ELSEVIER

Applied Numerical Mathematics 24 (1997) 365–378

APPLIED NUMERICAL MATHEMATICS

A Krylov projection method for systems of ODEs

E. Celledoni *, I. Moret

Dipartimento di Scienze Matematiche, Università di Trieste, Piazzale Europa 1, 34127 Trieste, Italy

Abstract

In this paper we consider approximations of solutions of IVPs obtained through projections into Krylov subspaces. Numerical experiments on parabolic equations illustrate the performance of the method. © 1997 Elsevier Science B.V.

Keywords: Ordinary differential equations; Initial value problems; Krylov subspace methods

1. Introduction

For solving initial value problems, numerical algorithms which make use of Krylov subspace techniques have been recently proposed by various authors in the literature. An approach, considered in [1,5] consists in reformulating the differential system as a Volterra integral equation, to which a Krylov approximation method is then directly applied. So, the arising method turns out to be an acceleration of the waveform relaxation. Other procedures, like those considered for instance in [2,3,9], hinge on approximations of the matrix exponential operator, obtained by projecting the exponential of a large matrix into small Krylov subspace. As shown both by theoretical analysis and by numerical experiences, these approximations work well. The approach presented in this paper can be viewed as a generalization of that idea since it allows to solve a large IVP through the solutions of small ones.

2. The Krylov method for IVPs

Given a scalar function $f(t)$ and a real vector $v = [v_1, v_2, \dots, v_m]^{\mathrm{T}} \in \mathbb{R}^m$ we denote by $f(t) \cdot v$ the vector function $[f(t)v_1, f(t)v_2, \dots, f(t)v_m]^{\mathrm{T}}$, moreover, given a vector function $z(t) = [z_1(t), z_2(t), \dots, z_m(t)]^{\mathrm{T}}$, we set

$$z^{\mathrm{T}}v = \left(z^{\mathrm{T}}v\right)(t) = \sum_{j=1}^{m} v_j z_j(t), \quad t \in [0, T].$$

* Corresponding author. E-mail: elena@dsmaosf.univ.trieste.it.

PII S0168-9274(97)00033-0

In particular, if $z(t) = f(t) \cdot u$, then $z^{\mathrm{T}}v = f(t)(u^{\mathrm{T}}v)$, where $u^{\mathrm{T}}v$ is the Euclidean scalar product. The Euclidean norm is denoted by $\|\cdot\|_2$.

Let us consider the linear IVP

$$\begin{cases} y' - Ay = f(t) \cdot v, & t \in [0,T], \\ y(0) = 0, \end{cases} \tag{1}$$

where A is an $m \times m$ real matrix, assumed to be time-independent, $f(t)$ is continuous on $[0,T]$, $v \in \mathbb{R}^m$, $y(t)$ is the unknown vector, for $t \in [0,T]$. As is well known, this problem has the solution

$$y(t) = \int_0^t \exp\big(A(t-s)\big) f(s) \cdot v \, \mathrm{d}s.$$

For any positive integer n, let

$$W_n(A,v) = \operatorname{span}\{v, Av, \dots, A^{n-1}v\} = \operatorname{span}\{v_1, v_2, \dots, v_n\}$$

be the nth Krylov subspace generated by A and v, where $v_1, v_2, \dots, v_n$ is the orthonormal basis given by the Arnoldi's algorithm [8]:

Algorithm 2.1.

Set $v_1 = v/\|v\|_2$.

For $k = 1, 2, \dots,$ do

$q_{k+1} := (I - P_k)Av_k,$

$v_{k+1} := q_{k+1}/\|q_{k+1}\|_2,$

end.

Here P_k denotes the orthogonal projection on $W_k(A,v) = \operatorname{span}\{v_1, \dots, v_k\}$. As is well known the $n \times n$ matrix $H_n = (h_{i,j})_{i,j=1,\dots,n}$, whose entries are

$$h_{i,j} = v_i^{\mathrm{T}} A v_j,$$

is an upper Hessenberg matrix. In particular $h_{j+1,j} = \|q_{j+1}\|_2$. Moreover, considering the $m \times n$ matrix $V_n = [v_1, v_2, \dots, v_n]$, for every n, the following relation holds:

$$AV_n = V_n H_n + h_{n+1,n} v_{n+1} e_n^{\mathrm{T}}, \tag{2}$$

where e_n is the nth vector of the canonical basis of $\mathbb{R}^n$.

Then, we consider the following approximation scheme.

Algorithm 2.2. Krylov Projection Method (KPM)

Solve the IVP

$$\begin{cases} z'(t) - H_n z(t) = f(t) \cdot \|v\|_2 \cdot e_1, & t \in [0,T], \\ z(0) = 0, \end{cases} \tag{3}$$

where e_1 is the first vector of the canonical basis of $\mathbb{R}^n$ and $z(t) = [z_1(t), z_2(t), \dots, z_n(t)]^{\mathrm{T}}$. Then approximate y by

$$y_n(t) = V_n z(t) = \sum_{j=1}^{n} z_j(t) \cdot v_j.$$

Remark 2.1. Let $u \in \mathbb{R}^m$. The computation of $\mathrm{e}^{At}u$ can be carried out by solving system (1) with $v = Au$, $f(t) = 1$, since its solution is $\mathrm{e}^{At}u - u$. In this case our method gives the approximation of $\mathrm{e}^{At}u$ already proposed and studied in [2,3,9], where for actual computation of y_n other types of algorithms are considered, instead of solving system (3).

Let us now consider the residual of system (1) at y_n, that is

$$r_n(t) = f(t) \cdot v - y_n'(t) + Ay_n(t), \quad t \in [0, T].$$

Since

$$r_n(t) = f(t) \cdot v - V_n z'(t) + AV_n z(t),$$

using (2) and (3) we get

$$r_n(t) = h_{n+1,n} e_n^{\mathrm{T}} z(t) \cdot v_{n+1} = h_{n+1,n} z_n(t) \cdot v_{n+1}. \tag{4}$$

Since $h_{n+1,n} = 0$ for some $n \leqslant m$, this shows the finite termination of the procedure.

Remark 2.2. It is important to observe that if one wants to restart the procedure by considering the IVP

$$\begin{cases} (y - y_n)' - A(y - y_n) = r_n, \\ (y - y_n)(0) = 0, \end{cases}$$

owing to (4), we have to deal with a problem of the same type of the original one. The method arising by restarting after k steps, for a fixed positive integer k, will be denoted by KPM(k).

Remark 2.3. Obviously the method can be applied even when the right-hand side of (1) has the form $\sum_{j=1}^{k} f_j(t) \cdot w_j$. Indeed in this case we have to solve, possibly in parallel, k independent problems of type (1).

3. Error estimates

In this section we give a priori estimates for the residuals r_n. Here we consider any index $n \geqslant 1$ and for simplicity, we set $H = H_n$. Of course we assume that $h_{j+1,j} \neq 0$, for $j = 1, 2, \ldots, n$. We set

- $\|f\| = \max_{t \in [0,T]} |f(t)|$,
- $\|r_n\| = \max_{t \in [0,T]} \|r_n(t)\|_2$,
- $\gamma_n = \max_{t \in [0,T]} \|\exp(Ht)\|_2$,
- $\gamma = \max_{t \in [0,T]} \|\exp(At)\|_2$,
- $C_k = \prod_{j=1}^{k} h_{j+1,j}$, for $k = 1, 2, \ldots, n$.

Moreover, e_n and e_1 are respectively the nth and the first vector of the canonical basis of $\mathbb{R}^n$. From now on we denote by I any identity operator. Before presenting an error analysis of the proposed method, we recall some facts of linear algebra which will be useful in the sequel.

The results contained in the following proposition follow by straightforward computation, taking into account the Hessenberg structure of H.

Proposition 3.1. *For* $k = 0, 1, 2, \ldots, n-2$, $e_n^{\mathrm{T}} H^k e_1 = 0$, *and*

$$e_n^{\mathrm{T}} H^{n-1} e_1 = C_{n-1}.$$

A bound for C_n can be provided using the following statement.

Proposition 3.2 [7, p. 125]. *Let* M *be an* $m \times m$ *matrix and let us denote by*

$$\sigma_1(M) \geqslant \sigma_2(M) \geqslant \cdots \geqslant \sigma_m(M)$$

the singular values of M *arranged in decreasing order. Let* $\{u_i\}_{i=1}^k$, $\{w_i\}_{i=1}^k$ *be two orthonormal families in* $\mathbb{R}^m$. *Let us consider the* $k \times k$ *matrix* B *having entries* $u_i^{\mathrm{T}} M w_j$ *for* $i, j = 1, \ldots, k$. *Then*

$$\left|\mathrm{Det}(B)\right| \leqslant \prod_{j=1}^{k} \sigma_j(M).$$

Proposition 3.3. *For every index* $k \geqslant 1$ *and for any scalar* λ, *we have*

$$C_k \leqslant \prod_{j=1}^{k} \sigma_j(\lambda I + A). \tag{5}$$

Proof. By definition $h_{j+1,j} = v_{j+1}^{\mathrm{T}} A v_j = v_{j+1}^{\mathrm{T}} (\lambda I + A) v_j$. Then considering the $k \times k$ upper triangular matrix B having entries $v_{j+1}^{\mathrm{T}} (\lambda I + A) v_i$, for $i, j = 1, \ldots, k$, by Proposition 3.2 we get the result. □

From [4, p. 341] we have:

Proposition 3.4. *There are entire functions* $\psi_0(x), \ldots, \psi_{n-1}(x)$ *such that*

$$\exp(Hx) = \sum_{j=1}^{n-1} \psi_j(x) H^j. \tag{6}$$

The following result concerning Hessenberg matrices is well known (see [6]).

Proposition 3.5. *Each eigenvalue of* H *has geometric multiplicity equal to* 1 *and the minimal polynomial of* H *is its characteristic polynomial.*

The following statement comes directly from the Cayley–Hamilton theorem.

Proposition 3.6. *Assume that* $(H - \lambda_0 I)$ *is nonsingular and let*

$$\mathrm{Det}\big(\lambda I - (H - \lambda_0 I)\big) = \lambda^n + \sum_{k=1}^{n} \alpha_k \lambda^{n-k}.$$

Then

$$(H-\lambda_0 I)^{-1} = -\frac{1}{\alpha_n}\big((H-\lambda_0 I)^{n-1} + \alpha_1 (H-\lambda_0 I)^{n-2} + \cdots + \alpha_{n-1}\big).$$

Now let λ_0 be such that $(A-\lambda_0 I)$ is nonsingular, so that we can consider the algebraic linear system $(A-\lambda_0 I)u = v$. Let us assume that $(H-\lambda_0 I)$ is nonsingular too, and let z be the solution of $(H-\lambda_0 I)z = \|v\|_2 \cdot e_1$ which yields the so-called FOM-approximation [8] of u given by

$$u_n = V_n z.$$

Then, according to (2), the residual $\rho_n(A,\lambda_0) := v - (A-\lambda_0 I)u_n$ is given by

$$\rho_n(A,\lambda_0) = -h_{n+1,n}\|v\|_2 \big(e_n^{\mathrm{T}}(H-\lambda_0 I)^{-1} e_1\big) \cdot v_{n+1}.$$

Proposition 3.7. *Under the assumptions above we have*

$$\rho_n(A,\lambda_0) = C_n \frac{1}{\mathrm{Det}(\lambda_0 I - H)} \|v\|_2 v_{n+1}. \tag{7}$$

Proof. By Propositions 3.6 and 3.1 we easily obtain

$$e_n^{\mathrm{T}}(H-\lambda_0 I)^{-1} e_1 = -\frac{1}{\mathrm{Det}(\lambda_0 I - H)} C_{n-1}.$$

Therefore (7) follows. □

Now let us consider the approximation method proposed in the previous section. Our first purpose is to give a suitable representation of z_n in the residual (4). From Proposition 3.4, recalling Proposition 3.1, we have

$$e_n^{\mathrm{T}} \exp(Hx) e_1 = C_{n-1}\psi_{n-1}(x). \tag{8}$$

Clearly $\psi_{n-1}(x) \to 0$, as $x \to 0$. We also observe that considering the derivatives $\psi_{n-1}^{(k)}(x)$, $k = 1, 2, \ldots,$

$$C_{n-1}\psi_{n-1}^{(k)}(x) = e_n^{\mathrm{T}} H^k \exp(Hx) e_1. \tag{9}$$

So, by Proposition 3.1,

$$\psi_{n-1}^{(k)}(0) = 0, \quad \text{for } k = 0, 1, \ldots, n-2. \tag{10}$$

Since

$$z(t) = \left(\int_0^t \exp\big(H(t-s)\big) f(s) \cdot e_1 \,\mathrm{d}s\right) \|v\|_2,$$

using (8), from (4), we easily get

$$r_n(t) = \|v\|_2 C_n \left(\int_0^t \psi_{n-1}(t-s) f(s) \,\mathrm{d}s\right) \cdot v_{n+1}. \tag{11}$$

Theorem 1.

$$\|r_n\| \leqslant \left(\frac{T^n}{n!}\right) h_{n+1,n}\|f\|\,\|v\|_2 \max_{t\in[0,T]} \left(e_n^{\mathrm{T}} H^{n-1} \exp(Ht)e_1\right). \tag{12}$$

Proof. Using the Taylor expansion, by (10), we have

$$\psi_{n-1}(x) = \psi_{n-1}^{(n-1)}(\xi)\frac{x^{n-1}}{(n-1)!}, \quad 0 \leqslant \xi \leqslant x. \tag{13}$$

Since

$$\left|\int_0^t \psi_{n-1}(t-s)f(s)\,\mathrm{d}s\right| \leqslant \|f\| \int_0^t \left|\psi_{n-1}(x)\right| \mathrm{d}x,$$

by (13) from (11) we easily get

$$\|r_n\| \leqslant \|v\|_2 h_{n+1,n}\|f\| \max_{\xi\in[0,T]} \left|C_{n-1}\psi_{n-1}^{(n-1)}(\xi)\right| \frac{T^n}{n!}.$$

Hence, by (9), (12) follows. □

Remark 3.1. Setting $F(A) = \{u^{\mathrm{T}}Au\colon\ \|u\|_2 = 1\}$, we have $F(H) \subset F(A)$. It is easy to see that $\|H\|_2 \leqslant \|A\|_2$ and $\gamma_n \leqslant \gamma$ (see [2]). Moreover, if $F(A)$ is contained in the left-half plane then $\gamma \leqslant 1$.

By these considerations, recalling that $h_{n+1,n} = v_{n+1}^{\mathrm{T}} A v_n$, from (12) we get:

Corollary 3.1.

$$\|r_n\| \leqslant \left(\frac{T^n}{n!}\right) \|A\|_2^n \gamma\|f\|\,\|v\|_2. \tag{14}$$

Although very general, the above estimate turns out to be rather pessimistic. Anyhow it points out the important fact that, for every fixed index k, it is possible to find T, such that the restarted procedure KPM(k), discussed in Remark 2.2, converges. Below we give further estimates for the symmetric case.

From now on let A be symmetric. Then H is symmetric too and by Proposition 3.5 each eigenvalue of H has algebraic multiplicity (and index) equal to 1. Let $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ be the eigenvalues of H.

Let $f(z)$ be a function defined on the spectrum of H. Then by [4, p. 314], we have

Proposition 3.8.

$$f(H) = \sum_{k=1}^{n} f(\lambda_k) Z_{k0}, \tag{15}$$

where

$$Z_{k0} = \prod_{j=1,\, j\neq k}^{s} (H - \lambda_j I) \Big/ \prod_{j=1,\, j\neq k}^{s} (\lambda_k - \lambda_j). \tag{16}$$

According to the above result, we have, for any scalar x,

$$\exp(Hx) = \sum_{k=1}^{n} \exp(\lambda_k x) Z_{k0}. \tag{17}$$

Then, referring to (6), in the symmetric case we have

$$\psi_{n-1}(x) = \sum_{k=1}^{n} \exp(\lambda_k x) \Big/ \prod_{j=1, j\neq k}^{n} (\lambda_k - \lambda_j)$$

and from interpolation theory

$$\psi_{n-1}(x) = \exp(\zeta x) x^{n-1}/(n-1)!, \quad \text{for } \zeta \in [\lambda_1, \lambda_n]. \tag{18}$$

Then, from (11) we easily get the bound

$$\|r_n\| \leqslant (T^n/n!) C_n \gamma_n \|f\| \, \|v\|_2. \tag{19}$$

Bounds for C_n can be obtained from Proposition 3.3 and for γ_n from Remark 2.3. For instance, from (19), we have

Proposition 3.9. *Assume that A is symmetric and negative definite with eigenvalues $\mu_1 \leqslant \mu_2 \leqslant \cdots \leqslant \mu_m < 0$. Then*

$$\|r_n\| \leqslant \frac{T^n \nu_n \|f\| \, \|v\|_2}{n!},$$

where

$$\nu_n = \min\left[\frac{(\mu_m - \mu_1)^n}{2^n}, \prod_{j=1}^{n} (\mu_m - \mu_j), \prod_{j=0}^{n-1} (\mu_{m-j} - \mu_1) \right].$$

Proof. Take into account that $\gamma < 1$ and take in (5) $\lambda = (\mu_m + \mu_1)/2$, $\lambda = \mu_m$ and $\lambda = \mu_1$. □

Now let a real λ_0 be such that $(A - \lambda_0 I)$ is negative definite. Since $(H - \lambda_0 I)$ is negative definite too, we can consider the FOM approximation and the corresponding residual $\rho_n(A, \lambda_0)$ represented by (7).

Proposition 3.10. *Assume that A is symmetric negative definite. Then*

$$\|r_n\| \leqslant (T^n/n!) \big\|\rho_n(A, 0)\big\|_2 \|f\| \operatorname{Det}(-H).$$

Proof. Clearly H is symmetric and negative definite. Then the result follows from (19), (7) and $\gamma_n < 1$. □

Theorem 2. *Let $(A - \lambda_0 I)$ be symmetric negative definite. There is a function $g(t)$ with*

$$\|g\| \leqslant \big\| \exp(\lambda_0 x) \big\|$$

such that

$$r_n(t) = g(t) \|f\| \rho_n(A, \lambda_0). \tag{20}$$

Proof. Referring to (11) let us set

$$g(t) = \mathrm{Det}(\lambda_0 I - H)\int_0^t \left(\psi_{n-1}(t-s)f(s)/\|f\|\right)\mathrm{d}s.$$

So that, from (11) and (7) we get (20).

Let us recall that, under our assumptions,

$$\psi_{n-1}(x) = \sum_{k=1}^{n}\frac{\exp(\lambda_k x)}{\prod_{j=1,j\neq k}^{n}(\lambda_k-\lambda_j)} = \exp(\lambda_0 x)\sum_{k=1}^{n}\frac{\exp((\lambda_k-\lambda_0)x)}{\prod_{j=1,j\neq k}^{n}(\lambda_k-\lambda_j)}.$$

Then

$$\left|\int_0^t \psi_{n-1}(t-s)f(s)\,\mathrm{d}s\right| \leqslant \max_{0\leqslant x\leqslant T}\left[\exp(\lambda_0 x)\right]\|f\|\int_0^t\sum_{k=1}^{n}\frac{\exp((\lambda_k-\lambda_0)x)}{\prod_{j=1,j\neq k}^{n}(\lambda_k-\lambda_j)}\,\mathrm{d}x.$$

Therefore,

$$\|g\| \leqslant \mathrm{Det}(\lambda_0 I - H)\max_{0\leqslant x\leqslant T}\left[\exp(\lambda_0 x)\right]\int_0^t\sum_{k=1}^{n}\frac{\exp((\lambda_k-\lambda_0)x)}{\prod_{j=1,j\neq k}^{n}(\lambda_k-\lambda_j)}\,\mathrm{d}x.$$

Observe that, for $t \geqslant 0$,

$$\int_0^t\sum_{k=1}^{n}\frac{\exp((\lambda_k-\lambda_0)x)}{\prod_{j=1,j\neq k}^{n}(\lambda_k-\lambda_j)}\,\mathrm{d}x \leqslant \int_0^\infty\sum_{k=1}^{n}\frac{\exp((\lambda_k-\lambda_0)x)}{\prod_{j=1,j\neq k}^{n}(\lambda_k-\lambda_j)}\,\mathrm{d}x = \frac{1}{\prod_{j=1}^{n}(\lambda_0-\lambda_j)}$$

so that $\|g\| \leqslant \max_{0\leqslant x\leqslant T}\exp(\lambda_0 x)$. □

By this result we can prove the following statement for the restarted method.

Proposition 3.11. *Assume that A is symmetric and negative definite. Then, for every $T > 0$ and for every index k,* KPM(k) *converges.*

Proof. Taking $\lambda_0 = 0$, by Theorem 2 we have, for every k,

$$r_k(t) = g(t)\|f\|\rho_k(A,0)$$

with $\|g\| \leqslant 1$. Now we know that $\rho_k(A,0)$ is the kth residual of the Conjugate Gradient Method applied to the linear system $Ax = v$, with starting guess the null vector. So, as is well known, there exists $\varepsilon_k < 1$ (independent of v) such that

$$\left\|\rho_k(A,0)\right\|_{A^{-1}} \leqslant \varepsilon_k\|v\|_{A^{-1}},$$

where

$$\|v\|_{A^{-1}} = \sqrt{-v^{\mathrm{T}}A^{-1}v}.$$

Then we have, for every $0 \leqslant t \leqslant T$,

$$\left\|r_k(t)\right\|_{A^{-1}} \leqslant \varepsilon_k\|f\|\,\|v\|_{A^{-1}}$$

and, setting $r_0(t) = f(t) \cdot v$, for every $T > 0$, we have

$$\max_{0 \leqslant t \leqslant T} \|r_k(t)\|_{A^{-1}} \leqslant \varepsilon_k \max_{0 \leqslant t \leqslant T} \|r_0(t)\|_{A^{-1}}.$$

This means that, for every k, the restarted procedure converges. □

4. Numerical experiments

In this section some numerical examples to illustrate the behavior of the scheme described above are presented. All numerical tests performed concern the solution of linear systems of ODEs arising from the semi-discretization of parabolic partial differential equations.

4.1. Symmetric tests

The first part of numerical tests comes from the discretization of the following bidimensional partial differential equation

$$\begin{cases} \dfrac{\partial u(x,y,t)}{\partial t} = \Delta u(x,y,t) + r(x,y,t), \\ u(x,y,0) = 0, \end{cases} \tag{21}$$

defined on the unit square and this analogous three-dimensional one

$$\begin{cases} \dfrac{\partial u(x,y,z,t)}{\partial t} = \Delta u(x,y,z,t) + r(x,y,z,t), \\ u(x,y,z,0) = 0, \end{cases} \tag{22}$$

defined on the unit cube, both with $u = 0$ on the boundary. The function r is defined in such a way that the exact solutions of the above partial differential equations are respectively

$$u(x,y,t) = \frac{tx(x-1)y(y-1)}{t+1}$$

and

$$u(x,y,z,t) = x(x-1)y(y-1)z(z-1)\cos(20\pi t)t.$$

The semi-discretization of these problems by the usual finite differences, leads to linear systems of ODEs of the form

$$\begin{cases} y' - Ay = g(t), \quad t \in [0,T], \\ y(0) = 0, \end{cases} \tag{23}$$

where A is symmetric negative definite. The right-hand side $g(t)$ is a vectorial function, due to the discretization of the function r and it turns out to be the sum of two different simpler terms of the form

$$f(t) \cdot v = \left[f(t)v_1, f(t)v_2, \ldots, f(t)v_m\right]^{\mathrm{T}},$$

where $f(t)$ is a scalar continuous function. As it has been pointed out by Remark 2.3, such problems can be solved considering as many different simpler IVPs of the kind (1) as the number of different terms of type $f(t) \cdot v$ in the right-hand side; for each term we report in the figures the behavior of the residual norm for both the full and the restarted algorithm. Discretizing with $m+1$ grid points in each space-direction, in the bidimensional case we have a block $m^2 \times m^2$ tridiagonal matrix and in the three-dimensional case we have block pentadiagonal matrix of size $m^3 \times m^3$. Discretization has been done with $m = 30$ for test (21) and $m = 10$ for test (22). For integrating the arising projected problems (3) the implicit midpoint rule has been used; the time-step in all tests is $h = 1/32$. We have applied the Krylov Projection Method in both the full (KPM) and the restarted version (KPM(k)) (see Proposition 3.11). A special experiment has been performed using the every step restarted algorithm, namely KPM(1). In the following figures it is reported the behavior of the maximum norm of the residual of (1) obtained using Eq. (4) (we got comparable values for the error norm): the dotted line refers to the restarted algorithm and the solid line to the full one. Fig. 1 concerns the bidimensional test (21) (test 1). The restart has been performed every 10 iterations.

Fig. 2 concerns the three-dimensional test (22) (test 2). In this case the restart has been performed every 5 iterations.

Nice results come out in the application of KPM(1), the step by step restated algorithm: in this case at every iteration we need to solve only a simple scalar linear ODE. Figs. 3 and 4 refer to these experiments.

In spite of slowing the convergence with respect to the full algorithm, restarting at every step is effective, as it drops to the minimum the computational effort per iteration.

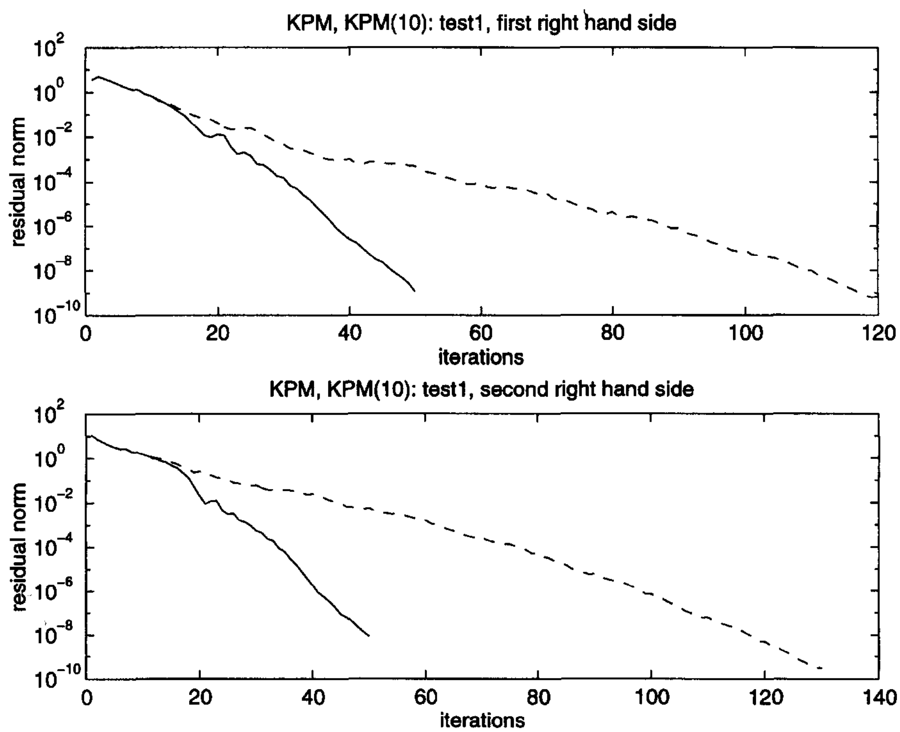

Fig. 1.

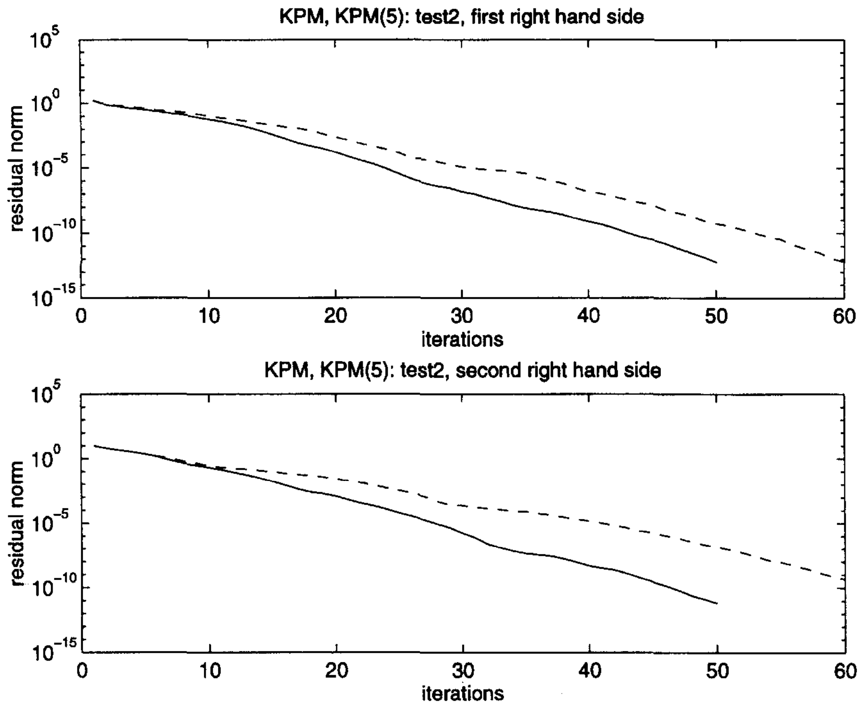
KPM, KPM(5): test2, first right hand side
residual norm
iterations
KPM, KPM(5): test2, second right hand side
residual norm
iterations

Fig. 2.

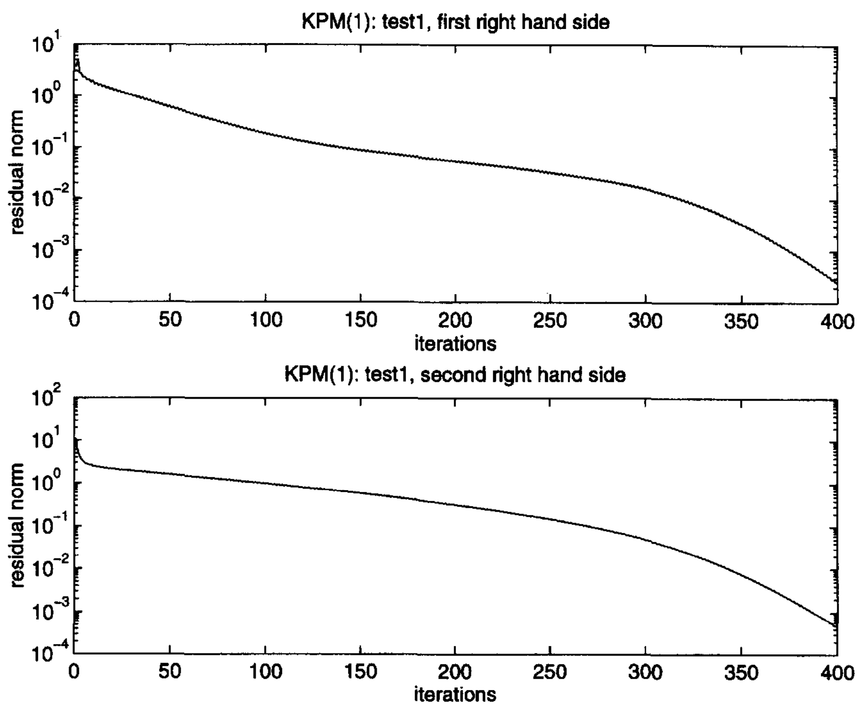
KPM(1): test1, first right hand side
residual norm
iterations
KPM(1): test1, second right hand side
residual norm
iterations

Fig. 3.

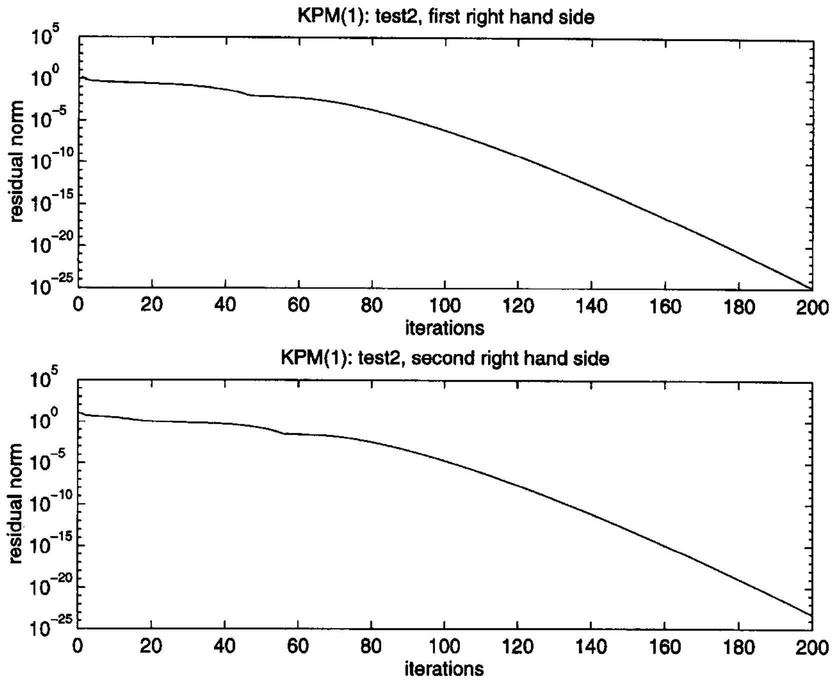

Fig. 4.

4.2. *Unsymmetric tests*

Unsymmetric tests are given by discretizing the following partial differential equation:

$$\frac{\partial u}{\partial t} = \Delta u + \gamma \frac{\partial u}{\partial x} + r(x, y, z, t)$$

defined on the unit cube, with homogeneous boundary conditions. Different tests are obtained fixing the parameter γ and the exact solution of the problem, from which the function r is deduced. The first unsymmetric test considers

$$u(x, y, z, t) = \frac{x(x-1)y(y-1)z(z-1)}{(t+1)} \tag{24}$$

and $\gamma = 50$. The second one is given by fixing the exact solution equal to

$$u(x, y, z, t) = \cos(20\pi t)x(x-1)y(y-1)z(z-1) \tag{25}$$

and $\gamma = 10$. Semi-discretization in space by finite differences leads to linear systems of ODEs whose matrices are unsymmetric. As in the symmetric tests the right-hand side turns out to be the sum of a few different terms of the form $f(t) \cdot v$. The spatial discretization has been done on a grid of $m = 10$ points in each direction and the time step in the implicit midpoint rule is $h = 1/32$. For both tests the right-hand side of (23) is composed of four different terms of the kind $f(t) \cdot v$. Fig. 5 refers to (24) (test 3) and Fig. 6 refers to (25) (test 4). The solid line represents KPM and the dotted one KPM(5).

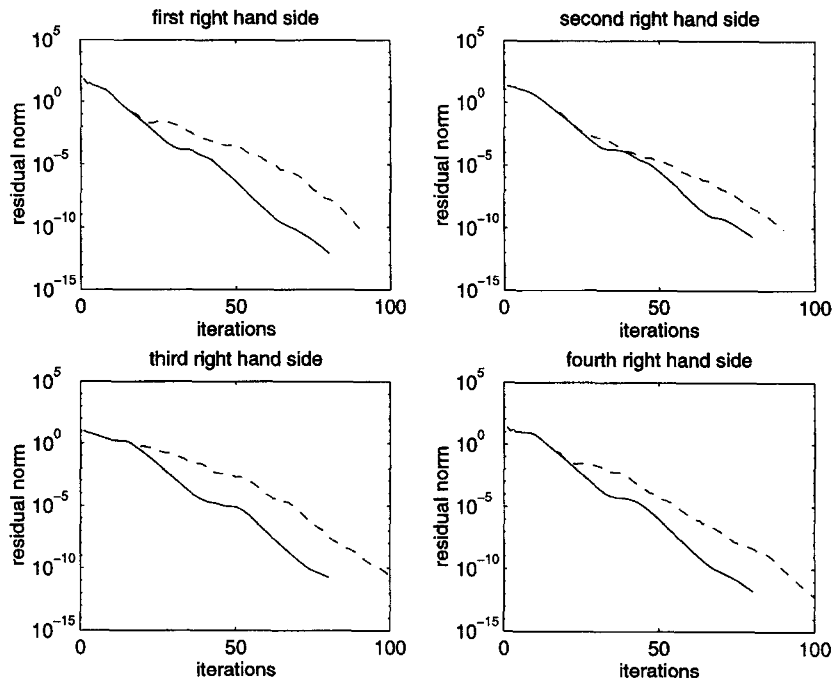

Fig. 5. KPM, KPM(5); test 3.

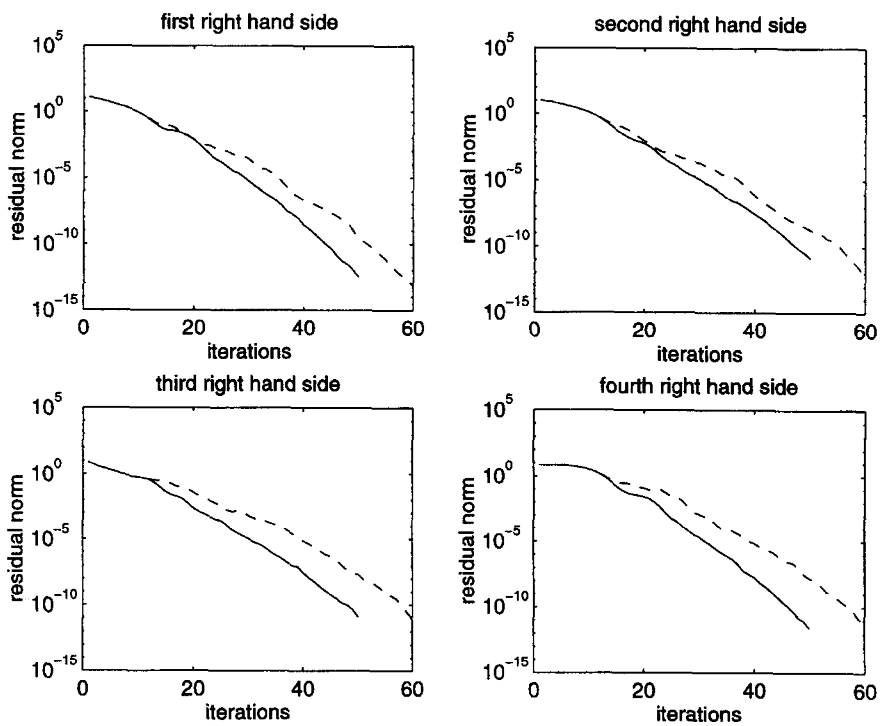

Fig. 6. KPM, KPM(5); test 4.

References

[1] E. Celledoni, A comparison of Krylov subspace methods for systems of ODEs, Quad. Mat. No. 317, Dipartimento di Scienze Matematiche Università degli Studi di Trieste (1994).
[2] E. Gallopoulos and Y. Saad, Efficient solution of parabolic equations by Krylov approximation methods, *SIAM J. Sci. Statist. Comput.* 13 (1992) 1236–1264.
[3] M. Hochbruck and C. Lubich, On Krylov subspace approximation to the matrix exponential operator, *SIAM J. Numer. Anal.*, submitted.
[4] P. Lancaster and M. Tismenetsky, *The Theory of Matrices* (Academic Press, New York, 1985).
[5] A. Lumsdaine and J.K. White, Accelerating waveform relaxation methods with applications to parallel semiconductor device simulation, *Numer. Funct. Anal. Optim.* 16 (1995) 395–414.
[6] B.N. Parlett, Global convergence of the basic QR algorithm on Hessenberg matrices, *Math. Comp.* 22 (1968) 803–817.
[7] A. Pietsch, *Eigenvalues and s-Numbers* (Cambridge University Press, Cambridge, 1987).
[8] Y. Saad, Krylov subspace methods for solving large unsymmetric linear systems, *Math. Comp.* 37 (1981) 105–126.
[9] Y. Saad, Analysis of some Krylov subspace approximations to the matrix exponential operator, *SIAM J. Numer. Anal.* 29 (1992) 209–228.