# A rational Krylov method for solving time-periodic differential equations

I. MORET
Dipartimento di Matematica e Informatica,
Università di Trieste,
Via Valerio 12/1, 34127 Trieste, Italy

P. NOVATI
Dipartimento di Matematica Pura ed Applicata
Unversità dell'Aquila,
Via Vetoio, Coppito, 67010, L'Aquila, Italy

**Abstract**

A rational Arnoldi method, for computing the action of a matrix function, is applied in the construction of a numerical scheme for solving time-periodic linear differential problems arising from the semidiscretization of parabolic problems. Theoretical and computational results illustrate the effectiveness of the approach.

## 1 Introduction

Restricted denominator (RD) rational forms of the type

$$R_{j,k}(x;\rho) = \frac{p_j(x)}{(1+\rho x)^k},$$

where $\rho$ is a real parameter and $p_j$ is a polynomial of degree $j$, were introduced by S. P. Nørsett in [24]. They are particularly attractive for the approximation of matrix functions, since all the needed inversions involve the same matrix. In [22], dealing with the computation of exponentials of large and sparse matrices, the authors studied approximations of this type obtained via projections onto Krylov subspaces. This approach has been discussed also by J. van den Eshof and M. Hochbruck in [9]. The fact that such methods turn out to be effective for initial value problems (cf. [25]) motivates the investigation on their application to the treatment of time-periodic problems, that is, systems of the type

$$\begin{aligned} &\frac{dy}{dt} + Ay = F(t), \quad t \in [0,T],\ F(t) \in \mathbb{R}^N, \\ &\qquad y(0) = y(T), \end{aligned} \tag{1}$$

where $F(t)$ is a vector valued function and $A$ is a $N \times N$ real matrix, which is *noncritical* with respect to $T$, that is

$$in\omega \notin \sigma(A), \text{ for every } n \in \mathbb{Z}, \ \omega = \frac{2\pi}{T},$$

where $\sigma(A)$ denotes the spectrum of $A$. This assumption on $A$ ensures that system (1) has a unique solution for any $F \in L^1(0,T)$ [28]. Moreover, we assume that, as it occurs in several important applications, all the eigenvalues of $A$ have positive real part. Referring to models by which various physical phenomena can be described, we can think to (1) as the semidiscretization of a parabolic boundary value problem.

For solving (1), when $A$ is of large dimension, the classical methods based on shooting or on global discretization (cf. [7]) are usually considered inadequate. Multigrid strategies [13] and waveform relaxation methods [28], [5], possibly combined together [19], [28], may be also adopted. Yet, they often present convergence problems as well as complexities in their implementation. An alternative procedure, here considered, can be built up employing the explicit formula of the solution of (1) that is

$$y(t) = \exp(-tA)(I - \exp(-TA))^{-1} \int_0^T \exp((s-T)A)F(s)ds + \int_0^t \exp((s-t)A)F(s)ds,$$

which suggests the following scheme:

**Step 1**. Compute $v(T)$, where $v(t)$ solves the IVP

$$\begin{aligned} &\frac{dv}{dt} + Av = F(t), \ t \in (0,T], \\ &v(0) = 0. \end{aligned} \tag{2}$$

**Step 2**. Compute

$$y_0 = v(T) + w(T), \tag{3}$$

where, by the periodicity of $y(t)$,

$$w(T) = g(A)v(T),$$

with $g$ defined by

$$g(a) = \exp(-Ta)(1 - \exp(-Ta))^{-1}. \tag{4}$$

**Step 3**. Solve the IVP

$$\begin{aligned} &\frac{dy}{dt} + Ay = F(t), \ t \in (0,T], \\ &y(0) = y_0. \end{aligned} \tag{5}$$

Because of its simplicity the scheme is attractive. Indeed, besides the computation of $w(T)$, it requires only to solve two IVPs. Clearly, the crucial point lies in the effective computation of $w(T)$. In this paper we focus the attention mainly on this problem. For our goal, taking into account that only the action of the matrix $g(A)$ on a vector is required, Krylov subspaces methods are surely well suited. In the recent years, several studies have been devoted to such kind of methods for approximating the action of matrix functions. Concerning the standard polynomial Krylov methods, we quote here, among the others, the papers [18], [26], [14], [23], [1], [21]. Although such methods appear suitable for large and sparse matrices, to work on Krylov subspaces generated by discretizations of unbounded operators presents some fundamental difficulties, which involve the own definition of such subspaces, given, in general, for bounded operators. Such structural problems are displayed by the convergence of the arising approximations. Infact it deteriorates as the discretizations are refined, even dealing with entire functions (see [14]). By the way, we point out that the case here treated presents further problems due to the presence of a singularity in the function $g$, so that, as it is well known (see [18]), the convergence of polynomial approximations may be very slow.

The remedy proposed in [22] and [9] consists in considering subspaces generated by an auxiliary matrix obtained by a suitable rational form. So doing, a sort of preconditioning of the problem is introduced (cf. [9]). In the applications of our interest, the auxiliary matrix reflects the properties of an underlying compact operator so that the associated Krylov subspaces are well defined.

The paper is organized as follows. In section 2 we introduce the RD-rational Krylov method for the computation of $w(T)$. In section 3 we analize the convergence and we give some error estimates which will be useful in the control of the procedure. Finally, in Section 4 we present some numerical experiments which illustrate various features of the method.

## 2 The Rational Krylov method

Let $\Pi_k$ be the set of the algebraic polynomials of degree $\leq k$. The Euclidean scalar product is denoted by $\langle \cdot, \cdot \rangle$ and $\|.\|$ represents the Euclidean vector norm as well as the corresponding induced matrix norm. The notation $W(M)$ indicates the *numerical range* of a square matrix $M$, i.e.,

$$W(M) := \left\{ \frac{< x, Mx >}{< x, x >}, x \in \mathbb{C}, x \neq 0 \right\}.$$

For $0 < \vartheta < \frac{\pi}{2}$ and $\beta \geq 0$, let us define the set

$$\Sigma_{\vartheta,\beta} = \{\lambda \in \mathbb{C} : |\arg(\lambda - \beta)| \leq \vartheta\},$$

and assume that

$$W(A) \subset \Sigma_{\vartheta,\beta}, \tag{6}$$

for some $\beta > 0$.

Let us introduce the matrix

$$Z = (I + hA)^{-1}, \tag{7}$$

where $h > 0$ is a real parameter. For $0 < \gamma < 1/2$ , let $D_\gamma$ be the disk centered at $(\gamma, 0)$ with radius $\gamma$, i.e.,

$$D_\gamma = \left\{\lambda : \Re(\lambda)^2 + \Im(\lambda)^2 - 2\gamma\Re(\lambda) \leq 0\right\}.$$

Then, we define the set

$$S_{\vartheta,\gamma} = \Sigma_{\vartheta,0} \cap D_\gamma. \tag{8}$$

The following statements can be easily proved.

**Lemma 1** *For any $h > 0$ let us set $\gamma(\beta) = \frac{1}{2(1+h\beta)}$ and $\beta(\gamma) = h^{-1}((2\gamma)^{-1}-1)$. We have:*

1. *the function $z(a) = (1 + ha)^{-1}$ maps any set $\Sigma_{\vartheta,\beta}$ $(\beta > 0)$ into $S_{\vartheta,\gamma(\beta)}$ and the inverse function $a(z) = h^{-1}(z^{-1} - 1)$ maps $S_{\vartheta,\gamma}$ $(0 < \gamma < 1/2)$ into $\Sigma_{\vartheta,\beta(\gamma)}$;*
2. *if $W(A) \subset \Sigma_{\vartheta,\beta}$ $(\beta > 0)$, then $W(Z) \subset S_{\vartheta,\gamma(\beta)}$.*

We seek for approximations, to $w(T) = g(A)v(T)$, belonging to the Krylov subspaces

$$K_m(Z, v) = span\left\{v, Zv, Z^2v, ..., Z^{m-1}v\right\},$$

where

$$v = v(T).$$

Such approximations have the form $R_{m-1,}(A; h)v$, where

$$R_{m-1}(a; h) = \frac{p_{m-1}(a)}{(1 + ha)^{m-1}}, \quad p_{m-1} \in \Pi_{m-1}, \quad m \geq 1, \quad \Re(a) > 0.$$

For building up the spaces $K_m(Z, v)$ we employ the classical Arnoldi's method (cf. [27]) which provides an orthonormal sequence of vectors $\{v_1, v_2, ..., v_j, ...\}$, with $v_1 = \frac{v}{\|v\|}$ , such that, for each $m \geq 1$, $K_m(Z, v) = span\left\{v_1, v_2, ..., v_m\right\}$. As it is well known we have

$$ZV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^T, \tag{9}$$

where $V_m = [v_1, v_2, ..., v_m]$ and $H_m = V_m^H Z V_m$ with entries $h_{i,j}$ Here and below $e_j$ is the $j$-th vector of the canonical basis of $\mathbb{R}^m$.

Let us go back to Step 2 in our scheme, namely to the computation of $w(T) = g(A)v$. For $a \in \Sigma_{\vartheta,\beta}$ $(\beta > 0)$, in light of Lemma 1, setting $z = (1+ha)^{-1}$ and $\tau = \frac{T}{h}$ we can write

$$g(a) = f(z), \tag{10}$$

where

$$f(z) = \frac{\exp(-\tau(z^{-1} - 1))}{1 - \exp(-\tau(z^{-1} - 1))}. \tag{11}$$

By statement 2 in Lemma 1 the function $f(z)$ is well defined in any set $S_{\vartheta,\gamma}$.

Accordingly, in $K_m(Z, v)$ we define the $m$-th RD-rational approximation to $w(T)$ as

$$w_m(T) = p(Z)v,$$

where $p \in \Pi_{m-1}$ interpolates in the Hermite sense the function $f$ in the eigenvalues of $H_m$. Using (9) one can see that

$$w_m(T) = V_m f(H_m) \|v\| e_1,$$

that is

$$w_m(T) = V_m g(B_m) \|v\| e_1, \tag{12}$$

where the matrix $B_m$ satisfies

$$H_m(I + hB_m) = I. \tag{13}$$

The following result can easily be proved, with the help of Lemma 1 and by the fact that $W(H_m) \subseteq W(Z)$.

**Lemma 2** *Let assumption (6) hold. For any $m \geq 1$, let the matrices $V_m$ and $H_m$ be generated by Arnoldi's method as above and let $B_m$ be defined by (13). Then*

$$\Re(W(B_m)) \geq \beta.$$

Owing to this result, $B_m$ is *noncritical* with respect to $T$, for every $T > 0$ and hence $g(B_m)$ is well defined. Assuming that $m$ is not large, the computation of $g(B_m)e_1$ can be carried out as follows. If $A$ is symmetric then $B_m$ is symmetric too and the Schur decomposition can be employed. In the general case, we can adopt one of the existing algorithms for the computation of the matrix exponential, as for instance the Matlab routine `expm`, based on the scaling and squaring process. This is what we did in our numerical tests. Anyhow, the computation of $g(B_m)e_1$ requires a certain amount of work, so that, in order to control the approximation process, error estimates which do not imply this computation are clearly welcome. Some of these will be obtained in the next section.

**Remark 3** *Formally equivalent alternatives to the scheme presented in the introduction can be considered. At first we note that, rather than to use (3), one can obtain $y_0$ as $y_0 = (1 - \exp(-TA))^{-1}v(T)$. For our purposes we prefer to use (3) because, in order to give the convergence statements, it allows us to use some known results (cf. [22]). We point out that the two representations yield equivalent numerical results, for both our RD-rational method and the polynomial Krylov method considered, for comparison, in the numerical tests of Section 4. Observe that, having computed the m-th approximation to $y_0$, say $y_{0,m} = v(T) + w_m(T)$, one can evaluate the residual $v(T) - (1 - \exp(-TA))y_{0,m}$ using again the RD-rational Krylov method, dealing now with the exponential function (see [22], [9]). Such method applies also when the solution $v(t)$ of (2) is maintained at disposal, so that, in alternative to Step 3, we can get $y(t) = v(t) + \exp(-tA)y_0$, for $t \in (0, T]$.*

**Remark 4** *The approach extends easily to more general cases. For instance when $A$ is replaced by a time-dependent matrix of type $A(t) = \alpha(t)A + \beta(t)I$, where $\alpha$ and $\beta$ are scalar functions, or even when the left-hand side of equation (1) has the more general form $M\frac{dy}{dt} + Ny$, with $M$ invertible. In this case $Z$ can be taken as $Z = (M + hN)^{-1}M$ . It is easy to realize that in this way we obtain again a reformulation of the type (10). Applications to this and other more general situations will be discussed in a forthcoming paper.*

# 3 Error estimates

In this section we give convergence results and error estimates for the approximations $w_m(T)$, with the aim of determining also suitable choices for the parameter $h$ in (7). In order to do this, some of the arguments used in [22], for the exponential function, can still be employed. Yet, now, dealing with the function $g$, because of the presence of the singularity (at $a = 0$), their application needs some care and some additional considerations are in order.

Referring to (9), we introduce the polynomial

$$q_m^{(0)}(z) = \det(zI - H_m).$$

One easily realizes that

$$\left\| q_m^{(0)}(Z)v \right\| = \|v\| \prod_{j=1}^{m} h_{j+1,j}. \tag{14}$$

We start our convergence analysis giving some statements which generalize Proposition 3.2 in [22].

**Theorem 5** *Let $G$ be a bounded domain (in the complex plane) whose boundary is a rectifiable Jordan curve $\Gamma$. Assume that $W(Z) \subset G$ and that there exixts $d_\Gamma > 0$ such that*

$$dist(\lambda, W(Z)) \geq |\lambda|\, d_\Gamma, \text{ for every } \lambda \in \Gamma.$$

*Let $f(z)$ be a function such that $f(z)/z$ is analytic in $G$ and continuous on $\Gamma$, then setting*

$$\varepsilon_k(\Gamma, f) = \min_{p_k \in \Pi_k} \max_{\lambda \in \Gamma} \int_\Gamma \left| \frac{f(\lambda) - p_k(\lambda)}{\lambda} \right| |d\lambda|,$$

*we have, by the Cauchy integral formula,*

$$\|f(Z)v - V_m f(H_m)\, \|v\|\, e_1\| \leq \frac{\varepsilon_{m-1}(\Gamma, f)\, \|v\|}{\pi d_\Gamma} \ (\to 0 \text{ as } m \to \infty).$$

*Moreover, the following bound holds*

$$\|f(Z)v - V_m f(H_m)\, \|v\|\, e_1\| \leq \frac{\left\| q_m^{(0)}(Z)v \right\|}{2\pi d_\Gamma} \int_\Gamma \frac{|f(\lambda)|}{|\lambda| \left| q_m^{(0)}(\lambda) \right|} |d\lambda|. \tag{15}$$

In order to apply this results to our case, we need the following lemma.

**Lemma 6** *Let $W(Z) \subseteq S_{\vartheta,\gamma(\beta)}$, with $\gamma(\beta) = (1+h\beta)^{-1}/2$. Let $h > 0$ and $\vartheta < \vartheta^* < \frac{\pi}{2}$ be such that*

$$2\gamma^* := (1+h\beta)^{-1} \frac{\cos\vartheta^* + \sin(\vartheta^* - \vartheta)}{\cos\vartheta^* \cos^2(\vartheta^* - \vartheta)} < 1. \tag{16}$$

*Then*

$$dist(\lambda, W(Z)) \geq |\lambda| \sin(\vartheta^* - \vartheta), \tag{17}$$

*for every $\lambda \in \Gamma$ , where $\Gamma$ is the the contour of the set $S_{\vartheta^*,\gamma^*}$ (cf. (8)).*

**Proof.** If $\lambda \in \Gamma$ belongs to the lines $\lambda = |\lambda| (\cos\vartheta^* \pm i \sin\vartheta^*)$, then clearly (17) holds. Now consider $\lambda (\neq 0) \in \Gamma \cap D_{\gamma^*}$ and observe that

$$|\lambda| > 2\gamma^* \cos\vartheta^*. \tag{18}$$

Let us set $\gamma = \gamma(\beta)$. By geometric arguments one can see that

$$dist(\lambda, W(Z)) \geq \left( \gamma^2 + \frac{(\gamma^* - \gamma)}{\gamma^*} |\lambda|^2 \right)^{\frac{1}{2}} - \gamma.$$

By this relation, using (18) one realizes that (16) implies (17). ■

It is easy to see that, since $\beta > 0$, for every $\vartheta$ there are $h$ and $\vartheta^*$ which verify the assumptions of the previous lemma and moreover that, for every given $\vartheta$ and $h$, there is $\vartheta^*$ which fulfils (16). Therefore that taking $G = S_{\vartheta^*,\gamma^*}$, with $\vartheta^*$ and $\gamma^*$ as in Lemma 6, for the function $f$ defined in (11), the assumptions of Theorem 5 are fulfilled with $d_\Gamma = \sin(\vartheta^* - \vartheta)$.

Moreover, from the bound (15) we are able to derive error estimates as follows..

**Proposition 7** *Let $W(Z) \subseteq S_{\vartheta,\gamma(\beta)}$. Taking any $\vartheta^*$ and $\gamma^*$ as in Lemma 6, the following error bound holds*

$$\|w(T) - w_m(T)\| \leq \frac{\|v\| \exp(\tau) C_\Gamma (m-1)!}{\pi\tau^m (\sin(\vartheta^* - \vartheta)\cos\vartheta^*)^m \sin(\vartheta^* - \vartheta)} \prod_{j=1}^{m} h_{j+1,j}. \tag{19}$$

*where*

$$C_\Gamma = \left(1 - \exp\left(-T\beta^*\right)\right)^{-1},$$

*with*

$$\beta^* = h^{-1}\left((2\gamma^*)^{-1} - 1\right). \tag{20}$$

**Proof.** Take $G = S_{\vartheta^*,\gamma^*}$, with $\vartheta^*$ and $\gamma^*$ as in Lemma 6. From (15) we obtain

$$\|w(T) - w_m(T)\| \leq \frac{\left\| q_m^{(0)}(Z) v \right\|}{2\pi \sin(\vartheta^* - \vartheta)} \int_\Gamma \frac{|f(\lambda)|}{|\lambda| \left| q_m^{(0)}(\lambda) \right|} |d\lambda|, \tag{21}$$

where $\Gamma$ is the contour of $S_{\vartheta^*,\gamma^*}$. Setting $\beta^*$ as in (20) we have

$$\max_{\lambda\in\Gamma}\left|\frac{1}{1-\exp(-\tau(\lambda^{-1}-1))}\right| = (1-\exp(-T\beta^*))^{-1} = C_\Gamma,$$

and from (21) we get

$$\int_\Gamma \frac{|f(\lambda)|}{|\lambda|\left|q_m^{(0)}(\lambda)\right|}\,|d\lambda| \le C_\Gamma \int_\Gamma \frac{\left|\exp(-\tau(\lambda^{-1}-1))\right|}{|\lambda|\left|q_m^{(0)}(\lambda)\right|}\,|d\lambda|\,.$$

Since $W(H_m)\subseteq W(Z)$ and recalling that the zeros of the (monic) polynomial $q_m^{(0)}$ are the eigenvalues of $H_m$, by Lemma 6 we have, for $\lambda\in\Gamma$,

$$\left|q_m^{(0)}(\lambda)\right| \ge (|\lambda|\sin(\vartheta^*-\vartheta))^m.$$

Therefore

$$\int_\Gamma \frac{\left|\exp(-\tau(\lambda^{-1}-1))\right|}{|\lambda|\left|q_m^{(0)}(\lambda)\right|}\,|d\lambda| \le \frac{\exp(\tau)}{(\sin(\vartheta^*-\vartheta))^m}\int_\Gamma \frac{\exp(-\tau\,|\lambda|^{-1}\cos\vartheta^*)}{|\lambda|^{m+1}}\,|d\lambda|\,.$$

One proves that (cf. [22] Corollary 3.4)

$$\int_\Gamma \frac{\exp(-\tau\,|\lambda|^{-1}\cos\vartheta^*)}{|\lambda|^{m+1}}\,|d\lambda| \le 2\int_0^\infty \frac{\exp(-\tau x^{-1}\cos\vartheta^*)}{x^{m+1}}dx = \frac{2(m-1)!}{(\tau\cos\vartheta^*)^m}. \tag{22}$$

Hence by (14) and (21) we obtain the desired result. ■

We emphasize the fact that (19) gives an error estimate which is computable without knowing the approximation $w_m(T)$. The quantities involved in (19) are $\tau$, $\vartheta^*$, $C_\Gamma$ and $\prod_{j=1}^m h_{j+1,j}$. These values depend on the parameter $h$. In general, the optimal value of $h$ is not easy to find. As a consequence each case under consideration must be examined separately, taking into account that the basic requirement for a good choice of $h$ is to satisfy (16) with a suitable $\vartheta^*$.

For instance, the minimization of the term

$$\frac{\exp(\tau)}{\tau^m(\sin(\vartheta^*-\vartheta)\cos\vartheta^*)^m},$$

in the right-hand side of (19), suggests the values $h=\frac{T}{m}$, and $\vartheta^*=(\pi+2\vartheta)/4$. If $\beta$ is sufficiently large, values around these are well suited. In light of various numerical experiments, we point out that, in this case and when $T$ is no too small, the method maintains a fast convergence for all the values of $h$ into the interval $[\frac{T}{10},T]$. Otherwise, if $\beta$ is small, in order to satisfy (16), a larger $h$ must be taken. The values of $\beta$ and $\vartheta$ can be obtained from informations on the location of the numerical range $W(Z)$. In several important situations such informations are available by theoretical results (see e.g., [10]). Otherwise, numerical techniques can be adopted. On this topics we refer to the book [12]. As it is well known, in general, for $m$ sufficiently large, the extreme eigenvalues of $H_m$ match sufficiently well the ones of $Z$. This suggests that, in practice, one could get the required informations by simply investigating $W(H_m)$.

# 4 Numerical experiments

In our numerical experiments the matrix $A$ is obtained by discretizing 1-D or 2-D second order differential operators by finite differences. Besides to stress the effectiveness of the proposed rational method (the RA (Rational Arnoldi)-method in the sequel) we want also to compare it with the classical polynomial method which seeks approximations to $w(T)$ in the Krylov subspaces $K_m(A, v)$ constructed by the Arnoldi method (the PA (Polynomial Arnoldi)-method) [18]. All methods are implemented in Matlab. In the examples, we consider the errors with respect to a reference solution.

The PA-method does not require any inversion, however, as we already pointed out, if $A$ arises from the discretization of an elliptic operator its convergence deteriorates refining the discretization (cf. [14]). The previous theoretical results, as well as the numerical experiences, emphasize that this does not occur for the RA method, which furthermore presents, in all the cases treated, a fast convergence. On the other hand, the rational method needs to solve a certain number of linear systems, which involve the same matrix $(I + hA)$, so that a factorization or a good preconditioner must be computed only once. There are various important cases in the applications where such systems can be solved cheaply, owing to the particular structure of $A$ (cf. [8]). In all our experiments we use the LU (or Cholesky) factorization. We refer to [9] for an ad-hoc use of iterative methods.

Furthermore we stress the fact that this preliminary work can turn out to be useful even in the solution of the initial value problems in our scheme, where methods that require the solution of linear systems involving matrices of the type $(I + \delta A)$, with $\delta$ close to $h$, are employed. For details on this point we refer to the recent paper [2]. This situation occurs both if we use an implicit method and if we adopt one of the various exponential integrators proposed in the recent literature (cf. [11], [15], [16], [6], [17], [20], [3], [25] ). As it is well known, the latter methods require several applications of $\exp(-tA)$ and of other related matrix functions. Such computations can be performed by the RD-Krylov rational method (cf. [22], [9] and [25]) working again on the subspaces $K_m(Z, w)$, as explained in Example 3 below.

The first two examples refer simply to the computation of $g(A)v$, with $g$ defined by (4).

**Example 1**. We begin by considering the one dimensional operator

$$L = -\frac{\partial^2}{\partial x^2} + c\frac{\partial}{\partial x}, \quad c \in \mathbb{R},$$

in the interval $(0, 1)$, with Dirichlet boundary conditions. It is discretized on a uniform meshgrid, with meshsize $\delta = 1/(N + 1)$ using central differences. Accordingly an usual tridiagonal matrix $A$ arises. Table 1 shows the fast convergence of the RA-method and the drastic work-reduction with respect to the PA-method. There we report the values of $m$ (number of iterations) and the related computational costs with respect to the dimension of the problem $ops/N$,

required for getting an Euclidean error norm less than $10^{-6}$. The LU factorization of $(I+hA)$ can be done very cheaply and we do not consider this in the costs reported in Table 1. The vector $v$ is taken as the (normalized) discretization of the function $v(x) = x(1-x)$.

| | RA | | PA | |
|---|---|---|---|---|
| $N$ | $m$ | $ops/N$ | $m$ | $ops/N$ |
| 100 | 5 | 76 | 74 | 5994 |
| 200 | 6 | 96 | 150 | 23550 |
| 300 | 6 | 96 | 227 | 53118 |

Table 1 - 1D- test with $T = 0.1, c = 5, h = T/10$.

**Example 2**. In 2-D we consider the second order differential operator

$$L = -\Delta + c_1 \frac{\partial}{\partial x} + c_2 \frac{\partial}{\partial y}, \quad c_1, c_2 \in \mathbb{R}$$

where $\Delta$ is the Laplacian on the unit square $(0,1) \times (0,1)$, again with Dirichlet boundary conditions. The spatial meshsize is $\delta = 1/(n+1)$ in both directions. Central differences are used for all the derivatives, obtaining a block-tridiagonal $N \times N$ ($N = n^2$) matrix $A$.

For the selected values of $n$ and $h$, the LU (or Cholesky) factorization of $(I+hA)$ has been computed. The vector $v$ is now the (normalized) discretization of the function $v(x,y) = x(1-x)y(1-y)$. We have performed various numerical tests with values of $h$ in the interval $[\frac{T}{10}, T]$, obtaining similar results. Some of them are reported in Tables 2-5.

Figure 1 shows a comparison between the true errors and the error estimates (19), in the case $c_1 = c_2 = 0$, $T = 0.1$, $N = 2500$. We notice that $\beta \approx 19$ because in this case the eigenvalues of $A$ are $\lambda_{i,j} = \frac{4}{\delta^2}(\sin^2 \frac{i\pi\delta}{2} + \sin^2 \frac{j\pi\delta}{2})$, $1 \le i, j \le n$. Accordingly, we can take $h = T$ and $\vartheta^* = \pi/5$, so that condition (16), with $\vartheta = 0$, remains fulfilled with $2\gamma^* \approx 0.9$.

In Tables 2-5 we compare the computational costs of the RA and the PA methods on a PC 3.5GHz. We report therein the values of $m$ in order to achieve a relative error less than $10^{-4}$, together with the speed-up ($SU$ = CPU time of PA / CPU time of RA), for two nonsymmetric cases, with different values of $T$. For the RA method the CPU time comprehends also the cost of the $LU$ factorization, which could be used also in solving problems (2) and (5), as we do in the next example. For both methods the CPU time considered includes the cost of the evaluation of the approximation $w_m(T)$ only at the final step. It is worth while nothing that this is a clear advantage for the polynomial method, for which the usual way for monitoring the error (cf. [15], [21]) requires several evaluations of the approximations.$w_m(T)$. Moreover, in light of the long recursions required, there are other facts which do not speak in favour of the PA-method, that is, the need of a large storage and the possible necessity of re-orthogonalization.

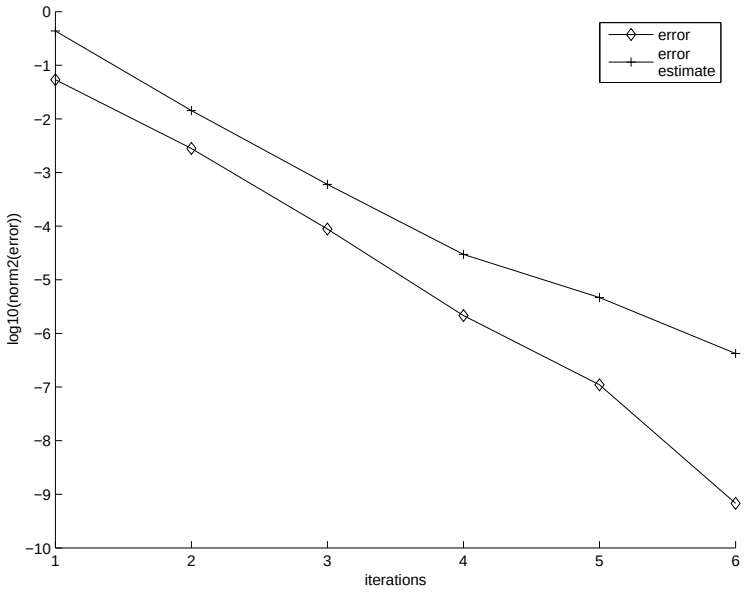


Figure 1: Example 2- Error and error bound in a 2-D symmetric case with $T = h = 0.1$, and $N = 2500$.

| | PA | RA | | | |
|---|---|---|---|---|---|
| | | $h = T/10$ | | $h = T$ | |
| $N$ | $m$ | $m$ | $SU$ | $m$ | $SU$ |
| 400 | 38 | 9 | 4.8 | 6 | 5.0 |
| 900 | 55 | 9 | 5.8 | 6 | 6.1 |
| 1600 | 72 | 9 | 4.5 | 6 | 4.9 |
| 2500 | 90 | 9 | 5.3 | 6 | 5.9 |

.

Table 2 - 2D-test with $T = 0.1, c1 = 10, c2 = 5$.

| | PA | RA | | | |
|---|---|---|---|---|---|
| | | $h = T/10$ | | $h = T$ | |
| $N$ | $m$ | $m$ | $SU$ | $m$ | $SU$ |
| 400 | 52 | 9 | 6.7 | 8 | 7.1 |
| 900 | 76 | 9 | 10.3 | 8 | 10.8 |
| 1600 | 89 | 9 | 7.1 | 8 | 7.6 |
| 2500 | 111 | 9 | 8.3 | 8 | 8.9 |

Table 3 - 2D-test with $T = 0.5, c1 = 10, c2 = 5$.

| | PA | RA | | | |
|---|---|---|---|---|---|
| | | $h = T/10$ | | $h = T$ | |
| $N$ | $m$ | $m$ | $SU$ | $m$ | $SU$ |
| 400 | 34 | 9 | 4.6 | 7 | 4.9 |
| 900 | 49 | 9 | 5.4 | 7 | 5.8 |
| 1600 | 64 | 9 | 4.3 | 7 | 4.7 |
| 2500 | 79 | 9 | 4.8 | 7 | 5.2 |

Table 4 - 2D-test with $T = 0.1, c1 = 20, c2 = 0$.

| | PA | RA | | | |
|---|---|---|---|---|---|
| | | $h = T/10$ | | $h = T$ | |
| $N$ | $m$ | $m$ | $SU$ | $m$ | $SU$ |
| 400 | 43 | 11 | 4.2 | 10 | 4.3 |
| 900 | 50 | 11 | 4.0 | 10 | 4.2 |
| 1600 | 75 | 11 | 3.7 | 10 | 4.0 |
| 2500 | 92 | 11 | 4.1 | 10 | 4.5 |

Table 5 - 2D-test with $T = 0.3, c1 = 20, c2 = 0$.

**Example 3**. Finally we present a numerical test where we solve a periodic problem (1) completely, making a comparison between the polynomial and the rational Arnoldi methods into the framework of an exponential integrator. Here we briefly describe such method. Let us consider (2) (or (5)) and assume that $F(t) = \sum_{j=0}^{p} t^j b_j$, for $t \in [t_0, t_0 + \Delta t]$, where $b_j \in \mathbb{R}^N$, $j = 0, 1, .., p$. Then $v(t_0 + \Delta t)$ can be expressed as

$$v(t_0 + \Delta t) = \varphi_0(-\Delta t A)v(t_0) + \sum_{j=1}^{p+1}(j-1)!(\Delta t)^j \varphi_j(-\Delta t A)b_{j-1}, \qquad (23)$$

where

$$\varphi_0(-\Delta t a) = \exp(-\Delta t a),$$
$$\varphi_j(-\Delta t a) = \frac{1}{(j-1)! t^j}\int_0^{\Delta t} \exp(-(\Delta t - s)a)s^{j-1}ds, \text{ for } j = 1, 2, ...$$

These functions can be represented in the recursive form

$$\varphi_{k+1}(u) = \frac{\varphi_k(u) - \frac{1}{k!}}{u}, \ \varphi_k(0) = \frac{1}{k!}, \text{ for } k = 0, 1, 2, ...$$

Accordingly we can employ the RD-rational Krylov method even for the computation of all the vectors $\varphi_j(-\Delta t A)b_{j-1}$.

In our example $A$ is the $N \times N$ (with $N = n^2$) block-tridiagonal matrix obtained by discretizing, as before in both the directions, the operator

$$L = -k_1 \frac{d^2}{dx^2} - k_2 \frac{d^2}{dy^2}, \qquad k_1, k_2 \in \mathbb{R},$$

on the square $(0,1) \times (0,1)$, with homogeneous Dirichlet conditions in $x$ and homogeneous Neumann conditions in $y$. We consider the values $k_1 = 1$, $k_2 = 10$, $N = 900$.

We take $F$ as the "sawtooth" function, namely the periodic function whose restriction to $[0,T)$ is $F(t) = te$, where $e = (1,...,1)^T$ and $T = 0.5$. Problem (1) is solved following the three steps explained in the introduction. Because of the particular choice of $F(t)$, following (23) we compute

$$v(T) = T^2 \varphi_2(-TA)e,$$

and then

$$w(T) = g(A)v(T),$$

where $g$ is defined by (4). Observe that, in this particular case, one could work with a single matrix function.

The exact solution is given by

$$y(t) = \exp(-tA)y_0 + t^2 \varphi_2(-tA)e, \tag{24}$$

where $y_0 = v(T) + w(T)$ (see (3)). We choose to compute $y(t)$ also at the points $t_j = j\Delta t$, where $\Delta t = 0.1$ and $j = 1,...,4$, using formula (24). In Fig.2 we consider a work-precision diagram obtained applying the PA- and the RA-method for the numerical approximation of of $v(T)$, $w(T)$ and $y(t_j)$, $j = 1,...,4$. The diagram is obtained increasing the values of $m$ for the computation of all the matrix functions involved (actually for the rational method $m \leq 6$). The error considered is

$$err = \max_j \|y(t_j) - \overline{y}(t_j)\|, \qquad j = 0,...,4,$$

where $\overline{y}$ is the numerical solution. Here, $Z$ is defined with $h = T/10$.

As expected, although the RA-method initially suffers because of the cost of the LU decomposition, it can rapidly achieve a great accuracy with respect to the PA-method, that requires a large number of iterations in order to reach the superlinear convergence (cf. [14]) and hence an accurate approximation of the solution.

# References


[1] L. Bergamaschi, M. Caliari and M. Vianello, Efficient approximation of the exponential operator for discrete 2D advection-diffusion problems, Numer. Linear Algebra Appl. 10 (2003), 271-289.

[2] M. Benzi and D. Bertaccini, Approximate inverse preconditioning for shifted linear systems, BIT 43 (2003), 231-244.

[3] H. Berland, B. Skaflestad and W. Wright, Expint — A Matlab package for exponential integrators. To appear on ACM Transactions on Mathematical Software.

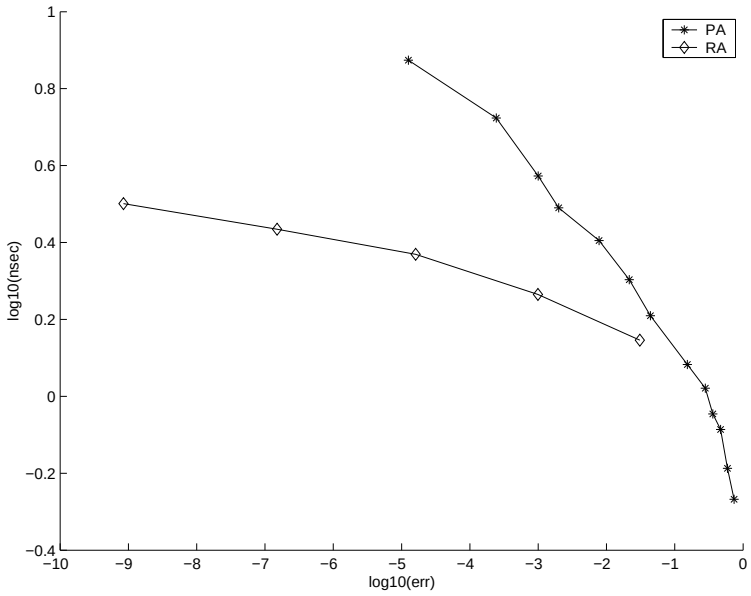


Figure 2: Work-precision diagram for Example 3.

[4] B. Beckermann, Image numérique, GMRES et polynômes de Faber. C. R. Math. Acad. Sci. Paris 340 (2005), no. 11, 855–860.

[5] E. Celledoni, Discrete QMR and BGC in the numerical solution of linear systems of ODEs, J. Comput. Appl. Math. 91 (1998), 159-177.

[6] S. M. Cox and P. C. Matthews, Exponential time differencing for stiff systems, J. Comput. Phys. 176 (2002), 430-455.

[7] P. Deuflhard and F. Bornemann, Scientific Computing with Ordinary Differential Equations, Springer, Berlin, 2002.

[8] V. Druskin and L. Knizhnerman, Extended Krylov subspaces: approximation of the matrix square root and related functions, SIAM J. Matrix Anal. Appl. 19 (1998), 755-771.

[9] J. van den Eshof and M. Hochbruck, Preconditioning Lanczos approximations to the matrix exponential, SIAM J. Sci. Comput. 27 (2005), 1438-1457

[10] M. Eiermann, Fields of values and iterative methods, Linear Algebra and its Applications 180 (1993), 167-197..

[11] E. Gallopoulos and Y. Saad, Efficient solution of parabolic equations by Krylov approximation methods, SIAM Sci. Stat. Comput. 13 (1992), 1236-1264.

[12] K.E. Gustafson and D.K.M. Rao, Numerical Range: The Field of Values of Linear Operators and Matrices, Springer-Verlag, New york, 1997.

[13] W. Hackbusch, Fast numerical solution of time-periodic parabolic problems by a multigrid method, SIAM Sci. Statist. Comput. 2 (1981), 198-206.

[14] M. Hochbruck and C. Lubich, On Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal. 34 (1997), 1911-1925.

[15] M. Hochbruck, C. Lubich and H. Selhofer, Exponential integrators for large systems of differential equations, SIAM J. Sci. Comput. 19 (1998), 1552-1574.

[16] M. Hochbruck and A. Ostermann, Exponential RungeKutta methods for parabolic problems, Appl. Numer. Math. 53 (2005), 323-339.

[17] A. K. Kassam and L. N. Trefethen, Fourth-order time stepping for stiff PDES, SIAM J. Sci. Comput. 26 (2005), 1214-1233.

[18] L. Knizhnerman, Calculation of functions of unsymmetric matrices using Arnoldi's method, U.S.S.R. Comput. Maths. Math. Phys. 31 (1991), 1-9.

[19] C. Lubich and A. Ostermann, Multigrid dynamic iteration for parabolic equations, BIT 27 (1987), 216-234.

[20] B. V. Minchev and W. M. Wright, A review of exponential integrators for first order semi-linear problems. Preprint Numerics 2/2005, Norwegian University of Science and Technology, Trondheim, Norway, www math.ntnu.no/preprint/numerics/2005/ N2-2005.ps.

[21] I. Moret, P. Novati, Interpolatig functions of matrices on zeros of quasi-kernel polynomials, Numerical Linear Algebra and Applications 12 (2005), 337-353

[22] I. Moret and P. Novati, RD-rational approximations of the matrix exponential, BIT 44 (2004), 595-615.

[23] I. Moret and P. Novati, The computation of functions of matrices by truncated Faber series, Numer. Funct. Anal. and Optimiz. 22 (2001), 697-719.

[24] S. P. Nørsett, Restricted Padé approximations to the exponential function, SIAM J. Numer. Anal. 15 (1978), 1008-1029.

[25] P. Novati, On the construction of Restricted-Denominator Exponential W-methods, Preprint 2006.

[26] Y. Saad, Analysis of some Krylov subspace approximations to the matrix exponential operator, SIAM J. Numer. Anal. 29 (1992), 209-228.

[27] L. N. Trefethen, D. Bau , Numerical Linear Algebra. SIAM, Philadelphia, 1997.

[28] S. Vandevalle and R. Piessens, On dynamic iteration methods for solving time-periodic differential equations, SIAM J. Numer. Anal. 30 (1993), 286-303.