ON THE CONVERGENCE OF KRYLOV SUBSPACE METHODS FOR MATRIX MITTAG-LEFFLER FUNCTIONS

IGOR MORET* AND PAOLO NOVATI†

Abstract. In this paper we analyze the convergence of some commonly used Krylov subspace methods for computing the action of matrix Mittag-Leffler functions. As it is well known, such functions find application in the solution of fractional differential equations. We illustrate the theoretical results by some numerical experiments.

1. Introduction. Krylov subspace methods represent nowadays a standard approach for approximating the action of a function of a large matrix on a vector, namely $y = f(A)v$. For a general discussion on the computation of matrix functions we refer to the book [16]. The convergence properties of Krylov methods have been widely studied in literature. Among the more recent papers, we cite [21], [13], [6], [2], [19], [10]. A particular attention has been devoted to the exponential and related functions involved in the solution of differential problems. Such functions belong to the large class of entire functions which take their name from Gösta Mittag-Leffler. A generalized Mittag-Leffler (ML) function is formally defined in correspondence of two given parameters $\alpha, \beta \in \mathbb{C}$, $\operatorname{Re}\alpha > 0$, by the series expansion

$$E_{\alpha,\beta}(z) = \sum\nolimits_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}, \quad z \in \mathbb{C}, \tag{1.1}$$

where Γ denotes the gamma function. The exponential-like functions, involved in certain modern integrators for evolution problems, correspond to the case of $\alpha = 1$ and $\beta = k = 1, 2, ...$, that is, $E_{1,1}(z) = \exp(z)$ and

$$E_{1,k}(z) = \frac{1}{z^{k-1}} \left(\exp(z) - \sum\nolimits_{j=0}^{k-2} \frac{z^j}{j!} \right), \quad \text{for } k \geq 2.$$

We have to point out that, till few years ago, even the computation of a ML function in the scalar case was a difficult task. Only recently efficient algorithms have been developed [28], [30] and nowadays we are able to treat the matrix case at a reasonable cost, at least for small matrices. This clearly suggests the use of Krylov projection methods for the treatment of larger ones. To the best of our knowledge, till now in literature there are not results concerning the convergence of such methods for generalized ML functions. Our study wants to give a contribute in order to fill this gap. Precisely we consider here two methods. The first one is the standard method (SKM) which seeks for approximations in the Krylov subspaces generated by A and v. The second one is the one-pole rational method, here denoted as RAM, sometimes named SI-method, since it works on the Krylov subspaces of a suitable shifted and inverted matrix. For the matrix exponential, it was introduced in [24] and [11]. Further results and applications can be found in [22], [26], [29], [15], [23], [20].

Our investigations are mainly motivated by the fact that the Mittag-Leffler functions are related to the solution of fractional differential equations (FDEs) arising in fractional generalizations of several mathematical models (see e.g. [27] for many

*Department of Mathematics and Informatics, University of Trieste, Via Valerio 12/1, 34127 Trieste, Italy (moret@units.it)
†Department of Pure and Applied Mathematics, Unversity of Padova, Via Trieste 63, 35121, Padova, Italy (novati@math.unipd.it)

examples). Considering real values of the parameters α and β, our analysis will try to emphasize the features of the examined methods in view of their application to such problems, with a particular attention to the case $0 < \alpha < 1$, which is that most frequently occurring in this context.

The paper is organized as follows. In section 2 we recall the basic properties of the ML functions, pointing out their connections with fractional differential problems. In section 3 and 4 we study the convergence of the SKM and RAM respectively. In section 5 we present some numerical experiments.

2. Mittag-Leffler functions and Fractional Differential Equations (FDEs). Here we outline some basic properties of the ML functions. During the paper $\alpha,\ \beta \in \mathbb{R}$. At first we notice that a ML function possesses also some integral representations. For our purposes we consider the following one. For any $\varepsilon > 0$ and $0 < \mu < \pi$, let us denote by

$$C(\varepsilon, \mu) = C_1(\varepsilon, \mu) \cup C_2(\varepsilon, \mu)$$

the contour in the complex plane where

$$C_1(\varepsilon, \mu) = \{\lambda : \lambda = \varepsilon \exp(i\varphi),\ \text{for} - \mu \leq \varphi \leq \mu\}$$

and

$$C_2(\varepsilon, \mu) = \{\lambda : \lambda = r \exp(\pm i\mu),\ \text{for } r \geq \varepsilon\}.$$

The contour $C(\varepsilon, \mu)$ divides the complex plane into two domains, denoted by $G^-(\varepsilon, \mu)$ and $G^+(\varepsilon, \mu)$, lying respectively on the left and on the right of $C(\varepsilon, \mu)$. Accordingly the following integral representation for $E_{\alpha,\beta}(z)$ holds (cf. [27] p. 30).

LEMMA 2.1. *Let $0 < \alpha < 2$ and β be an arbitrary complex number. Then for every $\varepsilon > 0$ and μ such that*

$$\frac{\alpha\pi}{2} < \mu \leq \min[\pi, \alpha\pi],$$

we have

$$E_{\alpha,\beta}(z) = \frac{1}{2\alpha\pi i} \int_{C(\varepsilon,\mu)} \frac{\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}}{\lambda - z} d\lambda, \text{ for } z \in G^-(\varepsilon, \mu), \tag{2.1}$$

Given a square matrix A we can define its ML function by (1.1). If the spectrum of $-A$ lies into the set $G^-(\varepsilon, \mu)$, for some ε and μ as in the previous lemma, then by (2.1) and by the Dunford-Taylor integral representation of a matrix function, we have

$$E_{\alpha,\beta}(-A) = \frac{1}{2\alpha\pi i} \int_{C(\varepsilon,\mu)} \exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}} (\lambda I + A)^{-1} d\lambda. \tag{2.2}$$

By means of ML functions we can represent the solution of several differential problems. In particular they are a basic tool dealing with fractional differential problems. In order to emphasize the importance of the ML functions in this context, we briefly outline some facts on fractional calculus.

The Caputo's fractional derivative of order $\alpha > 0$ of a function f with respect to the point t_0 is defined by [3]

$$_{t_0}D_t^{\alpha} f(t) = \frac{1}{\Gamma(q - \alpha)} \int_{t_0}^{t} f^{(q)}(u)(t - u)^{q-\alpha-1} du,$$

where q is the integer such that $q-1<\alpha<q$. Similarly to what happens for the Grünwald-Letnikov and the Riemann-Liouville definitions of fractional derivative, under natural conditions on the function f, for $\alpha \to q$ the Caputo's derivative becomes the conventional q-th derivative (see [27] p. 79) so that it provides an interpolation between integer order derivatives. As well known, the main advantage of Caputo's approach is that initial conditions for FDEs takes the same form of the integer order case.

A peculiar property which distinguishes the fractional derivative from the integer one is that it is not a local operator, that is the value of ${}_{t_0}D_t^{\alpha} f(t)$ depends on all the values of f in the interval $[t_0, t]$. This memory property allows to model various physical phenomena very well, but, on the other hand, it increases the complexity in the numerical treatment of the related differential problems, compared to the integer-order case. We refer to [7], [8], [14], [27] for discussions and references on numerical methods for FDEs.

As a simple model problem, we consider here a linear FDE of the type

$$\begin{gathered} {}_0D_t^{\alpha} y(t) + Ay(t) = g(t),\ t>0, \\ y(0) = y_0, \end{gathered}$$

where y_0 is a given vector, $g(t)$ is a suitable vector function and $0<\alpha<1$. The solution of this problem ([27] p. 140) is given by

$$y(t) = E_{\alpha,1}(-t^{\alpha}A)y_0 + \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-A(t-s)^{\alpha}) g(s) ds.$$

The above formula can be viewed as the generalization of the variation-of-constants formula to the non-integer order case.

By means of the following general formula ([27] p.25) concerning the integration of ML functions

$$\frac{1}{\Gamma(\nu)} \int_0^z (z-s)^{\nu-1} E_{\alpha,\beta}(\lambda s^{\alpha}) s^{\beta-1} ds = z^{\beta+\nu-1} E_{\alpha,\beta+\nu}(\lambda z^{\alpha}),$$

for $\beta>0$, $\nu>0$, one obtains, for $k \geq 0$,

$$\int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-(t-s)^{\alpha}A) s^k ds = \Gamma(k+1) t^{\alpha+k} E_{\alpha,\alpha+k+1}(-t^{\alpha}A).$$

Accordingly if $g(s) = \sum_{k=0}^{q} s^k v_k$, for some vectors v_k, $k=0,1,...q$, we get

$$y(t) = E_{\alpha,1}(-t^{\alpha}A)y_0 + \sum_{k=0}^{q} \Gamma(k+1) t^{\alpha+k} E_{\alpha,\alpha+k+1}(-t^{\alpha}A) v_k, \quad t>0.$$

In the general case, provided that we are able to compute efficiently the action of a matrix ML function on a vector, the above formula can be interpreted as an exponential integrator for FDEs.

3. The Standard Krylov Method (SKM). Throughout the paper, given a $N \times N$ complex matrix A, we denote its spectrum by $\sigma(A)$ and its *numerical range* by $W(A)$, i.e.,

$$W(A) = \left\{ \frac{\langle x, Ax \rangle}{\langle x, x \rangle},\ x(\neq 0) \in \mathbb{C}^N \right\},$$

where $\langle \cdot, \cdot \rangle$ represents the Euclidean inner product. The norm $\|\cdot\|$ will be the Euclidean vector norm, as well as its induced matrix norm. We denote by Π_k the set of the algebraic polynomials of degree $\leq k$. Moreover we assume that for some $a \geq 0$ and $0 \leq \vartheta < \frac{\pi}{2}$

$$W(A) \subset \Sigma_{\vartheta,a} = \{\lambda \in \mathbb{C} : |\arg(\lambda - a)| \leq \vartheta\}. \tag{3.1}$$

¿From now on let $v \in \mathbb{C}^N$ be a given vector with $\|v\| = 1$. Given a suitable function f, the SKM seeks for approximations to $y = f(A)v$ in the Krylov subspaces

$$K_m(A, v) = span\left\{v, Av, \ldots, A^{m-1}v\right\}$$

associated to A and v. By means of the Arnoldi method we generate a sequence of orthonormal vectors $\{v_j\}_{j\geq 1}$, $v_1 = v$, such that $K_m(A, v) = span\{v_1, v_2, \ldots, v_m\}$ for every m.

Setting $V_m = [v_1, v_2, ..., v_m]$ and $H_m = V_m^H A V_m$ we have

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^T. \tag{3.2}$$

In the sequel e_j denotes the j-th column of the $m \times m$ unit matrix and the $h_{i,j}$ are the entries of H_m. For every j, the entries $h_{j+1,,j}$, are real and non-negative. For $m \geq 2$, we have implicitly assumed that $h_{j+1,,j} > 0$, for $j = 1, ..., m-1$. Accordingly, the m-th standard Krylov approximation to y is given by $V_m f(H_m) e_1$.

Here we study the convergence of the method for approximating $E_{\alpha,\beta}(-A)v$. In this section, for $m \geq 1$ we set

$$R_m = E_{\alpha,\beta}(-A)v - V_m E_{\alpha,\beta}(-H_m) e_1.$$

3.1. General error estimates. Assumption 3.1. *Let (3.1) hold. Let $\beta > 0$ and $0 < \alpha < 2$ be such that $\frac{\alpha\pi}{2} < \pi - \vartheta$, $\varepsilon > 0$ and*

$$\frac{\alpha\pi}{2} < \mu \leq \min[\pi, \alpha\pi],\ \mu < \pi - \vartheta. \tag{3.3}$$

If Assumption 3.1 holds, since $W(H_m) \subseteq W(A)$ from the integral formula (2.2) we get

$$R_m = \frac{1}{2\alpha\pi i} \int_{C(\varepsilon,\mu)} \exp(\lambda^{\frac{1}{\alpha}}) \lambda^{\frac{1-\beta}{\alpha}} \delta_m(\lambda) d\lambda, \tag{3.4}$$

where $\delta_m(\lambda) = (\lambda I + A)^{-1} v - V_m(\lambda I + H_m)^{-1} e_1$. For each $\lambda \in C(\varepsilon, \mu)$, the following inequalities can be proved, using (3.2), by some standard arguments (cf. [17], Lemma 1 and [6], Lemmas 1 and 2):

$$\|\delta_m(\lambda)\| \leq \left\|(\lambda I + A)^{-1} - V_m(\lambda I + H_m)^{-1} V_m^H\right\| \|p_m(A)v\| \tag{3.5}$$

for every $p_m \in \Pi_m$ such that $p_m(-\lambda) = 1$ and

$$\|\delta_m(\lambda)\| = \frac{\prod_1^m h_{j+1,j}}{|\det(\lambda I + H_m)|} \left\|(\lambda I + A)^{-1} v_{m+1}\right\|. \tag{3.6}$$

By these facts, below we give some error bounds, with the aim of investigating the role of the parameters α and β. In the sequel, for $\lambda \in C(\varepsilon, \mu)$, we set

$$D(\lambda) = dist(\lambda, W(-A)). \tag{3.7}$$

We observe that, under the previous assumptions, we can find a function $\nu(\varphi)$ such that, for any $\lambda = |\lambda| \exp(\pm i\varphi) \in C(\varepsilon, \mu)$, it holds

$$D(\lambda) \geq \nu(\varphi) \, |\lambda| \, , \text{ with } \nu(\varphi) \geq \nu > 0. \tag{3.8}$$

With this notation, we give the following result.

THEOREM 3.2. *Let Assumption 3.1 hold. For $m \geq 1$ and for every $M > 0$, we have*

$$\|R_m\| \leq \frac{\exp(M) \prod_{j=1}^{m} h_{j+1,j}}{\pi \nu^{m+1} M^{m\alpha+\beta-1}} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(\left|\cos\frac{\mu}{\alpha}\right| + 1))}{(m\alpha - 1 + \beta)} \right). \tag{3.9}$$

Proof. Since $\left\|(\lambda I + A)^{-1}\right\| \leq D(\lambda)^{-1}$and $W(H_m) \subseteq W(A)$, by (3.4) and (3.6) we obtain

$$\|R_m\| \leq \frac{\prod_{j=1}^{m} h_{j+1,j}}{2\pi\alpha} \int\limits_{C(\varepsilon,\mu)} \frac{\left|\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}\right|}{D(\lambda)^{m+1}} \, |d\lambda| \, .$$

Let us set

$$I_1 = \int\limits_{C_1(\varepsilon,\mu)} \frac{\left|\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}\right|}{D(\lambda)^{m+1}} \, |d\lambda|$$

and

$$I_2 = \int\limits_{C_2(\varepsilon,\mu)} \frac{\left|\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}\right|}{D(\lambda)^{m+1}} \, |d\lambda| \, .$$

By (3.8) we get

$$I_1 \leq 2\varepsilon^{\frac{1-\beta}{\alpha} - m} \int\limits_{0}^{\mu} \frac{\exp(\varepsilon^{\frac{1}{\alpha}} \cos\frac{\varphi}{\alpha})}{\nu(\varphi)^{m+1}} d\varphi,$$

and, by simple computations,

$$\begin{aligned} I_2 &\leq \frac{2}{\nu^{m+1}} \int\limits_{\varepsilon}^{+\infty} \frac{r^{\frac{1-\beta}{\alpha}} \exp(-r^{\frac{1}{\alpha}} \left|\cos\frac{\mu}{\alpha}\right|)}{r^{m+1}} dr \\ &= \frac{2}{\nu^{m+1}} \int\limits_{\varepsilon^{\frac{1}{\alpha}}}^{+\infty} \frac{\exp(-s \left|\cos\frac{\mu}{\alpha}\right|)}{s^{m\alpha+\beta}} ds \qquad (3.10) \\ &\leq \frac{2\alpha \exp(-\varepsilon^{\frac{1}{\alpha}} \left|\cos\frac{\mu}{\alpha}\right|)}{(m\alpha + \beta - 1)\nu^{m+1}\varepsilon^{\frac{m\alpha+\beta-1}{\alpha}}}. \qquad (3.11) \end{aligned}$$

Setting $\varepsilon = M^{\alpha}$, the result easily follows. □

By the same arguments, below we derive a further bound, that seems to be more suited for small α.

Corollary 3.3. *Let Assumption 3.1 hold. Let $m \geq 1$ be such that*

$$m\alpha + \beta > 0,$$

then, for every $M > 0$, we have

$$\|R_m\| \leq \frac{\exp(M)\prod_{j=1}^{m} h_{j+1,j}}{4\nu^{m+1}M^{m\alpha}}\kappa(M). \tag{3.12}$$

where

$$\kappa(M) = \frac{4M^{1-\beta}}{\pi}\left(\frac{\mu}{\alpha} + \frac{\exp(-M(\left|\cos\frac{\mu}{\alpha}\right| + 1))}{M\left|\cos\frac{\mu}{\alpha}\right|}\right). \tag{3.13}$$

Proof. With respect to the previous proof we only change the last bound (3.11). Taking again $\varepsilon = M^{\alpha}$, we obtain

$$\int_{\varepsilon}^{+\infty} \frac{r^{\frac{1-\beta}{\alpha}}\exp(-r^{\frac{1}{\alpha}}\left|\cos\frac{\mu}{\alpha}\right|)}{r^{m+1}}dr \leq \frac{\alpha\exp(-M\left|\cos\frac{\mu}{\alpha}\right|)}{M^{m\alpha+\beta}\left|\cos\frac{\mu}{\alpha}\right|}.$$

□

Remark 3.4. *We notice that, by a more precise evaluation of the term* $|\det(\lambda I + H_m)|$ *in (3.6) sharper a posteriori estimates could be obtained following the lines of the recent paper [6], where exponential-like functions have been considered.*

As expected, since for $\alpha > 0$ the ML functions are entire, from (3.9) (as well as from 3.12) it is possible to recognize the superlinear convergence of the SKM, at least for sufficiently large m. To do this let us take in (3.9) $M = m\alpha + \beta - 1$ We realize that the bound depends essentially on the term

$$\left(\frac{\exp(1)}{M}\right)^{M} \nu^{-(m+1)} \prod_{j=1}^{m} h_{j+1,j}. \tag{3.14}$$

Clearly, such term decays only for very large m if α and ν are small and the products $\prod_{j=1}^{m} h_{j+1,j}$ are large. We recall that such products can be estimated by means of the well-known inequality

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

that holds for every monic polynomial q_m of exact degree m. Accordingly, taking q_m as the monic Faber polynomial associated to a closed convex subset Ω such that $W(A) \subseteq \Omega$, by a result of Beckermann ([1]) and the definition of Faber polynomials, we get the bound

$$\prod_{j=1}^{m} h_{j+1,j} \leq 2\gamma^m, \tag{3.15}$$

where γ denotes the logarithmic capacity of Ω. Recall that if $\Omega = [a, b] \subseteq [0, +\infty)$ then $\gamma = (b - a)/4$. In conclusion we can say that the term (3.14), for m sufficiently large, can behave like

$$\left(\frac{\exp(1)}{m\alpha}\right)^{m\alpha} \left(\frac{\gamma}{\nu}\right)^{m}. \tag{3.16}$$

This predicts that the convergence of the standard Krylov method for ML functions can deteriorate as α decreases as well as $W(A)$ enlarges. This fact is confirmed by the numerical tests.

Beside the above general bounds, in particular situations sharper a priori error estimates can be derived. Below we consider the Hermitian case. Since we are dealing with any generalized ML function, following the lines of [17], we use formula (3.5) choosing suitably the polynomials p_m. An interesting alternative approach could be that based on the Faber tranform, recently discussed in [2]. Such approach should be preferable in the treatment of specific ML functions when suitable bounds of such functions are available.

For our purposes, we recall that if Ω is the above subset of the complex plane, then there is a constant $\omega(\Omega)$, depending only on Ω, such that for every polynomial p, there holds

$$\|p(A)\| \leq \omega(\Omega) \max_{z\in\Omega} |p(z)| \,.$$

We notice that if Ω is a real interval then $\omega(\Omega) = 1$. Bounds for sectorial sets can be found in [5]. In the general case we have $\omega(\Omega) \leq 11.08$, as stated in [4]. We refer to that paper for discussions on this point. Therefore, the application of formula (3.5) is related to the classical problem of estimating

$$\zeta_m(\Omega; -z) := \min_{p_m\in\Pi_k, p_m(-z)=1} \max_{\xi\in\Omega} |p_m(\xi)| \tag{3.17}$$

for $-z \notin \Omega$. A general result, which makes use of the near-optimality properties of the Faber polynomials associated to Ω, can be found in [17].

3.2. The Hermitian case. We study the Hermitian case assuming that $\Omega = [a, b] \subseteq [0, +\infty)$. Together with (3.5), we employ a result given in [12], Theorem 1 (cf. also [21] Theorem 4.3) that gives the bound

$$\zeta_m(\Omega; -z) \leq \frac{2}{\Phi(u(z))^m}, \tag{3.18}$$

where

$$\Phi(u) = u + \sqrt{u^2 - 1} \tag{3.19}$$

is the inverse Zhukovsky function and

$$u(z) = \frac{|b + z| + |a + z|}{b - a}. \tag{3.20}$$

Thus, from (3.4) and (3.5) we obtain the bound

$$\|R_m\| \leq \frac{2}{\alpha\pi} \int_{C(\varepsilon,\mu)} \frac{\left|\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}\right|}{D(\lambda)\Phi(u(\lambda))^m} |d\lambda| \,. \tag{3.21}$$

Below we consider in detail the case, frequently occurring in the applications, where $0 < \alpha < 1$, $\beta \geq \alpha$. For the sequel, for $\lambda = r \exp(i\varphi) \in C(\varepsilon, \mu)$, we set

$$\rho(r, \varphi) = \sqrt{a^2 + r^2 + 2ar \cos\varphi}. \tag{3.22}$$

THEOREM 3.5. *Assume that A is Hermitian with $\sigma(A) \subseteq [a,b] \subset [0,+\infty)$. Assume that $0 < \alpha < 1$ and $\beta \geq \alpha$. Let us take $\mu \leq \frac{\pi}{2}$, with $\frac{\alpha\pi}{2} < \mu \leq \alpha\pi$. Then for every index $m \geq 1$ and for every $M > 0$ we have*

$$\|R_m\| \leq \kappa(M)\exp(M)\Phi^{*-m}, \tag{3.23}$$

where

$$\Phi^* = \Phi(u(M^\alpha \exp(i\mu)))$$

and $\kappa(M)$ is given by (3.13).

Proof. Referring to (3.21), Now let us set

$$I_1 = \int\limits_{C_1(\varepsilon,\mu)} \left| \frac{\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}}{D(\lambda)\Phi(u(\lambda))^m} \right| |d\lambda|$$

and

$$I_2 = \int\limits_{C_2(\varepsilon,\mu)} \left| \frac{\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}}{D(\lambda)\Phi(u(\lambda))^m} \right| |d\lambda|\,.$$

Since $\mu \leq \frac{\pi}{2}$, for $\lambda = r\exp(i\varphi) \in C(\varepsilon,\mu)$ we have

$$D(\lambda) \geq \rho(r,\varphi) \geq r$$

and by simple computations one easily observes that the function $\Phi(u(\lambda))$ is increasing w.r.t. r as well as decreasing w.r.t. $|\varphi|$. Let us set $\varepsilon = M^\alpha$. Therefore we easily get

$$I_1 \leq \frac{2M^{1-\beta}\mu\exp(M)}{\Phi^{*m}}. \tag{3.24}$$

Moreover

$$I_2 \leq 2\int\limits_{M^\alpha}^{+\infty} \frac{r^{\frac{1-\beta}{\alpha}}\exp(-r^{\frac{1}{\alpha}}\left|\cos\frac{\mu}{\alpha}\right|)}{\Phi(u(r\exp(i\mu)))^m r}dr. \tag{3.25}$$

Thus

$$I_2 \leq \frac{2\alpha}{\Phi^{*m}}\int\limits_{M}^{+\infty} s^{-\beta}\exp(-s\left|\cos\frac{\mu}{\alpha}\right|)ds \tag{3.26}$$

and

$$I_2 \leq \frac{2\alpha M^{-\beta}\exp(-M\left|\cos\frac{\mu}{\alpha}\right|)}{\Phi^{*m}\left|\cos\frac{\mu}{\alpha}\right|}.$$

Hence, from (3.21) and by (3.24), we obtain (3.23). □

In order to satisfy the assumptions of Theorem 3.5 , if $0 < \alpha < \frac{1}{2}$, we can take $\mu = \alpha\pi$. If $\frac{1}{2} \leq \alpha < 1$ then $\mu = \frac{\pi}{2}$ is allowed. Below, for $\mu = \frac{\pi}{2}$, we state further formulae for $\kappa(M)$ in (3.23) when α is close to 1

COROLLARY 3.6. *Let the assumptions of Theorem 3.5 hold and moreover assume* $\frac{3}{4} \le \alpha < 1$ *and* $\mu = \frac{\pi}{2}$. *Then (3.23) holds with*

$$\kappa(M) = 2M^{1-\beta}(\frac{1}{\alpha} + \frac{2\Phi^{*\frac{1}{2}} \exp(-M(\left|\cos\frac{\pi}{2\alpha}\right| + 1))(b-a)^{\frac{1}{2}}}{\pi M^{\frac{\alpha}{2}}(3\alpha - 2)}). \tag{3.27}$$

If in addition $\beta > 1$, *then (3.23) holds also with*

$$\kappa(M) = 2M^{1-\beta}(\frac{1}{\alpha} + \frac{2\exp(-M(\left|\cos\frac{\pi}{2\alpha}\right| + 1)}{\pi(\beta - 1)}). \tag{3.28}$$

Proof. With respect to the proof of Theorem 3.5 we modify only the bound for I_2. Observing that, for $r \ge \varepsilon$,

$$\Phi(u(ir)) \ge \frac{4r}{b-a},$$

from (3.25), setting again $\varepsilon = M^\alpha$, we obtain

$$\begin{aligned} I_2 &\le \frac{\alpha M^{\alpha-\beta}\sqrt{b-a}}{\Phi^{*(m-\frac{1}{2})}} \int_M^{+\infty} \frac{\exp(-s\left|\cos\frac{\pi}{2\alpha}\right|)}{s^{\frac{3\alpha}{2}}} ds \\ &\le \frac{2\alpha M^{-\frac{\alpha}{2}-\beta+1}\sqrt{b-a}}{\Phi^{*(m-\frac{1}{2})}} \frac{\exp(-M\left|\cos\frac{\pi}{2\alpha}\right|)}{(3\alpha - 2)}. \end{aligned} \tag{3.29}$$

Hence (3.27) follows.

If $\beta > 1$ from (3.26) now we obtain

$$I_2 \le \frac{2\alpha M^{1-\beta}}{\beta - 1}\Phi^{*-m} \exp(-M\left|\cos\frac{\pi}{2\alpha}\right|).$$

and hence (3.28). ∎

We observe that the well-known relationship

$$E_{\alpha,\beta}(z) = E_{\alpha,\alpha+\beta}(z)z + \frac{1}{\Gamma(\beta)},$$

allows to use formula (3.28) even when $\alpha + \beta > 1$.

In the practical use of (3.23), with (3.13), (3.27) or (3.28), one can take M that minimizes the corresponding right-hand sides. In the theorem below we show that, suitable choices of the parameter M, yield superlinear convergence results which can be viewed as extensions of those given for the exponential function in [9], [17] and [2].

For the sake of simplicity, below we assume $a = 0$. The notation $\kappa(M)$ refers to any of the three formulae stated above.

THEOREM 3.7. *Let the assumptions and notations of Theorem 3.5 and Corollary 3.6 hold. Assume* $\frac{1}{2} \le \alpha < 1$ *and let* $\mu = \frac{\pi}{2}$, $a = 0$. *If, for* $c = 0.98$,

$$M = \left(\frac{cm\alpha}{\sqrt{b}}\right)^{\frac{2}{2-\alpha}} \le \left(\frac{b}{4}\right)^{\frac{1}{\alpha}} \tag{3.30}$$

then

$$\|R_m\| \le g_m \kappa(M), \tag{3.31}$$

where

$$g_m := \exp(-M(\sqrt{2}\alpha^{-1} - 1)).$$

Moreover, if

$$\left(\frac{m\alpha}{\sqrt{b}}\right)^{\frac{2\alpha}{2-\alpha}} \ge \frac{b}{4} \tag{3.32}$$

then taking

$$M = \left(\frac{\sqrt{b}}{2}\right)^{\frac{1}{\alpha}} \left(\frac{m\alpha}{\sqrt{b}}\right)^{\frac{1}{2-\alpha}}, \tag{3.33}$$

(3.31) holds with

$$g_m := \exp(M)(\frac{\sqrt{b}}{2})^m \left(\frac{m\alpha}{\sqrt{b}}\right)^{-\frac{m\alpha}{2-\alpha}}.$$

Proof. Now, referring to (3.19), (3.20), we have $\Phi^* = \Phi(u(iM^\alpha))$ and we realize that

$$u(iM^\alpha) \ge 1 + \frac{M^\alpha}{b}. \tag{3.34}$$

If we take M as in (3.30) then $\frac{M^\alpha}{b} \le \frac{1}{4}$ and it can be verified (cf. [17] Th. 2) that this implies

$$\Phi(u(iM^\alpha) \ge \exp(c\sqrt{2\frac{M^\alpha}{b}}). \tag{3.35}$$

Thus we find

$$\frac{\exp(M)}{\Phi(u(iM^\alpha)^m} \le \exp(M - cm\sqrt{2\frac{M^\alpha}{b}}) = g_m.$$

and the first part of the statement is proved.

The second part follows from $\Phi(u(iM^\alpha)) \ge \frac{4M^\alpha}{b}$ that is

$$\Phi(u(iM^\alpha)) \ge \frac{2}{\sqrt{b}} \left(\frac{m\alpha}{\sqrt{b}}\right)^{\frac{\alpha}{2-\alpha}}.$$

□

Remark 3.8. *It is interesting to give a look to the case $\alpha \to 0$, for $\beta = 1$. We recall that $E_{0,1}(-z) = (1+z)^{-1}, |z| < 1$. From (3.23) and (3.13) setting $\mu = \alpha\pi$ and $M = 1$, letting $\alpha \to 0$ we find*

$$\|R_m\| \le \frac{4(\pi \exp(1) + \exp(-1))}{\pi \Phi(u(1))^m}.$$

So we have retrieved the classical bound for the CG method, where the convergence depends on the conditioning.

4. The Rational Arnoldi Method (RAM). Let Assumption 3.1 hold. Let $h > 0$ be a given real parameter and let us set

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

Now we approximate $y = f(A)v$ in the Krylov subspaces $K_m(Z, v)$. We assume again that such subspaces are constructed by the Arnoldi method. Accordingly, now we get a sequence $\{u_j\}_{j\geq 1}$ of orthonormal basis-vectors, with $u_1 = v$, such that, setting $U_m = [u_1, u_2, ..., u_m]$, we have

$$ZU_m = U_m S_m + s_{m+1,m} u_{m+1} e_m^T, \tag{4.1}$$

where $S_m = U_m^H Z U_m$, with entries $s_{i,j}$, is a $m \times m$ upper Hessenberg matrix. Moreover we have $W(S_m) \subseteq W(Z)$. The m-th approximation to $y = f(A)v$ is now defined by $y_m = V_m f(B_m) e_1$, where B_m satisfies

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

Here we call this the Rational Arnoldi Method (RAM). We notice that in general $B_m \neq U_m^H A U_m$.

Let us set

$$\gamma = \frac{1}{2}(1 + ha)^{-1},$$

and

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

where D_γ is the disk of center and radius γ.

Lemma 4.1. *Assume (3.1). Then*

$$W(Z) \subset S_{\vartheta,\gamma}, \tag{4.2}$$

and for every m

$$W(B_m) \subset \Sigma_{\vartheta,a}. \tag{4.3}$$

Proof. From

$$\frac{< x, Zx >}{< x, x >} = \frac{< (I + hA)y, y >}{< x, x >}, y = Zx (\neq 0),$$

we get

$$\frac{< x, Zx >}{< x, x >} = \frac{(1 + h\overline{\eta})}{1 + 2h \operatorname{Re} \eta + h^2 \frac{\|Ay\|^2}{\|y\|^2}} \tag{4.4}$$

with $\eta \in W(A)$. Therefore, since $|\eta|^2 \leq \frac{\|Ay\|^2}{\|y\|^2}$ we obtain

$$\left| \frac{< x, Zx >}{< x, x >} \right|^2 \leq (1 + h \operatorname{Re} \eta)^{-1} \operatorname{Re} \frac{< x, Zx >}{< x, x >}.$$

which, together with (4.4), gives (4.2).

In order to prove (4.3), recalling that $S_m = U_m^H Z U_m$, for any $x(\neq 0) \in \mathbb{C}^m$ we have

$$h\frac{< x, B_m x >}{< x, x >} = \frac{< S_m y, y >}{< S_m y, S_m y >} - 1$$
$$= \frac{< ZU_m y, U_m y >}{< ZU_m y, P_m ZU_m y >} - 1$$
$$= \frac{< ZU_m y, ZU_m y > +h < ZU_m y, AZU_m y >}{< ZU_m y, P_m ZU_m y >} - 1.$$

where $P_m = U_m U_m^H$ is an orthogonal projection. Therefore

$$h\frac{< x, B_m x >}{< x, x >} = \frac{(1 + h\eta) \|ZU_m y\|^2}{\|P_m ZU_m y\|^2} - 1, \tag{4.5}$$

for some $\eta \in W(A)$. Since

$$\frac{\|ZU_m y\|^2}{\|P_m ZU_m y\|^2} \geq 1,$$

by (4.5) we get (4.3). □

4.1. General error analysis. Now let us set

$$R_m = E_{\alpha,\beta}(-A)v - U_m E_{\alpha,\beta}(-B_m)e_1.$$

Let (3.1) hold. Clearly, by Lemma 4.1 we have $\sigma(-A) \cup \sigma(-B_m) \subset G^-(\varepsilon, \mu)$ for every m and therefore

$$R_m = \frac{1}{2\pi\alpha i} \int_{C(\varepsilon,\mu)} \exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}} b_m(\lambda) d\lambda, \tag{4.6}$$

where

$$b_m(\lambda) = (\lambda I + A)^{-1} v - U_m(\lambda I + B_m)^{-1} e_1.$$

From now on, for each $\lambda \in C(\varepsilon, \mu)$ let us set

$$w(\lambda) = (h\lambda - 1). \tag{4.7}$$

Since $(\lambda I + A)^{-1} = hZ(I + \omega(\lambda)Z)^{-1}$ and $(\lambda I + B_m)^{-1} = hS_m(I + \omega(\lambda)S_m)^{-1}$, for every $h > 0$ and for every $m \geq 1$, by (4.1), one realizes that for $c \in \mathbb{C}^m$, with $e_m^T c = 0$,

$$((\lambda I + A)^{-1} - U_m(\lambda I + B_m)^{-1} U_m^H)(I + \omega(\lambda)Z)U_m c = 0$$

and therefore

$$b_m(\lambda) = ((\lambda I + A)^{-1} - U_m(\lambda I + B_m)^{-1} U_m^H) p_{m-1}(Z)v, \tag{4.8}$$

for every $p_{m-1} \in \Pi_{m-1}$ such that $p_{m-1}(-w(\lambda)^{-1}) = 1$.

In order to state a convergence result for the case $0 < \alpha < 1$, we set

$$d(\lambda) = dist(-w(\lambda)^{-1}, W(Z)),$$

and we make use of the following lemma.

LEMMA 4.2. *Assume (3.1). Then, for every $h > 0$ and for every $0 < \mu < \frac{\pi}{2}$, setting*

$$\overline{\rho} = \min[\vartheta, \mu]$$

and taking

$$\varepsilon = \frac{1}{h \cos \overline{\rho}},$$

there exists a positive constant d_0 such that, for $\lambda \in C(\varepsilon, \mu)$, we have

$$d(\lambda) \geq d_0 \left| w(\lambda)^{-1} \right|. \tag{4.9}$$

Proof. The result follows by (4.2). In fact, at first one verifies that (4.9) holds true whenever $\mathrm{Re}(-w(\lambda)^{-1}) \leq 0$. For $\lambda \in C(\varepsilon, \mu)$, this condition is verified if $\overline{\rho} = \mu$ and moreover when $\overline{\rho} = \vartheta$, $\lambda = r \exp(\pm i\varphi), r \geq \varepsilon$ and $\frac{r}{\varepsilon} \cos \varphi \geq \cos \vartheta$. Furthermore, in all the remaining cases when $\overline{\rho} = \vartheta$, one easily realizes that there is $C > 1$ such that $\left| \frac{\mathrm{Im}(-w(\lambda)^{-1})}{\mathrm{Re}(-w(\lambda)^{-1})} \right| \geq C \tan \vartheta$. □

THEOREM 4.3. *Assume $0 < \alpha < 1$ and $\beta \geq \alpha$. Then, there exists a function $g(h)$, continuous in any bounded interval $0 < h_1 \leq h \leq h_2$, such that, for $m \geq 2$,*

$$\|R_m\| \leq \frac{g(h)}{m-1},$$

for every matrix A satisfying (3.1).

Proof. Since $0 < \alpha < 1$ we can take $\frac{\alpha\pi}{2} < \mu < \frac{\pi}{2}$, $\mu \leq \alpha\pi$. Let us take

$$\varepsilon = \frac{1}{h \cos \overline{\rho}},$$

with $\overline{\rho} = \min[\vartheta, \mu]$ so that (4.9) holds. Let $\chi > 0$ be the logarithmic capacity of $W(Z)$. Let $\{F_k\}_{k \geq 0}$, be the sequence of the ordinary Faber polynomials associated to $W(Z)$. Then, by results in [1] and [24], Lemma 4.3, for $-w(\lambda)^{-1} \notin W(Z)$ and for $k \geq 1$ one gets the bound

$$\frac{\|F_k(Z)\|}{|F_k(-w(\lambda)^{-1})|} \leq \frac{8\chi}{k d(\lambda)}.$$

By this inequality, from (4.6), (4.8) and using (3.8) we have

$$\|R_m\| \leq \frac{8\chi}{(m-1)\pi\alpha\nu} \int\limits_{C(\varepsilon,\mu)} \left| \frac{\exp(\lambda^{\frac{1}{\alpha}}) \lambda^{\frac{1-\beta}{\alpha}}}{\lambda d(\lambda)} \right| |d\lambda|. \tag{4.10}$$

It can be seen that the integral in (4.10) is bounded by a continuous function of h. Indeed, let

$$I_1 = \int\limits_{C_1(\varepsilon,\mu)} \left| \frac{\exp(\lambda^{\frac{1}{\alpha}}) \lambda^{\frac{1-\beta}{\alpha}}}{\lambda d(\lambda)} \right| |d\lambda|$$

and

$$I_2 = \int\limits_{C_1(\varepsilon,\mu)} \left| \frac{\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}}{\lambda d(\lambda)} \right| |d\lambda| .$$

By (4.9), we obtain

$$I_1 \leq \frac{2\varepsilon^{\frac{1-\beta}{\alpha}}}{d_0} \int\limits_0^{\mu} \exp\left(\varepsilon^{\frac{1}{\alpha}} \cos\frac{\varphi}{\alpha}\right) c(\varphi) d\varphi, \tag{4.11}$$

where

$$c(\varphi) = \frac{\sqrt{1+\cos^2\overline{\rho} - 2\cos\overline{\rho}\cos\varphi}}{\cos\overline{\rho}}.$$

Therefore from (4.11) we get

$$I_1 \leq \frac{2\mu(h\cos\overline{\rho})^{\frac{\beta-1}{\alpha}} \exp\left((h\cos\overline{\rho})^{-\frac{1}{\alpha}}\right) c(\mu)}{d_0}. \tag{4.12}$$

Furthermore for $\lambda = r\exp(\pm i\mu), r \geq \varepsilon$,

$$|w(\lambda)| = hr\sqrt{1+(hr)^{-2} - 2(hr)^{-1}\cos\mu} \leq \sqrt{2}hr. \tag{4.13}$$

Thus, by (4.9) and (4.13) we get

$$\begin{aligned} I_2 &\leq \frac{2\sqrt{2}h}{d_0} \int\limits_{\varepsilon}^{+\infty} \exp\left(-r^{\frac{1}{\alpha}} \left|\cos\frac{\mu}{\alpha}\right|\right) r^{\frac{1-\beta}{\alpha}} dr \\ &\leq \frac{2\alpha\sqrt{2}h}{d_0} \int\limits_{\varepsilon^{\frac{1}{\alpha}}}^{+\infty} \exp\left(-s\left|\cos\frac{\mu}{\alpha}\right|\right) s^{\alpha-\beta} ds \end{aligned}$$

and finally

$$I_2 \leq \frac{2\alpha\sqrt{2}h^{\frac{\beta}{\alpha}}(\cos\overline{\rho})^{\frac{\beta-\alpha}{\alpha}} \exp\left(-(h\cos\overline{\rho})^{-\frac{1}{\alpha}}\left|\cos\frac{\mu}{\alpha}\right|\right)}{d_0\left|\cos\frac{\mu}{\alpha}\right|}. \tag{4.14}$$

By (4.10), the last bound together with (4.12) gives the thesis. □

Even if only qualitative, the above result is important since it points out that, contrary to the SKM, the convergence of the RAM cannot deteriorate as $W(A)$ enlarges and moreover that it is uniform with respect to h in any positive bounded interval.

REMARK 4.4. *As in Remark 3.8 let us consider the case $\beta = 1$, for $\alpha \to 0$. We expect that making $h \to 1$ the error vanishes. In fact, looking at (4.14) and (4.12) one can realize that this actually occurs. Take for instance, for small α, $\mu = \alpha\pi$ and $h = \alpha^\alpha$ and observe that $c(\mu) \to 0$.*

As for the SKM, more precise a priori error bounds can be obtained taking into account of specific situations, as we show below dealing with the Hermitian case.

4.2. The Hermitian case. If A is Hermitian with $\sigma(A) \subseteq [a, +\infty)$, referring to (3.8) and (3.22), for $\lambda = r\exp(i\varphi)$ we have

$$D(\lambda) \geq \rho(r, \varphi) \geq \rho(\varepsilon, \mu), \text{ for } 0 \leq |\varphi| \leq \frac{\pi}{2}, \tag{4.15}$$

and

$$D(\lambda) \geq r\,|\sin\varphi|\,, \quad \text{for } \frac{\pi}{2} < |\varphi| < \pi. \tag{4.16}$$

THEOREM 4.5. *Assume that A is Hermitian with $\sigma(A) \subseteq [a, +\infty)$, $a \geq 0$. Assume $0 < \alpha \leq \frac{2}{3}$ and $\beta \geq \alpha$. Then for every $m \geq 1$ we have*

$$\|R_m\| \leq \frac{K_1 Q_m h^{\frac{\beta-1}{\alpha}}}{(1+\sqrt{2})^{m-1}} + \frac{K_2 h^{\frac{\beta}{\alpha}}}{(m-1)^2}\exp(-\frac{h^{-\frac{1}{\alpha}}}{\sqrt{2}}), \tag{4.17}$$

where

$$Q_m = \max_{0 \leq |\varphi| \leq \frac{3\alpha\pi}{4}} \exp\left(h^{-\frac{1}{\alpha}}\cos\frac{\varphi}{\alpha}\right)(1-\cos\varphi)^{\frac{m-1}{2}},$$

and K_1, K_2 are (small) constants.

Proof. At first we state a general bound that holds for $0 < \alpha < 1$ and $\frac{\alpha\pi}{2} < \mu \leq \alpha\pi$. Hence we will derive (4.17).

Let us take take $\varepsilon = h^{-1}$. In (4.8) we apply (3.18) with (3.20), with $z = w(\lambda)^{-1}$, taking into account that, by Lemma 4.1, $W(Z) \subset \Omega = (0, 1]$. So doing, referring to (3.17), we get

$$\zeta_m(\Omega; -w(\lambda)^{-1}) \leq \frac{2}{\Phi(u)^m},$$

where

$$u = \frac{|w(\lambda)+1|+1}{|w(\lambda)|}.$$

For $\lambda \in C_1(\varepsilon, \mu)$, we obtain

$$u = \left(\frac{2}{1-\cos\varphi}\right)^{\frac{1}{2}}.$$

Thus

$$\Phi(u)^{-1} = (\frac{\sqrt{1-\cos\varphi}}{\sqrt{2}+\sqrt{1+\cos\varphi}}). \tag{4.18}$$

Furthermore for $\lambda = r\exp(\pm i\mu)$, $r \geq h^{-1}$, we easily get

$$u = \frac{hr+1}{|w(\lambda)|},$$

and then

$$\Phi(u)^{-1} = f(r) = \frac{\sqrt{h^2r^2+1-2hr\cos\mu}}{hr+1+\sqrt{2hr(1+\cos\mu)}}. \tag{4.19}$$

Therefore, combining (4.18) and (4.19), from (4.6) and (4.8) we obtain the error bound

$$\|R_m\| \leq \frac{4}{\alpha\pi}(h^{\frac{\beta-1-\alpha}{\alpha}} I_m + J_m), \tag{4.20}$$

where

$$I_m = \int_0^{\mu} \frac{\exp\left(h^{-\frac{1}{\alpha}} \cos\frac{\varphi}{\alpha}\right) \left(\frac{\sqrt{1-\cos\varphi}}{\sqrt{2}+\sqrt{1+\cos\varphi}}\right)^{m-1}}{D(h^{-1}\exp(i\varphi))} d\varphi, \tag{4.21}$$

and

$$J_m = \int_{\frac{1}{h}}^{+\infty} \frac{r^{\frac{1-\beta}{\alpha}} \exp\left(-r^{\frac{1}{\alpha}} \left|\cos\frac{\mu}{\alpha}\right|\right) f(r)^{m-1}}{D(r\exp(i\mu))} dr.$$

Now let $0 < \alpha \leq \frac{2}{3}$ and let us take $\mu = \frac{3\alpha\pi}{4}$. Since $hr \geq 1$, $r \geq \varepsilon$, and since now $\mu \leq \frac{\pi}{2}$ we have $D(r\exp(i\varphi) \geq r$ and we find that

$$f(r) \leq (1 + \frac{\eta}{\sqrt{hr}})^{-1}, \text{ with } \eta = \frac{1}{\sqrt{2}}.$$

Accordingly,

$$\begin{aligned} J_m &\leq \int_{\frac{1}{h}}^{+\infty} r^{\frac{1-\beta-\alpha}{\alpha}} \exp\left(-\frac{r^{\frac{1}{\alpha}}}{\sqrt{2}}\right) \left(\frac{\sqrt{hr}}{\eta+\sqrt{hr}}\right)^{m-1} dr \\ &\leq \int_{\frac{1}{h}}^{+\infty} r^{\frac{1-\beta-\alpha}{\alpha}} \exp\left(-\frac{r^{\frac{1}{\alpha}}}{\sqrt{2}} - \frac{(m-1)\eta}{(\eta+\sqrt{hr})}\right) dr \\ &\leq \alpha \int_{h^{-\frac{1}{\alpha}}}^{+\infty} s^{-\beta} \exp\left(-\frac{s}{\sqrt{2}} - \frac{(m-1)\eta}{(1+\eta)(hs^{\alpha})^{\frac{1}{2}}}\right) ds. \end{aligned}$$

Now, one verifies that, for every $c > 0$, for $s > 0$ it holds that

$$s^{-\alpha}\exp(-cs^{-\frac{\alpha}{2}}) \leq \left(\frac{2}{c}\right)^2 \exp(-2).$$

Therefore we get

$$\frac{4}{\pi\alpha} J_m \leq \frac{16h(1+\eta)^2\exp(-2)}{\pi\eta^2(m-1)^2} \int_{h^{-\frac{1}{\alpha}}}^{+\infty} \frac{\exp\left(-\frac{s}{\sqrt{2}}\right)}{s^{\beta-\alpha}} ds.$$

By this inequality and by (4.21) with (4.15), from (4.20) the result follows, with $K_1 = 3, K_2 = \frac{16\sqrt{2}(1+\eta)^2\exp(-2)}{\pi\eta^2}$. □

We notice that, by the arguments used above, bounds like (4.17) can be derived from (4.20) also for $\frac{2}{3} < \alpha \leq 1$, $\beta \geq \alpha$.

We do not make here a detailed discussion on the choice of the parameter h in the general case.We are convinced that, as it occurs for the matrix exponential, a value working well in the Hermitian case should work well in most of the cases, at least when the angle ϑ is not close to $\frac{\pi}{2}$. In the Hermitian case a reasonable choice could be that minimizing the right-hand side of (4.17) (or of (4.20)). Below we suggest a value, independent of β, which ensures an exponential-type error bound.

COROLLARY 4.6. *Let the assumptions and notations of Theorem 4.5 hold. For any fixed $m \geq 2$, taking, for any $\tau \geq 2(1-\cos\frac{3\alpha\pi}{4})$,*

$$h = \left(\frac{\tau}{m-1}\right)^{\alpha}, \tag{4.22}$$

we have

$$\|R_m\| \leq \frac{K_1 \tau^{\beta-1}(m-1)^{-(\beta-1)}}{(1+\sqrt{2})^{m-1}} + \frac{K_2 \tau^{\beta}}{(m-1)^{(\beta+2)}} \exp(-\frac{(m-1)}{\tau\sqrt{2}}).$$

Proof. Put the assigned value of h into (4.17) and observe that $Q_m \leq 1$. □

In the case $\beta = 1$, one could also take h by an heuristic criterion that often turned out to be satisfactory in practice. We know that for $\beta = 1$ and $\alpha = 0$, $h = 1$ is the obvious choice. On the other hand, for the exponential case, i.e., $\beta = \alpha = 1$, optimal values of h are known, see [24], [11] and [29]. Thus it seems reasonable to adopt an interpolation between these values.

Referring to the applications described in Section 2, for the computation of $E_{\alpha,\beta}(-t^{\alpha}A)$, our arguments lead to take h depending on t. However, for obvious computational reasons, we are interested to maintain the same h for a (large enough) window of values of t. This aspect will be subject of future investigations.

We want to point out that, in many cases, as for instance dealing with discretizations of elliptic operators, the RAM exhibits a very fast convergence and the a priori error bounds turn out to be pessimistic. As it is well known this fact often occurs in the application of Krylov subspace methods and it is due to their "good" adaptation to the spectrum (see [18]). Thus, in order to detect the actual behavior of the RAM, a posteriori error estimates could be more suitable. For instance, working on the lines of [24] and [6], for $m \geq 2$ one can consider in (4.8) the polynomial

$$p_{m-1}(z) = \frac{\det(S_{m-1} - zI)}{\det(S_{m-1} + w(\lambda)^{-1}I)}.$$

Since, as it is well known,

$$\|p_{m-1}(Z)v\| = \frac{\prod_{j=1}^{m-1} s_{j+1,j}}{|\det(S_{m-1} + w(\lambda)^{-1}I)|},$$

taking into account of (3.8) we obtain the a posteriori bound

$$\|R_m\| \leq \frac{\prod_{j=1}^{m-1} s_{j+1,j}}{\pi\alpha\nu} \int_{C(\varepsilon,\mu)} \frac{\left|\exp(\lambda^{\frac{1}{\alpha}})\lambda^{\frac{1-\beta}{\alpha}}\right|}{|\lambda|\,|\det(S_{m-1} + w(\lambda)^{-1}I)|} |d\lambda| \,.$$

We recall (c.f. [25]) that in several important situations the products $\prod_{j=1}^{k} s_{j+1,j}$ have a rapid decay, so producing a very fast convergence. Here we do not dwell more upon the implementation of such error bounds, to which we plan to devote a detailed study in a future work.

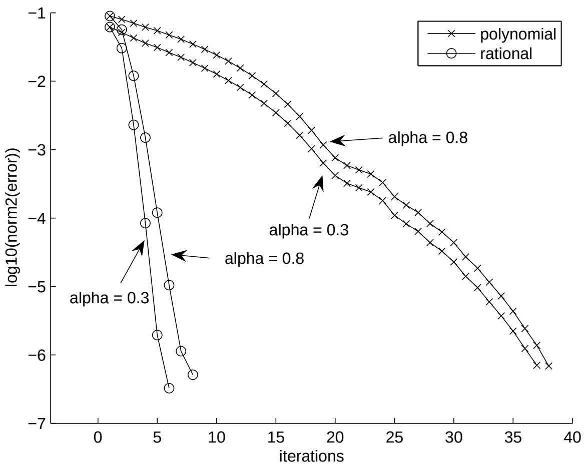

FIG. 5.1. *Approximation of (5.2) at* $t = 0.1$

5. Numerical experiments. In order to make some numerical comparisons between the rational and the polynomial method, we consider initial value problems of the type

$$\begin{aligned} {}_0D_t^{\alpha} y(t) + Ay(t) &= g(t), \quad t > 0, \\ y(0) &= y_0. \end{aligned} \tag{5.1}$$

Here we report some results concerning the two simple cases of $g(t) = 0$ and $g(t) = g \neq 0$ that lead respectively to the solutions

$$y(t) = E_{\alpha,1}(-t^{\alpha} A) y_0, \tag{5.2}$$

and

$$y(t) = y_0 + t^{\alpha} E_{\alpha,\alpha+1}(-t^{\alpha} A)(g - Ay_0). \tag{5.3}$$

Regarding the choice of the matrix $-A$ we discretize the 2-dimensional Laplacian operator in $(0,1) \times (0,1)$ with homogeneous Dirichlet boundary conditions using central differences on a uniform meshgrid of meshsize $\delta = 1/(n+1)$ in both directions. In this case equation (5.1) is a so-called Nigmatullin's type equation. In each example we have set $n = 30$ so that the dimension of the problem is $N = 900$. Moreover we set $y_0 = (1, ...1)^T/\sqrt{N}$.

In all the experiments, referring to the rational method, we have taken $h = 0.05$. In Figs. 5.1 and 5.2 we have plotted the error curves (with respect to a reference solution) of the rational and standard polynomial Arnoldi methods for the computation of $y(t)$ at $t = 0.1$ and $t = 1$, in the case of $g(t) = 0$ (i.e., with exact solution given by (5.2)) with $\alpha = 0.3$ and $\alpha = 0.8$.

Figs. 5.3 and 5.4 regard the approximation of $y(t)$ defined by (5.3) with $g = y_0/2$, again at $t = 0.1$ and $t = 1$. In both cases we have set $\alpha = 0.5$.

Finally in Fig. 5.5 we compare the error of the rational Arnoldi method with the error bound given by formula (4.20) for the computation of (5.2) in the case of

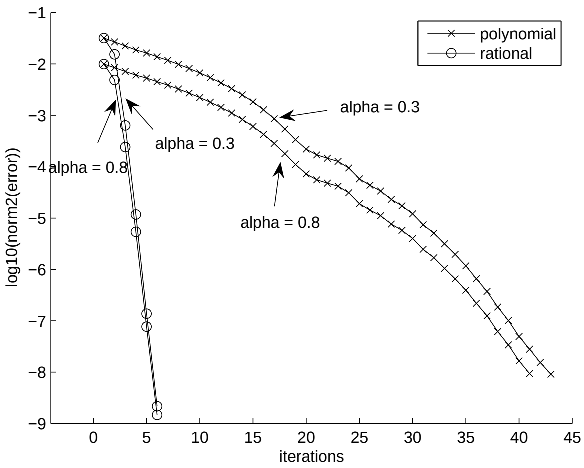

FIG. 5.2. *Approximation of (5.2) at* $t = 1$

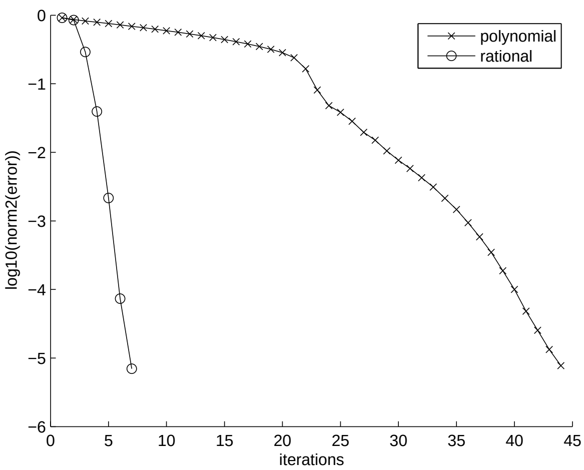

FIG. 5.3. *Approximation of (5.3) at* $t = 0.1$ *with* $\alpha = 0.5$.

$\alpha = 0.5$ and $t = 1$. For this case, accordingly with Corollary 4.6 we choose $h = 0.4$, that approximates the result of formula (4.22) when we set $m = 10$ as the expected number of iterations for the convergence.

REMARK 5.1. *The numerical experiments have been performed using Matlab. In particular for the computation of the projected functions of matrices we have used the classical approach based on the Schur decomposition together with the Matlab function* `MLF.m` *developed by Igor Podlubny and Martin Kacenak available at* `www.mathworks.com` *[28].*

REFERENCES

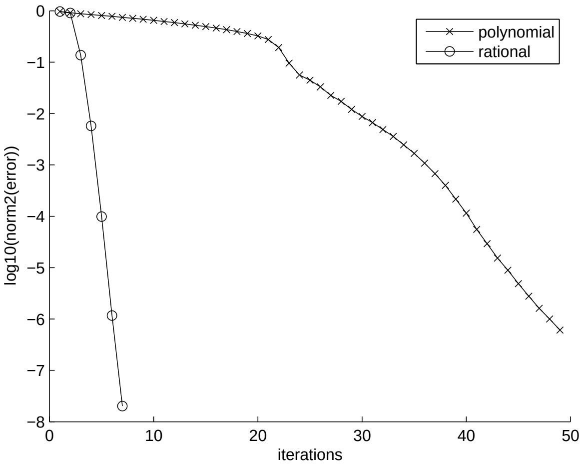

FIG. 5.4. *Approximation of (5.3) at $t = 1$ with $\alpha = 0.5$.*

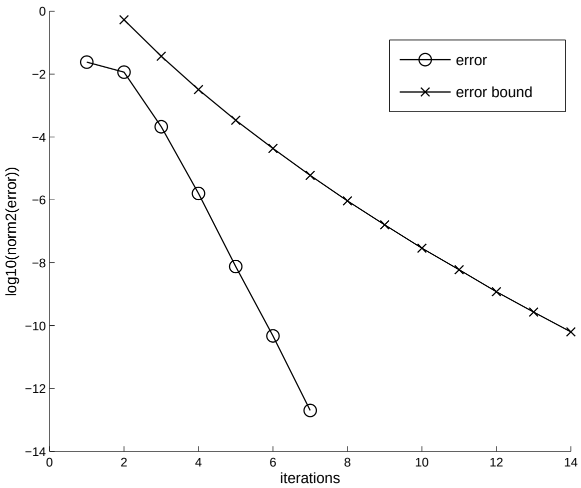

FIG. 5.5. *Error and error bound for the rational Arnoldi method applied to (5.2) with $\alpha = 0.5$, $t = 1$, $h = 0.4$.*

[1] B. BECKERMANN, *Image numérique, GMRES et polynômes de Faber*, C. R. Math. Acad. Sci., Paris 340 (2005), pp. 855–860.

[2] B. BECKERMANN AND L. REICHEL, *Error estimation and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883.

[3] M. CAPUTO, *Linear models of dissipation whose Q is almost frequency independent*, II. Geophys. J. Royal Astronom. Soc., 13 (1967), pp. 529–539.

[4] M. CROUZEIX, *Numerical range and numerical calculus in Hilbert space*, J. Functional Analysis, 244 (2007), pp. 668–690.

[5] M.CROUZEIX AND B.DELYON, *Some estimates for analytic functions of strip or sectorial operators*, Arch. Math., 81 (2003), pp. 553–566.

[6] F. DIELE, I. MORET AND S. RAGNI, *Error estimates for polynomial Krylov approximations to matrix functions*, SIAM J. Matrix Analysis and Appl., 30 (2008), pp. 1546–1565.

[7] K. Diethelm and N.J. Ford, *Analysis of fractional differential equations*, J. Math. Anal. Appl., 265 (2002), pp. 229–248.

[8] K. Diethelm, N.J. Ford, A.D. Freed and Yu. Luchko, *Algorithms for the fractional calculus: A selection of numerical methods*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 743–773.

[9] V. Druskin and L. Knizhnerman, *Two polynomial methods for calculating functions of symmetric matrices*, Comput. Maths. Math. Phys., 29(6) (1989), pp. 112–121.

[10] V. Druskin, L. Knizhnerman and M. Zaslavsky, *Solution of large scale evolutionary problems using rational Krylov subspaces with optimized shifts*, SIAM J. Sci. Comp., 31 (2009), pp. 3760–3780.

[11] J. v. d. Eshof and M. Hochbruck, *Preconditioning Lanczos approximations to the matrix exponential*, SIAM J. Sci. Comp., 27 (2005), pp. 1438–1457.

[12] R. W. Freund, *On Conjugate Gradient Type Methods and Polynomial Preconditioners for a Class of Complex Non-Hermitian Matrices*, Numer. Math., 57 (1990), pp. 285–312.

[13] A. Frommer and V. Simoncini, *Matrix Functions* in Model Order Reduction: Theory, Research Aspects and Applications, Mathematics in Industry, W. Schilders, and H. van der Vorst, eds., Springer, Heidelberg, in print.

[14] L. Galeone and R. Garrappa, *On multistep method for differential equations of fractional order*, Mediterr. J. Math., 99 (2006), pp. 1–16.

[15] V. Grimm and M. Hochbruch, *Rational approximations to trigonometric operators,* BIT, 48 (2008), pp. 215–229.

[16] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, 2008.

[17] M. Hochbruck and C. Lubich, *On Krylov subspace approximation to the matrix exponential operator*, SIAM J. Numer. Anal., 34 (1997), pp. 1911–1925.

[18] L. Knizhnerman, *Adaptation of the Lanczos and Arnoldi methods to the spectrum, or why the two Krylov subspace methods are powerful, Chebyshev Digest* 2002, 3 (2002), pp. 141–164.

[19] L. Knizhnerman and V. Simoncini, *A new investigation of the extended Krylov subspace method for matrix function evaluations*, Numer. Linear Algebra Appl., 17 (2010), pp.615-638.

[20] ST. Lee, HK. Pang and HW. Sun, *Shift-invert Arnoldi approximation to the Toeplitz matrix exponential,* SIAM Journal on Scientific Computing, 32 (2010), pp. 774-792.

[21] L. Lopez and V. Simoncini, *Analysis of projection methods for rational function approximation to the matrix exponential*, SIAM J. Numer. Anal., 44 (2006), pp. 613–635.

[22] I. Moret, *On RD-rational Krylov approximations to the core-functions of exponential integrators*, Numer. Linear Algebra Appl., 14 (2007), pp. 445–457.

[23] I. Moret, *Rational Lanczos approximations to the matrix square root and related functions*, Numer. Linear Algebra Appl., 16 (2009), pp. 431–445.

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

[25] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, 1993.

[26] P. Novati, *On the construction of Restricted-Denominator Exponential W-methods*, J. Comput. Appl. Math., 221 (2008), pp. 86–101.

[27] I. Podlubny, *Fractional Differential Equations*, Academic Press, New York, 1999.

[28] I. Podlubny and M. Kacenak, `MLF.m`, `www.mathworks.com`, 2001.

[29] M. Popolizio and V. Simoncini, *Acceleration techniques for approximating the matrix exponential operator*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 657–683.

[30] J. A. C. Weideman and L. N. Trefethen, *Parabolic and hyperbolic contours for computing the Bromwich integral,* Math. Comp., 76 (2007), pp. 1341–1356.