RD-Rational Approximations of the Matrix Exponential
*

I. MORET[1] and P. NOVATI[2]

[1] *Dipartimento di Matematica e Informatica*
Università degli Studi di Trieste, Via Valerio 12/1
34127, Trieste, Italy. email: moret@units.it

[2] *Dipartimento di Matematica Pura ed Applicata*
Università degli Studi dell'Aquila, Via Vetoio, Coppito
67010, L'Aquila, Italy. email: novati@univaq.it

Abstract.

Restricted Denominator (RD) rational approximations to the matrix exponential operator are constructed by interpolation in points related to Krylov subspaces associated to a rational transform of the particular matrix considered. Convergence analysis are provided. Numerical experiments show the effectiveness of the proposed methods in applications involving discretizations of differential operators.

AMS subject classification: 65F10, 65L05.

Key words: Krylov subspace methods, rational approximations, matrix exponential.

1 Introduction.

The approximation of the exponential by rational functions is a classical topic in the context of the numerical solution of differential equations. As pointed out by various authors, in view of the application to the matrix case, i.e. the computation of $exp(-tA)v$, where v is a given vector, a very important aspect to be considered is represented by the choice of the form of the denominator, in order to maintain the inversion process as simple as possible, exploiting possible special structures of the involved matrices. Moreover, there are various important cases where repeated actions of the inverse of a matrix on a vector can be carried out at a limited cost (see [8]). In light of these considerations, the so called Restricted Denominator RD-rational forms, introduced in [27] are of particular interest. Such functions are of the type $R_{j,k}(x) = \frac{q_j(x)}{(1+\rho x)^k}$, $\rho \in \mathbb{R}$, where q_j is a polynomial of degree $\leq j$. Therefore, in the matrix case, the linear systems to be solved share the same coefficient matrix, with obvious advantages, since matrix factorizations as well as preconditioners, like ILU or similar ones, can be computed once and for all.

In [27] such formulae are specialized in order to fulfil some order requirements, getting approximations which mainly fit the exponential and its derivatives at

*Received July 1997. Revised June 2003. Communicated by Axel Ruhe.

the origin, where they are Hermitian interpolators. However this fact may limit their practical effectiveness in the application to problems involving matrices which arise from discretizations of differential operators. We refer to [19] for a convincing discussion on this point. This fact motivated the research of formulae which are able to fit the exponential in certain given points [20] suggested by the particular problem we are dealing with. In the context of the RD methods the papers [28] and [19] work in this direction. More in general, this goal was pursued by all those authors who, for the numerical approximation of matrix functions, proposed polynomial methods which interpolate in points somehow related to the spectra of the matrices involved. The most popular example of such points is represented by the so-called *Ritz values* arising from the classical Arnoldi method for the construction of Krylov subspaces. The corresponding approximations, which are often viewed as projection methods, are discussed, among the others, in the papers [22], [7], [14], [34], [16]. Other interpolation points considered are zeros of Faber polynomials [25], or even Fejer points [30] or Leja points [4] associated to suitable subsets of the complex plane. Such procedures have recently found application in the context of the so-called exponential integrators (see e.g., [17], [29]). Unfortunately, they have the drawback that the convergence depends on the capacity of the numerical range of the matrix A and it deteriorates as this capacity increases.

All the above facts as well as some ideas contained in the classical book [44], have motivated the study of the RD-rational methods we propose in this paper. Referring to the more recent literature, they can be viewed as particular Extended Krylov Subspaces Methods [8] and more in general as Rational Krylov Methods [33]. From the projective point of view, they fall into the class of the so-called reduced-basis methods [32] or reduced-order models [13]. Finally we want to point out the recent paper [43], where various aspects concerning the present approach are discussed.

The paper is organized as follows. In Section 2, we discuss about the RD-rational approximations to the exponential function and we present the general RD-rational Krylov methods. In Section 3 and 4, we consider respectively the Arnoldi-Ritz values and the zeros of suitable Faber polynomials as interpolation points and we discuss the convergence properties. In Section 5, the effectiveness of the algorithms is illustrated by numerical experiments involving discretizations of differential operators.

2 RD-rational approximations.

In what follows the Euclidean scalar product is denoted by $< .,. >$. The Euclidean vector norm and its induced matrix norm is $\|.\|$. The notation $W(M)$ indicates the *numerical range* of a square matrix M, i.e.,

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

The spectrum of M is denoted by $\sigma(M)$. For any given $t > 0$, we consider the computation of $y(t) = \exp(-tA)v$, where A is a real $N \times N$ matrix and v is a

given real vector. For simplicity we assume from now on that

$$\|v\| = 1.$$

Before considering the matrix case, we want to discuss on the scalar case. As before mentioned, the RD-rational approximations to the exponential function were introduced in [27]. Here below we summarize some results given therein. In the sequel, we denote by Π_j the set of the algebraic polynomials of degree $\leq j$.

DEFINITION 2.1. *For any scalar $\varrho(\in \mathbb{R})$ a form of the type*

$$R_{j,k}(x;\rho) = \frac{q_j(x)}{(1+\rho x)^k}, \quad q_j \in \Pi_j, \quad k \geq 0, \tag{2.1}$$

is called an $RD(j,k)$-rational form.

In [27] it was proved that the highest attainable order of a $RD(j,k)$-approximation to $\exp(-x)$ is $j+1$. The order, at $x=0$, is q if the error is $O(x^{q+1})$ (cf. [28]). In particular ([27] Corollary 2.1, Th. 4.3, Th. 4.4) it was proved that, for any given $\varrho \in \mathbb{R}^+$, the $RD(k,k)$ -approximation to $\exp(-x)$, for $\Re(x) > 0$, of order at least k exists; it was constructed and, denoting it by $R_k^k(x;\varrho)$, an error formula is provided by which one proves that

$$\lim_{k\to\infty} R_k^k(x;\varrho) = \exp(-x),$$

uniformly for $x \in \mathbb{C}$ with $\Re(x) > 0$. Moreover, in [27] it was also shown how $RD(k,k)$-approximations of order $k+1$ can be built. We notice that this order can be achieved only for suitable values of ϱ, which change with k. Further results for the real case are also presented.

However, the statements on the order of a method should be consider with caution dealing with matrices whose eigenvalues grow rapidly, as in the case of discretizations of elliptic operators. In fact, the numerical experience shows that in such cases approximations like $R_k^k(x;\varrho)$, which fit the exponential and the derivatives at the origin, do not perform well, unless t is very small.

In order to obtain good RD-rational approximations, we try to construct them on the basis of information on the specific problem we are dealing with. For our purposes, let us formulate the scalar problem as the computation of $\exp(-ta)$, $\Re(a) > 0$, for some $t > 0$. Chosen a real parameter $h > 0$, we consider RD-rational forms of the type

$$R_k(a;h) = \frac{p_k(a)}{(1+ha)^k}, \quad p_k \in \Pi_k, \quad k \geq 0. \tag{2.2}$$

Setting $x = ha$ and $\varrho = \frac{h}{t}$, this can be viewed as a form (2.1). Accordingly we shall deal with polynomials of the shifted and inverted variable $z = (1+ha)^{-1}$. In fact, we can also express $\exp(-ta)$ as a function of z, namely

$$\exp(-ta) = f(z)$$

where

$$f(z) = \exp(-\tau(z^{-1} - 1)), \tag{2.3}$$

with $\tau = \frac{t}{h}$. Thus, we consider polynomial approximations to $f(z)$.

Now we introduce some notations. In the right-half plane we define the set

$$\Sigma_{\vartheta,\beta} = \left\{\lambda : |\arg(\lambda - \beta)| < \vartheta, 0 < \vartheta < \frac{\pi}{2}, \beta \geq 0\right\}.$$

and, for $r > 0$, the bounded sector

$$S_{\vartheta,r} = \{\lambda : \lambda \in \Sigma_{\vartheta,0}, 0 < |\lambda| \leq r\}. \tag{2.4}$$

Observe that the function $z(a) = (1+ha)^{-1}$ maps any set $\Sigma_{\vartheta,\beta}$ into $S_{\vartheta,(1+h\beta)^{-1}}$, more precisely into $\Sigma_{\vartheta,0} \cap D_{(1+h\beta)^{-1}/2}$, where by D_ρ we denote the disk of centre and radius ρ.

PROPOSITION 2.1. *Let τ be fixed. Given any sector $S_{\vartheta,r}$, for every integer $s \geq 0$ and for every $\varepsilon > 0$, there is $k > 0$ and a polynomial $p_k \in \Pi_k$, such that*

$$\left|\frac{f(\lambda) - p_k(\lambda)}{\lambda^s}\right| < \varepsilon, \tag{2.5}$$

for every $\lambda \in \overline{S_{\vartheta,r}}$.

PROOF. The function $\frac{f(\lambda)}{\lambda^s}$ is analytic in $S_{\vartheta,r}$ and continuous on $\overline{S_{\vartheta,r}}$, thus by Walsh's Theorem (cf. [24], Th 3.9) there is $k \geq s$ and a polynomial $q_{k-s} \in \Pi_{k-s}$, such that

$$\left|\frac{f(\lambda)}{\lambda^s} - q_{k-s}(\lambda)\right| < \varepsilon,$$

for every $\lambda \in \overline{S_{\vartheta,r}}$. Hence the thesis follows. □

Now consider the matrix case. We assume that

$$W(A) \subseteq \Sigma_{\vartheta,\beta}. \tag{2.6}$$

Accordingly A is a so-called *sectorial* operator, with vertex β and semi-angle θ. We refer to [21], Chap. V, for definitions and examples. Moreover, we refer to [11] for a discussion on the location of the numerical range.

Let us return to the computation of $y(t) = \exp(-tA)v$. Given a real parameter $h > 0$, by the *shift and invert* procedure we define the matrix

$$Z = (I + hA)^{-1}. \tag{2.7}$$

Note that if $A = M^{-1}P$, then $Z = (M + hP)^{-1}M$.

Under (2.6) we have (cf. [15])

$$\|Z\| \leq \frac{1}{1 + h\beta}, \tag{2.8}$$

more precisely one can see that $W(Z) \subseteq \Sigma_{\vartheta,0} \cap D_{(1+h\beta)^{-1}/2} \subset S_{\vartheta,(1+h\beta)^{-1}}$.

Our goal is to approximate $y(t)$ by vectors of the type $R_{m-1}(A;h)v$, for $m \geq 1$, which belong to the Krylov subspaces generated by Z and v, i.e., $K_m(Z,v) = span\left\{v, Zv, Z^2v, ..., Z^{m-1}v\right\}$. As it is usual, for generating such subspaces we consider a sequence of vectors $\{v_1, v_2, ..., v_j, ...\}$ with

$$v_1 = v,$$

such that

$$Zv_j = \sum_{i=1}^{j+1} h_{i,j} v_i, \quad j \geq 1,$$

We assume that $h_{j+1,j} > 0$, for $j = 1, 2, ..., m$.

Setting $h_{i,j} = 0$, for $j < i - 1$, denoting by H_m the $m \times m$ upper Hessenberg matrix having entries $h_{i,j}$, for $i, j = 1, 2, ..., m$, and considering the $N \times m$ matrix $V_m = [v_1, v_2, ..., v_m]$, we get the fundamental relationship

$$ZV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^T. \tag{2.9}$$

Here and below e_j is the j-th vector of the canonical basis of $\mathbb{R}^m$. Furthermore, setting

$$q_m^{(0)}(z) = \det(zI - H_m), \tag{2.10}$$

using (2.9) one finds that

$$(\prod_{j=1}^{m} h_{j+1,j}) v_{m+1} = q_m^{(0)}(Z)v. \tag{2.11}$$

For any given $t > 0$, we define the $m-$th approximation to $y(t) = \exp(tA)v$ by

$$y_m(t) = V_m f(H_m) e_1. \tag{2.12}$$

In other words,

$$y_m(t) = V_m \exp(-tB_m) e_1$$

where

$$H_m(I + hB_m) = I.$$

Observe that, if $\sigma(H_m)$ is contained in a set $\Sigma_{\vartheta,0} \cap D_\rho, 0 < \rho \leq \frac{1}{2}$, then $\Re\sigma(B_m) \geq \frac{(1-2\rho)}{2h\rho}$.

This approach extends that introduced, for self-adjoint operators, in [44], Chap. 5, where the transform $z = a^{-1}$ is employed. Some numerical experiments concerning that procedure can be found in the recent paper [31].

It is well known ([34]) that, as a consequence of (2.9), we have

$$y_m(t) = p_{m-1}^*(Z)v, \tag{2.13}$$

where $p_{m-1}^* \in \Pi_{m-1}$ interpolates f (in the Hermite sense) in the eigenvalues of H_m and moreover, for every $p_{m-1} \in \Pi_{m-1}$,

$$V_m p_{m-1}(H_m) e_1 = p_{m-1}(Z)v.$$

From this identity it follows that, for $m \geq 1$ and for every $p_{m-1} \in \Pi_{m-1}$,

$$y(t) - y_m(t) = f(Z)v - p_{m-1}(Z)v - V_m(f(H_m) - p_{m-1}(H_m))e_1, \tag{2.14}$$

and thus

$$\|y(t) - y_m(t)\| \leq \|(f(Z) - p_{m-1}(Z))v\| + \|V_m\| \, \|(f(H_m) - p_{m-1}(H_m))e_1\|, \tag{2.15}$$

If both Z and H_m are diagonalizable, i. e., $Z = XDX^{-1}, H_m = X_m D_m X_m^{-1}$, with D and D_m diagonal matrices, and if G is a suitable set containing both the spectra of Z and H_m, then from (2.15) we obtain

$$\|y(t) - y_m(t)\| \leq [cond_2(X) + \|V_m\| \, cond_2(X_m)] \max_{\lambda \in \overline{G}} |f(\lambda) - p_{m-1}(\lambda)| \qquad m \geq 1, \tag{2.16}$$

for every $p_{m-1} \in \Pi_{m-1}$. More detailed convergence analysis, based on (2.14), will be presented in the next sections.

3 The Arnoldi RD-method (RA approximation).

As previously mentioned, here below we will deal with two specific kinds of approximations of the type (2.12). We begin by considering that based on the classical Arnoldi algorithm.

Let us set

$$N^* = \max_m(\dim K_m(Z, v)).$$

As well known, the full Arnoldi method generates a sequence of orthonormal vectors $\{v_1, v_2, ..., v_{N^*}\}$ such that, for each $m \leq N^*$,

$$K_m(Z, v) = span\{v_1, v_2, ..., v_m\}.$$

These vectors satisfy (2.9) where the entries of the matrices H_m's are

$$h_{i,j} =< v_i, Zv_j >,$$

and $h_{j+1,j} > 0$, for $j < N^*$. For details on the implementation of the Arnoldi method we refer to [6].

As well known $W(H_m) \subseteq W(Z)$. Moreover, using $V_m^T Z V_m = H_m$ and the fact that $V_m V_m^T$ is an orthogonal projection, it is not difficult to see that $\Re(W(B_m)) \geq \beta$ and thus $\|y_m(t)\| \leq \exp(-t\beta)$.

Concerning the polynomials (2.10), denoting by $\Pi_m^{(0)}$ the set of the monic polynomials of exact degree m, they satisfy (see [42] p. 269) the minimization property:

$$\left\| q_m^{(0)}(Z)v \right\| \leq \|p_m(Z)v\|, \tag{3.1}$$

for all $p_m \in \Pi_m^{(0)}$. Observe also that

$$\left\| q_m^{(0)}(Z)v \right\| = \prod_{j=1}^{m} h_{j+1,j}. \tag{3.2}$$

In light of (3.1), a priori estimates of $\left\|q_m^{(0)}(Z)v\right\|$ can be obtained from well-known results on Chebyshev and Faber polynomials.

As well known (see e.g., [26], [12]), the rate of decay of the $h_{j+1,j}$' s plays a crucial role in the effectiveness of Krylov approximations methods. Various results in the literature relate the behavior of these quantities to the spectral properties of Z. In particular we stress the fact that, when Z represents a suitable discretization of a compact operator whose singular values have a rapid decay, then we may aspect such kind of decay even for the $h_{j+1,j}$' s and thus for the norms (3.2) (see [26], [18]). We recall that for a compact operator in a Hilbert space we have [26]

$$\left\|q_m^{(0)}(Z)v\right\|^{1/m} \to 0. \tag{3.3}$$

Here below, the corresponding approximations (2.12) will be denoted as *Rational Arnoldi* (RA). Observe that now (2.15) and (2.16) hold with $\|V_m\| = 1$. Clearly, in the symmetric case (2.16) reads

$$\|y(t) - y_m(t)\| \leq 2 \max_{\lambda \in J} |f(\lambda) - p_{m-1}(\lambda)|\,, \tag{3.4}$$

for every $p_{m-1} \in \Pi_{m-1}$, being J a closed real interval containing the spectrum of Z.

Here below, for the general case, we will derive some further convergence results.

PROPOSITION 3.1. *Let $W(Z) \subset G \subset S_{\vartheta,r}$, where G be a domain enclosed by a closed rectificable Jordan curve Γ. Assume that*

$$dist(\lambda, W(Z)) \geq d > 0, \tag{3.5}$$

for every $\lambda \in \Gamma$. Then the following error bounds hold:

$$\|y(t) - y_m(t)\| \leq \frac{\epsilon_{m-1}^{(0)}(\Gamma)}{\pi d}, \tag{3.6}$$

$$\|y(t) - y_m(t)\| \leq \frac{h_{m+1,m}}{2\pi d^2} \epsilon_{m-1}^{(0)}(\Gamma), \tag{3.7}$$

where

$$\epsilon_{m-1}^{(0)}(\Gamma) = \min_{p_{m-1} \in \Pi_{m-1}} \max_{\lambda \in \Gamma} \int_\Gamma |f(\lambda) - p_{m-1}(\lambda)|\,|d\lambda|\,. \tag{3.8}$$

PROOF. We can write (2.14) in the Dunford-Taylor integral form:

$$y(t) - y_m(t) = \frac{1}{2\pi i} \int_\Gamma (f(\lambda) - p_{m-1}(\lambda))[(\lambda I - Z)^{-1} v - V_m(\lambda I - H_m)^{-1} e_1] d\lambda. \tag{3.9}$$

Since $W(H_m) \subseteq W(Z)$, recalling the well-know inequality (see [37] Th.4.1)

$$||(\lambda I - Z)^{-1}|| \leq 1/dist(\lambda, W(Z)), \tag{3.10}$$

we obtain (3.6).

Moreover, using (2.9) we have

$$v - (\lambda I - Z)V_m(\lambda I - H_m)^{-1}e_1 = h_{m+1,m}(e_m^T(\lambda I - H_m)^{-1}e_1)v_{m+1},$$

thus from (3.9) we also obtain

$$y(t)-y_m(t) = \frac{h_{m+1,m}}{2\pi i}\int_\Gamma (f(\lambda)-p_{m-1}(\lambda))(e_m^T(\lambda I-H_m)^{-1}e_1)(\lambda I-Z)^{-1}v_{m+1}d\lambda, \tag{3.11}$$

from which by (3.10) we get (3.7). □

The next results allow us to control the convergence even in absence of condition (3.5).

PROPOSITION 3.2. *Let $W(Z) \subset S_{\vartheta,r}$. Let Γ^* be the contour of any sector S_{ϑ^*,r^*} with $\vartheta < \vartheta^* < \frac{\pi}{2}$ and $r^* \geq r(1-\sin(\vartheta^*-\vartheta))^{-1}$, and, for $s = 1, 2$, set*

$$\epsilon_{m-1}^{(s)}(\Gamma^*) = \min_{p_{m-1}\in\Pi_{m-1}} \max_{\lambda\in\Gamma^*} \int_{\Gamma^*} \left|\frac{f(\lambda)-p_{m-1}(\lambda)}{\lambda^s}\right| |d\lambda|\,. \tag{3.12}$$

Then

$$\|y(t) - y_m(t)\| \leq \frac{\epsilon_{m-1}^{(1)}(\Gamma^*)}{\pi\sin(\vartheta^*-\vartheta)}, \tag{3.13}$$

$$\|y(t) - y_m(t)\| \leq \frac{h_{m+1,m}}{2\pi\sin(\vartheta^*-\vartheta)^2}\epsilon_{m-1}^{(2)}(\Gamma^*). \tag{3.14}$$

PROOF. From (3.9) with $\Gamma = \Gamma^*$, which still holds (cf. [9], Chap 7, p. 601), we get

$$\|y(t) - y_m(t)\| \leq \frac{1}{2\pi}\left(\left\|\int_{\Gamma^*} d_{1,m}(\lambda)d\lambda\right\| + \left\|\int_{\Gamma^*} d_{2,m}(\lambda)d\lambda\right\|\right), \tag{3.15}$$

where

$$d_{1,m}(\lambda) = (f(\lambda) - p_{m-1}(\lambda))(\lambda I - Z)^{-1}v,$$

and

$$d_{2,m}(\lambda) = (f(\lambda) - p_{m-1}(\lambda))V_m(\lambda I - H_m)^{-1}e_1.$$

Owing to (3.10), we choose p_{m-1} such that, for $i = 1, 2$,

$$\left\|\int_{\Gamma^*} d_{i,m}(\lambda)d\lambda\right\| \leq \frac{\epsilon_{m-1}^{(1)}(\Gamma^*)\,|\lambda|}{dist(\lambda, W(Z))}. \tag{3.16}$$

If $\lambda \in \Gamma^*$ is of the type $\lambda = |\lambda|\exp(\pm i\vartheta^*)$, then

$$dist(\lambda, W(Z)) \geq |\lambda|\sin(\vartheta^*-\vartheta). \tag{3.17}$$

This holds trivially also for all $\lambda \in \Gamma^*$ with $|\lambda| = r^*$. Therefore (3.13) follows, from (3.15) and (3.16), by easy computations.

In order to prove (3.14), consider the identity (3.11), with $\Gamma = \Gamma^*$. Now, choose p_{m-1} such that

$$\left\| \int_{\Gamma^*} (f(\lambda) - p_{m-1}(\lambda))(e_m^T(\lambda I - H_m)^{-1}e_1)(\lambda I - Z)^{-1}v_{m+1}d\lambda \right\| \leq \frac{\epsilon_{m-1}^{(2)}(\Gamma^*)\left|\lambda^2\right|}{dist(\lambda, W(Z))^2}, \tag{3.18}$$

for $\lambda \in \Gamma^*$. Arguing as before we get the thesis. □

The results stated in the last Proposition, as well as (3.4), have an important meaning, in particular (3.14) which relates the behavior of the errors to the decay of the elements $h_{m+1,m}$'s, whose importance we have already pointed out. Owing to them, if $Z \in F = \{Z_N\}$, being F a family of matrices, of size possibly increasing with N, all approximating a compact operator having the numerical range inside a sector $S_{\vartheta,r}$, then we can ensure that the rate of convergence of the approximations cannot arbitrarily deteriorate as N increases.

By the arguments used above we can also state error bounds involving the norms $\left\|q_m^{(0)}(Z)v\right\|$.

PROPOSITION 3.3. *Under the assumptions of Proposition 3.1 we have*

$$\|y(t) - y_m(t)\| \leq \frac{\left\|q_m^{(0)}(Z)v\right\|}{2\pi d} \int_{\Gamma^*} \frac{|f(\lambda)|}{\left|q_m^{(0)}(\lambda)\right|} |d\lambda|, \tag{3.19}$$

and, under the assumptions of Proposition 3.2, we have

$$\|y(t) - y_m(t)\| \leq \frac{\left\|q_m^{(0)}(Z)v\right\|}{2\pi \sin(\vartheta^* - \vartheta)} \int_{\Gamma^*} \frac{|f(\lambda)|}{|\lambda|\left|q_m^{(0)}(\lambda)\right|} |d\lambda|. \tag{3.20}$$

PROOF. By the fact that H_m is a non-derogatory Hessenberg matrix, using the Cayley-Hamilton theorem, one proves (cf. [25]) that

$$\left\|h_{m+1,m}(e_m^T(\lambda I - H_m)^{-1}e_1)v_{m+1}\right\| = \frac{\left\|q_m^{(0)}(Z)v\right\|}{\left|q_m^{(0)}(\lambda)\right|}, \lambda \in \Gamma(\Gamma^*). \tag{3.21}$$

Introducing this in (3.11) and using (3.5) we get (3.19) and (3.20). □

By (3.2), (3.19) and (3.21) provides also applicable error bounds, available without the knowledge of $y_m(t)$. In light of our numerical experiences, such bounds turned out to match with the errors in a satisfactory way, for the values of m of interest. Let us consider in particular (3.20), from which, using (3.17), we get the bound

$$\|y(t) - y_m(t)\| \leq \frac{\left\|q_m^{(0)}(Z)v\right\|}{2\pi[\sin(\vartheta^* - \vartheta)]^{m+1}} \int_{\Gamma^*} \frac{|f(\lambda)|}{|\lambda|^{m+1}} |d\lambda|, \tag{3.22}$$

which does not require the polynomial $q_m^{(0)}$. The integrals $\int_{\Gamma^*} \frac{|f(\lambda)|}{|\lambda|^{m+1}} |d\lambda|$ can be computed employing well-known formulae ([1], p. 228). The following corollary gives a result in this direction.

COROLLARY 3.4. *Under the assumptions above, we have*

$$\|y(t) - y_m(t)\| \leq C_m \left\| q_m^{(0)}(Z)v \right\|, \tag{3.23}$$

where

$$C_m = \frac{\exp(\tau)(m-1)!}{\pi[\sin(\frac{\pi-2\vartheta}{4})]^{m+1}(\tau\cos(\frac{\pi+2\vartheta}{4}))^m}.$$

PROOF. Let $r^* \to \infty$. Using the Euler's integral ([1], pp. 255), by simple computation we find

$$\int_0^\infty \frac{\exp(-\frac{\tau}{x}\cos\vartheta^*)}{x^{m+1}} dx = \frac{(m-1)!}{(\tau\cos\vartheta^*)^m}.$$

Using this for evaluating the integral in (3.22) (for $\lambda = |\lambda|\exp(\pm i\vartheta^*)$ and taking $\vartheta^* = \frac{\pi+2\vartheta}{4}$ we obtain (3.23). □

In the self-adjoint case (3.23) gives

$$\|y(t) - y_m(t)\| \leq \frac{\exp(\tau)2^{m+1}(m-1)! \left\| q_m^{(0)}(Z)v \right\|}{\sqrt{2}\pi\tau^m}. \tag{3.24}$$

The growing factor C_m in (3.23) can be controlled by τ. The optimal choice in this sense is $\tau = m$. This is in accordance with existing results concerning RD-rational approximations of the exponential function. See the discussion in [43], where other applicable error estimates are proposed. Since we must avoid to change h as m increases, in practice, having to carry out the integration over a time-window, it is reasonable to fix the value of h, taking into account both of the conditioning of $(I + hA)$ and, if possible, of the decay of $\left\| q_m^{(0)}(Z)v \right\|$, and then to select suitably the time-steps. Dealing with matrices arising from elliptic operators, a good strategy, suggested by the mesh-independence, may consist in testing the convergence on matrices of lower dimension.

4 The Faber RD-method (RF approximation).

As an alternative to the Arnoldi method, we discuss here below the use of Faber polynomials associated to a suitable set in the complex plane. Faber polynomials represent a very useful tool in approximation theory and their application in the construction of Krylov subspaces has been already considered in the literature. For instance, see [38], [10], [25] and the references therein.

Referring to (2.4) let us consider a convex set $\Omega \subset \Sigma_{\vartheta,0} \cap D_\rho$, for some $0 < \rho \leq \frac{1}{2}$, with Ω containing the spectrum of Z. So we also assume that Ω is symmetric with respect to the real axis. We notice that the theory presented below holds, with some minor changes, also under less restrictive assumptions on Ω (see [36], [38]).

Let γ be the *logarithmic capacity* (cf. [5], p. 332) of Ω and let the function

$$\psi(w) = \gamma w + c_0 + c_1 w^{-1} + c_2 w^{-2} + ... \tag{4.1}$$

map one-to-one conformally the domain $|w| \geq 1$ onto $\overline{\mathbb{C}} \backslash \Omega$. In some simple but important cases, as for instance ellipses and thus circles and segments, or even lemniscates, the development (4.1) can be easily constructed (see [24], [38], [10]) and the logarithmic capacity γ easily evaluated. In fact, if the boundary of Ω is an ellipse having axes with lengths $2a$ and $2b$, then $\gamma = \frac{a+b}{2}$ and $c_j = 0$ for $j \geq 2$. For sets having a less simple shape we can resort to some available numerical techniques (see [38], [41]). Anyhow we notice that in our case it is $\gamma < \rho$ (cf. [5], p. 334).

Referring to (4.1), let us consider the sequence of polynomials $\{F_m(z)\}_{m \geq 0}$ defined by

$$\begin{gathered} F_0(z) = 1, \qquad \gamma F_1(z) = z - c_0, \\ \gamma F_m(z) = (z - c_0) F_{m-1}(z) - (c_1 F_{m-2}(z) + ... + c_{m-1} F_0(z)) - (m-1) c_{m-1}. \end{gathered} \tag{4.2}$$

for $m \geq 2$. These are the "ordinary" Faber polynomials associated to Ω.

If the boundary of Ω is an ellipse or it is an interval (in the complex plane), the associated Faber polynomials are scaled and translated Chebyshev polynomials, whose formulae can be found in [38] and [10].

Under our assumptions on Ω, for $|w| \geq 1$, we have (cf. [38])

$$|w|^m - 1 \leq |F_m(\psi(w))| \leq 2 |w|^m . \tag{4.3}$$

Now, for our purposes, let us define

$$v_j = F_j(z)\, v.$$

By (4.2) we have that (2.9) holds with

$$h_{j,j} = c_0, \quad h_{j+1,j} = \gamma, \quad h_{1,j} = j c_{j-1}, \quad for \quad j > 1, \tag{4.4}$$

$$h_{i,j} = c_{j-i}, \quad for \quad i \geq 2, j \geq 3, j > i. \tag{4.5}$$

Thus, in light of (2.10) and (2.11), it is

$$q_m^{(0)}(z) = \gamma^m F_m(z),$$

and the eigenvalues of H_m are now the zeros of $F_m(z)$, which (cf. [23]) are contained in Ω. General results upon interpolation on such points can be found in [39]. The approximation method (2.12) will be denoted as *Rational Faber* (RF).

Here below we state some convergence results for the RF approximations. We refer to the notation used in Propositions 3.1 and 3.2. Since, in general, now we cannot ensure that $W(H_m) \subseteq W(Z)$, then it is convenient to work by means of Kreiss constants. In alternative one can also consider $\varepsilon-$pseudospectra. For

analyses on the connections between Kreiss constants and Faber polynomials we refer to [40].

LEMMA 4.1. *(see [40], Th. 1.1) For every $m \geq 0$ it is*

$$\|F_m(Z)\| \leq 2(m+1)\exp(1)\widetilde{K}(\Omega), \tag{4.6}$$

where $\widetilde{K}(\Omega)$ is the Kreiss constant *defined by*

$$\widetilde{K}(\Omega) = \inf\left\{C : ||(zI - Z)^{-1}|| \leq \frac{C}{dist(z,\Omega)}, z \notin \Omega\right\}. \tag{4.7}$$

LEMMA 4.2. *([40], Th. 3.1) For any $|w| > 1$ we have*

$$\frac{1}{2}(|w| - 1) \leq \frac{dist(\psi(w), \Omega)}{|\psi'(w)|} \leq 2(|w| - 1).$$

LEMMA 4.3. *For every $\lambda \notin \Omega$ it is*

$$dist(\lambda, \Omega) \leq \frac{4\gamma}{m} |F_m(\lambda)|. \tag{4.8}$$

PROOF. Clearly any $\lambda \notin \Omega$ can be written as $\lambda = \psi(w)$, for some w with $|w| > 1$. Since $\left|\psi'(w)\right| \leq 2\gamma$, for $|w| > 1$, (cf. [23], p. 195), by Lemma 4.2 we have

$$dist(\lambda, \Omega) \leq 4\gamma(|w| - 1).$$

Hence, by (4.3),

$$m \cdot dist(\lambda, \Omega) \leq 4\gamma(|w|^m - 1) \leq 4\gamma |F_m(\lambda)|.$$

□

PROPOSITION 4.4. *Let $\Omega \subset G \subset S_{\vartheta,r}$, where G is a domain enclosed by a rectificable Jordan curve Γ. Assume that*

$$dist(\lambda, \Omega) \geq d > 0, \quad \lambda \in \Gamma.$$

Then, referring to (3.8), for the RF-approximation we have

$$\|y(t) - y_m(t)\| \leq \frac{4(m+1)\gamma\exp(1)\widetilde{K}(\Omega)^2}{m\pi d^2}\epsilon_{m-1}^{(0)}(\Gamma), \quad m \geq 1. \tag{4.9}$$

PROOF. Since

$$h_{m+1,m}(e_m^T(\lambda I - H_m)^{-1}e_1)v_{m+1} = \frac{F_m(Z)v}{F_m(\lambda)},$$

from (3.11) we get

$$\|y(t) - y_m(t)\| = \frac{1}{2\pi}\left\|\left(\int_\Gamma \frac{f(\lambda) - p_{m-1}(\lambda)}{F_m(\lambda)}(\lambda I - Z)^{-1}d\lambda\right)F_m(Z)v\right\|, \tag{4.10}$$

for every $p_{m-1} \in \Pi_{m-1}$. From this, by (4.6), (4.7) and (4.8) we get (4.9). ∎

Furthermore, also results similar to those given in Proposition 3.2 can be derived, as shown here below.

PROPOSITION 4.5. *Under the assumptions of Proposition 3.2, for the RF-approximation we have*

$$\|y(t) - y_m(t)\| \leq \frac{4\gamma(m+1)\exp(1)\widetilde{K}(\Omega)^2}{m\pi\sin(\vartheta^* - \vartheta)^2}\epsilon_{m-1}^{(2)}(\Gamma^*). \tag{4.11}$$

PROOF. Take the sector S_{ϑ^*,r^*} as in Proposition 3.2. Reconsider (4.10), from which, by (4.6), (4.7) and (4.8) we get

$$\|y(t) - y_m(t)\| \leq \frac{4\gamma\exp(1)(m+1)\widetilde{K}(\Omega)^2}{m\pi}\int_{\Gamma^*}\left|\frac{f(\lambda) - p_{m-1}(\lambda)}{dist(\lambda,\Omega)^2}\right||d\lambda|\,,$$

for every $p_{m-1} \in \Pi_{m-1}$.

Arguing as in the proof of Proposition 3.2, we obtain

$$\int_{\Gamma^*}\left|\frac{f(\lambda) - p_{m-1}(\lambda)}{dist(\lambda,\Omega)^2}\right||d\lambda| \leq \frac{\epsilon_{m-1}^{(2)}(\Gamma^*)}{\sin(\vartheta^* - \vartheta)^2}.$$

∎

In absence of information on the spectrum of Z, algorithms based on Faber polynomials need a preliminary phase where estimates of the eigenvalues of Z are achieved, in order to construct in a suitable way the set Ω. For this purpose various techniques have been proposed, mainly in the context of hybrid methods for linear systems. For a discussion on this point and for related references see [38], [25].

5 Numerical experiments.

For our numerical experiments we consider classical test-matrices obtained by discretization of differential operators of the type

$$-\Delta + \tau_1\frac{\partial}{\partial x} + \tau_2\frac{\partial}{\partial y}, \qquad \tau_1, \tau_2 \in \mathbb{R}, \tag{5.1}$$

where Δ denotes the 2-dimensional Laplacian. Dirichlet boundary conditions on the square $(0,1)\times(0,1)$ are considered. We discretize using central differences with uniform meshsize $\delta = 1/(n+1)$ along each direction. Thus, we get a matrix A of order $N = n^2$ which we represent by means of a sum of Kronecker products as follows,

$$A := \frac{-1}{\delta^2}\left\{I_n \otimes C_1 + C_2 \otimes I_n\right\},$$

where I_n is the identity matrix of order n, and

$$C_i := \begin{bmatrix} -2 & 1-\tau_i\frac{\delta}{2} & & \\ 1+\tau_i\frac{\delta}{2} & -2 & 1-\tau_i\frac{\delta}{2} & \\ & 1+\tau_i\frac{\delta}{2} & \ddots & \ddots \\ & & \ddots & \ddots \end{bmatrix} \in \mathbb{R}^{n\times n}, \qquad i = 1, 2.$$

The spatial mesh-size δ will be always taken in such a way that the Péclet numbers $\tau_i \frac{\delta}{2}$ are less then one, so that centered differences yield stable discretizations. Numerical tests on polynomial methods on these problems, as well as interesting discussions on various related computational aspects, can be found in the recent papers [2], [3], [4]. In our examples we take

$$v = (1, 1, ..., 1)^T / n.$$

As previously mentioned, the Faber method needs in general a preliminary phase where estimates of the eigenvalues are achieved, in order to construct the set Ω. For our model problem we can avoid this, taking into consideration the known fact that the spectrum of A is in the rectangle

$$R_n = -(n+1)^2 \left\{ [4 - 2\Re\upsilon_n, 4 + 2\Re\upsilon_n] \times [-2i\Im\upsilon_n, 2i\Im\upsilon_n] \right\},$$

with

$$\upsilon_n := \cos\left(\frac{\pi}{n+1}\right)\left(\sqrt{1 - \frac{\tau_1^2}{4(n+1)^2}} + \sqrt{1 - \frac{\tau_2^2}{4(n+1)^2}}\right).$$

Accordingly, in our experiments with Faber method, we use this information for constructing the required set Ω. More precisely, since for the values of τ_1 and τ_2 here considered, all the eigenvalues of A are real, then in the RF method we actually use Chebyshev polynomials.

In our numerical experiments a matrix factorization (LU or LL^T) of $(I + hA)$ is computed, once and for all. It can be used in every application of the methods, i.e., for different values of t and, overall, for different v. Thus the weight of the cost of this initial phase will be diluted as long as repeated applications will be carried on. In our numerical experiences we take h such that $\tau = \frac{t}{h}$ varies in the range $[1, 10]$.

The tables illustrating the numerical experiments are organized in four groups, depending on the different features of the methods we are interested to point out. Comparison will be made in particular with the classical polynomial method which interpolates the exponential in the Arnoldi-Ritz values associated to the Krylov subspaces $K_m(A, v)$, denoted here as the PA (Polynomial Arnoldi) method.

Group 1. By Tables 1 and 2, here below, we want to stress the mesh-independence of the RD- rational approximations (RA and RF) with respect to the mesh-dependent behavior of the PA method. For our purpose we limit here the attention to the iteration number m. In Tables 1 and 2, for increasing values of n, we report the corresponding values of m required for getting an error-norm $\leq 10^{-6}$. In the RD-methods we take, for every n, $h = 0.01$. The results here below refer to the case $t = h$.

n	PA	RA	RF
20	19	11	19
30	27	11	20
40	36	12	20
50	45	12	21
60	53	12	22
70	62	12	22
80	70	12	22

Table 1: $\tau_1 = \tau_2 = 0$.

n	PA	RA	RF
20	22	17	23
30	32	17	23
40	42	18	23
50	52	19	23
60	62	19	23
70	72	19	26
80	82	19	26

Table 2: $\tau_1 = 10$, $\tau_2 = 5$.

Group 2. In figures 5.1 and 5.2 we report a comparison between the errors (solid lines) and the error estimates (dotted lines) given by (3.23) and (3.24) for the RA method.

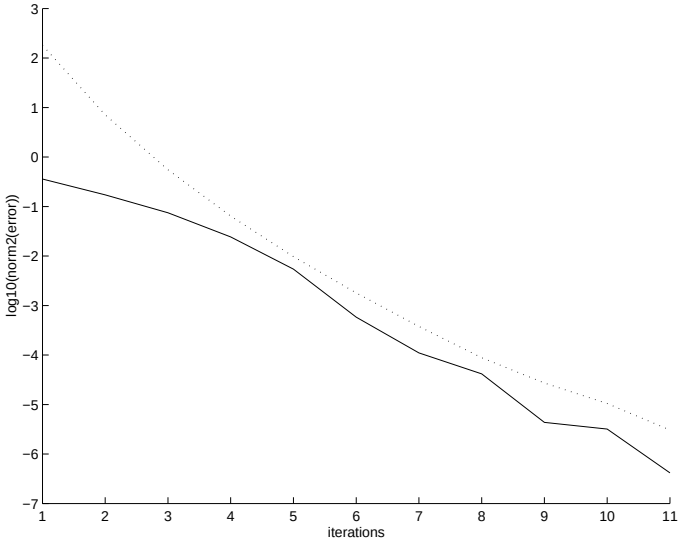

Figure 5.1: $n = 30, \tau_1 = \tau_2 = 0, t = 0.01, h = 0.001$

Group 3. The third group of numerical experiments compares our RD-methods (for simplicity, here we consider only the method RA) with other types of RD-rational approximations and with a classical Rosenbrock stiff solver.

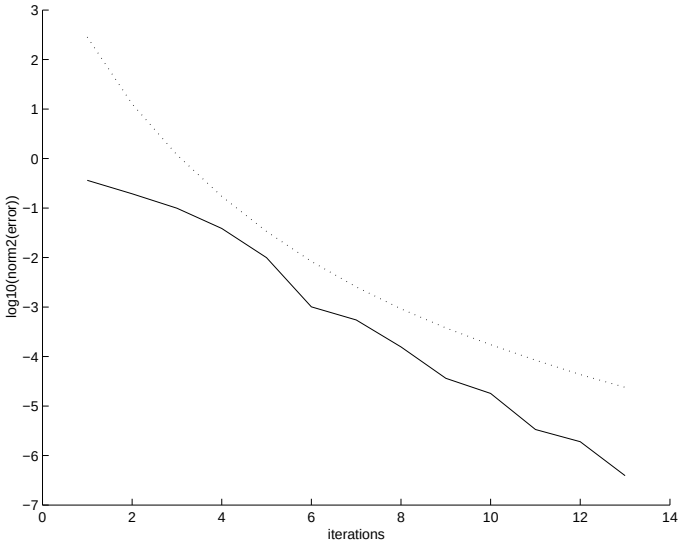

Figure 5.2: $n = 30, \tau_1 = 10, \tau_2 = 5, t = 0.01, h = 0.001$

In Tables 3 and 4 we consider the RA approximations of degree 2 (RA(2)) and 3 (RA(3)) together with some *restricted-Padè approximations* proposed in [27], example 2.1. These are RD(j, k) formulae of order $j + 1$ and in the tables below we denote them by S_k^j. The tables contain the errors for various values of t. In the RA methods we take always $h = t$.

t	RA(2)	RA(3)	S_2^2	S_3^2	S_3^3	S_4^3
0.01	0.031	0.013	0.101	0.058	0.086	0.052
0.02	0.040	0.005	0.156	0.070	0.133	0.064
0.05	0.028	0.001	0.240	0.088	0.206	0.077
0.1	0.008	0.0005	0.308	0.130	0.264	0.119

Table 3: $n = 20$; $\tau_1 = \tau_2 = 0$.

t	RA(2)	RA(3)	S_2^2	S_3^2	S_3^3	S_4^3
0.01	0.060	0.053	0.120	0.098	0.103	0.095
0.02	0.111	0.033	0.195	0.140	0.170	0.134
0.05	0.032	0.011	0.306	0.205	0.267	0.193
0.1	0.001	0.0002	0.406	0.132	0.349	0.125

Table 4: $n = 20$; $\tau_1 = 20$, $\tau_2 = 10$.

In order to make a comparison with a standard stiff solver of common use, even if not the best for such problems, we consider the Rosenbrock method implemented by the Matlab function ODE23S. It requires the solution of, at least, a linear system at each step. Due to the fact that this method uses a step-size control strategy, then the coefficient matrices of the linear systems to be solved could be different one from the other. Table 5 reports the results, in the symmetric case $\tau_1 = \tau_2 = 0$, for three values of t. Regarding the parameters

AbsTol and *RelTol* of the step-size control in the Rosenbrock method (see [35]), we define $AbsTol = RelTol$. The number of steps is denoted by *nstepROS* (*errorROS* is the corresponding error). The terms *nstepRA* (value of m) and *errorRA* refer to to the RA method with $h = t$.

t	*AbsTol*	*nstepROS*	*errorROS*	*nstepRA*	*errorRA*
0.005	1E-3	10	1.28E-4	7	8.11E-5
	1E-5	19	6.39E-5		
0.01	1E-3	11	1.45E-4	7	5.98E-5
	1E-5	24	7.87E-5		
0.02	1E-3	13	1.67E-4	7	5.52E-6
	1E-5	30	9.00E-5		

Table 5: $n = 20$; $\tau_1 = \tau_2 = 0, h = t$.

Group 4. In the following four pictures we take into account of the whole computational costs, measured in terms of scalar products. The comparison is made with the Arnoldi polynomial method PA. The pictures below show the behavior of $\log_{10}$. The cost of the initial matrix factorization is not considered.

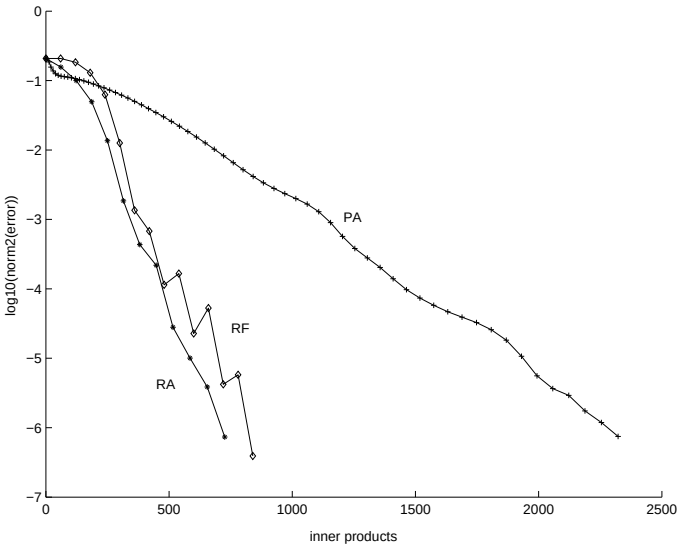

Figure 5.3: $n = 30, \tau_1 = 10, \tau_2 = 5, t = 0.05, h = 0.01$

REFERENCES

1. M. Abramovitz, A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, 1965.
2. L. Bergamaschi, M. Vianello, *Efficient computation of the exponential operator for large, sparse, symmetric matrices*, Numer. Linear Algebra Appl., 7 (2000), pp. 27–45 .

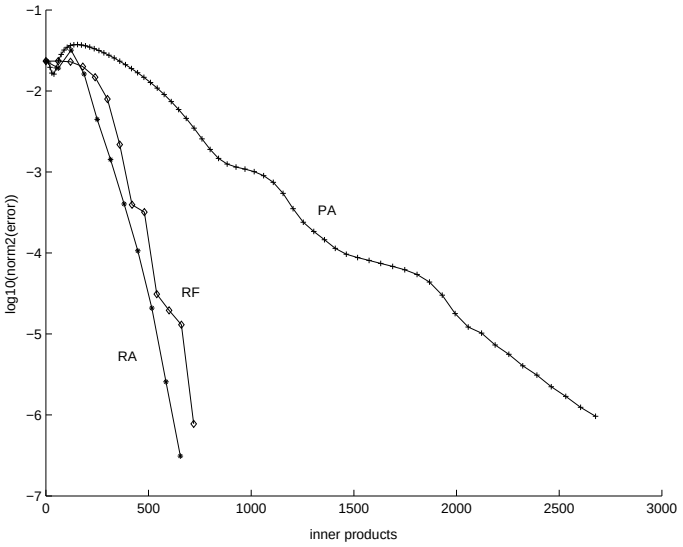

Figure 5.4: $n = 30, \tau_1 = 10, \tau_2 = 5, t = 0.1, h = 0.01$

3. L. Bergamaschi, M. Caliari, M. Vianello, *Efficient approximation of the exponential operator for discrete 2D advection-diffusion problems*, Numer. Linear Algebra Appl., 10 (2003), pp. 271–289.

4. M. Caliari, M. Vianello, L. Bergamaschi, *Interpolating discrete advection-diffusion propagators at spectral Leja sequence*. J. Comput. Appl. Math., in press, 2004.

5. J. B. Conway, *Functions of One Complex Variable II* Springer, Berlin, 1995.

6. J. Drkosova', A. Greenbaum, M. Rozloznik, Z. Strakos, *Numerical stability of GMRES*, BIT, 35 (1995), pp. 309–330.

7. V. Druskin, A. Greenbaum, L. Knizhnerman, *Using nonorthogonal Lanczos vectors in the computation of matrix functions*, SIAM J. Sc. Comput., 19 (1998), pp. 38–54.

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

9. N. Dunford, J. T. Schwartz, *Linear Operators, Part I*, John Wiley& Sons, New York, 1963.

10. M. Eiermann, *On semiiterative methods generated by Faber polynomials*, Numer. Math., 56 (1989), pp. 139–156 .

11. M. Eiermann, *Fields of values and iterative methods*, Linear Algebra Appl., 180 (1993) pp. 167–197.

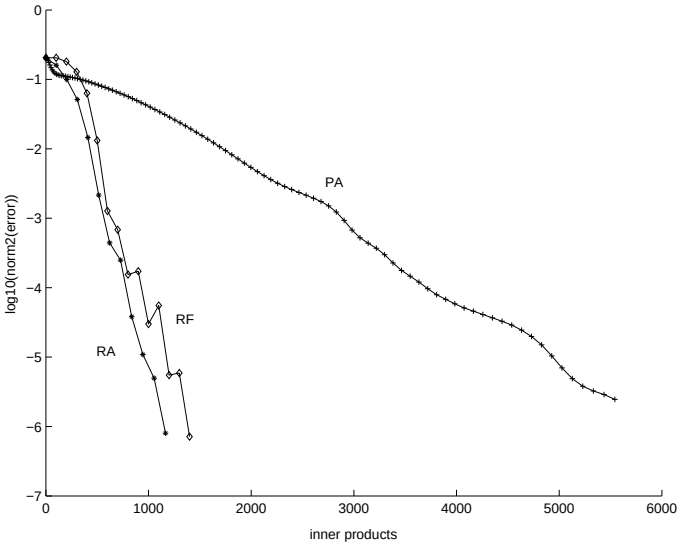

Figure 5.5: $n = 50, \tau_1 = 10, \tau_2 = 5, t = 0.05, h = 0.01$

12. M. Eiermann, O. G. Ernst, *Geometric aspects of the theory of Krylov subspace methods*, Acta Numerica, 10 (2001), pp. 251–312.

13. R.W. Freund, *Krylov-subspace methods for reduced-order modeling in circuit simulation*, J. Comput. Appl. Math., 123 (2000), pp. 395–421.

14. E. Gallopoulos, Y. Saad, *Efficient solution of parabolic equations by Krylov approximation methods*, SIAM Sci. Stat. Comput., 13 (1992), pp. 1236–1264.

15. E. Hairer, G. Wanner, *Solving Ordinary Differential Equations II* Springer, Berlin, 1991.

16. M. Hochbruck, C. Lubich, *On Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 34 (1997), pp. 1911–1925.

17. M. Hochbruck, C. Lubich, H. Selhofer, *Exponential integrators for large systems of differential equations*, SIAM J. Sci. Comput., 19 (1998), pp. 1552–1574.

18. Hyvönen, O. Nevanlinna, *Robust bounds for Krylov methods*, BIT, 40 (2000), pp. 267–290.

19. A. Iserles, *Rational interpolation to exp(-x) with application to certain stiff problems*, SIAM J. Numer. Anal., 18 (1981), pp. 1–12 .

20. A. Iserles, S. P. Norsett, *Order Stars* Chapman&Hall, London, 1991).

21. T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin, 1976)

22. L. Knizhnerman, *Calculation of functions of unsymmetric matrices using Arnoldi's method*, U.S.S.R. Comput. Maths. Math. Phys., 31 (1991), pp. 1–9.

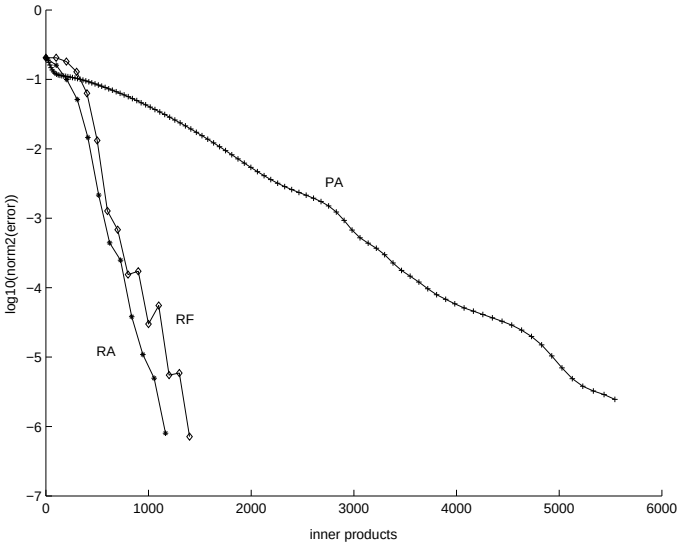

Figure 5.6: $n = 50, \tau_1 = 10, \tau_2 = 5, t = 0.1, h = 0.01$

23. T. Kovari, C. Pomerenke, *On Faber polynomials and Faber expansions*, Math. Z., 99 (1967), pp. 193–206.

24. A. I. Markushevich, *Theory of Functions of a Complex Variable, vol. III*, Prentice-Hall, Englewood Cliffs, N. J., 1967.

25. I. Moret, P. Novati, *An interpolatory approximation of the matrix exponential based on Faber polynomials*, Journal C.A.M., 131 (2001), pp. 361–380.

26. O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, 1993.

27. S.P. Norsett, *Restricted Padé approximations to the exponential function*, SIAM J. Numer. Anal., 15 (1978), pp. 1008–1029.

28. S. P. Norsett, S.R.Trickett, *Exponential fitting of restricted rational approximations to the exponential function*, in Rational Approximation and Interpolation, 466-476, Graves -Morris, Saff and Varga, Eds., Lecture Notes in Mathematics, 1105, Springer, Berlin, 1984.

29. P. Novati, *Solving linear initial value problems by Faber polynomials*, Numer. Linear Algebra Appl., 10 (2003), pp. 247–270.

30. P. Novati, *A polynomial method based on Fejer points for the computation of functions of unsymmetric matrices*, Appl. Numer. Math., 44 (2003), pp. 201–224.

31. G. Pini,G. Gambolati, *Arnoldi and Crank-Nicholson methods for integration in time of transport equation*, Int. J. Numer. Meth. Fluids, 35 (2001), pp. 25–38.

32. T. A. Porsching, M. Lin Lee, *The reduced basis method for initial value problems*, SIAM J. Numer. Anal., 24 (1987), pp. 1277–1287.

33. A. Ruhe, *Rational Krylov algorithms for nonsymmetric eigenvalue problems*, in Recent advances in iterative methods, G. Golub, A. Greenbaum and M. Luskin, eds., IMA Vols in Mathematics and its Applications, 60, 149-164, Springer, Berlin, 1993.

34. Y. Saad, *Analysis of some Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 29 (1992), pp. 209–228.

35. L. F. Shampine, M. W. Reichelt, *The Matlab ODE suite*, SIAM J. Sci. Comput., 18 (1997), pp. 1–22.

36. V. I. Smirnov, N. A. Lebedev, *Functions of a Complex Variable-Constructive Theory*, Iliffe Books, London, 1968.

37. M. N. Spijker, *Numerical ranges and stability estimates*, Appl. Numer. Math., 13 (1993), pp. 241–249.

38. G. Starke, R. S. Varga, *A hybrid Arnoldi-Faber iterative method for nonsymmetric systems of linear equations*, Numer. Math., 64 (1993), pp. 213–240.

39. H. Tal-Ezer, *High degree polynomial interpolation in Newton form*, SIAM J. Sci. Stat. Comput., 12 (1991), pp. 648–667.

40. K. Toh, L. N. Trefethen, *The Kreiss matrix theorem on a general domain*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 145–165.

41. L. N. Trefethen, *Numerical computation of the Schwarz-Christoffel transformation*, SIAM J. Sci. Statist. Comput., 1 (1980), pp. 82–102.

42. L. N. Trefethen, D. Bau , Numerical Linear Algebra (SIAM, Philadelphia, 1997).

43. J. Van Den Eshof, M. Hochbruck, *Preconditioning Lanczos approximations to the matrix exponential*. Preprint (2004).

44. Yu. V. Vorobyev, *Method of Moments in Applied Mathematics*, Gordon and Breach, New York, 1965.