Numerical Algorithms (2020) 84:1329–1352
https://doi.org/10.1007/s11075-019-00846-z

ORIGINAL PAPER

On the residual norms, the Ritz values and the harmonic Ritz values that can be generated by restarted GMRES

Jurjen Duintjer Tebbens[1,2] · **Gérard Meurant**[3]

Received: 24 July 2019 / Accepted: 5 November 2019 / Published online: 12 December 2019

Abstract

The paper gives a characterization of all linear systems such that when restarted GMRES is applied, prescribed admissible residual norms and (harmonic) Ritz values for all iterations inside the individual cycles are generated. Additionally, the system matrices can have any nonzero eigenvalues. The total number of GMRES iterations inside all cycles considered is assumed to be smaller than the system size. It is shown that stagnation at the end of a restart cycle must be mirrored at the beginning of the next cycle and that this is the only restriction for prescribed residual norms of restarted GMRES. The relation between prescribed residual norms of restarted GMRES and those of the corresponding full GMRES process is studied and linear systems are given where full and restarted GMRES give the same convergence history.

Keywords Restarted GMRES · (harmonic) Ritz values · GMRES stagnation · Prescribed convergence

✉ Jurjen Duintjer Tebbens
duintjertebbens@cs.cas.cz

Gérard Meurant
gerard.meurant@gmail.com

1 Institute of Computer Science, Czech Academy of Sciences, Pod Vodárenskou věží 2, 18 207 Prague 8 - Libeň, Czech Republic

2 Department of Biophysics and Physical Chemistry, Faculty of Pharmacy, Charles University, Akademika Heyrovského 1203, 500 05 Hradec Králové, Czech Republic

3 30 rue du sergent Bauchat, 75012 Paris, France

1 Introduction

The GMRES method [30] is a popular Krylov subspace method for the solution of linear systems

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

with nonsingular, non-Hermitian, sparse and possibly very large matrices A. The kth GMRES iterate x_k minimizes the norm of the residual vector $r_k = b - Ax_k$ over all vectors in the kth Krylov subspace span$\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$, leading to non-increasing residual norms. The employed basis for the kth Krylov subspace being orthogonal in GMRES, it must in general be generated with long recurrences. Storage and computational costs therefore grow with the iteration number k. In practice one usually attempts to find a preconditioner to obtain a low number of iterations necessary to find a sufficiently accurate approximation to the solution. But powerful preconditioners may be expensive or difficult to find. Another popular way to keep storage and computational costs low (possibly in combination with preconditioning) is to restart the GMRES method. After a small number of m iterations, the current approximation x_m is used as the initial approximation for a new series of m GMRES iterations and this process is repeated. We will denote it by GMRES(m) and will call the non-restarted process full GMRES. GMRES(m) produces non-increasing residual norms just as full GMRES, but there is no guarantee that the solution of (1) has been found, in exact arithmetic, by the nth iteration. In fact, there is no guarantee that the solution will be found at all, as iterates may stagnate. GMRES(m) may produce identical approximations during an entire cycle of m iterations and consequently, all subsequent cycles behave the same way.

Part of the available convergence results for GMRES(m) consider sufficient non-stagnation conditions (see, e.g., [12, 29, 33, 39, 40]) and a variety of techniques has been proposed to reduce the risk of stagnation. Especially strategies using approximate eigenspaces to deflate the matrix or augment the Krylov subspaces have been successful for many problems (for overviews, see, e.g., [3, 11, 17, 29]). The main idea is to eliminate the influence of eigenvalues that are assumed to hamper convergence, for instance through moving eigenvalues close to the origin towards unity (this holds for some preconditioners as well). Information about such eigenvalues are mostly obtained from the Ritz values or the harmonic Ritz values generated by GMRES during the restart cycles (see, e.g., the series of papers [24–26]). While this seems to often be a very efficient strategy in practice, in theory it needs not be universally applicable. The precise influence of eigenvalues on GMRES convergence can be very complicated, depending strongly on the right-hand side vector b and on eigenvectors [23], even with unitary matrices [9]. Moreover, by a theorem by Greenbaum, Pták and Strakoš, any non-increasing convergence curve can be generated by full GMRES with matrices having prescribed (nonzero) eigenvalues [1, 18, 19, 21]. Counterintuitive cases from practice exist [20] and a partial analogue for GMRES(m) under certain restrictive assumptions was formulated by Vecharynski and Langou in [36]. Another reason why deflation strategies need not work in general is that even if some eigenvalues do hamper convergence, it might in theory be impossible to approximate them anyhow with Ritz or harmonic Ritz values in a GMRES or, equivalently, Arnoldi process. This was shown for non-restarted Arnoldi processes in [5] and [6], but not for restarted Arnoldi processes (in fact, it was shown that (harmonic) Ritz values can be chosen independently not only from the spectrum but from residual norms in full GMRES as well).

In the present paper we are interested in the questions whether in the restarted GMRES method, any non-increasing convergence curve is possible with any spectrum, any Ritz values (in all iterations of restart cycles) and any harmonic Ritz values (in all iterations of restart cycles). At first sight, it may seem straightforward that the answers are positive if they are for full GMRES. But several theoretical properties cannot be inherited. For example, we will show that some non-increasing convergence curves are *not* admissible for GMRES(m). In fact, the restarting mechanism in GMRES complicates convergence analysis considerably (more so than in the restarted FOM method [31]); detailed investigations of the restart mechanism and its consequences for convergence behavior can be found, among others, in [4], [32], [35] and [38]. To answer the questions formulated above as completely as possible, we will focus on prescribing all the entries of the Hessenberg matrices generated in the individual restart cycles. When this is possible, we can prescribe in particular the residual norms inside the cycles (which extends the results in [36]) as well as the Ritz values or the harmonic Ritz values at every iteration of the cycles. We will attempt to explicitly construct linear systems with such prescribed behavior. We also address the relation with the convergence curves that are generated by full GMRES applied to such systems. This will offer some insight into how it is possible that for some linear systems encountered sometimes in practice, the speed of convergence seems to be inversely proportional to the restart length, i.e., a larger m yields slower convergence of GMRES(m) (see, e.g., [11, 13]).

The paper is organized as follows. In the remainder of this section we introduce some further notation and recall relevant results related to full GMRES. The second section describes inadmissible convergence curves for GMRES(m) and the next section constructs linear systems generating any admissible curve with prescribed spectrum for the system matrix and prescribed (harmonic) Ritz values in the individual restart cycles. Section 4 addresses the relation of the constructed linear systems with full GMRES. Further comments and conclusions are given in Section 5. Throughout the paper we assume exact arithmetic and we assume that the initial guess x_0 in (restarted) GMRES processes is zero. With "the subdiagonal" and "subdiagonal entries" we will mean the (entries on the) first diagonal under the main diagonal. We will denote by e_j the jth column of the identity matrix of appropriate order and $\overline{\alpha}$ denotes the complex conjugate of a complex number α. $B^{\dagger}$ denotes the Moore-Penrose pseudoinverse of a matrix B and $\|\cdot\|$ denotes the Euclidean norm for vectors and the induced norm for matrices.

1.1 Further notation and preliminaries

We consider in total no more than $N-1$ restarts where $Nm < n$. In words, the total number of GMRES iterations inside all cycles (including the initial cycle) is smaller than the system size. This does not represent a restriction for most practical situations, though a complete theoretical study of restarted GMRES should address the question of convergence behavior after more than n iterations as well (a result for this situation was given in [31]). The first $N-1$ restarts, where $Nm < n$, are sometimes referred to as the initial cycles (see, e.g., [36]), but here we will denote by 'initial cycle' only the very first m iterations, before the first restart.

Our goal is to construct matrices A and right-hand sides b such that if GMRES(m) is applied, it exhibits some prescribed behavior for the residual norms and the (harmonic) Ritz values. A and b will always be constructed as

$$A = VHV^*, \qquad b/\|b\| = Ve_1, \tag{2}$$

for an upper Hessenberg matrix H and a unitary matrix V. For the purposes of this paper it suffices to restrict to unreduced Hessenberg matrices H; thus the constructed matrices A are non-derogatory. Generalizations to the case of early termination might be derived along the same lines as it was done in [7] for full GMRES (see also the end of the paper [36]).

GMRES residual norms are unitarily invariant, i.e., GMRES applied to B and c gives the same residual norms as GMRES applied to WBW^* and Wc for any unitary matrix W. The same holds for the (harmonic) Ritz values obtained from the Arnoldi process. With respect to (2) this means that it suffices to study GMRES for H with right-hand side e_1. To construct matrices and right-hand sides yielding prescribed residual norms and (harmonic) Ritz values, we will therefore concentrate in (2) on the choice of H and consider V a free parameter matrix.

Any product of the form UCU^{-1} where U is nonsingular upper triangular and C is a companion matrix yields an unreduced upper Hessenberg matrix. Conversely, any unreduced upper Hessenberg matrix H can be decomposed as

$$H = UCU^{-1}, \tag{3}$$

where U is nonsingular upper triangular and C is the companion matrix for the polynomial whose roots are the eigenvalues of H. To find U, it suffices to equate consecutively the columns 1 until $n-1$ of the equation $HU = UC$ starting with Ue_1 being a nonzero multiple of e_1. The decomposition (3), which we call the triangular Hessenberg decomposition, is useful for creating Hessenberg matrices with a given spectrum yielding prescribed residual norms and (harmonic) Ritz values when GMRES is applied with right-hand side e_1. The next theorem shows that in *full* GMRES, the choice of the first row of U^{-1} can force prescribed residual norms and the columns of U^{-1} determine ordinary Ritz values. It is a slight modification of [7, Theorem 1], formulated in terms of the Hessenberg matrix that is actually generated by GMRES.

Theorem 1 *Consider a set of tuples of complex numbers*

$$\begin{aligned}
\mathcal{R} = \{ \;& \rho_1^{(1)}, \\
& (\rho_1^{(2)}, \rho_2^{(2)}), \\
& \vdots \\
& (\rho_1^{(n-1)}, \ldots, \rho_{n-1}^{(n-1)}), \\
& (\lambda_1, \ldots\ldots\ldots, \lambda_n)\},
\end{aligned}$$

such that $(\lambda_1, \ldots, \lambda_n)$ *contains no zero number and consider n positive numbers*

$$f(0) \geq f(1) \geq \cdots \geq f(n-1) > 0,$$

such that for $i = 1, \ldots, n-1$, $f(i-1) = f(i)$ *if and only if* $(\rho_1^{(i)}, \ldots, \rho_i^{(i)})$ *contains a zero number.*

If A is a matrix of order n and b a nonzero n-dimensional vector, then the following assertions are equivalent:

1. *The GMRES method applied to A and right-hand side b with zero initial guess yields residual vectors* r_i, $i = 0, \ldots, n-1$ *such that*

$$\|r_i\| = f(i), \quad i = 0, \ldots, n-1,$$

A has eigenvalues $\lambda_1, \ldots, \lambda_n$ *and* $\rho_1^{(i)}, \ldots, \rho_i^{(i)}$ *are the eigenvalues of the ith leading principal submatrix of the generated Hessenberg matrix (the Ritz values) for all* $i = 1, \ldots, n-1$.

2. *The GMRES method applied to A and right-hand side b with zero initial guess generates an upper Hessenberg matrix of the form*

$$H = \begin{bmatrix} \chi^T \\ 0 \ \Sigma \end{bmatrix}^{-1} C^{(n)} \begin{bmatrix} \chi^T \\ 0 \ \Sigma \end{bmatrix},$$

where $C^{(n)}$ *is the companion matrix of the polynomial with roots* $\lambda_1, \ldots, \lambda_n$,

$$\chi_0 = \frac{1}{f(0)}, \qquad |\chi_i| = \sqrt{\frac{1}{f(i)^2} - \frac{1}{f(i-1)^2}}, \qquad i = 1, \ldots, n-1$$

and the columns of the nonsingular upper triangular matrix Σ *of size* $n-1$ *contain the coefficients of polynomials* $p_i(\rho)$ *with roots* $\rho_j^{(i)}$:

$$p_i(\rho) \equiv \prod_{j=1}^{i}(\rho - \rho_j^{(i)}) = \frac{1}{\sigma_{i,i}}\left(\chi_i + \sum_{j=1}^{i}\sigma_{j,i}\rho^j\right), \quad i = 1, \ldots, n-1$$

with $\sigma_{i,i} \in \mathbb{R}, \quad \sigma_{i,i} > 0, \quad i = 1, \ldots, n-1.$

Please note the relation between stagnating iterations and zero Ritz values in the previous theorem. Further results on GMRES stagnation in relation to prescribed convergence behavior can be found in [22]. We remark that instead of prescribing all ordinary Ritz values, it is also possible to prescribe all *harmonic* Ritz values by an appropriate choice of the entries of Σ. For details, including the situation where GMRES stagnates, we refer to [5, Theorem 4.4].

As mentioned, all relevant values in GMRES are obtained from the Hessenberg matrix and this clearly also holds for the size $(m+1) \times m$ upper Hessenberg matrix generated in some cycle of GMRES(m). In the sequel we will therefore focus on the choices of the small Hessenberg matrices of the individual restart cycles. In fact, we will attempt to prescribe all their entries. Representing the first m steps of a GMRES process, they can be written in the form of the upper left submatrix

$$\begin{aligned}\begin{bmatrix} I_{m+1} & 0 \end{bmatrix} H \begin{bmatrix} I_m \\ 0 \end{bmatrix} &= \begin{bmatrix} I_{m+1} & 0 \end{bmatrix} \begin{bmatrix} \chi^T \\ 0 \;\; \Sigma \end{bmatrix}^{-1} C^{(n)} \begin{bmatrix} \chi^T \\ 0 \;\; \Sigma \end{bmatrix} \begin{bmatrix} I_m \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \chi_0 \; \dots \; \chi_m \\ 0 \;\; \Sigma_m \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I_m \end{bmatrix} \begin{bmatrix} \chi_0 \; \dots \; \chi_{m-1} \\ 0 \;\; \Sigma_{m-1} \end{bmatrix}\end{aligned}$$

of H in the previous theorem (with Σ_m denoting the size m leading principal submatrix of Σ). It follows directly from that theorem that the size $(m+1) \times m$ upper Hessenberg matrix generated in the kth cycle of GMRES(m) is

$$\underline{H}_m^{(k)} = \begin{bmatrix} \chi_0^{(k)} \; \dots \; \chi_m^{(k)} \\ 0 \;\; \Sigma_m^{(k)} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I_m \end{bmatrix} \begin{bmatrix} \chi_0^{(k)} \; \dots \; \chi_{m-1}^{(k)} \\ 0 \;\; \Sigma_{m-1}^{(k)} \end{bmatrix} \in \mathbb{C}^{(m+1)\times m} \tag{4}$$

with Σ_m a nonsingular upper triangular matrix of size m with leading principal submatrix Σ_{m-1} and

$$\chi_0^{(k)} = \frac{1}{f_0^{(k)}}, \qquad |\chi_i^{(k)}| = \sqrt{\left(\frac{1}{f_i^{(k)}}\right)^2 - \left(\frac{1}{f_{i-1}^{(k)}}\right)^2}, \qquad i = 1, \dots, m \tag{5}$$

if and only if the $m+1$ positive numbers

$$f_0^{(k)} \geq f_1^{(k)} \geq \dots \geq f_{m-1}^{(k)} \geq f_m^{(k)} > 0$$

correspond to the residual norms generated in that cycle. By the choice of the entries of $\Sigma_m^{(k)}$ we can prescribe either the admissible Ritz or the admissible harmonic Ritz values of the cycle.

2 Inadmissible non-increasing convergence curves

In this section we investigate whether, as is the case for full GMRES, any non-increasing convergence curve is possible for restarted GMRES. We will see that the answer is negative and describe a type of inadmissible convergence curve.

Let the residual vectors generated in the first two cycles of GMRES(m) be denoted as

$$r_0^{(1)} = b, r_1^{(1)}, \dots, r_m^{(1)}, \qquad r_0^{(2)} = r_m^{(1)}, r_1^{(2)}, \dots, r_m^{(2)}.$$

We wish to construct A and b of the form (2) such that in the m iterations of the initial cycle the generated Arnoldi decomposition is of the form

$$AV_m^{(1)} = V_{m+1}^{(1)} \underline{H}_m^{(1)}, \quad \text{where} \quad V_{m+1}^{(1)*} V_{m+1}^{(1)} = I_{m+1}, \quad V_{m+1}^{(1)} e_1 = b/\|b\|. \tag{6}$$

Here, the upper Hessenberg matrix $\underline{H}_m^{(1)}$ of size $(m+1) \times m$ is a *given* matrix whose entries have been chosen such that prescribed residual norms and (harmonic) Ritz values are obtained. It can be constructed using (4). $V_{m+1}^{(1)}$ is an arbitrary matrix of size $n \times (m+1)$ with orthonormal columns. Because no restart took place in the initial cycle, it is trivial that we can choose the first m columns of H and the first $m+1$ columns $v_1, \dots, v_{m+1}$ of V in (2) as

$$H \begin{bmatrix} I_m \\ 0 \end{bmatrix} = \begin{bmatrix} \underline{H}_m^{(1)} \\ 0 \end{bmatrix}, \qquad V \begin{bmatrix} I_{m+1} \\ 0 \end{bmatrix} = [v_1, \dots, v_{m+1}] = V_{m+1}^{(1)}. \tag{7}$$

At the beginning of the restart, the new initial Arnoldi vector is the normalized last residual vector of the previous (initial) cycle and can be written as a linear combination of the Arnoldi vectors of the previous (initial) cycle. The coefficients of this linear combination are given in the lemma below; it turns out they depend only on the residual norms (modulo phase angles) generated before the restart.

Lemma 1 *Let the initial cycle of GMRES(m) have generated in its mth iteration Arnoldi vectors $v_1, \dots, v_{m+1}$ and an upper Hessenberg matrix $\underline{H}_m^{(1)}$ with decomposition* (4) *satisfying*

$$\chi_0^{(1)} = \frac{1}{\|r_0^{(1)}\|}, \qquad |\chi_i^{(1)}| = \sqrt{\frac{1}{\|r_i^{(1)}\|^2} - \frac{1}{\|r_{i-1}^{(1)}\|^2}}, \qquad i = 1, \dots, m.$$

If

$$g^{(1)} = \|r_m^{(1)}\| \begin{bmatrix} \overline{\chi}_0^{(1)} \\ \vdots \\ \overline{\chi}_m^{(1)} \end{bmatrix} \in \mathbb{C}^{m+1}, \tag{8}$$

then

$$\frac{r_m^{(1)}}{\|r_m^{(1)}\|} = [v_1, \dots, v_{m+1}]\, g^{(1)}. \tag{9}$$

Moreover,

$$(g^{(1)})^* \underline{H}_m^{(1)} e_j = 0, \qquad j = 1, \dots, m. \tag{10}$$

.

Proof Because $\underline{H}_m^{(1)}$ has the form (4), we have

$$(g^{(1)})^* \underline{H}_m^{(1)} =$$

$$\|r_m^{(1)}\|\, e_1^T \begin{bmatrix} \chi_0^{(1)} & \dots\ \chi_m^{(1)} \\ 0 & \Sigma_m^{(1)} \end{bmatrix} \begin{bmatrix} \chi_0^{(1)} & \dots\ \chi_m^{(1)} \\ 0 & \Sigma_m^{(1)} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I_m \end{bmatrix} \begin{bmatrix} \chi_0^{(1)} & \dots\ \chi_{m-1}^{(1)} \\ 0 & \Sigma_{m-1}^{(1)} \end{bmatrix} = 0 \cdot e_1^T,$$

which proves (10). After m iterations of GMRES we have

$$r_m^{(1)} = V_{m+1}^{(1)} \left(\|r_0^{(1)}\|\, e_1 - \underline{H}_m^{(1)} y_m \right), \qquad y_m = \arg\min \left\| \|r_0^{(1)}\|\, e_1 - \underline{H}_m^{(1)} y \right\|,$$

and with $y_m = \|r_0^{(1)}\| (\underline{H}_m^{(1)})^\dagger e_1$, we have

$$r_m^{(1)} = \|r_0^{(1)}\| [v_1, \dots, v_{m+1}] \left(I_{m+1} - \underline{H}_m^{(1)} (\underline{H}_m^{(1)})^\dagger \right) e_1.$$

Taking norms on both sides gives

$$\|r_m^{(1)}\| = \|r_0^{(1)}\| \left\| \left(I_{m+1} - \underline{H}_m^{(1)} (\underline{H}_m^{(1)})^\dagger \right) e_1 \right\|. \tag{11}$$

Thus,

$$\frac{r_m^{(1)}}{\|r_m^{(1)}\|} = [v_1, \ldots, v_{m+1}]c, \quad c = \frac{\left(I_{m+1} - \underline{H}_m^{(1)}(\underline{H}_m^{(1)})^\dagger\right)e_1}{\left\|\left(I_{m+1} - \underline{H}_m^{(1)}(\underline{H}_m^{(1)})^\dagger\right)e_1\right\|}. \tag{12}$$

Furthermore,

$$(g^{(1)})^* c = (g^{(1)})^* \frac{\left(I_{m+1} - \underline{H}_m^{(1)}(\underline{H}_m^{(1)})^\dagger\right)e_1}{\left\|\left(I_{m+1} - \underline{H}_m^{(1)}(\underline{H}_m^{(1)})^\dagger\right)e_1\right\|} = \|r_m^{(1)}\| \frac{\chi_0^{(1)}}{\|r_m^{(1)}\|}\|r_0^{(1)}\| = 1,$$

where we used (12), (11), and (8). It is easily checked from (5) that $\|g^{(1)}\| = 1$. By the Cauchy-Schwarz inequality $1 = (g^{(1)})^* c \leq \|g^{(1)}\|\,\|c\| = 1$, with equality if and only if $g^{(1)}$ and c are collinear. □

As mentioned, the choice (7) of the first m columns of H trivially guarantees that $\underline{H}_m^{(1)}$ is generated in the first m iterations of GMRES(m) applied to (2). We next wish to define the entries of columns $m+1$ until $2m$ of H such that when GMRES(m) is applied to (2), then, for a *given* Hessenberg matrix $\underline{H}_m^{(2)}$, the Arnoldi decomposition built in the m iterations of the second cycle is of the form

$$AV_m^{(2)} = V_{m+1}^{(2)}\underline{H}_m^{(2)}, \quad V_{m+1}^{(2)*}V_{m+1}^{(2)} = I_{m+1}, \quad V_{m+1}^{(2)}e_1 = V_{m+1}^{(1)}g^{(1)}. \tag{13}$$

Unfortunately, the Hessenberg matrix $\underline{H}_m^{(2)}$ cannot be chosen fully arbitrarily: The next two results show that if the initial cycle has some stagnating iterations at its end, this puts restrictions on the residual norms (and thus on $\underline{H}_m^{(2)}$) for the second cycle.

Theorem 2 *Let A be a non-derogatory matrix. Assume that for the first cycle of GMRES(m) applied to A with a vector b we have*

$$\|r_{m-j-1}^{(1)}\| > \|r_{m-j}^{(1)}\| = \cdots = \|r_m^{(1)}\|.$$

Then, $[r_m^{(1)}]^* A^i r_m^{(1)} = 0,\ i = 1, \ldots, j.$

Proof Let A be of the form (2). With Lemma 1, $r_m^{(1)}$ can be written as

$$r_m^{(1)} = \|r_m^{(1)}\|^2 [v_1, \ldots, v_{m+1}] \begin{bmatrix} \overline{\chi}_0^{(1)} \\ \vdots \\ \overline{\chi}_m^{(1)} \end{bmatrix}.$$

The stagnation assumption implies that $\chi_\ell^{(1)} = 0$ for $\ell = m+1-j, \ldots, m$. We observe that the $m+1$ first columns of V are equal to those of $V_{m+1}^{(1)}$. Therefore,

$$r_m^{(1)} = \|r_m^{(1)}\|^2 V z^{(1)},$$

where $z^{(1)}$ is a vector whose only nonzero components are $z_{\ell+1}^{(1)} = \overline{\chi}_\ell$, $\ell = 0, \ldots, m-j$. Let us denote $z^{(1)} = [\xi^{(1)}, 0, \cdots, 0]^T$ where $\xi^{(1)}$ is a vector of length $m+1-j$.

Let us consider $[r_m^{(1)}]^* A^i r_m^{(1)}$. Obviously, we have $A^i = V H^i V^*$ and

$$[r_m^{(1)}]^* A^i r_m^{(1)} = \|r_m^{(1)}\|^4 [z^{(1)}]^* V^* V H^i V^* V z^{(1)} = \|r_m^{(1)}\|^4 [z^{(1)}]^* H^i z^{(1)}.$$

For $i = 1$, we have that $z^{(2)} = Hz^{(1)}$ is a vector with nonzero components on positions 1 until $m + 2 - j$ and $z^{(2)}$ belongs to the span of columns $He_1, \ldots, He_{m+1-j}$. Using (10), this gives $[z^{(1)}]^* Hz^{(1)} = 0$. Similarly, for $i = 2$, we have that $z^{(3)} = Hz^{(2)} = H^2 z^{(1)}$ is a vector with nonzero components on positions 1 until $m + 3 - j$ and $z^{(3)}$ belongs to the span of columns $He_1, \ldots, He_{m+2-j}$. Using (10), this gives $[z^{(1)}]^* H^2 z^{(1)} = 0$. For $i = j$, we have that $z^{(j)} = Hz^{(j-1)} = H^j z^{(1)}$ is a vector with nonzero components on positions 1 until $m + 1$ and $z^{(j)}$ belongs to the span of columns $He_1, \ldots, He_m$, giving $[z^{(1)}]^* H^j z^{(1)} = 0$, using (10). □

Corollary 1 *With the notation of Theorem 2 and assuming that*

$$\|r_{m-j-1}^{(1)}\| > \|r_{m-j}^{(1)}\| = \cdots = \|r_m^{(1)}\|,$$

we have

$$\|r_0^{(2)}\| = \|r_1^{(2)}\| = \cdots = \|r_j^{(2)}\|,$$

which means that we have stagnation for the first j iterations of the second cycle.

Proof The result of the corollary is obtained from Theorem 2 using the results of [37] on partial initial stagnation. □

The previous result can be easily generalized for any pair of two consecutive restart cycles. Thus, a phase of stagnation at the end of some cycle is literally mirrored, with the same length, at the beginning of the next cycle. Equivalently, if a zero Ritz value appears during the last j subsequent iterations of a cycle, it must reappear in the first j iterations of the cycle that follows. As for harmonic Ritz values, stagnation corresponds to an infinite harmonic Ritz value being added to the harmonic Ritz values of the previous iteration (see [5, Theorem 4.4] for details). We are not aware of any reference to this property in the literature, though the discussion in [32, Section 4] suggests that after stagnation in a cycle, restarting will produce little new information to overcome slow convergence. Our result shows that, even though we do not orthogonalize against the basis vectors computed in the previous cycle, there is a lot of information in the residual vector which is used to build the new basis. In the case of stagnation this can be considered as unfortunate. In this case it may be better to restart from the last FOM(m) approximate solution (if it is available). Other possibilities are to go back to an iteration before stagnation started or to increase m hoping that, at some near iteration, we can get out of stagnation.

Summarizing, it is not possible to prescribe for GMRES(m) *any* non-increasing convergence curve. As soon as we prescribe stagnation at the end of some cycle, the beginning of the next cycle must stagnate too; prescribing strictly decreasing residual norms for these iterations is inadmissible. The next section will show that this is the *only* type of inadmissible convergence curve for GMRES(m).

3 Prescribing admissible convergence curves and (harmonic) Ritz values

We will now assume that the very last iteration of each cycle does not stagnate. For that case, we will construct linear systems that generate in every cycle fully prescribed Hessenberg matrices (all their entries are prescribed values). This will allow to fix residual norms inside cycles and we can prescribe the Ritz values or the harmonic Ritz values generated in the restart cycles as well. To begin with, we again consider the first two cycles. We have the following simple lemma.

Lemma 2 *Let GMRES(m) be applied to A and b of the form* (2) *and let it have generated after m iterations the Arnoldi decomposition of the form* (6). *Then, for a given $(m+1) \times m$ Hessenberg matrix $\underline{H}_m^{(2)}$, the restarted GMRES method has generated, at the end of the second cycle, an Arnoldi decomposition of the form* (13) *if and only if m iterations of the Arnoldi process with input matrix H and initial vector $\left[\, (g^{(1)})^T \;\; 0 \,\right]^T$ generate the decomposition*

$$H Z_m^{(2)} = Z_{m+1}^{(2)} \underline{H}_m^{(2)}, \quad Z_{m+1}^{(2)} e_1 = \left[\, (g^{(1)})^T \;\; 0 \,\right]^T, \tag{14}$$

where the matrix $Z_{m+1}^{(2)} = V^* V_{m+1}^{(2)}$ has orthogonal columns.

Proof For the claim it suffices to use $A = VHV^*$ and to define $Z_{m+1}^{(2)} \equiv V^* V_{m+1}^{(2)}$. □

Our goal is to find entries of H which ensure that (14) holds for a given matrix $\underline{H}_m^{(2)}$. However, the involved matrix $Z_{m+1}^{(2)}$ with orthonormal columns is not fixed. The probably easiest way to satisfy (14) is to assume that $Z_{m+1}^{(2)}$ has the particularly simple structure

$$Z_{m+1}^{(2)} = \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \\ 0 & 0 \end{bmatrix}. \tag{15}$$

We can find the columns $m+1$ until $2m$ of the matrix H satisfying (14) for this specific choice directly: Equating the first column of

$$H Z_m^{(2)} = Z_{m+1}^{(2)} \underline{H}_m^{(2)} \tag{16}$$

gives, with the notation $(g^{(1)})^T = [\hat{g}^{(1)}, g_{m+1}^{(1)}]^T$,

$$H Z_m^{(2)} e_1 = \begin{bmatrix} \underline{H}_m^{(1)} \\ 0 \end{bmatrix} \hat{g}^{(1)} + g_{m+1}^{(1)} H e_{m+1} = h_{1,1}^{(2)} \begin{bmatrix} g^{(1)} \\ 0 \\ \vdots \end{bmatrix} + h_{2,1}^{(2)} e_{m+2},$$

where $h_{i,j}^{(2)}$ are the prescribed entries of $\underline{H}_m^{(2)}$. Thus, the nonzero entries of the $m+1$st column of H satisfy

$$\begin{bmatrix} h_{1,m+1} \\ \vdots \\ h_{m+1,m+1} \end{bmatrix} = \frac{1}{g_{m+1}^{(1)}}\left(h_{1,1}^{(2)} g^{(1)} - \underline{H}_m^{(1)} \hat{g}^{(1)}\right), \quad h_{m+2,m+1} = \frac{h_{2,1}^{(2)}}{g_{m+1}^{(1)}}. \tag{17}$$

For the columns $m+2$ until $2m$ of H we obtain directly, using "MATLAB notation," from

$$H Z_m^{(2)}[e_2, \ldots, e_m] = H_{:,m+2:2m} = Z_{m+1}^{(2)} \underline{H}_m^{(2)} [e_2, \ldots, e_m]$$

that

$$H_{:,m+2:2m} = \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \underline{H}_m^{(2)} \begin{bmatrix} 0 \\ I_{m-1} \end{bmatrix} \\ 0 \\ \vdots \end{bmatrix} = \begin{bmatrix} g^{(1)} e_1^T \underline{H}_m^{(2)} \begin{bmatrix} 0 \\ I_{m-1} \end{bmatrix} \\ \\ \begin{bmatrix} 0 & I_m \end{bmatrix} \underline{H}_m^{(2)} \begin{bmatrix} 0 \\ I_{m-1} \end{bmatrix} \\ 0 \\ \vdots \end{bmatrix}. \tag{18}$$

The obtained values (17) and (18) for the entries in the columns $m+1$ until $2m$ of H can be represented as is done in the following theorem. In the statement we use that a nonsingular $m \times m$ leading principal submatrix of the generated Hessenberg matrix guarantees that there is no stagnation at the end of the corresponding GMRES(m) cycle (see, e.g., [2]). Please note that the displayed matrix

$$\begin{bmatrix} \underline{H}_m^{(1)} & 0 \\ 0 & \underline{H}_m^{(2)} \\ \vdots & \vdots \end{bmatrix}$$

is unreduced upper Hessenberg; its $m+1$st row contains m zeros, the entry $h_{m+1,m}^{(1)}$ and then the first row of $\underline{H}_m^{(2)}$.

Theorem 3 *Given two $(m+1) \times m$ unreduced Hessenberg matrices $\underline{H}_m^{(1)}$ and $\underline{H}_m^{(2)}$ where $\underline{H}_m^{(1)}$ has nonsingular $m \times m$ leading principal submatrix, these two Hessenberg matrices are consecutively generated in the first two cycles of the GMRES(m) method applied to H and e_1 if the first $2m$ columns of H are of the form*

$$H_{:,1:2m} = \begin{bmatrix} \underline{H}_m^{(1)} & 0 \\ 0 & \underline{H}_m^{(2)} \\ \vdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \underline{H}_0^{(2)} \\ \vdots & \vdots \\ & 0 \end{bmatrix},$$

where the $(m+2) \times m$ matrix $\underline{H}_0^{(2)}$ is the rank-two matrix

$$\underline{H}_0^{(2)} = \begin{bmatrix} \hat{g}^{(1)} \\ g_{m+1}^{(1)} - 1 \\ 0 \end{bmatrix} e_1^T \underline{H}_m^{(2)} - \frac{1}{g_{m+1}^{(1)}}\left(\begin{bmatrix} (g_{m+1}^{(1)} - 1) h_{1,1}^{(2)} g^{(1)} \\ (1 - g_{m+1}^{(1)}) h_{2,1}^{(2)} \end{bmatrix} + \begin{bmatrix} \underline{H}_m^{(1)} \hat{g}^{(1)} \\ 0 \end{bmatrix}\right) e_1^T.$$

Proof We would like to write the entries in columns $m+1$ until $2m$ found in (17) and (18) in the form

$$H_{:,m+1:2m} = \begin{bmatrix} 0 \\ \underline{H}_m^{(2)} \\ 0 \end{bmatrix} + \begin{bmatrix} \underline{H}_0^{(2)} \\ \vdots \\ 0 \end{bmatrix}.$$

As can be seen from (18), in H the rows $m+2$ until $2m+1$ of columns $m+2$ until $2m$ are just the trailing $m \times (m-1)$ block of $\underline{H}_m^{(2)}$. As for the first $m+1$ rows of columns $m+2$ until $2m$, they can be written as

$$g^{(1)} \cdot e_1^T \underline{H}_m^{(2)} \begin{bmatrix} 0 \\ I_{m-1} \end{bmatrix} = \begin{bmatrix} 0 \\ e_1^T \underline{H}_m^{(2)} \begin{bmatrix} 0 \\ I_{m-1} \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \hat{g}^{(1)} \\ \delta^{(1)} \end{bmatrix} e_1^T \underline{H}_m^{(2)} \begin{bmatrix} 0 \\ I_{m-1} \end{bmatrix},$$

with $\delta^{(1)} = g_{m+1}^{(1)} - 1$. The $m+1$st column of H, according to (17), has leading $m+1$ entries which can be written in the form

$$\begin{aligned} \begin{bmatrix} h_{1,m+1} \\ \vdots \\ h_{m+1,m+1} \end{bmatrix} &= \frac{1}{g_{m+1}^{(1)}} \left(h_{1,1}^{(2)} g^{(1)} - \underline{H}_m^{(1)} \hat{g}^{(1)} \right) \\ &= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ h_{1,1}^{(2)} \end{bmatrix} + h_{1,1}^{(2)} \begin{bmatrix} \hat{g}^{(1)} \\ \delta^{(1)} \end{bmatrix} - \frac{1}{d_{m+1}^{(1)}} \left(h_{1,1}^{(2)} \delta^{(1)} \begin{bmatrix} d_1^{(1)} \\ \vdots \\ d_{m+1}^{(1)} \end{bmatrix} + \underline{H}_m^{(1)} \hat{g}^{(1)} \right). \end{aligned}$$

The last row of $\underline{H}_0^{(2)}$ follows from the fact that $h_{m+2,m+1} = \frac{h_{2,1}^{(2)}}{g_{m+1}^{(1)}}$ (see (17)). □

We now generalize the previous theorem for other cycles. During the second restart cycle, let an Arnoldi decomposition denoted

$$AV_m^{(3)} = V_{m+1}^{(3)} \underline{H}_m^{(3)} \tag{19}$$

be generated where the matrix $V_{m+1}^{(3)}$ has orthogonal columns. It follows from Lemma 1 that

$$V_{m+1}^{(3)} e_1 = \frac{r_m^{(2)}}{\|r_m^{(2)}\|} = \|r_m^{(2)}\| V_{m+1}^{(2)} \begin{bmatrix} \overline{\chi}_0^{(2)} \\ \vdots \\ \overline{\chi}_m^{(2)} \end{bmatrix},$$

where $\chi_0^{(2)} = \frac{1}{\|r_0^{(2)}\|}$, $\quad |\chi_j^{(2)}| = \sqrt{\frac{1}{\|r_j^{(2)}\|^2} - \frac{1}{\|r_{j-1}^{(2)}\|^2}}, \qquad j = 1, \ldots, m.$

For the same reasons as explained in Lemma 2, the decomposition (19) with $\underline{H}_m^{(3)}$ *given* is generated by restarted GMRES(m) applied to $A = VHV^*$ and $b = Ve_1$ if (and only if)

$$HZ_m^{(3)} = Z_{m+1}^{(3)} \underline{H}_m^{(3)}, \tag{20}$$

with a matrix $Z_{m+1}^{(3)}$ with orthonormal columns and with initial vector

$$Z_{m+1}^{(3)}e_1 = V^*V_{m+1}^{(3)}e_1 = V^*V_{m+1}^{(2)}\|r_m^{(2)}\| \begin{bmatrix} \overline{\chi}_0^{(2)} \\ \vdots \\ \overline{\chi}_m^{(2)} \end{bmatrix}$$

$$= \|r_m^{(2)}\| Z_{m+1}^{(2)} \begin{bmatrix} \overline{\chi}_0^{(2)} \\ \vdots \\ \overline{\chi}_m^{(2)} \\ 0 \\ \vdots \end{bmatrix} = \|r_m^{(2)}\| \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \overline{\chi}_0^{(2)} \\ \vdots \\ \overline{\chi}_m^{(2)} \\ 0 \\ \vdots \end{bmatrix}.$$

Let us use the notation

$$\begin{bmatrix} g^{(2)} \\ 0 \\ \vdots \end{bmatrix} \equiv \|r_m^{(2)}\| \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \overline{\chi}_0^{(2)} \\ \vdots \\ \overline{\chi}_m^{(2)} \\ 0 \\ \vdots \end{bmatrix}, \qquad (g^{(2)})^T = [\hat{g}^{(2)}, g_{2m+1}^{(2)}]^T \tag{21}$$

and let us choose the matrix $Z_{m+1}^{(3)}$ with orthonormal columns in (20) to have the simple form

$$Z_{m+1}^{(3)} = \begin{bmatrix} g^{(2)} & 0 \\ 0 & I_m \\ 0 & 0 \end{bmatrix}. \tag{22}$$

Let $H_{1:2m+1,1:2m}$ be as defined in Theorem 3, which guarantees that the initial cycle and the first restart applied to H and e_1 generate the given matrices $\underline{H}_m^{(1)}$ and $\underline{H}_m^{(2)}$, respectively. To generate in the next cycle the prescribed matrix $\underline{H}_m^{(3)}$, we can apply Theorem 3 analogously. It gives that the first $3m$ columns of H must be of the form

$$H_{:,1:3m} = \begin{bmatrix} H_{1:2m+1,1:2m} & 0 \\ 0 & \underline{H}_m^{(3)} \\ \vdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \underline{H}_0^{(3)} \\ \vdots & \vdots \\ & 0 \end{bmatrix},$$

where the size $(2m+2) \times m$ matrix $\underline{H}_0^{(3)}$ is the rank-two matrix

$$\underline{H}_0^{(3)} = \begin{bmatrix} \hat{g}^{(2)} \\ g_{2m+1}^{(2)} - 1 \\ 0 \end{bmatrix} e_1^T \underline{H}_m^{(3)}$$
$$- \frac{1}{g_{2m+1}^{(2)}} \left(\begin{bmatrix} (g_{2m+1}^{(2)} - 1)h_{1,1}^{(3)} g^{(2)} \\ (1 - g_{2m+1}^{(2)})h_{2,1}^{(3)} \end{bmatrix} + \begin{bmatrix} H_{1:2m+1,1:2m}\hat{g}^{(2)} \\ 0 \end{bmatrix} \right) e_1^T.$$

All further columns of H, determining the Hessenberg matrices generated in all cycles after the third cycle, can be defined analogously. We summarize this result in the following theorem.

Theorem 4 *Let us for N such that $Nm < n$ have $Nm+1$ given positive decreasing numbers satisfying*

$$\begin{aligned} f_0^{(1)} &\geq f_1^{(1)} \geq \dots f_{m-1}^{(1)} > f_m^{(1)} = f_0^{(2)} \\ &\geq f_1^{(2)} \geq \dots f_{m-1}^{(2)} > f_m^{(2)} = f_0^{(3)} \\ &\vdots \\ &\geq f_1^{(N)} \geq \dots f_{m-1}^{(N)} > f_m^{(N)} > 0 \end{aligned} \tag{23}$$

and N given size $(m+1)\times m$ upper Hessenberg matrices $\underline{H}_m^{(k)}$, $1 \leq k \leq N$ of the form

$$\underline{H}_m^{(k)} = \begin{bmatrix} \chi_0^{(k)} & \dots & \chi_m^{(k)} \\ 0 & \Sigma_m^{(k)} & \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I_m \end{bmatrix} \begin{bmatrix} \chi_0^{(k)} & \dots & \chi_{m-1}^{(k)} \\ 0 & \Sigma_{m-1}^{(k)} & \end{bmatrix} \tag{24}$$

where

$$\chi_0^{(k)} = \frac{1}{f_0^{(k)}}, \qquad |\chi_i^{(k)}| = \sqrt{\left(\frac{1}{f_i^{(k)}}\right)^2 - \left(\frac{1}{f_{i-1}^{(k)}}\right)^2}, \qquad i = 1, \dots, m,$$

and where $\Sigma_m^{(k)}$ is a nonsingular upper triangular matrix with leading principal submatrix $\Sigma_{m-1}^{(k)}$. Let for $1 \leq k \leq N$

$$\begin{bmatrix} g^{(k)} \\ 0 \\ \vdots \end{bmatrix} = f_m^{(k)} \begin{bmatrix} g^{(k-1)} & 0 \\ 0 & I_m \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \overline{\chi}_0^{(k)} \\ \vdots \\ \overline{\chi}_m^{(k)} \\ 0 \\ \vdots \end{bmatrix}$$

where $g^{(0)} \equiv 1$ and let

$$(g^{(k)})^T = [\hat{g}^{(k)}, g_{km+1}^{(k)}]^T.$$

Then GMRES(m) applied to the unreduced upper Hessenberg matrix H and e_1 generates consecutively the residual norms satisfying $\|r_j^{(k)}\| = f_j^{(k)}$ and the upper Hessenberg matrices $\underline{H}_m^{(1)}, \dots, \underline{H}_m^{(N)} \in \mathbb{C}^{(m+1)\times m}$ if the first m columns of H are

$$H_{:,1:m} = \begin{bmatrix} \underline{H}_m^{(1)} \\ 0 \\ \vdots \end{bmatrix}$$

and for every k, $1 \leq k < N$, the first $(k+1)m$ columns of H are of the form

$$H_{:,1:(k+1)m} = \begin{bmatrix} H_{1:km+1,1:km} & 0 \\ 0 & \underline{H}_m^{(k+1)} \\ \vdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \underline{H}_0^{(k+1)} \\ \vdots & \vdots \\ & 0 \end{bmatrix},$$

where the size $(km+2)\times m$ *matrix* $\underline{H}_0^{(k+1)}$ *is the rank-two matrix*

$$\underline{H}_0^{(k+1)} = \begin{bmatrix} \hat{g}^{(k)} \\ g_{km+1}^{(k)} - 1 \\ 0 \end{bmatrix} e_1^T \underline{H}_m^{(k+1)}$$
$$- \quad \frac{1}{g_{km+1}^{(k)}} \left(\begin{bmatrix} (g_{km+1}^{(k)} - 1)h_{1,1}^{(k+1)} g^{(k)} \\ (1 - g_{km+1}^{(k)})h_{2,1}^{(k+1)} \end{bmatrix} + \begin{bmatrix} H_{1:km+1,1:km}\hat{g}^{(k)} \\ 0 \end{bmatrix} \right) e_1^T.$$

The previous theorem shows that any non-increasing convergence curve is possible for restarted GMRES provided strict decrease is prescribed at the end of all cycles. This is proved for the first N cycles, where $Nm < n$. More importantly, the theorem shows how to prescribe *all* the Hessenberg matrices generated during all the N restarts. By the choice of the entries of these Hessenberg matrices, we can prescribe as well the Ritz values or the harmonic Ritz values of the cycles. We remark that other choices of the entries of the Hessenberg matrices might prescribe other interesting values. For example, in [10, Section 6.3] it is shown that singular values can be prescribed. As for the spectrum of the system matrix, we can use the following.

Lemma 3 *Let* H *be an unreduced upper Hessenberg matrix of size* n *whose leading* $n-1$ *columns are given and consider* n *complex numbers* $\lambda_1, \ldots, \lambda_n$. *The last column of* H *can be chosen such that* H *has the eigenvalues* $\lambda_1, \ldots, \lambda_n$.

Proof See [6, Theorem 2.2] or [28, Theorem 3]. □

4 Relation with full GMRES

The construction in Theorem 4 represents only one particular way to prescribe the Hessenberg matrices generated in GMRES(m), because it assumes a specific chosen structure of the bases $Z^{(k)}$ in the subsequent restart cycles (see, e.g., (15), (22)). It is a construction with an interesting relation to the full GMRES process. To prove this relation we need the following lemma, which has its own interest.

Lemma 4 *Let full GMRES generate the Hessenberg matrix* H *and consider the recursion*

$$\chi_0 = \frac{1}{\|r_0\|}, \qquad \chi_i = \frac{-1}{h_{i+1,i}} [\chi_0, \ldots, \chi_{i-1}]^T \begin{bmatrix} h_{1,i} \\ \vdots \\ h_{i,i} \end{bmatrix}, \qquad i = 1, 2, \ldots, n-1. \quad (25)$$

Then $\chi = [\chi_0, \ldots, \chi_{n-1}]^T$ *is the first row of* U^{-1} *in the triangular Hessenberg decomposition (3) of* H *(with* $Ue_1 = \|r_0\|e_1$*), and the* i*th residual norm satisfies*

$$\|r_i\|^2 = \left(\sum_{j=0}^{i} |\chi_j|^2 \right)^{-1}, \qquad i = 1, 2, \ldots, n-1. \quad (26)$$

Proof In [27, Section 4] it was proved that if $H = UCU^{-1}$ is the triangular Hessenberg decomposition (3) and $\nu_{i,j}$ are the entries of U^{-1}, then

$$\begin{bmatrix} \nu_{1,i+1} \\ \vdots \\ \vdots \\ \nu_{i+1,i+1} \end{bmatrix} = \frac{1}{h_{i+1,i}} \left(\begin{bmatrix} 0 \\ \nu_{1,i} \\ \vdots \\ \nu_{i,i} \end{bmatrix} - \begin{bmatrix} U_i^{-1} h_i \\ \vdots \\ 0 \end{bmatrix} \right), \quad i = 1, \ldots, n-1, \tag{27}$$

where $h_i = [h_{1,i}, \cdots, h_{i,i}]^T$ and where U_i is the leading principal submatrix of size i of U. The recursion in the claim follows, with the notation $\chi_j = \nu_{1,j+1}$, from equating the first row in (27). Because of Theorem 1, the residual norms satisfy

$$\frac{1}{\|r_0\|} = \chi_0, \qquad \sqrt{\frac{1}{\|r_i\|^2} - \frac{1}{\|r_{i-1}\|^2}} = |\chi_i|, \qquad i = 1, \ldots, n-1.$$

Equality (26) follows from a straightforward induction argument applied to $\frac{1}{\|r_i\|^2} - \frac{1}{\|r_{i-1}\|^2} = |\chi_i|^2$, $i = 1, \ldots, n-1$. □

Formula (26) gives with (25) a recursion to compute the GMRES residual norms directly from the entries of the generated Hessenberg matrix, which is different from the standard computation using a QR-decomposition of the Hessenberg matrix.

Theorem 5 *Consider the matrix H constructed in Theorem 4. The residual norms generated when GMRES(m) is applied to the linear system with H and e_1 are the same as when full GMRES is applied to this linear system.*

Proof The first m residual norms are trivially identical. For the rest of the proof we will compare the sizes of the values χ_i corresponding to full GMRES with those corresponding to restart cycles; if they are equal, the residual norms are equal as well. The values of the χ_i for full GMRES can be computed from the recurrence (25) in Lemma 4.

The first $\chi_i^{(2)}$ in the first restart are, with Lemma 4,

$$\chi_0^{(2)} = \frac{1}{\|r_0^{(2)}\|}, \qquad \chi_1^{(2)} = -\frac{\chi_0^{(2)} h_{1,1}^{(2)}}{h_{2,1}^{(2)}}. \tag{28}$$

The value χ_{m+1} for full GMRES is

$$\chi_{m+1} = \frac{-1}{h_{m+2,m+1}} (\chi_0, \ldots, \chi_m)^T \begin{bmatrix} h_{1,m+1} \\ \vdots \\ h_{m+1,m+1} \end{bmatrix}, \tag{29}$$

where the entries $h_{1,m+1} \ldots, h_{m+1,m+1}$ are given by

$$\begin{bmatrix} h_{1,m+1} \\ \vdots \\ h_{m+1,m+1} \end{bmatrix} = \frac{1}{g_{m+1}^{(1)}} \left(h_{1,1}^{(2)} g^{(1)} - \underline{H}_m^{(1)} \begin{pmatrix} g_1^{(1)} \\ \vdots \\ g_m^{(1)} \end{pmatrix} \right), \quad h_{m+2,m+1} = \frac{h_{2,1}^{(2)}}{g_{m+1}^{(1)}},$$

see (17). If we multiply the first equality in the previous equation with the row vector $(\chi_0, \ldots, \chi_m)^T$ we have for the first term on the right-hand side $(\chi_0, \ldots, \chi_m)^T h_{1,1}^{(2)} g^{(1)} = \chi_0^{(2)} h_{1,1}^{(2)}$ because

$$(\chi_0, \ldots, \chi_m)^T g^{(1)} = 1/\|r_m^{(1)}\| = \chi_0^{(2)}, \tag{30}$$

see (8). The second term is zero because of (10).

Substituting in (29) gives $\chi_{m+1} = \frac{-\chi_0^{(2)} h_{1,1}^{(2)}}{h_{2,1}^{(2)}}$, which equals the second expression in (28).

The next value $\chi_i^{(2)}$ after the first restart is, with Lemma 4,

$$\chi_2^{(2)} = \frac{-(\chi_0^{(2)}, \chi_1^{(2)})^T \begin{bmatrix} h_{1,2}^{(2)} \\ h_{2,2}^{(2)} \end{bmatrix}}{h_{3,2}^{(2)}}.$$

For full GMRES we have

$$\chi_{m+2} = \frac{-(\chi_0, \ldots, \chi_m)^T g^{(1)} h_{1,2}^{(2)} - \chi_{m+1} h_{2,2}^{(2)}}{h_{3,2}^{(2)}},$$

see (18). Because of (30), $\chi_{m+2} = \chi_2^{(2)}$. The equality of the remaining χ_i, and therefore of the remaining residual norms follows by induction. The same proof can be used for subsequent restart cycles. $\square$

The linear system we have constructed in Theorem 4 to generate prescribed GMRES(m) residual norms represents in fact a best case scenario for restarted GMRES: It converges as fast as full GMRES. An analogue result for the FOM method, where the situation is somewhat simpler, was given in [31]. In other words, the GMRES minimization process can for this system be carried out with $m+1$-term recurrences. This property is due to the special structure of Arnoldi vectors in the individual restart cycles used to construct the linear systems (see e.g., (15) and (22)). It implies that all Arnoldi vectors generated in subsequent restart cycles except the initial ones are orthogonal to each other (this is in fact the situation where $\kappa(\left[V_{m+1}^{(1)}, V_{m+1}^{(2)} \begin{bmatrix} 0 \\ I_m \end{bmatrix} \right]) = 1$ in [32, Corollary 6.3]).

The matrix H constructed in Theorem 4 is in general not a normal matrix, but the GMRES minimization process can be carried out with short recurrences. This does not contradict the Faber-Manteuffel theorem [14, 15], [21, Chapter 4], because in our case the orthogonal basis can be generated with short recurrences for a particular right-hand side vector only (the first unit vector).

An interesting consequence is that no other restart length than m can give faster convergence (as no restart length can give faster convergence than full GMRES). In particular, not even *larger* restart lengths produce convergence faster than GMRES(m) for the systems constructed in Theorem 4. Thus we have found a class of systems exhibiting the counterintuitive behavior encountered sometimes in practice, where a larger restart parameter slows down convergence speed. This behavior has for instance been observed for sparse matrices resulting from standard five-point stencils [34].

We give an example for illustration. Suppose we wish to construct a linear system $Ax = b$ with $A \in \mathbb{R}^{16\times 16}$, $b \in \mathbb{R}^{16}$, such that the residual norm history for GMRES(5) is

$$\begin{aligned}\left[\|r_0^{(1)}\|, \|r_1^{(1)}\|, \ldots, \|r_5^{(1)}\|\right] &= [1, 0.7, 0.4, 0.1, 0.07, 0.04],\\ \left[\|r_0^{(2)}\|, \|r_1^{(2)}\|, \ldots, \|r_5^{(2)}\|\right] &= [0.04, 0.01, 0.007, 0.004, 0.001, 7\cdot 10^{-4}],\\ \left[\|r_0^{(3)}\|, \|r_1^{(3)}\|, \ldots, \|r_5^{(3)}\|\right] &= [7\cdot 10^{-4}, 4\cdot 10^{-4}, 10^{-4}, 7\cdot 10^{-5}, 4\cdot 10^{-5}, 10^{-5}].\end{aligned} \tag{31}$$

The residual norms for the three restart cycles can be obtained by defining three appropriate Hessenberg matrices of size 6×5 using (4), where the values χ_i are determined by the prescribed residual norms except for the phase angles. We will choose all these values to be positive and we will choose all three matrices Σ_5 in (4) to be the upper triangular matrix of ones (we are not interested in forcing specific (harmonic) Ritz values here). The corresponding Hessenberg matrices $\underline{H}_5^{(1)}, \underline{H}_5^{(2)}, \underline{H}_5^{(3)}$ will be generated by GMRES(5) if we use the construction of Theorem 4. In our example we use $V \equiv I_{16}$ (though any other unitary V would give the same behavior reported below).

GMRES(5) applied to the linear systems constructed in this way produces the solid convergence curve in Fig. 1. Of course, it corresponds with the residual norm history given in (31). Our experiment also confirmed that full GMRES applied to the system yields the same solid curve. The other curves represent residual norms generated with larger restart lengths (6, 7, and 8). As explained, they do not show faster convergence behavior than GMRES (5).

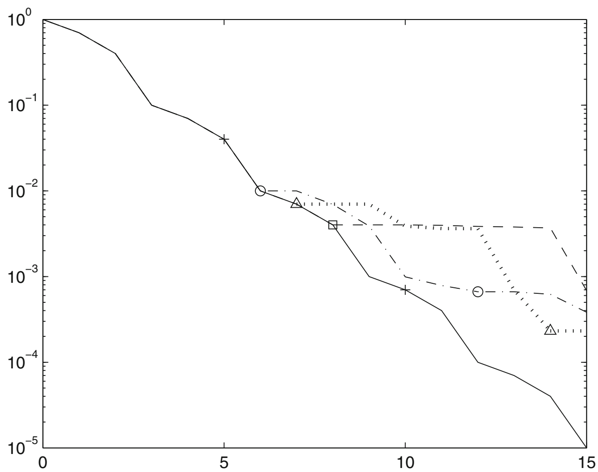

Fig. 1 GMRES residual norm curves for GMRES(5) (solid, crosses indicate restart), GMRES(6) (dash-dotted, circles indicate restart), GMRES(7) (dotted, triangles indicate restart) and GMRES(8) (dashed, squares indicate restart)

As mentioned, we have so far constructed particular linear systems generating prescribed Hessenberg matrices based on the assumption that the Arnoldi vectors of the restart cycles have an especially simple structure (see (15), (22)). Other choices of these Arnoldi vectors might be possible and give additional ways to prescribe the Hessenberg matrices of the individual cycles. But the resulting linear systems will in general not satisfy Theorem 5 with the consequences for larger restart lengths that we just discussed.

To investigate how to obtain alternative ways to prescribe the Hessenberg matrices of the individual cycles, let us focus on the first two cycles. Let $\underline{H}_{2m} = H_{1:2m+1,1:2m}$ denote the left upper block of the Hessenberg matrix H defined in Theorem 4, which guarantees that the given Hessenberg matrices $\underline{H}_m^{(1)}$ and $\underline{H}_m^{(2)}$ are generated in the first two cycles. Let $\underline{\hat{H}}_{2m} \in \mathbb{C}^{(2m+1)\times 2m}$ be the left upper block of another Hessenberg matrix $\hat{H}$ such that the same Hessenberg matrices $\underline{H}_m^{(1)}$ and $\underline{H}_m^{(2)}$ are generated as well, in the first two cycles of GMRES(m) applied to $\hat{H}$ and e_1. The unreduced upper Hessenberg matrix $\underline{\hat{H}}_{2m}$ can be transformed into $\underline{H}_{2m}$ with a nonsingular upper triangular matrix as follows. Let us decompose $\underline{H}_{2m}$ and $\underline{\hat{H}}_{2m}$ as

$$\underline{H}_{2m} = U_{2m+1} C_0 U_{2m}^{-1}, \quad \underline{\hat{H}}_{2m} = \hat{U}_{2m+1} C_0 \hat{U}_{2m}^{-1}, \qquad C_0 = \begin{bmatrix} 0 \\ I_{2m} \end{bmatrix} \tag{32}$$

(with U_{2m} resp. $\hat{U}_{2m}$ being the leading principal submatrix of size $2m$ of U_{2m+1} resp. $\hat{U}_{2m+1}$) by equating consecutively the columns 1 until $2m$ of the equations $\underline{H}_{2m} U_{2m} = U_{2m+1} C_0$ and $\underline{\hat{H}}_{2m} \hat{U}_{2m} = \hat{U}_{2m+1} C_0$ with $U_{2m} e_1 = \hat{U}_{2m} e_1 = e_1$. Then

$$\underline{\hat{H}}_{2m} = X_{2m+1}^{-1} \underline{H}_{2m} X_{2m}, \quad X_{2m+1} = U_{2m+1} \hat{U}_{2m+1}^{-1}, \quad X_{2m} = U_{2m} \hat{U}_{2m}^{-1}. \tag{33}$$

In the next theorem we give necessary and sufficient conditions for the entries of the matrix X_{2m+1} such that GMRES(m) applied to $\hat{H}$ and e_1 generates the given Hessenberg matrices $\underline{H}_m^{(1)}$ and $\underline{H}_m^{(2)}$.

Theorem 6 *Let $\underline{H}_m^{(1)} \in \mathbb{C}^{(m+1)\times m}$ and $\underline{H}_m^{(2)} \in \mathbb{C}^{(m+1)\times m}$ be given unreduced upper Hessenberg matrices with real positive subdiagonal and nonsingular leading $m \times m$ principal submatrix. The first two restart cycles of GMRES(m) applied to $\hat{H} \in \mathbb{C}^{n\times n}$ and $e_1 \in \mathbb{C}^n$ generate, subsequently, $\underline{H}_m^{(1)}$ and $\underline{H}_m^{(2)}$ if and only if the left upper block $\underline{\hat{H}}_{2m} \in \mathbb{C}^{(2m+1)\times 2m}$ of $\hat{H}$ is of the form*

$$\underline{\hat{H}}_{2m} = \begin{bmatrix} I_{m+1} & \underline{H}_m^{(1)} S_m \\ 0 & R_m \end{bmatrix}^{-1} \underline{H}_{2m} \begin{bmatrix} I_{m+1} & \underline{H}_m^{(1)} S_{m-1} \\ 0 & R_{m-1} \end{bmatrix}, \tag{34}$$

where $\underline{H}_{2m} \in \mathbb{C}^{(2m+1)\times 2m}$ is the left upper block of the Hessenberg matrix H defined in Theorem 4, $R_m \in \mathbb{C}^{m\times m}$ is a nonsingular upper triangular matrix with leading principal submatrix R_{m-1} such that $R_m^* R_m - I_m$ is positive semidefinite and $S_m \in \mathbb{C}^{m\times m}$ is a square matrix with first $m-1$ columns denoted by S_{m-1} such that

$$(\underline{H}_m^{(1)} S_m)^* \underline{H}_m^{(1)} S_m = R_m^* R_m - I_m. \tag{35}$$

Proof GMRES(m) applied to $\hat{H}$ and e_1 generates in the initial cycle the Hessenberg matrix $\underline{H}_m^{(1)}$ if and only if

$$\begin{bmatrix} I_{m+1} 0 \end{bmatrix} \underline{\hat{H}}_{2m} \begin{bmatrix} I_m \\ 0 \end{bmatrix} = \underline{H}_m^{(1)} \quad \Leftrightarrow \quad \begin{bmatrix} I_{m+1} 0 \end{bmatrix} X_{2m+1}^{-1} \underline{H}_{2m} X_{2m} \begin{bmatrix} I_m \\ 0 \end{bmatrix} = \underline{H}_m^{(1)}.$$

Using the facts that the leading $(m+1) \times m$ submatrix of $\underline{H}_{2m}$ is equal to $\underline{H}_m^{(1)}$ and that X_{2m+1} is upper triangular, we obtain the equivalent condition that the leading principal submatrix of X_{2m+1} of size $m+1$ must be the identity matrix.

GMRES(m) applied to $\hat{H}$ and e_1 generates in the second cycle the Hessenberg matrix $\underline{H}_m^{(2)}$ if and only if $\hat{H}\hat{Z}_m = \hat{Z}_{m+1}\underline{H}_m^{(2)}$ with the columns of $\hat{Z}_{m+1}$ orthonormal to each other and $\hat{Z}e_{m+1} = \begin{bmatrix} (g^{(1)})^T & 0 \end{bmatrix}^T$ (see Lemma 2). As this Arnoldi decomposition lives only on the left upper $(2m+1) \times 2m$ block of $\hat{H}$ it can also be written, with a slight abuse of notation, as $\underline{\hat{H}}_{2m}\hat{Z}_m = \hat{Z}_{m+1}\underline{H}_m^{(2)}$. Similarly, $\underline{H}_{2m} Z_m = Z_{m+1}\underline{H}_m^{(2)}$, where H is defined as in Theorem 4 and

$$Z_{m+1} = \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \end{bmatrix} \in \mathbb{C}^{(2m+1)\times(m+1)}.$$

Thus $X_{2m+1}\underline{\hat{H}}_{2m} X_{2m}^{-1} Z_m = Z_{m+1}\underline{H}_m^{(2)}$ (see (33)), and by comparison with $\underline{\hat{H}}_{2m}\hat{Z}_m = \hat{Z}_{m+1}\underline{H}_m^{(2)}$ we have $\hat{Z}_{m+1} = X_{2m+1}^{-1} Z_{m+1}$. The matrix X_{2m+1} has the form

$$X_{2m+1} = \begin{bmatrix} I_{m+1} & Y \\ 0 & R_m \end{bmatrix}$$

and since the columns of Z_{m+1} must be orthonormal, we have

$$\begin{aligned} \hat{Z}_{m+1}^* \hat{Z}_{m+1} &= Z_{m+1}^* X_{2m+1}^{-*} X_{2m+1}^{-1} Z_{m+1} \\ &= \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \end{bmatrix}^* \begin{bmatrix} I_{m+1} & -Y R_m^{-1} \\ 0 & R_m^{-1} \end{bmatrix}^* \begin{bmatrix} I_{m+1} & -Y R_m^{-1} \\ 0 & R_m^{-1} \end{bmatrix} \begin{bmatrix} g^{(1)} & 0 \\ 0 & I_m \end{bmatrix} \\ &= \begin{bmatrix} 1 & -\left(g^{(1)}\right)^* Y R_m^{-1} \\ -R_m^{-*} Y^* g^{(1)} & (Y R_m^{-1})^* Y R_m^{-1} + R_m^{-*} R_m^{-1} \end{bmatrix} = I_{m+1}, \end{aligned}$$

where we used that $\|g^{(1)}\| = 1$. The orthogonal complement of $g^{(1)}$ is the space generated by the columns of $\underline{H}_m^{(1)}$ (see (10)). Therefore, the off-diagonal blocks are zero if and only if Y is of the form $\underline{H}_m^{(1)} S_m$. Then, the trailing principal submatrix equals I_m if and only if (35) is satisfied. □

Theorem 6 can be generalized for more than two cycles, thus giving a description of all possible ways for constructing linear systems with prescribed GMRES(m) residual norms and prescribed (harmonic) Ritz values, provided there is no stagnation at the end of the restart cycles. In combination with Theorem 4, it would be possible to formulate a complete parametrization of the entire class of linear systems yielding prescribed Hessenberg matrices for the first N cycles when GMRES(m) is applied. The freedom allowed by the parametrization is in the choice of the unitary matrix V, in the individual upper triangular matrices R_m in (34) and of course in possibly undefined columns of H corresponding to cycles after the Nth cycle (which can be used to prescribe the spectrum).

We may, by the choice of R_m in (34), try to modify the residual norms for full GMRES while leaving the convergence of GMRES(m) unchanged. In the remainder of this section we merely outline how this could be done and some difficulties that can arise, without actually answering the question of whether the convergence of GMRES(m) and full GMRES can be prescribed simultaneously.

Let us decompose $\underline{H}_{2m}$ and $\hat{\underline{H}}_{2m}$ in (34) as in (32). Because H is the matrix from Theorem 4, Theorem 5 tells us that it generates the same residual norms as GMRES(m) when full GMRES is applied. Therefore, the first row of $(U_{2m})^{-1}$ contains entries χ_i satisfying

$$\chi_0 = \frac{1}{\|r_0^{(1)}\|}, \qquad |\chi_i| = \sqrt{\frac{1}{\|r_i^{(1)}\|^2} - \frac{1}{\|r_{i-1}^{(1)}\|^2}}, \quad i = 1, \dots m,$$
$$|\chi_{m+i}| = \sqrt{\frac{1}{\|r_i^{(2)}\|^2} - \frac{1}{\|r_{i-1}^{(2)}\|^2}}, \quad i = 1, \dots m-1.$$

For the decomposition of the matrix $\hat{\underline{H}}_{2m}$ in (32), we have

$$\hat{U}_{2m}^{-1} = U_{2m}^{-1} \begin{bmatrix} I_{m+1} & \underline{H}_m^{(1)} S_{m-1} \\ 0 & R_{m-1} \end{bmatrix}$$

and the first row of $\hat{U}_{2m}^{-1}$ denoted as $[\chi_0^F, \dots, \chi_{2m-1}^F]^T$ determines the residual norms generated when we apply full GMRES to $\hat{H}$ with right-hand side e_1 through

$$\chi_0^F = \frac{1}{\|r_0^F\|}, \quad |\chi_i^F| = \sqrt{\frac{1}{\|r_i^F\|^2} - \frac{1}{\|r_{i-1}^F\|^2}}, \quad i = 1, \dots 2m-1.$$

For example, the $(m+2)$nd entry of that row is

$$\chi_{m+1}^F = [\chi_0, \dots, \chi_{2m-1}] \begin{bmatrix} I_{m+1} & \underline{H}_m^{(1)} S_{m-1} \\ 0 & R_{m-1} \end{bmatrix} e_{m+2} = \chi_{m+1} r_{1,1}, \tag{36}$$

where $r_{1,1}$ is the leading entry of R_m and where we used the property (10). We have to distinguish two cases.

First, let $\chi_{m+1} = 0$, i.e., the first iteration of the restart cycle stagnates. Then necessarily $\chi_{m+1}^F = 0$, i.e., full GMRES stagnates as well, regardless of the choice of R_m. Once more, the phenomenon of stagnation puts a restriction on residual norms that can be prescribed. Clearly if we prescribe stagnation in iterations $m+1$ until $2m$ for GMRES(m), i.e., $\chi_{m+1} = \chi_{m+2} = \cdots = \chi_{2m} = 0$, then full GMRES must stagnate during these iterations, too. And conversely, using (36) again, it is easy to see that if we prescribe stagnation in iterations $m+1$ until $2m$ for full GMRES, GMRES(m) must stagnate as well in these iterations. Thus, we have the following theorem.

Theorem 7 *Let there be no stagnation at the end of the $k-1$st restart cycle of GMRES(m), $km < n$. GMRES(m) stagnates during the first j iterations of the kth restart cycle, if and only if full GMRES applied to the same system stagnates as well in the corresponding iterations (i.e., in iterations $km+1, \dots, km+j$).*

Let us now assume that $\chi_{m+1} \neq 0$ in (36). Then, the entry χ^F_{m+1} in (36) can be made any value equal or larger than χ_{m+1} (which corresponds to the inequality $\|r^F_{m+1}\| \leq \|r^{(2)}_1\|$, which must hold) with a number $r_{1,1}$ satisfying $|r_{1,1}| \geq 1$. This is the first step in finding an upper triangular matrix R_m such that $R^*_m R_m - I_m$ is positive semidefinite and such that

$$\chi^F_{m+i} = [\chi_0, \dots, \chi_{2m-1}] \begin{bmatrix} I_{m+1} & \underline{H}^{(1)}_m S_{m-1} \\ 0 & R_{m-1} \end{bmatrix} e_{m+i+1}, \qquad i = 1, \dots, m-1.$$

for given values of χ^F_{m+i} and of χ_{m+i}. Using property (10), this amounts to finding an upper triangular matrix R_m such that

$$[\chi_{m+1}, \dots, \chi_{2m-1}] R_{m-1} = [\chi^F_{m+1}, \dots, \chi^F_{2m-1}]$$

and such that $R^*_m R_m - I_m$ is positive semidefinite.

We remark that in the case where $\chi^F_{m+i} = \chi_{m+i}$, $i = 1, \dots, m-1$, R_m can be trivially chosen as the identity matrix, but other appropriate choices of R_m might result in $\chi^F_{m+i} = \chi_{m+i}$, $i = 1, \dots, m-1$, as well. Hence the systems constructed in Theorem 4 seem not to be the only systems where the full GMRES process can produce the same behavior as GMRES(m) and be computed with $m+1$-term recurrences (with prescribed upper Hessenberg matrices $\underline{H}^{(k)}_m$ for the individual cycles).

5 Conclusions and open questions

We showed that the admissible non-increasing convergence curves for restarted GMRES satisfy precisely one restriction: Stagnation at the end of a cycle is always repeated at the beginning of the next cycle. Thus it does not seem to be a good idea to restart GMRES with the current approximation if stagnation is observed. This is a strong motivation for restarting with other types of approximations or with incorporation of a deflation strategy. We further showed that neither the Ritz nor the harmonic Ritz values generated in restarted GMRES need be useful approximations of the eigenvalues. Moreover, the convergence history of restarted GMRES need not be governed by the eigenvalues.

Our parametrization of the entire class of matrices and right-hand sides yielding prescribed residual norms, eigenvalues, and (harmonic) Ritz values reveals, as a by-product, a class of matrices and right-hand sides for which full GMRES can be carried out with short, $m + 1$-term recurrences. It yields some examples of the counterintuitive behavior of restarted GMRES where a larger restart length gives slower convergence. Another result related to the connection with full GMRES is that restarted GMRES stagnates at the beginning of a cycle if and only if full GMRES stagnates in the corresponding iterations as well (where we assume restarted GMRES did not stagnate at the end of the previous cycle). Of course, in practical problems, we almost never encounter exact stagnation.

An interesting question is whether our results can be formulated for matrices with a given sparsity pattern, like those arising in finite differences or elements discretizations. Another question is whether the obtained results are valid for other restarted

Krylov subspace methods. Generalizations of analogue results for full GMRES to other (non-restarted) methods like the QMR method [16] were given in [8].

We did not show here some additional results when prescribing stagnation at the end of cycles and, necessarily, stagnation at the beginning of the next cycles. We did not either address prescribing the behavior for iteration numbers higher than the system size. This case, though not very relevant for practice, leads to an interesting theoretical challenge for possible future work.

Acknowledgments We thank Eric de Sturler for a discussion that has stimulated the search for some of the results in Section 4.

Funding information The work of J. Duintjer Tebbens was supported by the long-term strategic development financing of the Institute of Computer Science (RVO: 67985807) of the Czech Academy of Sciences.

References

1. Arioli, M., Pták, V., Strakoš, Z.: Krylov sequences of maximal length and convergence of GMRES. BIT **38**(4), 636–643 (1998)
2. Brown, P.N.: A theoretical comparison of the Arnoldi and GMRES algorithms. SIAM J. Sci. Statist. Comput. **12**(1), 58–78 (1991)
3. Chapman, A., Saad, Y.: Deflated and augmented Krylov subspace techniques. Numer. Linear Algebra Appl. **4**(1), 43–66 (1997)
4. De Sturler, E.: Truncation strategies for optimal Krylov subspace methods. SIAM J. Numer. Anal. **36**(3), 864–889 (1999)
5. Du, K., Duintjer Tebbens, J., Meurant, G.: Any admissible harmonic Ritz value set is possible for GMRES. Electron. Trans. Numer. Anal. **47**(SI), 37–56 (2017)
6. Duintjer Tebbens, J., Meurant, G.: Any Ritz value behavior is possible for Arnoldi and for GMRES. SIAM J. Matrix Anal. Appl. **33**(3), 958–978 (2012)
7. Duintjer Tebbens, J., Meurant, G.: Prescribing the behavior of early terminating GMRES and Arnoldi iterations. Num. Algor. **65**(1), 69–90 (2014)
8. Duintjer Tebbens, J., Meurant, G.: On the convergence of Q-OR and Q-MR Krylov methods for solving nonsymmetric linear systems. BIT **56**(1), 77–97 (2016)
9. Duintjer Tebbens, J., Meurant, G., Sadok, H., Strakoš, Z.: On investigating GMRES convergence using unitary matrices. Lin. Alg. Appl. **450**, 83–107 (2014)
10. Eiermann, M., Ernst, O.G.: Geometric aspects of the theory of Krylov subspace methods. Acta Numer. **10**, 251–312 (2001)
11. Eiermann, M., Ernst, O.G., Schneider, O.: Analysis of acceleration strategies for restarted minimal residual methods. J. Comput. Appl. Math. **123**(1–2), 261–292 (2000)
12. Elman, H.: Iterative Methods for Large Sparse Nonsymmetric Systems of Linear Equations. PhD thesis. Department of Computer Science. Yale University, New Haven (1982)
13. Embree, M.: The tortoise and the hare restart GMRES. SIAM Rev. **45**(2), 259–266 (2003)
14. Faber, V., Liesen, J., Tichý, P.: The Faber-Manteuffel theorem for linear operators. SIAM J. Numer. Anal. **46**(3), 1323–1337 (2008)
15. Faber, V., Manteuffel, T.: Necessary and sufficient conditions for the existence of a conjugate gradient method. SIAM J. Numer. Anal. **21**(2), 352–362 (1984)
16. Freund, R.W., Nachtigal, N.M.: QMR: A quasi-minimal residual method for non-Hermitian linear systems. Numer. Math. **60**(3), 315–339 (1991)
17. Gaul, A., Gutknecht, M.H., Liesen, J., Nabben, R.: A framework for deflated and augmented Krylov subspace methods. SIAM J. Matrix Anal. Appl. **34**(2), 495–518 (2013)
18. Greenbaum, A., Pták, V., Strakoš, Z.: Any nonincreasing convergence curve is possible for GMRES. SIAM J. Matrix Anal. Appl. **17**(3), 465–469 (1996)

19. Greenbaum, A., Strakoš, Z.: Matrices that generate the same Krylov residual spaces. In: Recent Advances in Iterative Methods, IMA Vol. Math. Appl., vol. 60, pp. 95–118. Springer, New York (1994)
20. Liesen, J., Strakoš, Z.: GMRES convergence analysis for a convection-diffusion model problem. SIAM J. Sci. Comput. **26**(6), 1989–2009 (2005). (electronic)
21. Liesen, J., Strakoš, Z.: Krylov Subspace Methods, Principles and Analysis. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford (2013)
22. Meurant, G.: Necessary and sufficient conditions for GMRES complete and partial stagnation. Submitted
23. Meurant, G., Duintjer Tebbens, J.: The role eigenvalues play in forming GMRES residual norms with non-normal matrices. Numer. Algor. **68**, 143–165 (2015)
24. Morgan, R.B.: A restarted GMRES method augmented with eigenvectors. SIAM J. Matrix Anal. Appl. **16**(4), 1154–1171 (1995)
25. Morgan, R.B.: Implicitly restarted GMRES and Arnoldi methods for nonsymmetric systems of equations. SIAM J. Matrix Anal. Appl. **21**(4), 1112–1135 (2000)
26. Morgan, R.B.: GMRES with deflated restarting. SIAM J. Sci. Comput. **24**(1), 20–37 (2002)
27. Nachtigal, N.M., Reichel, L., Trefethen, L.N.: A hybrid GMRES algorithm for nonsymmetric linear systems. SIAM J. Matrix Anal. Appl. **13**(3), 796–825 (1992)
28. Parlett, B., Strang, G.: Matrices with prescribed Ritz values. Linear Algebra Appl. **428**(7), 1725–1739 (2008)
29. Saad, Y.: Analysis of augmented Krylov subspace methods. SIAM J. Matrix Anal. Appl. **18**(2), 435–449 (1997)
30. Saad, Y., Schultz, M.H.: GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Statist. Comput. **7**(3), 856–869 (1986)
31. Schweitzer, M.: Any finite convergence curve is possible in the initial iterations of restarted FOM. Electron. Trans. Numer. Anal. **45**, 133–145 (2016)
32. Simoncini, V.: On the convergence of restarted Krylov subspace methods. SIAM J. Matrix Anal. Appl. **22**(2), 430–452 (2000)
33. Simoncini, V., Szyld, D.B.: New conditions for non-stagnation of minimal residual methods. Numer. Math. **109**(3), 477–487 (2008)
34. de Surler, E.: Personal communication (2013)
35. Vecharynski, E., Langou, J.: The cycle-convergence of restarted GMRES for normal matrices is sublinear. SIAM J. Sci. Comput. **32**(1), 186–196 (2010)
36. Vecharynski, E., Langou, J.: Any admissible cycle-convergence behavior is possible for restarted GMRES at its initial cycles. Num. Lin. Algebr. Appl. **18**, 499–511 (2011)
37. Zavorin, I., O'Leary, D.P., Elman, H.: Complete stagnation of GMRES. Linear Algebra Appl. **367**, 165–183 (2003)
38. Zhong, B., Morgan, R.B.: Complementary cycles of restarted GMRES. Numer. Linear Algebra Appl. **15**(6), 559–571 (2008)
39. Zítko, J.: Generalization of convergence conditions for a restarted GMRES. Numer. Linear Algebra Appl. **7**(3), 117–131 (2000)
40. Zítko, J.: Some remarks on the restarted and augmented GMRES method. Electron. Trans. Numer. Anal. **31**, 221–227 (2008)