s being the coefficients of the respective polynomials, $\alpha_0^{(k)} \neq 0$. Each polynomial $p_k(x)$ in Equation (14) can be thought of as the characteristic polynomial of an $m \times m$ matrix with spectrum Λ_k.

Let us also introduce an arbitrary $(t+1) \times (t+1)$ matrix C with the spectrum Λ_{q+1}:

$$C = (\beta_{ij}), \quad \Lambda(C) = \Lambda_{q+1}, \quad i, j = 1, \ldots, t+1 = n-mq. \tag{15}$$

We define the operator $\mathscr{A} : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ as follows:

$$
\begin{aligned}
\mathscr{A} r_{k-1} &= w_1^{(k)}, \\
\mathscr{A} w_1^{(k)} &= w_2^{(k)}, \\
&\vdots \\
\mathscr{A} w_{m-2}^{(k)} &= w_{m-1}^{(k)}, \\
\mathscr{A} w_{m-1}^{(k)} &= -\alpha_0^{(k)} r_k + \alpha_0^{(k)} r_{k-1} + \alpha_1^{(k)} w_1^{(k)} + \cdots + \alpha_{m-1}^{(k)} w_{m-1}^{(k)}, \quad k=1, \ldots q \\
\mathscr{A} r_q &= \beta_{11} r_q + \beta_{21} \widehat{s_1} + \cdots + \beta_{t+1,1} \widehat{s_t}, \\
\mathscr{A} \widehat{s_1} &= \beta_{12} r_q + \beta_{22} \widehat{s_1} + \cdots + \beta_{t+1,2} \widehat{s_t}, \\
&\vdots \\
\mathscr{A} \widehat{s_t} &= \beta_{1,t+1} r_q + \beta_{2,t+1} \widehat{s_1} + \cdots + \beta_{t+1,t+1} \widehat{s_t},
\end{aligned}
\tag{16}
$$

where $\alpha_j^{(k)}$'s are the coefficients of polynomials (14) and β_{ij}'s are the elements of the matrix C in Equation (15).

The following Lemma shows that, given vectors r_k and orthonormal sets $\mathscr{W}_m^{(k)}$ constructed according to Lemma 1, the linear operator $\mathscr{A}$, defined by (16) and represented by a matrix A in the canonical basis, generates the desired Krylov residual subspaces given in Equation (2); and the spectrum of $\mathscr{A}$ can be arbitrarily chosen.

Lemma 2
Let the initial residual vector r_0, $\|r_0\| = f(0)$, as well as the residual vectors r_k and orthonormal sets $\mathscr{W}_m^{(k)}$ be constructed according to Lemma 1. Let $\mathscr{S}$ be the basis of $\mathbb{C}^n$ as defined by Equation (12) and Λ be an arbitrary set of n nonzero complex numbers. Then the linear operator $\mathscr{A}$ defined according to Equation (13)–(16) generates the Krylov residual subspaces given in Equation (2), where the matrix A is a representation of $\mathscr{A}$ in the canonical basis. Moreover, the spectrum of $\mathscr{A}$ is Λ.

Proof
From the definition (16) of the linear operator $\mathscr{A}$ one can notice that

$$
\begin{aligned}
\mathscr{A} r_{k-1} &= w_1^{(k)}, \\
\mathscr{A}^2 r_{k-1} &= w_2^{(k)},
\end{aligned}
$$

$$\vdots$$

$$\mathscr{A}^{m-1}r_{k-1}=w_{m-1}^{(k)},$$

$$\mathscr{A}^m r_{k-1}=-\alpha_0^{(k)}r_k+\alpha_0^{(k)}r_{k-1}+\alpha_1^{(k)}w_1^{(k)}+\cdots+\alpha_{m-1}^{(k)}w_{m-1}^{(k)}, \quad k=1,\ldots q.$$

Since, by construction in Lemma 1, i.e. Equation (7), $0\neq-\alpha_0^{(k)}r_k+\alpha_0^{(k)}r_{k-1}\in\mathrm{span}\{w_m^{(k)}\}$, the above relations immediately imply that for each $k=1,\ldots q$,

$$\mathrm{span}\{\mathscr{A}r_{k-1},\ldots,\mathscr{A}^j r_{k-1}\}=\mathrm{span}\,\mathscr{W}_j^{(k)}, \quad j=1,\ldots,m.$$

Thus, given the representation A of the linear operator $\mathscr{A}$ in the canonical basis, we proved that $\mathscr{A}$ generates the Krylov residual subspaces given in Equation (2).

To prove that the arbitrarily chosen set Λ is the spectrum of $\mathscr{A}$, let us consider the matrix $[\mathscr{A}]_{\mathscr{S}}$ of the operator $\mathscr{A}$ in the basis $\mathscr{S}$ (see Equation (12) and (16)):

$$[\mathscr{A}]_{\mathscr{S}}=\begin{bmatrix} 0 & 0 & \ldots & 0 & \alpha_0^{(1)} & & & & & & & \\ 1 & 0 & \ldots & 0 & \alpha_1^{(1)} & & & & & & & \\ 0 & 1 & \ldots & 0 & \alpha_2^{(1)} & & & & & & & \\ \vdots & \vdots & \ddots & \vdots & \vdots & & & & \mathbf{0} & & & \\ 0 & 0 & \ldots & 1 & \alpha_{m-1}^{(1)} & & & & & & & \\ & & & & -\alpha_0^{(1)} & \ddots & & & & & & \\ & & & & & \ddots & \ddots & & & & & \\ & & & & & & -\alpha_0^{(q-1)} & 0 & 0 & \ldots & 0 & \alpha_0^{(q)} \\ & & & & & & & 1 & 0 & \ldots & 0 & \alpha_1^{(q)} \\ & & & & & & & 0 & 0 & \ldots & 0 & \alpha_2^{(q)} \\ & & & & & & & \vdots & \vdots & \ddots & \vdots & \vdots \\ & & \mathbf{0} & & & & & 0 & 0 & \ldots & 1 & \alpha_{m-1}^{(q)} \\ & & & & & & & -\alpha_0^{(q)} & \beta_{11} & \ldots & & \beta_{1,t+1} \\ & & & & & & & & \vdots & \ddots & & \vdots \\ & & & & & & & & \beta_{t+1,1} & \ldots & & \beta_{t+1,t+1} \end{bmatrix}. \quad (17)$$

The matrix $[\mathscr{A}]_{\mathscr{S}}$ has a block lower triangular structure, hence $[\mathscr{A}]_{\mathscr{S}}$'s spectrum is the union of the eigenvalues of all diagonal blocks. The first q blocks are the companion matrices corresponding to the sets Λ_k, $k=1,\ldots,q$, with characteristic polynomials defined in Equation (14). The last block is exactly the matrix $C=(\beta_{ij})$ from Equation (15) with the spectrum Λ_{q+1}. Thus, given the partition (13) of the set Λ, we conclude that the spectrum of $\mathscr{A}$ is Λ. $\qquad\square$

2.2.3. Step 3: Conclusion of the proof of Theorem 2 for the case of the strictly decreasing cycle-convergence. Finally, we define A as the representation of the operator $\mathscr{A}$ in the canonical

basis $\mathscr{E} = \{e_1, e_2, \ldots, e_n\}$,

$$A = S[\mathscr{A}]_{\mathscr{S}} S^{-1}, \tag{18}$$

where the square matrix S is formed by the vectors given in Equation (12) written as columns and $[\mathscr{A}]_{\mathscr{S}}$ is defined by Equation (17). The constructed matrix A provides the prescribed (strictly decreasing) norms of residual vectors at the first q GMRES(m) cycles when starting with r_0 and its spectrum is Λ. Note that the field over which the resulting matrix is defined depends heavily on the partition (13) of the set Λ, e.g. A turns out to be (non-Hermitian) complex when a conjugate pair from Λ is not included into the same subset Λ_k.

2.3. Extension to the case of stagnation

In Sections 2.2.1–2.2.3 we proved Theorem 2 only for the case of the strictly decreasing positive sequence $\{f(k)\}_{k=0}^q$. Now, in order to conclude the rest of Theorem 2, we will consider the case of stagnation: $f(0) > f(1) > \cdots > f(s) > 0$ and $f(s) = f(s+1) = \cdots = f(q)$. The latter fits well (after a minor modification) into the framework presented above.

Let us set $q = s+1$ and without loss of generality reduce the problem to constructing a matrix A with a spectrum Λ and an initial residual vector r_0, $\|r_0\| = f(0)$, for which GMRES(m) produces the following sequence of residual norms: $f(0) > f(1) > \cdots > f(q-1) = f(q) > 0$. We observe that the sequence is strictly decreasing up to the last cycle q. Thus, by Lemma 1, at the initial $q-1(=s)$ cycles we are able to construct sets $\mathscr{W}_m^{(k)}$ and vectors r_k, such that $\|r_k\| = f(k)$ and the set

$$\{r_0, \mathscr{W}_{m-1}^{(1)}, \ldots, r_{q-1}, \mathscr{W}_{m-1}^{(q-1)}, r_{q-1}\} \tag{19}$$

is linearly independent. Then, formally following the construction in Lemma 1 at the cycle q, we come up with an orthonormal set $\mathscr{W}_m^{(q)}$ from the orthogonal complement of (19) and the residual vector $r_q = r_{q-1}$. This leads to the set (3), which is no longer linearly independent due to the above-mentioned equality of residual vectors. To enforce the linear independence we substitute in (3) the 'inconvenient' vector r_q by $w_m^{(q)} + r_{q-1}$ and obtain the set

$$\{r_0, \mathscr{W}_{m-1}^{(1)}, \ldots, r_{q-2}, \mathscr{W}_{m-1}^{(q-1)}, r_{q-1}, \mathscr{W}_m^{(q)}, w_m^{(q)} + r_{q-1}\}, \tag{20}$$

which is linearly independent, due to the fact that the orthonormal set $\mathscr{W}_m^{(q)}$ is chosen, by construction, from the orthogonal complement of (19).

The rest of the proof exactly follows the pattern described in Sections 2.2.1–2.2.3 with r_q replaced by $w_m^{(q)} + r_{q-1}$, $q = s+1$; see (12)–(18). The resulting matrix A has the prescribed spectrum Λ and with the initial residual vector r_0, $\|r_0\| = f(0)$, provides the desired cycle-convergence of GMRES(m) with a stagnation starting after cycle s.

This concludes the proof of Theorem 2.

3. REMARKS AND GENERALIZATIONS

In this section, we present several remarks and generalizations related to the main result of the paper.

3.1. Difference with the work of Greenbaum et al. [12]

For the reader familiar with the work of Greenbaum *et al.* [12], it might be tempting to obtain the present result by pursuing the following scheme: fix r_0 and then consider the first restarted GMRES cycle as the initial part of a full GMRES run where the convergence is prescribed for the first m iterations (and set arbitrarily for the remaining $n-m$ iterations). Then, similarly, given the starting residual vector r_1 provided by this first cycle, construct the next Krylov residual subspace which provides the desired convergence following the scheme of Greenbaum *et al.* [12]. Proceed identically for the remaining cycles. This approach, however, does not guarantee the linear

independence of the set $\overline{\mathscr{S}}$ in Equation (3) and, hence, one meets the problem of defining the linear operator $\mathscr{A}$. These considerations were the reason for Assumptions (7), (8) on the residual reduction inside a cycle, which allowed to quite easily justify the linear independence of the set $\overline{\mathscr{S}}$ and, as well, to control the spectrum.

3.2. Generating examples with nonzero r_{q+1}

In Section 2.2, when constructing a matrix A and an initial residual vector r_0 which provide the prescribed strictly decreasing cycle-convergence of GMRES(m), we note from the definition (16) of the operator $\mathscr{A}$ that span$\{r_q, \widehat{s}_1, \ldots, \widehat{s}_t\}$ is an invariant subspace of $\mathscr{A}$ and, hence,

$$r_q \in A\mathscr{K}_{t+1}(A, r_q),$$

where A is the representation of the operator $\mathscr{A}$ in the canonical basis and $t = n - mq - 1$ ($<m$, by the assumption that $q = \max\{z \in \mathbb{Z} : z < n/m\}$ at the beginning of Section 2.2). This implies that at the end of the $(q+1)$th cycle GMRES(m) converges to the exact solution of Equation (1), i.e. $r_{q+1} = 0$. This fact might seem unnatural and undesirable, e.g. for constructing theoretical examples. The 'drawback', however, can be easily fixed by a slight correction of the basis $\mathscr{S}$ in Equation (12)—somewhat similarly to how we handled the stagnation case in Theorem 2.

Given residuals r_k and orthonormal sets $\mathscr{W}_m^{(k)}$ constructed according to Lemma 1, instead of considering the set $\mathscr{S}$, we consider the following basis of $\mathbb{C}^n$:

$$\tilde{\mathscr{S}} = \{r_0, w_1^{(1)}, \ldots, w_{m-1}^{(1)}, \ldots, r_{q-1}, w_1^{(q)}, \ldots, w_{m-1}^{(q)}, r_q + \gamma r_{q-1}, \widehat{s}_1, \ldots, \widehat{s}_t\}, \tag{21}$$

where $\gamma \neq -1, 0$. Here, we substituted the basis vector r_q in Equation (12) by $r_q + \gamma r_{q-1}$. The vector $r_q + \gamma r_{q-1}$ cannot be represented as a linear combination of other vectors in $\tilde{S}$, since it contains the component r_q, which is not represented by these vectors. Hence, $\tilde{\mathscr{S}}$ is indeed a basis of $\mathbb{C}^n$. Thus, we can define the operator $\mathscr{A}$ by its action on $\tilde{\mathscr{S}}$

$$\begin{aligned}
\mathscr{A}r_{k-1} &= w_1^{(k)}, \\
\mathscr{A}w_1^{(k)} &= w_2^{(k)}, \\
&\vdots \\
\mathscr{A}w_{m-2}^{(k)} &= w_{m-1}^{(k)}, \\
\mathscr{A}w_{m-1}^{(k)} &= -\alpha_0^{(k)} r_k + \alpha_0^{(k)} r_{k-1} + \alpha_1^{(k)} w_1^{(k)} + \cdots + \alpha_{m-1}^{(k)} w_{m-1}^{(k)}, \quad k=1,\ldots,q-1, \\
\mathscr{A}r_{q-1} &= w_1^{(q)}, \\
\mathscr{A}w_1^{(q)} &= w_2^{(q)}, \\
&\vdots \\
\mathscr{A}w_{m-2}^{(q)} &= w_{m-1}^{(q)}, \\
\mathscr{A}w_{m-1}^{(q)} &= \frac{-\alpha_0^{(q)}}{1+\gamma}(r_q + \gamma r_{q-1}) + \alpha_0^{(q)} r_{q-1} + \alpha_1^{(q)} w_1^{(q)} + \cdots + \alpha_{m-1}^{(q)} w_{m-1}^{(q)}, \\
\mathscr{A}(r_q + \gamma r_{q-1}) &= \beta_{11}(r_q + \gamma r_{q-1}) + \beta_{21}\widehat{s}_1 + \cdots + \beta_{t+1,1}\widehat{s}_t, \\
\mathscr{A}\widehat{s}_1 &= \beta_{12}(r_q + \gamma r_{q-1}) + \beta_{22}\widehat{s}_1 + \cdots + \beta_{t+1,2}\widehat{s}_t, \\
&\vdots \\
\mathscr{A}\widehat{s}_t &= \beta_{1,t+1}(r_q + \gamma r_{q-1}) + \beta_{2,t+1}\widehat{s}_1 + \cdots + \beta_{t+1,t+1}\widehat{s}_t,
\end{aligned} \tag{22}$$

where $\alpha_j^{(k)}$'s are the coefficients of the corresponding characteristic polynomials (14) and β_{ij}'s are the elements of the matrix C in Equation (15). The fact that the operator $\mathscr{A}$ produces the correct Krylov residual subspace at the cycle q, i.e.

$$\operatorname{span}\{\mathscr{A}r_{q-1},\ldots,\mathscr{A}^m r_{q-1}\}=\operatorname{span}\ \mathscr{W}_m^{(q)},$$

can be observed from the following equalities:

$$\mathscr{A}w_{m-1}^{(q)}=\frac{-\alpha_0^{(q)}}{1+\gamma}(r_q+\gamma r_{q-1})+\alpha_0^{(q)}r_{q-1}+\alpha_1^{(q)}w_1^{(q)}+\cdots+\alpha_{m-1}^{(q)}w_{m-1}^{(q)}$$

$$=\frac{-\alpha_0^{(q)}}{1+\gamma}(r_q-r_{q-1}+(1+\gamma)r_{q-1})+\alpha_0^{(q)}r_{q-1}+\alpha_1^{(q)}w_1^{(q)}+\cdots+\alpha_{m-1}^{(q)}w_{m-1}^{(q)}$$

$$=\frac{-\alpha_0^{(q)}}{1+\gamma}(r_q-r_{q-1})+\alpha_1^{(q)}w_1^{(q)}+\cdots+\alpha_{m-1}^{(q)}w_{m-1}^{(q)},$$

where, by Equation (22), $\mathscr{A}w_{m-1}^{(q)}=\mathscr{A}^m r_{q-1}$ and, by Equation (7), $0\neq r_q-r_{q-1}\in\operatorname{span}\{w_m^{(q)}\}$.

The matrix $[\mathscr{A}]_{\tilde{\mathscr{S}}}$ of the operator $\mathscr{A}$, defined by Equation (22), in the basis $\tilde{\mathscr{S}}$ is identical to Equation (17) with the only change of the subdiagonal element $-\alpha_0^{(q)}$ to $-\alpha_0^{(q)}/(1+\gamma)$, $\gamma\neq-1,0$. Hence, $\mathscr{A}$ has the desired spectrum Λ. The representation A of the operator $\mathscr{A}$ in the canonical basis is then determined by the similarity transformation in Equation (18), with the matrix S formed by vectors from $\tilde{\mathscr{S}}$ in Equation (21) written as columns.

Finally, to see that with the new definition of the operator $\mathscr{A}$, the residual vector r_{q+1} is generally nonzero, we notice from Equation (22) that now $\operatorname{span}\{r_q+\gamma r_{q-1},\widehat{s_1},\ldots,\widehat{s_t}\}$ is an invariant subspace of $\mathscr{A}$ and, hence,

$$r_q+\gamma r_{q-1}\in A\mathscr{K}_{t+1}(A,r_q+\gamma r_{q-1}),\quad \gamma\neq-1,0,$$

or,

$$r_q\in A\mathscr{K}_{t+1}(A,r_q)+\mathscr{K}_{t+2}(A,r_{q-1}),\tag{23}$$

where $t+1\leqslant m$ by the assumption that $q=\max\{z\in\mathbb{Z}:z<n/m\}$. Owing to the fact that $r_{q+1}=0$ if and only if $r_q\in A\mathscr{K}_m(A,r_q)$, it suffices to show that the component of the vector r_q from $\mathscr{K}_{t+2}(A,r_{q-1})$ in the representation (23) does not generally belong to $A\mathscr{K}_m(A,r_q)$. To see this, since $\gamma\neq0$, we observe that the term from $\mathscr{K}_{t+2}(A,r_{q-1})$ in (23) contains a nonzero component in the direction r_{q-1}, which is not in $A\mathscr{K}_m(A,r_q)$ unless the initial residual r_0 is chosen from a specific subspace of $\mathbb{C}^n$, i.e. expressing r_{q-1} in terms of r_0 and vectors $w_m^{(k)}$ by Equation (7),

$$r_0\notin\mathscr{I},\ \ \mathscr{I}=\operatorname{span}\{w_m^{(1)},\ldots,w_m^{(q-1)}\}+A\mathscr{K}_m(A,r_q),$$

where $\dim\mathscr{I}\leqslant q+m-1<n/m+m-1\leqslant n$, provided that $0<m<n$.

3.3. Any admissible convergence behavior is possible for full and restarted GMRES (at its q initial cycles)

As we pointed out in Introduction, the convergence behavior of *full* GMRES in Theorem 1 is restricted to the class of convergence sequences which allow convergence to the exact solution only at step n, i.e. $f(0)\geqslant f(1)\geqslant\cdots\geqslant f(n-1)>0$ ($f(n)=0$). Similarly, the cycle-convergence behavior of *restarted* GMRES in Theorem 2 is restricted to cycle-convergence sequences which exclude the possibility of convergence to the exact solution within the initial q cycles, i.e. $f(0)>f(1)>\cdots>f(s)>0$ and $f(s)=f(s+1)=\cdots=f(q)$. It turns out that assumptions in Theorems 1 and 2 are sufficient for the theorems also to hold when $f(0)\geqslant f(1)\geqslant\cdots\geqslant f(n-1)\geqslant0$ and $f(0)>f(1)>\cdots>f(s)\geqslant0$, $f(s)=f(s+1)=\cdots=f(q)$, respectively.

Given a matrix order n, assume that we want to construct a matrix A with a prescribed spectrum Λ and an initial residual vector r_0 (or, equivalently, a right-hand side b, since $r_0 = b$ after setting x_0 to 0) such that *full* GMRES applied to the corresponding system (1) results in the following convergence pattern: $f(0) \geqslant f(1) \geqslant \cdots \geqslant f(s-1) > f(s) = f(s+1) = \cdots = f(n-1) = 0$, $s < n$, $\|r_k\| = f(k)$. The construction is straight-forward. We first split the set Λ into two disjoint subsets, say, $\Lambda = \Lambda_s \cup \Lambda_{n-s}$, where Λ_s contains s elements from Λ while the rest $n-s$ elements are included into Λ_{n-s}. Next, by Theorem 1 we construct a matrix $A_s \in \mathbb{C}^{s \times s}$ and a right-hand side $b_s \in \mathbb{C}^s$ ($x_0 = 0 \in \mathbb{C}^s$), such that *full* GMRES applied to the system $A_s x = b_s$ produces the convergence sequence $f(0) \geqslant f(1) \geqslant \cdots \geqslant f(s-1) > 0$ ($f(s) = 0$), moreover the spectrum of A_s is Λ_s. Finally, we define the resulting matrix $A \in \mathbb{C}^{n \times n}$ and the right-hand side vector $b \in \mathbb{C}^n$ ($x_0 = 0 \in \mathbb{C}^n$) as follows:

$$A = \begin{pmatrix} A_s & 0 \\ 0 & A_{n-s} \end{pmatrix}, \quad b = \begin{pmatrix} b_s \\ 0 \end{pmatrix}, \tag{24}$$

where $A_{n-s} \in \mathbb{C}^{(n-s) \times (n-s)}$ is an arbitrary matrix with a spectrum Λ_{n-s}. It is easy to see that *full* GMRES applied to the system of equations defined by Equation (24) will produce the desired sequence of residual norms $f(0) \geqslant f(1) \geqslant \cdots \geqslant f(s-1) > f(s) = f(s+1) = \cdots = f(n-1) = 0$, $\|r_k\| = f(k)$, $r_0 = b$. Clearly, the matrix A in Equation (24) has the prescribed spectrum $\Lambda = \Lambda_s \cup \Lambda_{n-s}$.

For the *restarted* GMRES, the construction of a matrix A with the spectrum Λ and a right-hand side b ($x_0 = 0$) that provide the cycle-convergence sequence $f(0) > f(1) > \cdots > f(s-1) > f(s) = f(s+1) = \cdots = f(q) = 0$ is analogous, $s \leqslant q$, $\|r_k\| = f(k)$. Following Theorem 2 one constructs a matrix $A_s \in \mathbb{C}^{ms \times ms}$ and a right-hand side vector $b_s \in \mathbb{C}^{ms}$, such that GMRES(m) applied to the corresponding linear system produces the cycle-convergence curve $f(0) > f(1) > \cdots > f(s-1) > f(s) = 0$. The spectrum of A_s is chosen to coincide with a subset of ms elements of Λ. The construction of the resulting matrix $A \in \mathbb{C}^{n \times n}$ and a right-hand side $b \in \mathbb{C}^n$ is then accomplished by introducing an $(n-ms) \times (n-ms)$ diagonal block with eigenvalues from Λ which are not in the spectrum of A_s and expanding the vector b with $(n-ms)$ zeros, similarly to Equation (24).

3.4. Restarted GMRES with variable restart parameter

The result of Theorem 2 generalizes to the case when the restart parameter m is not fixed, but varies over successive cycles according to *a priori* prescribed parameter sequence $\{m_k\}_{k=1}^q$. The proof, basically, repeats the one in Section 2 with the difference that the constructed operator $\mathscr{A}$ in the corresponding basis will have block lower triangular structure with varying diagonal block sizes m_k, rather than the constant size $m_k = m$ as in Equation (17).

Corollary 1

Given a matrix order n, a sequence $\{m_k\}_{k=1}^q$, $0 < m_k < n$ and a positive sequence $\{f(k)\}_{k=0}^q$, such that $f(0) > f(1) > \cdots > f(s) > 0$ and $f(s) = f(s+1) = \cdots = f(q)$, where q is defined by the condition $\sum_{k=1}^q m_k < n$, $0 \leqslant s \leqslant q$. There exists an $n \times n$ 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 a restart parameter varying according to the sequence $\{m_k\}_{k=1}^q$ 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.

4. CONCLUSION

In this paper we show that, regardless of the eigenvalue distribution of a coefficient matrix, any admissible cycle-convergence behavior is possible for restarted GMRES at q initial cycles. This leads to the conclusion that no estimates which rely solely on the matrix spectrum can be derived to characterize the cycle-convergence of restarted GMRES at the first q cycles when the method is applied to a linear system with a general nonsingular non-Hermitian matrix. Though in practice q tends to be reasonably large ($q < n/m$), it remains an open question if the above-mentioned estimates can or cannot be available at cycles that follow the n/mth GMRES(m) cycle.

REFERENCES

1. Saad Y, Schultz MH. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* 1986; **7**(3):856–869.
2. Eisenstat SC, Elman HC, Schultz MH. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM Journal on Numerical Analysis* 1983; **20**:345–357.
3. Elman HC. Iterative methods for large sparse nonsymmetric systems of linear equations. *Ph.D. Thesis*, Yale University, New Haven, CT, 1982.
4. Greenbaum A. *Iterative Methods for Solving Linear Systems*. SIAM: Philadelphia, PA, 1997.
5. Beckermann B, Goreinov SA, Tyrtyshnikov EE. Some remarks on the Elman estimate for GMRES. *SIAM Journal on Matrix Analysis and Applications* 2005; **27**(3):772–778.
6. Simoncini V, Szyld D. New conditions for non-stagnation of minimal residual methods. *Numerische Mathematik* 2008; **109**(3):477–487.
7. Zítko J. Generalization of convergence conditions for a restarted GMRES. *Numerical Linear Algebra with Applications* 2000; **7**(3):117–131.
8. Simoncini V, Szyld DB. On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods. *SIAM Review* 2005; **47**:247–272.
9. van der Vorst HA, Vuik C. The superlinear convergence behaviour of GMRES. *Journal of Computational and Applied Mathematics* 1993; **48**(3):327–341.
10. Baker AH, Jessup ER, Jessup ER, Kolev TzV. A simple strategy for varying the restart parameter in GMRES(m). *Journal of Computational and Applied Mathematics* 2009; **230**(2):751–761.
11. Vecharynski E, Langou J. The cycle-convergence of restarted GMRES for normal matrices is sublinear. *SIAM Journal on Scientific Computing* 2010; **32**(1):186–196.
12. Greenbaum A, Pták V, Z. Strakoš. Any nonincreasing convergence curve is possible for GMRES. *SIAM Journal on Matrix Analysis and Applications* 1996; **17**(3):465–469.
13. Arioli M, Pták V, Strakoš Z. Krylov sequences of maximal length and convergence of GMRES. *BIT* 1998; **38**(4):636–643.
14. Nachtigal NM, Reddy SC, Trefethen LN. How fast are nonsymmetric matrix iterations? *SIAM Journal on Matrix Analysis and Applications* 1992; **13**(3):778–795.
15. Embree M. How descriptive are GMRES convergence bounds? *Technical Report 99/08*, Oxford University Computing Laboratory, 1999.
16. Zhong B, Morgan RB. Complementary cycles of restarted GMRES. *Numerical Linear Algebra with Applications* 2008; **15**(6):559–571.
17. Embree M. The tortoise and the hare restart GMRES. *SIAM Review* 2003; **45**(2):259–266.
18. Eiermann M, Ernst OG, Schneider O. Analysis of acceleration strategies for restarted minimal residual methods. *Journal of Computational and Applied Mathematics* 2000; **123**(1–2):261–292.
19. Baglama J, Calvetti D, Golub GH, Reichel L. Adaptively preconditioned GMRES algorithms. *SIAM Journal on Scientific Computing* 1999; **20**(1):243–269.
20. Chapman A, Saad Y. Deflated and augmented Krylov subspace techniques. *Numerical Linear Algebra with Applications* 1997; **4**(1):43–66.
21. Duff IS, Giraud L, Langou J, Martin É. Using spectral low rank preconditioners for large electromagnetic calculations. *International Journal for Numerical Methods in Engineering* 2005; **62**(3):416–434.
22. Erhel J, Burrage K, Pohl B. Restarted GMRES preconditioned by deflation. *Journal of Computational and Applied Mathematics* 1996; **69**:303–318.
23. Morgan RB. GMRES with deflated restarting. *SIAM Journal on Scientific Computing* 2002; **24**(1):20–37.
24. Zítko J. Some remarks on the restarted and augmented GMRES method. *Electronic Transactions on Numerical Analysis* 2008; **31**:221–227.