Comments on the GMRES Convergence for Preconditioned Systems

Nabil Gmati[1] and Bernard Philippe[2]

[1] ENIT — LAMSIN. B.P. 37, 1002 Tunis Belvédère, Tunisie
[2] INRIA/IRISA, Campus de Beaulieu, 35042 RENNES Cedex, France

Abstract. The purpose of this paper is to comment a frequent observation by the engineers studying acoustic scattering. It is related to the convergence of the GMRES method when solving systems $Ax = b$ with $A = I - B$. The paper includes a theorem which expresses the convergence rate when some eigenvalues of B have modulus larger than one; that rate depends on the rate measured when solving the system obtained by spectral projection onto the invariant subspace corresponding to the other eigenvalues. The conclusion of the theorem is illustrated on the Helmholtz equation.

1 Introduction

The purpose of this paper is to comment a frequent observation by the engineers studying acoustic scattering. It is related to the convergence of the GMRES method when solving systems $Ax = b$ with $A = I - B \in \mathbb{R}^{n\times n}$. If the spectral radius $\rho(B)$ is smaller than one, it is easy to see that GMRES converges better than the convergence obtained by the Neumann series $A^{-1} = \sum_{k\geq 0} B^k$; engineers usually claim that when some eigenvalues of B lie out the unit disk, GMRES still converges. Our attempt is to explain that effect.

First, let us define the context precisely. For any matrix $A \in \mathbb{R}^{n\times n}$ and from an initial guess x_0, the GMRES method [12] iteratively builds a sequence of approximations x_k $(k \geq 1)$ of the solution x such that the residual $r_k = b - Ax_k$ satisfies

$$\|r_k\| = \min_{\substack{q \in \mathcal{P}_k \\ q(0)=1}} \|q(A)r_0\| \,, \tag{1}$$

where $\mathcal{P}_k$ is the set of polynomials of degree k or less. Throughout this paper, the considered norms are the 2-norms.

Since the implementation of the method involves the construction by induction of an orthonormal system V_k which makes its storage mandatory, the method needs to be restarted at some point $k = m$ to be tractable. The corresponding method is denoted $GMRES(m)$.

The studied decomposition, $A = I - B$, arises from the introduction of a preconditioner as usually done to improve the convergence. This is a non singular

I. Lirkov, S. Margenov, and J. Waśniewski (Eds.): LSSC 2007, LNCS 4818, pp. 40–51, 2008.
© Springer-Verlag Berlin Heidelberg 2008

matrix M that is easy to invert (i.e. solving systems $My = c$ is easy) and such that M^{-1} is some approximation of A^{-1}. It can be applied on the left side of the system by solving $M^{-1}Ax = M^{-1}b$ or on the right side by solving $AM^{-1}y = b$ with $x = M^{-1}y$. By considering the corresponding splitting $A = M - N$, we shall investigate the behavior of GMRES on the system

$$(I - B)x = b \ , \tag{2}$$

where B is respectively $B = M^{-1}N$ (left side) or $B = NM^{-1}$ (right side). By rewriting condition (1) in that context, the sequence of residuals satisfies (see the proof of Theorem 1) :

$$\|r_k\| = \min_{\substack{q \in \mathcal{P}_k \\ q(1)=1}} \|q(B)r_0\| \ . \tag{3}$$

In this paper, after some comments on various interpretations of the convergence of GMRES, we first consider the situation where $\rho(B) < 1$. We demonstrate that it is hard to have bounds that illustrate the usual nice behavior of GMRES in such situations. Therefore, for the situation where some eigenvalues of B lie outside the unit disk, we propose to link the residual evolution to the residual evolution of a projected system, which discards all the outer eigenvalues.

2 About the Convergence of GMRES

In exact arithmetic and in finite dimension n, GMRES should be considered as a direct method since for every x_0, there exists a step $k_{end} \leq n$ such that $x_{k_{end}} = x$. Unfortunately, computer arithmetic is not exact and in most situations $k_{end} = n$ is much too big to be reached. However, when B is rank deficient, $k_{end} < n$. For instance, when the preconditioner M is defined by a step of the Multiplicative Schwarz method obtained from an algebraic overlapping 1-D domain decomposition, it can be proved that k_{end} is no bigger than the sum of the orders of the overlaps [1].

When the residuals does not vanish at some point $k \leq m$ where m is the maximum feasible size for the basis V_k, restarting becomes necessary. In that situation, the method GMRES(m) builds a sequence of approximations $x_0^{(K)}$ corresponding to all the initial guesses of the outer iterations. It is known that, unless the symmetric part of $I - B$ is positive or negative definite, stagnation may occur which prevents convergence.

In this paper, we limit the study of the convergence to the observation of the residual decrease during one outer iteration.

2.1 Existing Bounds

In [5], Embree illustrates the behavior of several known bounds. He proves that, for non normal matrices, any of the bounds can be overly pessimistic in some situations. Here, we only consider two bounds in the case where $\rho(B) < 1$.

By selecting the special polynomial $q(Z) = z^k$ in (3), we get the bound

$$\|r_k\| \leq \|B^k\| \, \|r_0\| \, . \tag{4}$$

The bound indicates that asymptotically there should be at least a linear convergence rate when $\rho = \rho(B) < 1$. For instance, if B can be diagonalized by $B = XDX^{-1}$, then the following bound holds :

$$\|r_k\| \leq cond(X)\rho^k \, \|r_0\| \, , \tag{5}$$

where $cond(X)$ is the condition number of the matrix of eigenvectors. It is clear that the bound might be very poor for large values of the condition number. For instance, it might happen that $cond(X)\rho^n > 1$, in which case no information can be drawn from (5).

The bound can be improved by an expression involving the condition numbers of the eigenvalues [5]. These bounds can be generalized when matrix B involves some Jordan blocks [13]. However, as mentioned earlier, there are always situations in which the bounds blow up.

In order to get a bound $\|r_k\| \leq K\gamma^k$ with a constant K smaller than the condition numbers of the eigenvalues, other expressions have been obtained from the field $W(A)$ of values of A ($W(A) = \{u^H Au | u \in \mathbb{C}^n, \ \|u\| = 1\}$). That set is convex and includes the eigenvalues (and therefore their convex hull). It is contained in the rectangle limited by the extremal eigenvalues of the symmetric and skew-symmetric parts of A. It is easy to see that $W(A) = 1 - W(B)$. When A is positive definite (or negative definite), which is equivalent to assuming that $1 \notin W(B)$, Beckermann [2] proved that

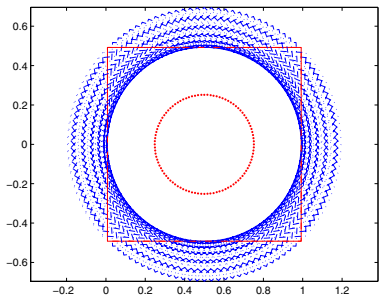
$$\|r_k\| \leq (2 + \gamma)\gamma^k \|r_0\| \tag{6}$$

where $\gamma = 2\sin\left(\frac{\beta}{4-2\beta/\pi}\right) < \sin\beta$ with β defined by : $\cos\beta = \frac{dist(0,W(A))}{\max|W(A)|}$. That result slightly improved earlier results [4].

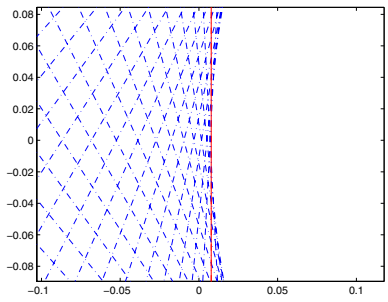
The behaviors of the two bounds (5) and (6) are illustrated on a special test matrix. The chosen matrix is built by the following instruction (in MATLAB syntax): `A=0.5*eye(100)+0.25*gallery('smoke',100)`. The parameters which are involved in the two bounds are displayed in Table 1. In Figure 1, the field of values of A is displayed (inner domain of the shaded region) as well as a zoom of the set at the origin. Below, the explicit bounds are plotted and compared to the residuals of GMRES applied to $Ax = b$ with $b = (1, \cdots, 1)^T/\sqrt{n}$. Although that matrix may be considered rather special, the behavior is common to the situation of non-normal matrices. The expressions of the bounds are of

Table 1. Characteristics of the matrix used in Figure 1

Spectral radius of $B = I - A$	: 0.752
Condition number of the eigenvectors	: 6.8×10^{13}
Parameter γ (from the field of values)	: 0.996



Field of values of A
(Eigenvalues plotted inside)



Zoom near the origin.

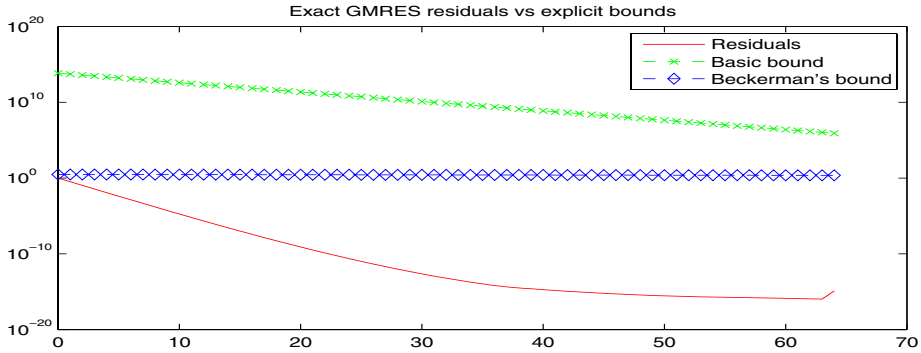


Fig. 1. Classical behavior of explicit bounds

the shape $K\gamma^k$, and the non-normality will impact either the constant K which blows up or the rate γ which will be very close to 1.

From that experience, it is clear that nothing can be expected from explicit bounds to illustrate the claim of the engineers mentioned in the introduction. The obstacle is twofold: (i) even for $\rho(B) < 1$ the bounds are often useless; (ii) the bounds have a linear behavior. Several authors have attempted to exhibit superlinear bounds (see for instance [9,10]) but the advantage is more theoretical than practical.

In that perspective, it is better to show how some eigenvalues of B outside the unit disk may deteriorate the convergence with respect to the situation $\rho(B) < 1$.

2.2 When Some Eigenvalues of B Lie Outside the Unit Disk

Let us denote by $D = \{z \in \mathbb{C},\ |z| < 1\}$ the open unit disk and by $(\lambda_i)_{i=1,n}$ the eigenvalues of B.

Theorem 1. *If $A = I - B \in \mathbb{R}^{n\times n}$ is non singular, with p eigenvalues of B outside the open unit disk D, then for GMRES and for $k \geq p$:*

$$\|r_k\| \leq K\|r_{k-p}^{red}\|$$

where r_k^{red} is the residual corresponding to the use of GMRES on the system projected onto the invariant subspace excluding the exterior eigenvalues; the constant K only depends on the eigenvalues exterior to D and on the operator projected onto the invariant subspace of all other eigenvalues. All projections considered are spectral.

Proof. When applied to the system $(I - B)x = b$ and starting from an initial guess x_0, the GMRES iteration builds a sequence of iterates x_k such that the corresponding residual can be expressed in polynomial terms: $r_k = \pi(I - B)\ r_0$ where the polynomial $\pi \in P_k(\mathbb{C})$ is a polynomial of degree no larger than k and such that $\pi(0) = 1$. The polynomial π minimizes the Euclidean norm $\|r_k\|$ and it is uniquely defined as long as the exact solution of the system is not obtained. Through a change of variable $\tau(z) = \pi(1 - z)$, the residual is expressed as $r_k = \tau(B)\ r_0$ where the normalizing condition becomes $\tau(1) = 1$. Therefore, for any polynomial $\tau \in \mathcal{P}_k(\mathbb{C})$ such that $\tau(1) = 1$, the following bound stands:

$$\|r_k\| \leq \|\tau(B)\ r_0\|. \tag{7}$$

We shall now build a special residual polynomial. For that purpose, let us decompose the problem onto the two supplementary invariant subspaces $span(X_1) \oplus span(X_2)$ where X_1 and X_2 are bases of the invariant subspaces corresponding to respectively the p largest eigenvalues of B and the $n - p$ other eigenvalues. By denoting $X = [X_1, X_2]$, $Y = X^{-T} = [Y_1, Y_2]$ and $P_i = X_i Y_i^T$ the spectral projector onto $span(X_i)$ for $i = 1, 2$, the matrix B can be decomposed into $B = B_1 + B_2$ where $B_i = P_i B P_i$ for $i = 1, 2$. Therefore the nonzero eigenvalues of B_1 lie outside D whereas the spectrum of B_2 is included in D. Moreover, for any polynomial π, the decomposition $\pi(B) = \pi(B_1)P_1 + \pi(B_2)P_2$ holds.

Let π_1 the polynomial of degree p that vanishes at each eigenvalue $\lambda_i \notin D$ and such that $\pi_1(1) = 1$. The polynomial π_1 is uniquely defined by

$$\pi_1(z) = \prod_{\lambda_i \notin D} \frac{z - \lambda_i}{1 - \lambda_i}. \tag{8}$$

By construction, that polynomial satisfies the following property

$$\pi_1(B_1)P_1 = 0. \tag{9}$$

Let τ_2 be the residual polynomial corresponding to the iteration $k - p$ when solving the system projected onto the invariant subspace $span(X_2)$ (spectral projection):

$$\begin{cases} (P_2 - B_2)x = P_2 b, \\ P_1 x = 0. \end{cases}$$

Therefore the residual of the reduced system is $r_{k-p}^{red} = \tau_2(B_2)P_2 r_0$. By considering the polynomial $\tau = \pi_1 \tau_2$, since $\tau(B_1)P_1 = 0$ we get

$$\begin{aligned} \tau(B)\ r_0 &= \tau(B_1)\ P_1 r_0 + \tau(B_2)\ P_2 r_0\ , \\ &= \pi_1(B_2)\ \tau_2(B_2)\ P_2 r_0\ , \\ &= \pi_1(B_2)\ r_{k-p}^{red}\ , \\ &= \pi_1(B_2)P_2\ r_{k-p}^{red}\ . \end{aligned}$$

The last transformation is not mandatory but highlights that the operator which defines the constant K is zero on $span(X_1)$. By inequality (7), the conclusion of the theorem holds with $K = \|\pi_1(B_2)P_2\|$.

The result of the theorem may be interpreted informally as saying that p eigenvalues of B outside the unit disk only introduce a delay of p steps in the convergence of GMRES. This fact will be illustrated in the experiments of the next section.

3 Illustration

3.1 Solving a Helmholtz Problem

Let $\Omega_i \subset \mathbb{R}^3$ be a bounded obstacle with a regular boundary Γ and Ω_e be its unbounded complementary. The Helmholtz problem models a scattered acoustic wave propagating through Ω_e; it consists in determining u such that

$$\begin{cases} \Delta u + \kappa^2 u = 0 \text{ in } \Omega_e, \\ \partial_n u = f \text{ on } \Gamma, \\ (\frac{x}{|x|}.\nabla - i\kappa)u = e^{i\kappa|x|}O(\frac{1}{|x|^2}) \quad x \in V_\infty, \end{cases} \tag{10}$$

where κ is the wave number and where V_∞ is a neighborhood of infinity. The last condition represents the Sommerfeld radiation condition. To solve the boundary value problem (10), we may consider an integral equation method (eventually coupled to finite element method for non-constant coefficients). The efficiency of this approach has been investigated by several authors, e.g. [8,11]. An alternative approach consists in using a coupled method which combines finite elements and integral representations [6]. This approach avoids the singularities of the Green function. The idea simply amounts to introducing a fictitious boundary Σ surrounding the obstacle. The Helmholtz problem is posed in the truncated domain Ω_c (delimited by Γ and Σ) with a non-standard outgoing condition using the integral formula which is specified on Σ,

$$\Delta u + \kappa^2 u = 0 \text{ in } \Omega_c, \partial_n u = f \text{ on } \Gamma, \tag{11}$$

$$(\partial_n - i\kappa)u(x) = \int_\Gamma (u(y)\partial_n K(x-y) - f(y)K(x-y))\,\mathrm{d}\gamma(y), \forall x \in \Sigma, \tag{12}$$

where $K(x) = (\frac{x}{|x|}.\nabla - i\kappa)G_\kappa(x)$ and G_κ is the Green function. Observe that the integral representation is used only on Σ which avoids occurrences of singularities. We suppose that this problem is discretized by a Lagrange finite element method. Let N_{Ω_c} be the total number of degrees of freedom on Ω_c and N_Σ (resp. N_Γ) be the number of degrees of freedom on Σ (resp. Γ). The shape function associated with node x_α is denoted w_α. Let u_α be the approximation of the solution u at x_α and $v = (u_\alpha)_\alpha$.

The linear system can be formulated as follows:

$$(A - C)v = b \tag{13}$$

where

$$A_{\alpha,\beta} = \int_{\Omega_c} \left(\nabla w_\alpha(x)\nabla\overline{w}_\beta(x) - \kappa^2 w_\alpha(x)\overline{w}_\beta(x)\right) dx - i\kappa \int_\Sigma w_\alpha(x)\overline{w}_\beta(x)d\sigma(x)$$

$$C_{\alpha,\beta} = \int_\Sigma \int_\Gamma w_\alpha(x)\partial_n K(x-y)\ d\gamma(y)\overline{w}_\beta(x)\ d\sigma(x)$$

The matrix of the system (13) is complex, non Hermitian and ill-conditioned. Matrix A is a sparse matrix but matrix C, which represents non-local coupling terms enforced by the integral term, is dense. Since, solving a system in A is easier than solving system (13), the matrix A is chosen as a preconditioner. In other words, the linear system is formulated as follow:

$$(I_{N_{\Omega_c}} - B)v = c, \tag{14}$$

where $B = A^{-1}C$ and $c = A^{-1}b$. Direct methods adapted to sparse matrices are good candidates for solving the preconditioning step.

3.2 Definition of the Test Problems

The numerical results deal with examples of acoustic scattering with an incident plane wave. Four tests are defined (see Figure 2).

The scatterer is considered to be a cavity. In the first test, the computational domain is rectangularly shaped while in the second and third tests, the domains are nonconvex. The cavity of Test 2 is the same as in Test 1. Cavities and domains are the same for Test 3 and Test 4. The meshing strategy implies two rows of finite elements in the non convex cases (Tests 2, 3 and 4). The characteristic parameters for every test are listed in Table 2.

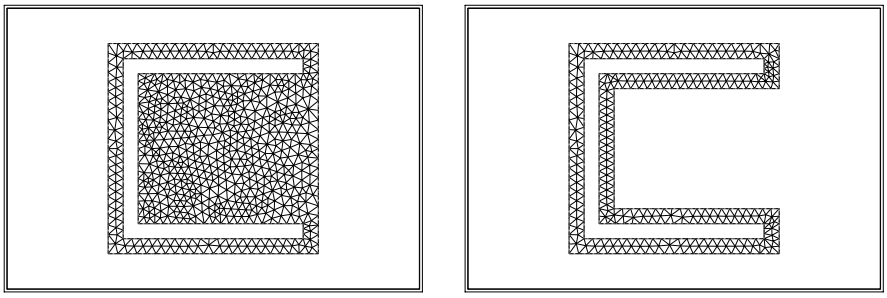
Table 2. Characteristics of the tests

	N_{el}	N_{Ω_c}	N_Σ	h	κ
Test # 1	1251	734	88	0.06	1
Test # 2	543	403	134	0.06	1
Test # 3	2630	1963	656	0.005	1
Test # 4	2630	1963	656	0.005	62

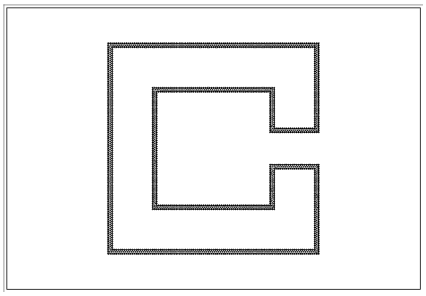
N_{el} : number of finite elements;
N_{Ω_c} : number of degrees of freedom in Ω_c;
N_Σ : number of degrees of freedom in Σ;
h : mean diameter of the mesh;
κ : wave number.

3.3 Numerical Results

For the four tests, system (13) is solved by GMRES preconditioned by A which corresponds to solving system (14). In Figures 3 and 4, for every test, the plot



Test problem # 1 Test problem # 2



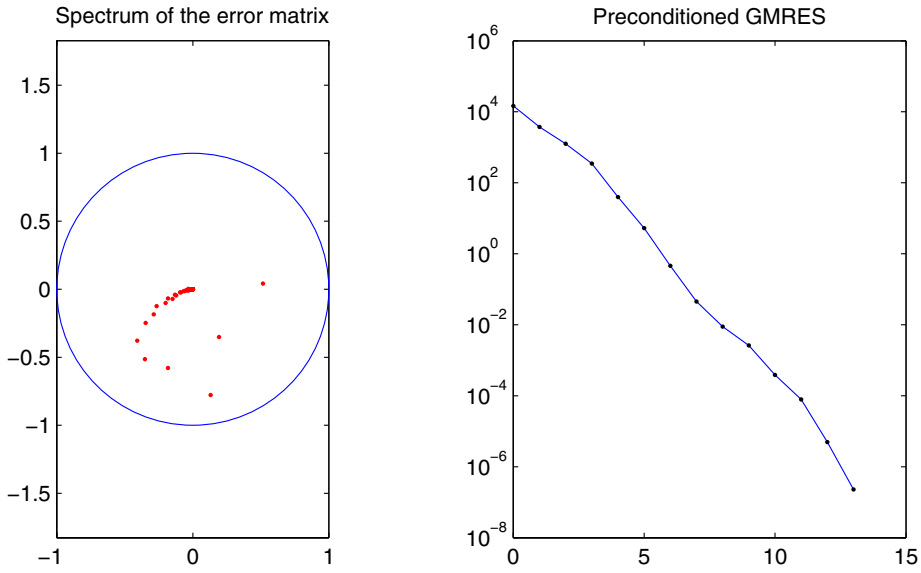
Test problems # 3 & # 4

Fig. 2. Cavities, domains and characteristics for the test problems

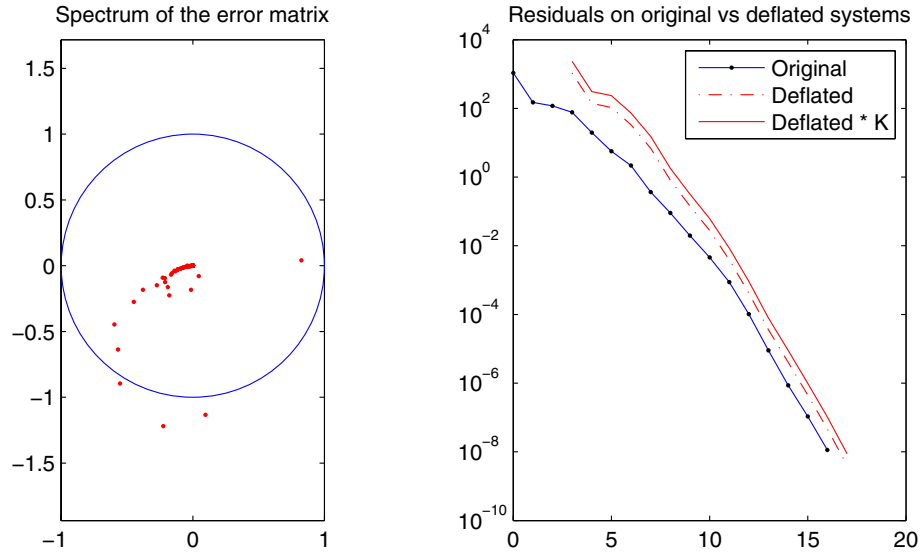
of the eigenvalues of B is reported, with respect to the unit disk (left figure), as well as the sequence of the GMRES residuals (right figure). Parameter p denotes the number of eigenvalues outside the unit disk. In order to illustrate Theorem 1, when $p > 0$ the residuals of the projected system are plotted with a shift (delay) p on the iteration numbers (dashed line). The same curve is plotted (solid line) with the residuals multiplied by the constant K in the bound given by the theorem.

The study of the spectrum of $B = A^{-1}C$ in Test 1 and Test 2 shows that when the fictitious boundary is located far from the cavity (Test 1), the whole spectrum is included in the unit disk. In that situation, GMRES converges in a few iterations. In Test 2, three eigenvalues exit the unit disk which deteriorates the convergence. This is foreseeable by reminding the connection between the method coupling Finite Element Method with an integral representation and the Schwarz method: the larger computational domain (which corresponds to the overlapping in the Schwarz method) — the faster the convergence [3,7].

In Test 3 and Test 4, the same cavity with the same domain and the same discretization are considered. Only the wave numbers differ. The mesh size corresponds to twenty finite elements by wave length when $\kappa = 62$. When κ increases, the number of eigenvalues outside the unit disk decreases. However, it can be

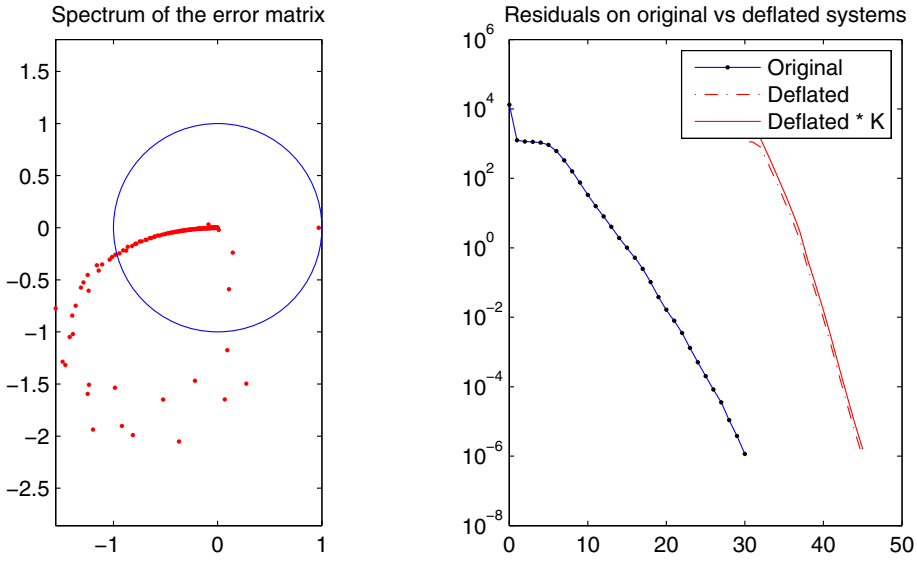


Helmholtz Test 1 – $p = 0$.

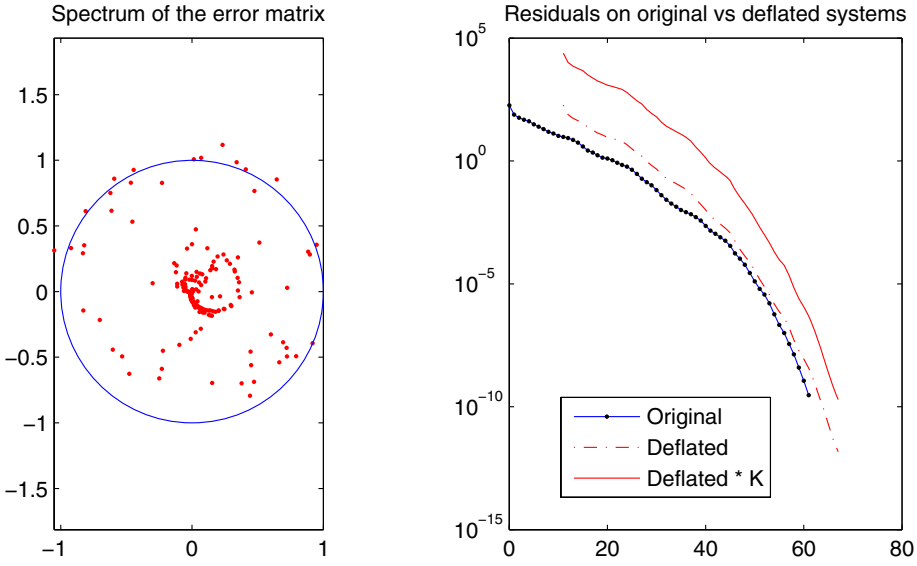


Helmholtz Test 2 – $p = 3$.

Fig. 3. Spectra of B and GMRES residual evolutions



Helmholtz Test 3 – $p = 29$.



Helmholtz Test 4 – $p = 11$.

Fig. 4. Spectra of B and GMRES residual evolutions

seen that the convergence is not necessarily improved: the number of iterations is smaller when $\kappa = 1$ inducing $p = 29$ compared to $p = 11$ when $\kappa = 62$. This illustrates the difficulty of characterizing the convergence by only considering the modulus of the eigenvalues of B. Clustering of eigenvalues and non-normality of the operator play an important role.

In the situations where $p > 0$ (Test 2, 3 and 4), the bound given by the theorem appears to be often quite accurate. The worst case arises in Test 3 in which the number of outer eigenvalues is high ($p = 29$); in that situation the delay introduced by the shift p in the iteration number is too high because the convergence occurs at the same time that all the outer eigenvalues are killed.

4 Conclusion

Exhibiting realistic bounds for the successive residuals obtained when solving a linear system $Ax = b$ by GMRES is hopeless except in special situations. The knowledge of the spectrum of the operator is not even sufficient to fully understand the behavior of the convergence. These comments were already discussed by several authors but in this paper we have tried to explain the favorable situation where only a few eigenvalues of B ($A = I - B$) are of modulus larger than 1.

References

1. Atenekeng-Kahou, G.A., Kamgnia, E., Philippe, B.: An explicit formulation of the multiplicative Schwarz preconditioner. Research Report RR-5685, INRIA (2005) (Accepted for publication in APNUM)
2. Beckermann, B.: Image numérique, GMRES et polynômes de Faber. C. R. Acad. Sci. Paris, Ser. I 340, 855–860 (2005)
3. Ben Belgacem, F., et al.: Comment traiter des conditions aux limites à l'infini pour quelques problèmes extérieurs par la méthode de Schwarz alternée. C. R. Acad. Sci. Paris, Ser. I 336, 277–282 (2003)
4. Eisenstat, S.C., Elman, H.C., Schultz, M.H.: Variational iterative methods for non-symmetric systems of linear equations. SIAM Journal on Numerical Analysis 20(2), 345–357 (1983)
5. Embree, M.: How descriptive are GMRES convergence bounds?, `http://citeseer.ist.psu.edu/196611.html`
6. Jami, A., Lenoir, M.: A new numerical method for solving exterior linear elliptic problems. In: Lecture Notes in Phys., vol. 90, pp. 292–298. Springer, Heidelberg (1979)
7. Jelassi, F.: Calcul des courants de Foucault harmoniques dans des domaines non bornés par un algorithme de point fixe de cauchy. Revue ARIMA 5, 168–182 (2006)
8. Johnson, C., Nédélec, J.-C.: On the coupling of boundary integral and finite element methods. Math. Comp. 35(152), 1063–1079 (1980)
9. Kerhoven, T., Saad, Y.: On acceleration methods for coupled nonlinear elliptic systems. Numer. Maths 60(1), 525–548 (1991)

10. Moret, I.: A note on the superlinear convergence of GMRES. SIAM Journal on Numerical Analysis 34(2), 513–516 (1997)
11. Nédélec, J.-C.: Acoustic and electromagnetic equations Integral Representations for Harmonic Problems. In: Applied Mathematical Sciences, vol. 144, Springer, NewYork (2001)
12. Saad, Y., Schultz, M.H.: GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Stat. Comput. 7(3), 856–869 (1986)
13. Tichý, P., Liesen, J., Faber, V.: On worst-case GMRES, ideal GMRES, and the polynomial numerical hull of a Jordan block. Electronic Transactions on Numerical Analysis (submitted, June 2007)