Efficient computation of the core functions of exponential itegrators

Paolo Novati

Department of Pure and Applied Mathematics, University of Padova, Italy

Abstract. In this paper we investigate some practical aspects concerning the use of the Restricted-Denominator (RD) rational Arnoldi method for the computation of the core functions of exponential integrators for parabolic problems. We derive a useful a-posteriori bound that exploits the fast convergence of the Arnoldi method for compact operators. Some numerical experiments arising from the discretization of sectorial operators are presented.

INTRODUCTION

For the solution of large stiff problems of the type

$$u'(t) = f(y(t)) = Lu(t) + N(u(t)), \tag{1}$$

where $L \in \mathbb{R}^{M \times M}$ arises from the discretization of unbounded sectorial operators and N is a nonlinear function, in recent years much work has been done on the construction of exponential integrators that might represent a promising alternative to classical solvers (see e.g. [8] for a comprehensive survey). As well known the computation of the matrix exponential or related matrix functions is at the core of this kind of integrators. Under the hypothesis that such matrix functions are exactly evaluated, the linear stability can be trivially achieved for both Runge-Kutta and multistep based exponential integrators and hence highly accurate and stable integrators can be constructed.

An exponential integrator based on the variation-of-constants formula requires at each time step the evaluation of a certain number (depending on the accuracy) marix functions of the type $\varphi_k(hL)v$, where

$$
\begin{aligned}
\varphi_0(h\lambda) &= \exp(h\lambda), \\
\varphi_{k+1}(h\lambda) &= \frac{\varphi_k(h\lambda) - \frac{1}{k!}}{h\lambda}, \text{ for } k = 0, 1, 2, \dots,
\end{aligned} \tag{2}
$$

being h the time step. Among the existing techniques for the computation of matrix functions (we quote here the recent book of Higham [7] for a survey), in this context the Restricted-Denominator (RD) Rational Arnoldi algorithm introduced independently in [15] and [10] for the computation of the matrix exponential seems to be a reliable approach. It is based on the use of the so called RD rational forms

$$R_{i,j}(\lambda) = \frac{q_i(\lambda)}{(1 - \delta\lambda)^j}, \quad \delta \in \mathbb{R},$$

where q_i is a polynomial of degree $\leq i$. While in the matrix case, the use of these approximants requires the solution of linear systems with the matrix $(I - \delta L)$, as shown in [12] in the context of the solution of (1) when L is sectorial so typically sparse and well structured this linear algebra drawback can be almost completely overtaken organizing suitably the step-size control strategy and exploiting the properties of the RD Arnoldi method concerning the choice of the parameter δ. In other words the number of linear systems to be solved can be drastically reduced with respect to the total number of computations of matrix functions required by the integrator. Therefore the mesh independence property of the method, that leads to a very fast convergence with respect to a standard polynomial approach (see again [10]), can be fully exploited for the construction of competitive integrators.

A problem still open is that inside the integrator the rational Arnoldi algorithm (responsible for most of the computational cost) has to be supported by a robust and sharp error estimator. In the self-adjoint case the problem

has been treated in [11] where the author presents effective a-posteriori error estimates, even in absence of information on the location of the spectrum of L. Anyway, in the general case, when (1) arises for instance from the discretization of parabolic problems with advection terms and/or non-zero boundary conditions, the numerical range of L, that we denote by $F(L)$, may not reduce to a line segment. In this sense the basic aim of this paper is to fill this gap providing error estimates for the non-symmetric case using as few as possible information about the location of $F(L)$. It is necessary to keep in mind that a competitive code for (1) should also be able to update L (interpreted as the Jacobian of f, [14], [2]) so that $F(L)$ may be not fixed during the integration, and so it is important to reduce as much as possible any pre-processing technique to estimate $F(L)$. In particular assuming that $F(L) \subseteq \mathbb{C}^-$ we shall provide a-posteriori error estimates for the RD Arnoldi process using only information about the angle of the sector containing $F(L)$, angle that is typically independent of the sharpness of the discretization and hence computable working in small dimension.

THE RD RATIONAL ARNOLDI METHOD

In what follows we denote by $\|\cdot\|$ the Euclidean vector norm and its induced matrix norm. As already mentioned, the notation $F(L)$ indicates the *numerical range* of L, that is,

$$F(L) := \left\{ \frac{x^H L x}{x^H x}, x \in \mathbb{C}^M \setminus \{0\} \right\}.$$

The notation Π_m indicates the space of the algebraic polynomials of degree $\leq m$.

Given $0 \leq \theta < \frac{\pi}{2}$, let

$$S_\theta = \{\lambda : |\arg(-\lambda)| \leq \theta\} \subset \mathbb{C}^- \tag{3}$$

be the unbounded sector of the left half complex plane, symmetric with respect to the real axis with vertex in 0 and semiangle θ. Moreover, let Γ_θ be the boundary of S_θ. Throughout the paper we assume that $F(L) \subset int(S_\theta)$, the interior of S_θ. Accordingly, L is a so-called *sectorial* operator.

Given a vector $v \in \mathbb{R}^M$, with $\|v\| = 1$, consider the problem of computing

$$y^{(k)} = \varphi_k(hL)v, \tag{4}$$

where φ_k is defined by (2). The RD rational approach seeks for approximations to $\varphi_k(h\lambda)$ of the type

$$R_{m-1,m-1}(\lambda) = \frac{p_{k,m-1}(\lambda)}{(1 - \delta\lambda)^{m-1}}, \quad p_{k,m-1}(\lambda) \in \Pi_{m-1}, \quad m \geq 1,$$

where $\delta > 0$ is a suitable parameter. Turning to the matrix case, $y^{(k)}$ is approximated by elements of the Krylov subspaces

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

with respect to v and $Z = (I - \delta L)^{-1}$. In this sense the idea is to use a polynomial method to compute $y^{(k)} = f_k(Z)v$, where

$$f_k(z) := \varphi_k\left(\frac{h}{\delta}\left(1 - \frac{1}{z}\right)\right)$$

is singular at 0.

For the construction of the subspaces $K_m(Z,v)$ we employ the classical Arnoldi method. As is well known it generates an orthonormal sequence $\{v_j\}_{j \geq 1}$, with $v_1 = v$, such that $K_m(Z,v) = span\{v_1, v_2, ..., v_m\}$. Moreover, for every m,

$$ZV_m = V_m H_m + h_{m+1,m}v_{m+1}e_m^H, \tag{5}$$

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

The m-th RD-rational Arnoldi approximation to $y^{(k)}$ is defined as (see [9])

$$y_m^{(k)} = V_m f_k(H_m)e_1. \tag{6}$$

It can be seen that

$$y_m^{(k)} = \overline{p}_{k,m-1}(Z)v, \tag{7}$$

where $\overline{p}_{k,m-1} \in \Pi_{m-1}$ interpolates, in the Hermite sense, the function $f_k(z)$ in the eigenvalues of H_m (see [13]).

As mentioned in the Introduction, this technique has been introduced independently in [15] and [10]. Anyway, the idea of using rational Krylov approximations to matrix functions was originally introduced in [5].

ERROR BOUNDS

Before deriving an error bound for the method, we need to locate $F(Z)$. Consider the function $\chi(\lambda) = (1 - \delta\lambda)^{-1}$. Denoting by $D_{1/2,1/2}$ the disk centered at $1/2$ with radius $1/2$, let

$$G_\theta = \{z : z = \chi(\lambda), \lambda \in S_\theta\} \subseteq D_{1/2,1/2}. \tag{8}$$

Its boundary, Σ_θ, is made by two circular arcs meeting with angle 2θ at 0 and 1. Regarding the field of values of Z, $F(Z)$, we can state the following result that will be used by Theorem 2.

Proposition 1 *If $F(L) \subset int(S_\theta)$ then $F(Z) \subset int(G_\theta)$.*

Going back to the RD Arnoldi method, the corresponding error

$$E_{k,m} := y^{(k)} - y_m^{(k)}$$

can be expressed and bounded in many ways (we quote here the recent papers [1] and [4] for a background on the error estimates for both polynomial and rational Arnoldi approximation to matrix functions). In particular, exploiting the interpolatory nature of the standard Arnoldi method, we notice, as pointed out also in [6], that the error can be expressed in the form

$$E_{k,m} = g_{k,m}(Z)q_m(Z)v, \tag{9}$$

where $q_m(z) := \det(zI - H_m)$ and (cf. (7))

$$g_{k,m}(z) := \frac{f_k(z) - \overline{p}_{k,m-1}(z)}{\det(zI - H_m)}. \tag{10}$$

In order to provide a-posteriori error bounds we need to introduce the generalized Laguerre polynomials, defined by

$$L_n^{(\alpha)}(z) = \sum_{j=0}^n (-1)^j \binom{n+\alpha}{n-j} \frac{z^j}{j!}.$$

We can state the following main result that makes use of the relation (see [10])

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

Theorem 2 *Assume that $F(L) \subset S_\theta$, with $\theta < \frac{\pi}{3}$, and let $\tau := h/\delta$. Setting $c_j(\theta) := \left(1 + \sqrt{2(1 - \cos(\theta))}\right)^j$ we have*

$$\|E_{k,m}\| \leq K \frac{e^{\tau\left(\cos\theta - \frac{1}{2}\right) - m - k - 1}}{\tau^{m+k}} \left(\frac{2(m+k+1)}{2\cos\theta - 1}\right)^{m+k+1} C_{k,m}(\tau,\theta) \prod_{i=1}^m h_{i+1,i}, \tag{11}$$

$$\leq K \frac{e^{\tau\cos\theta - m - k - 1}}{\tau^{m+k}} \left(\frac{2(m+k+1)}{2\cos\theta - 1}\right)^{m+k+1} C'_{k,m}(\theta) \prod_{i=1}^m h_{i+1,i}, \tag{12}$$

where

$$C_{k,m}(\tau,\theta) \quad : \quad = \frac{(m-1)!}{(m+k)!} \sum_{j=0}^{m-1} \left|L_{m-1-j}^{(k)}(\tau)\right| c_j(\theta), \tag{13}$$

$$C'_{k,m}(\theta) \quad : \quad = \frac{(m-1)!}{(m+k)!} \sum_{j=0}^{m-1} \binom{m+k-j-1}{k} c_j(\theta), \tag{14}$$

and $2 \leq K \leq 11.08$ [3]. In the symmetric case we can take $K = 1$.

The reason for which we consider two bounds in Theorem 2 is that the second one (12) allows us to define suitably the parameter τ (and then δ) while the first one (11) should be used whenever τ has been defined. Indeed, assuming $\prod_{i=1}^{m} h_{i+1,i}$ independent of δ and then of τ by (12), looking for the minimum of $e^{\tau\cos\theta}\tau^{-(m+k)}$ we easily find that the optimal value for τ is given by

$$\tau = (m+k)/\cos\theta. \tag{15}$$

A COMPUTED EXAMPLE

In order to show the quality of the bound (11) let us consider the operator

$$Lu = -u'' + cu', \quad c \geq 0, \tag{16}$$

discretized with central differences in $[0,1]$ with uniform mesh $h = 1/(M+1)$, and Dirichelet boundary conditions. We consider the computation of the functions $\varphi_k(hL)v$, with $v = (1,...,1)^T/\sqrt{M}$, and $k = 0,1,2$, for $h = 0.5$. We take $\tau = 8/\cos\theta$ for all the experiments. The results for $M = 1000$ are displayed in Figure 1.
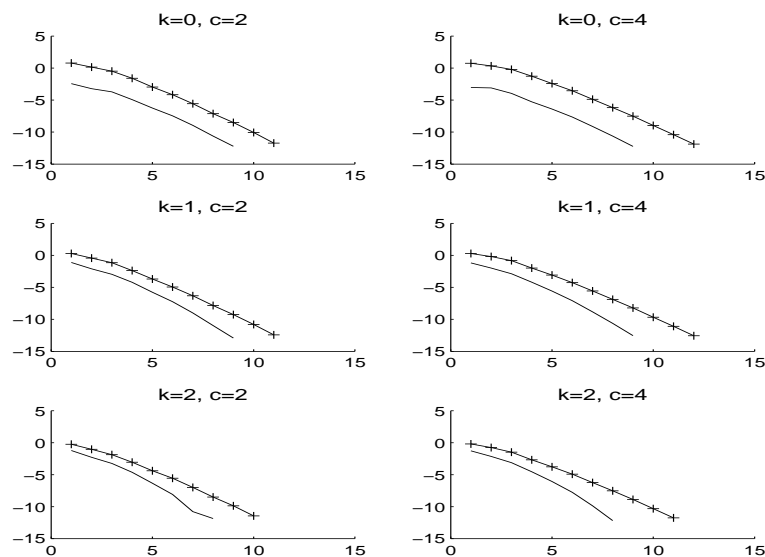

FIGURE 1. Error (-) and error bound (11) (+), in logarithmic scale, for $k = 0,1,2$, $h = 0.5$, L arising from (16) with $c = 2,4$.

REFERENCES

1. B. Beckermann, and L. Reichel, *SIAM J. Numer. Anal.* **47**, 3849–3883 (2009).
2. M. Caliari, and A. Ostermann, *Appl. Numer. Math.* **59**, 568–581 (2009).
3. M. Crouzeix, *J. Func. Anal.* **244**, 668–690 (2007).
4. F. Diele, I. Moret, and S. Ragni, *SIAM J. Matrix Anal. Appl.* **30**, 1546–1565 (2008).
5. V. Druskin, and L. Knizhnerman, *SIAM J. Matrix Anal. Appl.* **19**, 755–771 (1998).
6. M. Eiermann, and O. G. Ernst, *SIAM J. Numer. Anal.* **44**, 2481–2504 (2006).
7. N. J. Higham, *Functions of matrices. Theory and computation*. SIAM, Philadelphia, PA, 2008.
8. M. Hochbruck, and A. Ostermann, *Acta Numer.* **19**, 209–286 (2010).
9. L. Knizhnerman, *U.S.S.R. Comput. Maths. Math. Phys.* **31**, 1–9 (1991).
10. I. Moret, and P. Novati, *BIT* **44**, 595–615 (2004).
11. I. Moret, *Numer. Lin. Alg. Appl.* **14**, 445–457 (2007).
12. P. Novati, *J. Comput. Appl. Math.* **221**, 86–101 (2008).
13. Y. Saad, *SIAM J. Numer. Anal.* **29**, 209–228 (1992).
14. M. Tokman, *J. Comput. Phys.* **213**, 748–776 (2006).
15. J. v. d. Eshof, and M. Hochbruck, *SIAM J. Sci. Comp.* **27**, 1438–1457 (2005).