Numer. Math. 70: 73–89 (1995)

Numerische Mathematik
© Springer-Verlag 1995

Analysis of some vector extrapolation methods for solving systems of linear equations

K. Jbilou, H. Sadok

Laboratoire d'Analyse Numérique et d'Optimisation, UFR IEEA - M3, Université des Sciences et Technologies de Lille, F-59655 Villeneuve d'Ascq Cedex, France

Received January 11, 1994

Summary. In this paper, we consider some vector extrapolation methods for solving nonsymmetric systems of linear equations. When applied to sequences generated linearly, these methods namely the minimal polynomial extrapolation (MPE) and the reduced rank extrapolation (RRE), are Krylov subspaces methods and are respectively equivalent to the method of Arnoldi and to the GCR and GMRES. By considering the geometrical aspect of these methods, we derive new expressions for their residual norms and give a relationship between them; this allows us to compare the two methods. Using this new approach, we will show that for nonsingular skew symmetric matrices the GMRES stagnates every two iterations and the restarted version GMRES(m) ($m \geq 2$) is always convergent. Finally, the incomplete forms are considered and some convergence results are given.

Mathematics Subject Classification (1991): 65F10

1. Introduction

The purpose of the present work is to consider some vector extrapolation methods and their incomplete forms for solving linear systems of equations. The most popular vector extrapolation methods are the reduced rank extrapolation (RRE) of Eddy [6] and Mesina [14], the minimal polynomial extrapolation (MPE) of Cabay and Jackson [5], the topological epsilon algorithm (TEA) of Brezinski [3] and the modified minimal polynomial extrapolation (MMPE) of Pugatchev [15], Brezinski [3] and Sidi, Ford and Smith [22]. An extensive survey of these extrapolation methods has been carried out in [18], [21], [23] and [24]. In this paper, we will restrict ourselves to the first two methods.

It has been shown in [18], [10] and [1] that, when applied to sequences generated linearly, the MPE and the RRE are Krylov subspaces methods. In [19], Sidi showed that they are respectively equivalent to the method of Arnoldi considered by Saad [16] and the GCR of Eisenstat, Elman and Schultz [7] which is equivalent to the GMRES of Saad and Schultz [17]. An efficient implementation of the RRE and the MPE has been given in [20] and recently, in [11], new algorithms for the computation of the RRE and the MMPE were given and connections with other known projection algorithms have been developed.

As these methods could be regarded as oblique projection methods, we will see that their analysis is more suitable by using projector operators and angles between vectors and subspaces. In particular, we will derive useful expressions for the residual norms of the two methods and give a relationship between them. These results will be used to give a comparison of the performance of the two methods. We will also treat the problem of stagnation of the RRE which is mathematically equivalent to the GMRES. With regard to this problem, we give two important results for skew-symmetric matrices. We first show that for this class of matrices, we have stagnation of the full GMRES (or RRE) every two iterations. The second result is the fact that, when the matrix is skew symmetric, the restarted version of GMRES denoted GMRES(m) converges for any starting vector when $m \geq 2$.
In the last section, we will consider the incomplete forms of the MPE and RRE and give some convergence results.

Notations: For a matrix C, let $C_{\mathrm{s}} = \dfrac{C + C^{\mathrm{T}}}{2}$ and $C_{\mathrm{t}} = -\dfrac{C - C^{\mathrm{T}}}{2}$ be respectively the symmetric and the skew symmetric part of the matrix $C = C_{\mathrm{s}} - C_{\mathrm{t}}$. For any square matrix X, let $\lambda_{\min}(X)$ and $\lambda_{\max}(X)$ denote the eigenvalues of X of smallest and largest modulus,
$\rho(X) = \mid \lambda_{\max}(X) \mid$ the spectral radius of X. We shall denote by $\mathrm{span}\{v_1, v_2, \ldots, v_k\}$ the subspace spanned by $v_1, v_2, \ldots, v_k$ and for any vector v, $K_k(v, C)$ will denote the Krylov subspace spanned by $v, Cv, \ldots, C^{k-1}v$.

2. Vector extrapolation methods

Let (s_n) be a sequence of vectors of $\mathbb{R}^N$ and let us define the first and the second forward differences of s_n by

$$\Delta s_n = s_{n+1} - s_n \quad \text{and} \quad \Delta^2 s_n = \Delta s_{n+1} - \Delta s_n.$$

It has been shown in [18] that all the four extrapolation methods, namely the MPE, the RRE, the MMPE and the TEA, when applied to the sequence (s_n), produce an approximation $t_k^{(n)}$ of the limit or the antilimit of (s_n). This approximation is defined by

$$t_k^{(n)} = \sum_{j=0}^{k} \alpha_j s_{n+j} \tag{2.1}$$

such that

$$\sum_{j=0}^{k} \alpha_j = 1 \tag{2.2}$$

and

$$\sum_{j=0}^{k} \eta_{ij} \alpha_j = 0, \qquad i = 0, \ldots, k-1 \tag{2.3}$$

where the scalars η_{ij} are defined by

$$\begin{aligned}
\eta_{ij} &= (\Delta s_{n+i}, \Delta s_{n+j}) && \text{for the MPE}\\
\eta_{ij} &= (\Delta^2 s_{n+i}, \Delta s_{n+j}) && \text{for the RRE}\\
\eta_{ij} &= (y_{i+1}, \Delta s_{n+j}) && \text{for the MMPE}\\
\eta_{ij} &= (y, \Delta s_{n+i+j}) && \text{for the TEA}
\end{aligned}$$

and $\{y_1, \dots, y_k\}$ is a set of linearly independent vectors of $\mathbb{R}^N$ and y an arbitrary fixed vector.

Now, from (2.1), (2.2) and (2.3), $t_k^{(n)}$ can be expressed as a ratio of two determinants

$$(2.4) \quad t_k^{(n)} = \begin{vmatrix} s_n & s_{n+1} & \dots & s_{n+k} \\ \eta_{0,0} & \eta_{0,1} & \dots & \eta_{0,k} \\ \vdots & \vdots & \vdots & \\ \eta_{k-1,0} & \eta_{k-1,1} & \dots & \eta_{k-1,k} \end{vmatrix} \Big/ \begin{vmatrix} 1 & 1 & \dots & 1 \\ \eta_{0,0} & \eta_{0,1} & \dots & \eta_{0,k} \\ \vdots & \vdots & \vdots & \\ \eta_{k-1,0} & \eta_{k-1,1} & \dots & \eta_{k-1,k} \end{vmatrix} .$$

Let us denote Y_k and $\Delta^i S_n (i = 1, 2)$ the matrices whose columns are respectively $y_1, \dots, y_k$ and $\Delta^i s_n, \dots, \Delta^i s_{n+k-1}$. With these notations, $t_k^{(n)}$ can be written in a matrix form as

$$(2.5) \qquad t_k^{(n)} = s_n - \Delta S_n (Y_{k,n}^{\mathrm{T}} \Delta^2 S_n)^{-1} Y_{k,n}^{\mathrm{T}} \Delta s_n$$

where $Y_{k,n} = Y_k$ for the MMPE, $Y_{k,n} = \Delta^2 S_n$ for the RRE and $Y_{k,n} = \Delta S_n$ for the MPE.

It is clear that $t_k^{(n)}$ exists if and only if $\det(Y_{k,n}^{\mathrm{T}} \Delta^2 S_n) \neq 0$, the conditions to be satisfied for this have been given in [19].

The computation of $t_k^{(n)}$ could be done by some algorithms proposed in [8] and [12].

Let us consider now the system of linear equations

$$(2.6) \qquad Ax = f$$

where A is an $N \times N$ real nonsingular matrix and b is a vector of $\mathbb{R}^N$.

If we express the matrix A in the form

$$A = M - L$$

where M and L are two matrices of order N such that M is nonsingular, the system (2.6) is written as

$$(2.7) \qquad x = M^{-1} L x + M^{-1} f.$$

Such a technique includes the Jacobi, Gauss-Seidel and S.O.R methods. We define the sequence $\{s_j\}$ starting from s_0 by

$$(2.8) \qquad s_{j+1} = M^{-1} L s_j + M^{-1} f, \quad j = 0, 1 \dots .$$

Setting $B = M^{-1} L$ and $b = M^{-1} f$, (2.8) becomes

$$(2.9) \qquad s_{j+1} = B s_j + b, \quad j = 0, 1 \dots .$$

Let $C = I - B$ and define the residual $r(x)$ for a vector x by

$$r(x) = b - Cx.$$

Note that for $j = 0, 1, \dots$

$$r(s_j) = b - C\,s_j = \Delta s_j$$

and

$$\Delta^2 s_j = \Delta s_{j+1} - \Delta s_j = -C\Delta s_j. \tag{2.10}$$

It is known that, when applied to sequences generated linearly, the four extrapolation methods give the exact solution of the system (2.5).

Let us consider the following algorithm

$$A1 \quad \begin{cases} x_0 \text{ is a given vector, } s_0 = x_0,\ s_1 = Bs_0 + b \\ \text{for } k = 1, 2, \dots \\ \qquad s_{j+1} = Bs_j + b, \qquad j = \varphi(k-1)+1, \dots, \varphi(k) \\ \\ \qquad x_k = t_k^{(0)} \end{cases}$$

where

$\varphi(k) = k$ for the MPE, MMPE and RRE,
$\varphi(k) = 2k - 1$ for the TEA.

Remarks. 1. The convergence of the sequence (x_k) produced by the algorithm $A1$ is achieved in a finite number of iterations.

2. The linear system that we shall consider is the preconditionned one

$$C\,x = b \tag{2.12}$$

which is equivalent to (2.5).

In a matrix form, the iterate x_k can be expressed as

$$x_k = s_0 - \Delta S_0 (Y_{k,0}^{\mathrm{T}} \Delta^2 S_0)^{-1}\, Y_{k,0}^{\mathrm{T}} \Delta s_0. \tag{2.13}$$

Before closing this section, let us notice that since $\Delta^2 s_j = -C\,\Delta s_j = -Cr(s_j)$ for $j = 0, 1, \dots$, we have $\Delta^2 S_0 = -C\,\Delta S_0$.

In the sequel, let V_k be the subspace of $\mathbb{R}^N$ of dimension k spanned by $\Delta s_0, \dots, \Delta s_{k-1}$ and set $W_k = C\,V_k$. Thus the subspace W_k is spanned by $\Delta^2 s_0, \dots, \Delta^2 s_{k-1}$.

3. Analysis of the RRE

In order to obtain the RRE, we take $Y_{k,0} = \Delta^2 S_0$. In this case, the iterate x_k can be written as

$$x_k = s_0 - \Delta S_0 \Delta^2 S_0^{+} \Delta s_0 \tag{3.1}$$

where $\Delta^2 S_0^{+}$ is the pseudo-inverse [2] defined by

$$\Delta^2 S_0^{+} = (\Delta^2 S_0^{\mathrm{T}} \Delta^2 S_0)^{-1} \Delta^2 S_0^{\mathrm{T}}. \tag{3.2}$$

We also have

$$
\begin{aligned}
r_k = b - C\,x_k &= b - Cs_0 + C\Delta S_0 \Delta^2 S_0^+\,\Delta s_0 \\
&= \Delta s_0 - \Delta^2 S_0 \Delta^2 S_0^+\,\Delta s_0 \\
&= (I - \Delta^2 S_0\,\Delta^2 S_0^+)\,\Delta s_0. \qquad (3.3)
\end{aligned}
$$

Let us give now the next result to be used later

Theorem 1. *For any subspace S of $\mathbb{R}^N$, let P_S be the l_2-orthogonal projector onto S. Let F be a real $N \times p$ matrix and let us denote by $\Re(F)$ and $\aleph(F)$ the range and the null spaces of F, then*

a) $P_{\Re(F)} = F\,F^+$

b) $P_{\aleph(F^{\mathrm{T}})} = I - F\,F^+$

where $F^+ = (F^{\mathrm{T}}\,F)^{-1}\,F^{\mathrm{T}}$.

Proof. see [2]. □

We can state now the following result

Theorem 2. *Let P_k be the l_2-orthogonal projector onto W_k, then*

$$r_k = (I - P_k)\,r_0. \qquad (3.4)$$

Proof. Setting $F = \Delta^2 S_0$ in the preceding theorem we have $P_k = \Delta^2 S_0\,\Delta^2 S_0^+$, thus using (3.3), the result follows. □

Note that if $k = N$, then $W_k = \mathbb{R}^N$ and $P_N r_0 = r_0$, hence $r_N = 0$. Let us mention here that since $I - P_k$ is also an orthogonal projector onto $W_k^\perp$, the RRE is an orthogonal projection method while the other extrapolation methods are oblique projection methods (see [19]).

The acute angle between a vector x and a subspace S is defined by

$$\cos(x, S) = \max_{y \in S - \{0\}} \cos(x, y) = \max_{y \in S - \{0\}} \frac{|(x, y)|}{||x||\,||y||}.$$

Theorem 3. *Let θ_k be the acute angle between r_0 and the subspace W_k, then*

$$\| r_k \|^2 = (1 - \cos^2\theta_k)\ \| r_0 \|^2 . \qquad (3.5)$$

Proof. From Theorem 2 we have $r_k = r_0 - P_k\,r_0$, then

$$
\begin{aligned}
(r_k, r_k) &= (r_k, r_0 - P_k\,r_0) \\
&= (r_k, r_0) - (r_k, P_k\,r_0).
\end{aligned}
$$

But, since $(r_k, P_k\,r_0) = 0$, it follows that $\| r_k \|^2 = (r_k, r_0)$.
Therefore

$$\| r_k \|^2 = \| r_k \|\ \| r_0 \| \cos(r_k, r_0)$$

and consequently

$$\| r_k \| = \| r_0 \| \ \sin \theta_k .$$

Finally

$$\| r_k \|^2 = \| r_0 \|^2 \ (1 - \ \cos^2 \theta_k). \quad \Box$$

Remarks. 1. One can observe from this result that x_k exists unconditionnally and thus the RRE cannot break down.

2. From the preceding result, we have $\| r_k \| = \sin(\theta_k) \ \| r_0 \|$ thus $\| r_k \| \leq \| r_0 \|$. Now, as $W_{k-1} \subset W_k$, we have $\theta_k \leq \theta_{k-1}$, and then $\| r_k \| \leq \| r_{k-1} \|$. We will see that this is not always satisfied for the MPE.

3. As the RRE is mathematically equivalent to the GMRES, all the results stated for the RRE are also valid for the GMRES.

As we observed in the previous remark, the RRE (or the GMRES) cannot break down but could present a stagnation. On this problem, we have the following result

Theorem 4. *Let r_0 be a given vector. Then $\forall k < N$, we have*

$$r_k = r_0 \ \textit{ if and only if } \ (r_0, C^j r_0) = 0, \ \textit{for } j = 1, \ldots, k. \tag{3.6}$$

Proof. From Theorem 2, we see that $r_k = r_0$ if and only if $P_k r_0 = 0$. But $P_k r_0 = 0$ is equivalent to the fact that $r_0 \perp W_k$ and this is equivalent to $(r_0, C^j r_0) = 0$ for $j = 1, \ldots, k$. $\Box$

To illustrate this, consider the linear system $Cx = b$, where C is the circulant matrix

$$C = \begin{pmatrix} 0 & 0 & \ldots & 0 & 1 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{pmatrix} \ \text{ and } \ b = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This example has been mentioned before in [4] and [25]. Setting $r_0 = b$, we can verify that $(r_0, Cr_0) = (r_0, C^2 r_0) = \ldots = (r_0, C^{N-1} r_0) = 0$. It follows from Theorem 4 that $r_0 = r_1 = \ldots = r_k$. Hence, there is no convergence at all for the first $N - 1$ steps and as it is a finite method, we have $r_N = 0$. We notice that, in this case, the restarted version of the GMRES denoted GMRES(m) never converge for any value of $m \leq N - 1$.

Such a stagnation can never occur for (SPD) (symmetric positive definite) matrices. In fact, if the matrix C is SPD, it can be decomposed as $C = L^{\mathrm{T}} L$ where L is nonsingular. Then for any vector r_0, we have $(r_0, C^j r_0) = (r_0, (L^{\mathrm{T}} L)^j r_0) = ||L^{\mathrm{T}j} r_0||^2$ for $j = 1, \ldots, k$. Hence, $(r_0, C^j r_0) \neq 0$ which implies that $r_0 \neq r_k$.

For linear systems with skew symmetric matrices, we have the following result

Theorem 5. *If the matrix C is skew symmetric and nonsingular, then*

$$\forall r_0, \ \ \textit{we have} \ \ ||r_{2j}|| = ||r_{2j+1}||, \ j = 0, 1, \ldots .$$

Proof. First, note that if C is skew symmetric, then $\forall r_0,\ \forall k,\ (r_0, C^{2k+1}r_0) = 0$. Using the expression (3.4) of r_k, we deduce

$$||r_k||^2 = ||r_{k-1}||^2 - (r_{k-1}, P_k r_0). \tag{3.7}$$

From (3.7), we see that if $(r_{k-1}, P_k r_0) = 0$ then $||r_k|| = ||r_{k-1}||$. Let us show now that for $j = 0, \dots,$ we have $(r_{2j}, P_{2j+1}r_0) = 0$.
Remark that as $r_{2j} = r_0 - P_{2j}r_0$ and as P_{2j} is an orthogonal projector onto $W_{2j} = K_{2j}(Cr_0, C)$, it follows that $r_{2j} \perp W_{2j}$. Hence, from (3.4), r_{2j} can be written as

$$r_{2j} = r_0 + \sum_{i=1}^{2j} a_i C^i r_0.$$

Now, since $r_{2j} \perp W_{2j}$, it is orthogonal to $Cr_0, C^2r_0, \dots, C^{2j-1}r_0$ and then for $p = 0, \dots, j-1$,

$$(r_{2j}, C^{2p+1}r_0) = (r_0, C^{2p+1}r_0) + \sum_{i=1}^{2j} a_i(C^i r_0, C^{2p+1}r_0) = 0. \tag{3.8}$$

On the other hand, since C is skew symmetric, we have

$$(C^{2l}r_0, C^{2p+1}r_0) = 0 \ \text{ for } l = 0, \dots . \tag{3.9}$$

It follows from (3.8) that

$$\sum_{l=0}^{j-1} a_{2l+1}(C^{2l+1}r_0, C^{2p+1}r_0) = 0 \ \text{ for } p = 0, \dots, j-1.$$

The determinant of this linear system is the Gram determinant of $\{Cr_0, C^3r_0, \dots, C^{2j-1}r_0\}$ and since those vectors are linearly independent, this determinant is different from zero. Hence $a_1 = a_3 = \dots = a_{2j-1} = 0$ and then

$$r_{2j} = r_0 + \sum_{l=1}^{j} a_{2l}C^{2l}r_0.$$

Multiplying by $C^{2j+1}r_0$ and using (3.9) we obtain

$$(r_{2j}, C^{2j+1}r_0) = 0.$$

Finally, since $W_{2j+1} = W_{2j} \cup \{C^{2j+1}r_0\}$ and as $r_{2j} \perp W_{2j}$, it follows that $r_{2j} \perp W_{2j+1}$ and then $(r_{2j}, P_{2j+1}r_0) = 0$. Hence, invoking (3.7), we get $||r_{2j}|| = ||r_{2j+1}||$. □

Before closing this section, let us give the following result to be used in the last section.

Corollary 1. *Let f_k be any nonzero vector of W_k and let* $\gamma_k = \left(\dfrac{r_0}{\| r_0 \|}, \dfrac{f_k}{\| f_k \|}\right)$, *then*

$$\| r_k \|^2 \leq \| r_0 \|^2 \ (1 - \ \gamma_k^2). \tag{3.10}$$

Proof. The acute angle θ_k between r_0 and W_k is defined by

$$\cos\theta_k = \max_{t_k \in W_k - \{0\}} \left| \left(\frac{r_0}{\| r_0 \|}, \frac{t_k}{\| t_k \|} \right) \right| .$$

Now, since f_k is a vector of W_k, we have $\gamma_k^2 \leq \cos^2\theta_k$,
thus, using Theorem 3, (3.10) follows. $\square$

4. Analysis of the MPE

In order to obtain the MPE, we set in (2.13) $Y_{k,0} = \Delta S_0$. Then the iterate x_k is given by

$$x_k = s_0 - \Delta S_0 (\Delta S_0^{\mathrm{T}} \Delta^2 S_0)^{-1} \Delta S_0^{\mathrm{T}} \Delta s_0 \tag{4.1}$$

and the corresponding residual by

$$\rho_k = r_0 - \Delta^2 S_0 (\Delta S_0^{\mathrm{T}} \Delta^2 S_0)^{-1} \Delta S_0^{\mathrm{T}} r_0. \tag{4.2}$$

Of course we assume here that the matrix $(\Delta S_0^{\mathrm{T}} \Delta^2 S_0)$ is nonsingular. A sufficient condition ensuring this will be given later.

We shall treat now the MPE as an oblique projection method. Let us define the projector Q_k onto W_k and orthogonal to V_k by

$$Q_k x \in W_k \quad \text{and} \quad x - Q_k x \perp V_k.$$

The vector $Q_k x$ is uniquely defined if no vector of V_k is orthogonal to W_k which is equivalent to the assumption $\det(\Delta S_0^{\mathrm{T}}\, C\, \Delta S_0) \neq 0$. We notice that the interpretation of the four extrapolation methods, namely the RRE, the MPE, the TEA and the MMPE, as oblique projection methods has been given in [19], [10] and [1]. Other oblique projection methods have been studied by Saad in [16].

Theorem 6. *Let Q_k be the oblique projector as defined before. Then*

$$\rho_k = r_0 - Q_k r_0. \tag{4.3}$$

Proof. As W_k is spanned by the columns of the matrix $\Delta^2 S_0$, we have for some $y \in \mathbb{R}^N$, $Q_k x = \Delta^2 S_0\, y$.
Now since $x - Q_k x \perp V_k$, it follows that $\Delta S_0^{\mathrm{T}} (x - \Delta^2 S_0\, y) = 0$, hence

$$y = (\Delta S_0^{\mathrm{T}} \Delta^2 S_0)^{-1} \Delta S_0^{\mathrm{T}} x.$$

Consequently

$$Q_k x = \Delta^2 S_0\, (\Delta S_0^{\mathrm{T}} \Delta^2 S_0)^{-1} \Delta S_0^{\mathrm{T}} x. \tag{4.4}$$

Using (4.2), it follows that

$$\rho_k = r_0 - Q_k r_0. \quad \square$$

An important consequence of this theorem is the following result.

Theorem 7. *Let φ_k be the acute angle between r_0 and $Q_k r_0$, then*

$$\| \rho_k \|^2 = \frac{1 - \cos^2 \varphi_k}{\cos^2 \varphi_k} \| r_0 \|^2 . \tag{4.5}$$

Proof. We have shown that $\rho_k = r_0 - Q_k r_0$, then

$$\begin{aligned} (\rho_k, \rho_k) &= (\rho_k, r_0 - Q_k r_0) \\ &= (\rho_k, r_0) - (\rho_k, Q_k r_0). \end{aligned}$$

Now, since $\rho_k \in V_k^{\perp}$ and $r_0 \in V_k$, it follows that $(\rho_k, r_0) = 0$, thus

$$(\rho_k, \rho_k) = -(\rho_k, Q_k r_0).$$

Then

$$\| \rho_k \|^2 = \| \rho_k \| \| Q_k r_0 \| \sin \varphi_k . \tag{4.6}$$

On the other hand

$$\| r_0 \| = \| Q_k r_0 \| \cos \varphi_k . \tag{4.7}$$

Replacing in (4.6), we obtain

$$\| \rho_k \|^2 = \| \rho_k \| \| r_0 \| \frac{\sin \varphi_k}{\cos \varphi_k}$$

and finally

$$\| \rho_k \|^2 = \frac{1 - \cos^2 \varphi_k}{\cos^2 \varphi_k} \| r_0 \|^2 . \quad \square$$

This last result leads to some remarks:

1. It shows that x_k exists for the MPE if and only if $\cos \varphi_k \neq 0$.
2. It is clear that $\| \rho_k \| \leq \| r_0 \|$ if and only if $\varphi_k \in [0, \pi/4]$. If it is not the case, the residual norm for the MPE cannot decrease as it does for the RRE. This is the second disadvantage of the MPE.

Theorem 8. *If the matrix C has a positive definite symmetric part, then*

i) $$\cos \varphi_k \geq \frac{\lambda_{\min}(C_s)}{\| C \|} > 0$$

ii) *x_k exists and is unique for the MPE for $k < d$, where d is the degree of the minimal polynomial of C for r_0.*

Proof. i) From (4.7), we have

$$\cos \varphi_k = \frac{\| r_0 \|}{\| Q_k r_0 \|} \geq \frac{1}{\| Q_k \|}. \tag{4.8}$$

On the other hand, the matrix representation of the projector Q_k is defined by

$$Q_k = C \Delta S_0 (\Delta S_0^{\mathrm{T}} C \Delta S_0)^{-1} \Delta S_0^{\mathrm{T}}.$$

Note here that Q_k denotes at the same time a matrix and its associated operator.
Let $\Delta S_0 = Q_k^1 R_k^1$ where Q_k^1 is an orthogonal $n \times k$ matrix and R_k^1 is a nonsingular $k \times k$ matrix. Then

$$Q_k = CQ_k^1({Q_k^1}^{\mathrm{T}} CQ_k^1)^{-1} {Q_k^1}^{\mathrm{T}}.$$

Therefore, since Q_k^1 is orthogonal, we get

$$\| Q_k \| \leq \| ({Q_k^1}^{\mathrm{T}} CQ_k^1)^{-1} \| \| C \| .$$

Now, as C_{s} is positive definite, we have $\| ({Q_k^1}^{\mathrm{T}} CQ_k^1)^{-1} \| \leq \dfrac{1}{\lambda_{\min}(C_{\mathrm{s}})}$ (see [13]).
It follows from (4.8) that

$$\cos \varphi_k \; \geq \; \frac{\lambda_{\min}(C_{\mathrm{s}})}{\| C \|} \; > 0.$$

ii) One can see from Theorem 7 that for MPE, x_k exists if and only if $\cos \varphi_k \neq 0$ and this was proved in i). □

This theorem have been already given in [19]. We gave here another proof.

Theorem 9. *Assume that $r_0 \in W_k^{\perp}$. Then x_k^{MPE} does not exist for the MPE, and $x_k^{\mathrm{RRE}} = s_0$ for the RRE,*
where x_k^{MPE} and x_k^{RRE} denote respectively the iterates at the step k for the MPE and the RRE.

Proof. Since $r_0 \in W_k^{\perp}$, we have $\cos \varphi_k = 0$. Then Theorem 7 shows that x_k^{MPE} cannot exist for the MPE. On the other hand, as r_0 is orthogonal to W_k, $\theta_k = \pi/2$, then, from Theorem 2, we obtain $r_k^{\mathrm{RRE}} = r_0$. Therefore $x_k^{\mathrm{RRE}} = s_0$. □

This theorem gives an indication on the correlation of the performance of the two methods. It shows clearly that a breakdown in the MPE corresponds to a stagnation of the RRE ($x_k^{\mathrm{RRE}} = s_0$).

We want now to compare theoretically the residual norms of the MPE and the RRE and this is the subject of the following result

Theorem 10. *Let r_k and ρ_k denote respectively the residuals of x_k for the RRE and the MPE. If $r_0 = \rho_0$, then*

$$i) \quad \| r_k \|^2 \; = \; \frac{1 - \cos^2 \theta_k}{1 - \cos^2 \varphi_k} \cos^2 \varphi_k \; \| \rho_k \|^2$$

$$ii) \quad \| r_k \| \; \leq \; \cos \varphi_k \; \| \rho_k \| .$$

Proof. i) By using the results of Theorem 3 and Theorem 7.
ii) From the definitions of the angles θ_k and φ_k, we have

$$\cos \varphi_k \; \leq \; \cos \theta_k, \text{then} \quad 1 - \cos^2 \theta_k \; \leq \; 1 - \cos^2 \varphi_k. \tag{4.9}$$

Using (4.9) in i) we obtain

$$\| r_k \|^2 \leq \| \rho_k \|^2 \; \cos^2 \varphi_k. \quad \square$$

Let us mention that another result on the comparison between the method of Arnoldi and the GMRES was derived in [4]. This result is based on the implementation of the GMRES by using Givens rotations. Our result is based on the geomerical aspect of the two methods.

5. Incomplete forms of the MPE and the RRE

Let p_k and q_k be two integers and let us assume that $p_k < d$ where d is the degree of the minimal polynomial of B with respect to Δs_0.

We denote $\wp\,(p_k, q_k)$ the algorithm defined by

$$(5.1) \qquad \wp\,(p_k, q_k) \qquad \begin{cases} x_0 \text{ is a given vector} \\ \text{for } k = 1, 2, \ldots \\ \qquad s_0 = x_{k-1} \\ \qquad s_{j+1} = Bs_j + b, \qquad j = 0, 1, \ldots, \varphi(p_k, q_k) \\ x_k = t_{p_k}^{(q_k)}. \end{cases}$$

Where
$\varphi(p_k, q_k) = p_k + q_k$ for the incomplete MPE, MMPE and RRE,
$\varphi(p_k, q_k) = 2p_k + q_k - 1$ for the incomplete TEA.

The algorithm $\wp\,(p_k, q_k)$ was first introduced by Beuneu [1] for the RRE and recently, numerical results have been given by Gander, Golub and Gruntz [9] for the TEA.
In the sequel, we will consider only the RRE method. In this case, the algorithm (5.1) will be denoted IRRE(p_k,q_k).

Remarks. 1. As $p_k < d$, where d is the degree of the minimal polynomial for Δs_0, the convergence of the sequence (x_k) is not always achieved in a finite number of iterations which was the case for the complete methods.
2. p_k and q_k could be fixed integers or depending on k and their values could be changed during the computation.

Let us consider now the case where $q_k = 0$ and $p_k = m$ with m a fixed integer. In this case the obtained algorithm IRRE(m,0) is mathematically equivalent to GMRES(m). Here, the iterate x_k represents the vector obtained at the end of the k-th cycle in the GMRES(m). We denote by $y_j,\ j = 1, \ldots, m$ the j-th iterate in the cycle and defined by $y_0 = s_0$ and $y_j = t_j^{(0)}$.
We have the following result.

Theorem 11. *If C is a nonsingular skew symmetric matrix, then $\forall m \geq 2$, GMRES(m) converges for any initial vector x_0.*

Proof. We have $y_0 = x_{k-1}$, then $r(y_0) = r_{k-1}$. From Theorem 3 we get

$$(5.2) \qquad ||r(y_2)||^2 = (1 - \cos^2(\theta_2^{(k)}))\,||r_{k-1}||^2,$$

where $\theta_2^{(k)}$ is the acute angle between y_0 and the subspace spanned by $Cr_{k-1}, C^2 r_{k-1}$. Hence $r(y_2)$ can be written as $r(y_2) = r_{k-1} + \alpha C r_{k-1} + \beta C^2 r_{k-1}$. Now since $r(y_2)$ is orthogonal to Cr_{k-1} and as C is skew symmetric we have $\alpha = 0$, consequently

$$\cos(\theta_2^{(k)}) = \frac{|(r_{k-1}, C^2 r_{k-1})|}{||r_{k-1}||\,||C^2 r_{k-1}||} = \frac{||C r_{k-1}||^2}{||r_{k-1}||\,||C^2 r_{k-1}||}. \tag{5.3}$$

By using the fact that $\dfrac{||C r_{k-1}||}{||r_{k-1}||} \geq \sigma_{\min}(C)$ and $\dfrac{||C^2 r_{k-1}||}{||C r_{k-1}||} \leq \sigma_{\max}(C)$, where $\sigma_{\max}(C)$ and $\sigma_{\min}(C)$ are the largest and the smallest singular values of the matrix C, it follows that

$$\cos(\theta_2^{(k)}) \geq \frac{\sigma_{\min}(C)}{\sigma_{\max}(C)}. \tag{5.4}$$

Now, since $||r_k|| = ||r(y_k)|| \leq ||r(y_2)||$, we obtain from (5.2), (5.3) and (5.4) the relation

$$||r_k||^2 \leq \left(1 - \frac{\sigma^2_{\min}(C)}{\sigma^2_{\max}(C)}\right) ||r_{k-1}||^2.$$

It follows that

$$||r_k||^2 \leq \left(1 - \frac{\sigma^2_{\min}(C)}{\sigma^2_{\max}(C)}\right)^k ||r_0||^2.$$

Consequently, $\lim_{k\to\infty} ||r_k|| = 0$, which completes the proof. □

Let us come back to the general case and let $\tilde{V}_{p_k}$ and $\tilde{W}_{p_k}$ be the Krylov subspaces of dimension p_k of $\mathbb{R}^N$ defined by $V_{p_k} = \text{span}\{\Delta s_{q_k}, C\,\Delta s_{q_k}, \ldots, C^{p_k-1}\,\Delta s_{q_k}\}$ and $\tilde{W}_{p_k} = C\,\tilde{V}_{p_k}$. In the sequel, r_k^{IR} denotes the residual of x_k produced by the IRRE(p_k,q_k).

Theorem 12. *Let P_k be the orthogonal projector onto $\tilde{W}_{p_k}$ and let g_k be any nonzero vector of $\mathbb{R}^N$, then*

$$\begin{aligned} r_k^{\text{IR}} &= \Delta s_{q_k} - P_k \Delta s_{q_k} \\ \|\, r_k^{\text{IR}} \,\|^2 &= (1 - \cos^2\alpha_k)\ \|\, \Delta s_{q_k} \,\|^2 \\ \|\, r_k^{\text{IR}} \,\|^2 &\leq (1 - \delta_k^2)\ \|\, \Delta s_{q_k} \,\|^2 \end{aligned}$$

where $\delta_k = \dfrac{(\Delta s_{q_k}, g_k)}{\|\, \Delta s_{q_k} \,\|\,\|\, g_k \,\|}$, and α_k is the acute angle between Δs_{q_k} and $\tilde{W}_{p_k}$.

Proof. Similar to the proofs given in Theorem 2, Theorem 3 and Corollary 1. □

Next, we shall give some results on the convergence of the incomplete RRE method. First, we need the following lemma

Lemma 1. *[7] For any nonzero vector x of $\mathbb{R}^N$, we have*

$$\frac{(x, Cx)}{(Cx, Cx)} \geq \frac{\lambda_{\min}(C_s)}{\lambda_{\min}(C_s)\lambda_{\max}(C_s) + \rho^2(C_t)}.$$

We can state now the convergence result

Theorem 13. *If the symmetric part of C is positive definite, then*

$$\| r_k^{\mathrm{IR}} \| \leq \sqrt{1 - \frac{\lambda_{\min}(C_{\mathrm{s}})^2}{\lambda_{\max}(C^{\mathrm{T}}C)}} \; \| \Delta s_{q_k} \| \tag{5.5}$$

and

$$\| r_k^{\mathrm{IR}} \| \leq \sqrt{1 - \frac{\lambda_{\min}(C_{\mathrm{s}})^2}{\lambda_{\min}(C_{\mathrm{s}})\,\lambda_{\max}(C_{\mathrm{s}}) + \rho^2(C_{\mathrm{t}})}} \; \| \Delta s_{q_k} \| \;. \tag{5.6}$$

Proof. We have shown that

$$\| r_k^{\mathrm{IR}} \|^2 \leq \| \Delta s_{q_k} \|^2 \; (1 - \delta_k^2) \tag{5.7}$$

where δ_k is defined in Theorem 12. If we set $g_k = C\Delta s_{q_k}$, then δ_k^2 is defined by

$$\delta_k^2 = \frac{(\Delta s_{q_k}, C\Delta s_{q_k})^2}{\| \Delta s_{q_k} \|^2 \, \| C\Delta s_{q_k} \|^2}.$$

Therefore

$$\delta_k^2 = \frac{(\Delta s_{q_k}, C\Delta s_{q_k})}{(\Delta s_{q_k}, \Delta s_{q_k})} \, \frac{(\Delta s_{q_k}, C\Delta s_{q_k})}{(C\Delta s_{q_k}, C\Delta s_{q_k})}. \tag{5.8}$$

On the other hand

$$(\Delta s_{q_k}, C\Delta s_{q_k}) \;\geq \lambda_{\min}(C_{\mathrm{s}})(\Delta s_{q_k}, \Delta s_{q_k})$$

and

$$(C\Delta s_{q_k}, C\Delta s_{q_k}) \;\leq \lambda_{\max}(C^{\mathrm{T}}C)(\Delta s_{q_k}, \Delta s_{q_k}).$$

Invoking these two relations in (5.8), we obtain

$$\delta_k^2 \;\geq\; \frac{\lambda_{\min}(C_{\mathrm{s}})^2}{\lambda_{\max}(C^{\mathrm{T}}C)}.$$

Hence, using (5.7), inequality (5.5) follows.
Using Lemma 1 with $x = \Delta s_{q_k}$, one obtains

$$\frac{(\Delta s_{q_k}, C\,\Delta s_{q_k})}{(C\,\Delta s_{q_k}, C\,\Delta s_{q_k})} \;\geq\; \frac{\lambda_{\min}(C_{\mathrm{s}})}{\lambda_{\min}(C_{\mathrm{s}})\,\lambda_{\max}(C_{\mathrm{s}}) + \rho^2(C_{\mathrm{t}})}. \tag{5.9}$$

Now, if we use (5.9) in (5.8), we get the relation

$$\delta_k^2 \;\geq\; \frac{\lambda_{\min}(C_{\mathrm{s}})^2}{\lambda_{\min}(C_{\mathrm{s}})\,\lambda_{\max}(C_{\mathrm{s}}) + \rho^2(C_{\mathrm{t}})}.$$

Then, replacing in (5.7), the inequality (5.6) holds. □

Remarks. 1. From (2.8), we have $\Delta s_{q_k} = B^{q_k}\,\Delta s_0$, thus if $\| B \| < 1$, the convergence of $\| r_k^{\mathrm{IR}} \|$ to 0 will be faster for largest values of q_k.

2. Let us notice here that when $q_k = 0$, inequalities (5.5) and (5.6) are identical to those obtained for the GCR in [7] but in a different way. Other results on the error analysis for the RRE and MPE were also derived in [19].

3. Numerical experiments of some incomplete vector extrapolation methods can be found in [9] and [20].

6. Numerical experiments

We will give now some numerical examples to illustrate our theoretical results. The following experiments were run using MATLAB on Macintosh SE/30. The iterations were stopped as soon as $||r_k||/||r_0|| \leq 10^{-12}$.
For all the tests, the right-hand side b was set to $C\,\tilde{x}$, where $\tilde{x} = (1, \ldots, 1)^{\mathrm{T}}$.

Example 1

To illustrate the result of Theorem 5, we consider the next example given earlier in [4]. The skew symmetric matrix C is of order 40 and given by

$$C = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}.$$

In the following Table, we give the norm of the residuals obtained with GMRES, versus the number of iterations. The initial vector was chosen randomly.

Table 1.

k	$\|\|r_k\|\|$	k	$\|\|r_k\|\|$	k	$\|\|r_k\|\|$	k	$\|\|r_k\|\|$
1	$2.85e+00$	11	$3.39e-01$	21	$2.34e-01$	31	$1.90e-01$
2	$1.05e+00$	12	$3.15e-01$	22	$2.26e-01$	32	$1.84e-01$
3	$1.05e+00$	13	$3.15e-01$	23	$2.26e-01$	33	$1.84e-01$
4	$6.81e-01$	14	$2.83e-01$	24	$2.19e-01$	34	$1.79e-01$
5	$6.81e-01$	15	$2.83e-01$	25	$2.19e-01$	35	$1.79e-01$
6	$4.57e-01$	16	$2.62e-01$	26	$2.04e-01$	36	$1.72e-01$
7	$4.57e-01$	17	$2.62e-01$	27	$2.04e-01$	37	$1.72e-01$
8	$3.82e-01$	18	$2.50e-01$	28	$1.95e-01$	38	$1.67e-01$
9	$3.82e-01$	19	$2.50e-01$	29	$1.95e-01$	39	$1.67e-01$
10	$3.39e-01$	20	$2.34e-01$	30	$1.90e-01$	40	$1.64e-15$

Although GMRES(m) for $m \geq 2$ is always convergent for skew symmetric matrices as we proved (Theorem 11), this algorithm can have a poor performance.

Example 2

The next example is a model problem used by many authors before. The matrix A is given by the block tridiagonal matrix

$$A = \begin{pmatrix} E & -I & & & \\ -I & E & -I & & \\ & \ddots & \ddots & \ddots & \\ & & -I & E & -I \\ & & & -I & E \end{pmatrix}$$

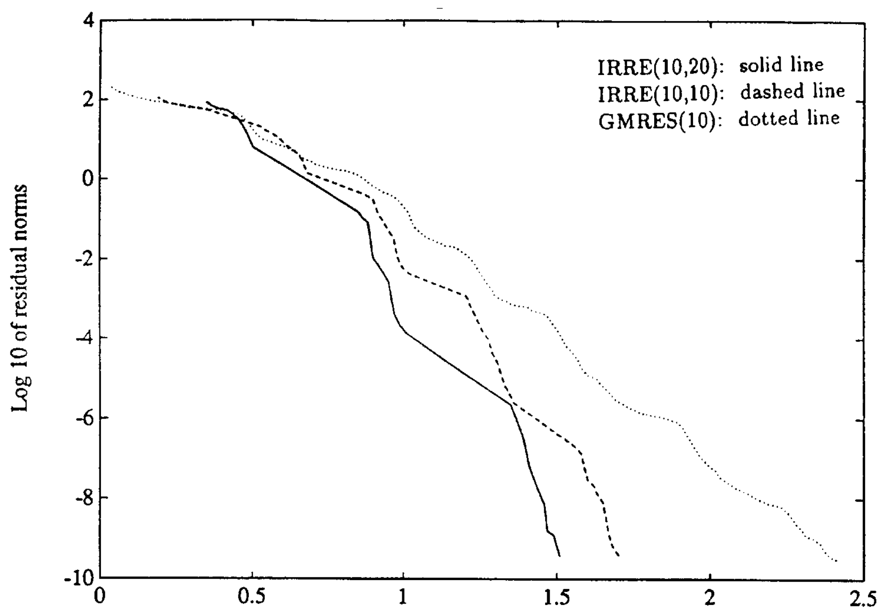

Fig. 1. N = 200, convergence curves for experiment 1

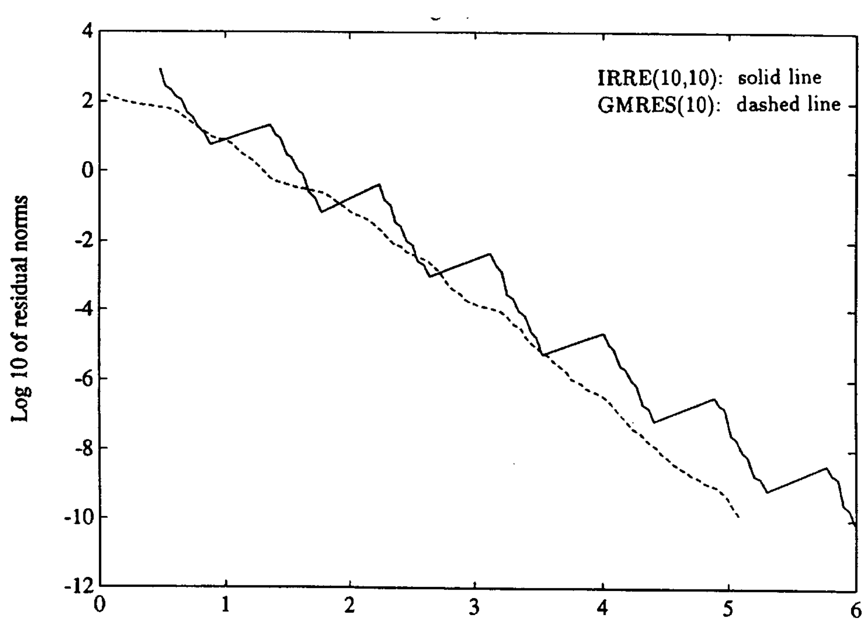

Fig. 2. N = 100, convergence curves for experiment 1

with

$$E = \begin{pmatrix} 4 & \alpha & & & \\ \beta & 4 & \alpha & & \\ & \ddots & \ddots & \ddots & \\ & & \beta & 4 & \alpha \\ & & & \beta & 4 \end{pmatrix}$$

and $\alpha = -1 + \delta$, $\beta = -1 - \delta$.
The matrix A represents the 5-point discretization of the operator $-\partial^2/\partial x^2 - \partial^2/\partial y^2 + \gamma\partial/\partial x$ on a rectangular region. The initial guess $x_0 = 0$.
We have chosen two values of δ:
$\delta = 1.5$, $n = \dim(E) = 20$ and $N = 200$ (Fig.1).
$\delta = 2.5$, $n = \dim(E) = 10$ and $N = 100$ (Fig.2).
Figure 1 and Fig. 2 show the behaviour of the residual norms, in a logarithmic scale, versus the number of multiplications for the IRRE(10,20) (Algorithm 5.1 with $p_k = 10$ and $q_k = 20$)(solid line), the IRRE(10,10)(dashed line) and GMRES(10)(dotted line). The two experiments were carried out with the three methods applied to the preconditioned system $Cx = D^{-1}Ax = D^{-1}f = b$ where D is the diagonal matrix of the matrix A, which corresponds to the Jacobi preconditioner.
If the basic iteration ($s_{j+1} = Bs_j + b$ with $B = I - D^{-1}A$ and $b = D^{-1}f$) converges as it is the case in Fig. 1, the Algorithm (5.1) performs better than the restarted GMRES. For IRRE(10,10), we first performed $q_k = 10$ Jacobi steps and then start the extrapolation of width $p_k = 10$, we again compute basic iteration steps starting with the new approximation t_{10}^{10} and so on. From Fig. 1, remark that the more Jacobi steps are performed ($q_k = 20$), the more effective the extrapolation works.
In Fig. 2, we give an example where the basic iteration is not convergent, hence oscillations appear in the residual norms for the incomplete RRE. For this experiment, we compared the behaviour of the residual norms, in a logarithmic scale, for IRRE(10,10)(solid line) and GMRES(10)(dashed line). We observed that altought the basic iterations are not convergent, the extrapolation works. However, a more effectif basic iteration such as Gauss-Seidel or SOR will give much more powerful results.

Acknowledgements. We would like to thank the referees for their valuable remarks and helpful suggestions. We also thank the Professor A. Sidi for pointing out an error in an earlier version of Theorem 11 and for several valuable remarks and corrections.

References

1. Beuneu, J. (1982): Méthodes de projection à convergence finie remarques sur leur forme incomplète, publication ANO 80, université de Lille 1
2. Bjorck, A. (1989): Least squares methods, in Handbook of Numerical Analysis, P. G. Ciarlet and J. L. Lions, eds., Elsevier/ North Holland, Amsterdam, The Netherlands
3. Brezinski, C. (1975): Généralisation de la transformation de Shanks, de la table de Padé et de l'epsilon-algorithm, Calcolo 12, 317–360
4. Brown, P.N. (1991): A theoretical comparison of the Arnoldi and GMRES algorithms, SIAM J. Sci. Stat. Comput 12, 58–76
5. Cabay, S., Jackson, L.W. (1976): A polynomial extrapolation method for finding limits and antilimits for vector sequences, SIAM J. Numer. Anal. 13, 734–752
6. Eddy, R.P. (1979): Extrapolation to the limit of a vector sequence, In Wang, P.C.C., ed., Information Linkage Between Applied Mathematics and Industry, pp. 387–396, New-York, Academic Press

7. Eisenstat, S.C., Elman, H.C., Schultz, M.H. (1983): Variational iterative methods for nonsymmetric systems of linear equations, SIAM J. Numer. Anal. 20, 345–357
8. Ford, W.D., Sidi, A. (1988): Recursive algorithms for vector extrapolation methods, Appl. Numer. Math. 4, 477–489
9. Gander, W., Golub, G.H., Gruntz, D. (1990): Solving linear equations by extrapolation, in: Kovalic, J.S. (ed.) Supercomputing. Nato ASI Series, Springer
10. Jbilou, K. (1988): Méthodes d' Extrapolation et de Projection. Applications aux suites de vecteurs, Thèse de 3ème cycle, Université de Lille 1
11. Jbilou, K. (1993): A general projection algorithm for solving linear systems of equations, Numer. Algorithms 4, 361–377
12. Jbilou, K., Sadok, H. (1991): Some results about vector extrapolation methods and related fixed point iterations, J. Comput. Appl. Math. 36, 385–398
13. Joubert, W. (1990): Iterative methods for the solution of nonsymmetric systems of linear equations, Report CNA-242, Center for Numerical Analysis, The University of Texas at Austin, February
14. Meŝina, M. (1977): Convergence acceleration for the iterative solution of x=Ax+f , Comput. Methods Appl. Mech. Eng. 10, 165–173
15. Pugatchev, B.P. (1978): Acceleration of the convergence of iterative processes and a method for solving systems of nonlinear equations, U.S.S.R. Comput. Math. Phys. 17, 199–207
16. Saad, Y. (1982): The Lanczos biorthogonalisation algorithm and other oblique projection methods for solving large unsymmetric systems, SIAM J. Numer. Anal. 19, 485–506
17. Saad, Y., Schultz, M.H. (1986): GMRES: A generalized minimal residual algorithm for solving non-symmetric linear systems, SIAM J. Sci. Statist. Comput. 7, 856–869
18. Sidi, A. (1986): Convergence and stability property of minimal polynomial and reduced rank extrapolation algorithms, SIAM J. Numer. Anal. 23, 197–209
19. Sidi, A. (1988): Extrapolation vs. projection methods for linear systems of equations, J. Comput. Appl. Math. 22, 71–88
20. Sidi, A. (1991): Efficient implementation of minimal polynomial and reduced rank extrapolation methods J. Comput. Appl. Math. 36, 305–337
21. Sidi, A., Bridger, J. (1988): Convergence and stability analyses of some vector extrapolation methods in the presence of defective iteration matrices J. Comput. Appl. Math., 22, 35–61
22. Sidi, A., Ford, W.F., Smith, D.A. (1986): Acceleration of convergence of vector sequences, SIAM J. Numer. Anal. 23, 178–196
23. Smith, D.A., Ford, W.F., Sidi, A. (1987): Extrapolation methods for vector sequences. SIAM Rev. 29, 199–233
24. Smith, D.A., Ford, W.F., Sidi, A. (1987): Correction to "Extrapolation methods for vector sequences". SIAM Rev. 30, 623–624
25. Van Der Vorst, H. (1989): The convergence behavior of some iterative solution methods, in: Gruber, R., Periaux, J., Shaw, R.P. (eds.) Proc. Fifth Internat. Symposium Numer. Meth. Engrg., Vol. 1, Springer-Verlag

This article was processed by the author using the LaTeX style file *pljour1* from Springer-Verlag.