Numerical approximation to the fractional derivative operator

P. Novati
Department of Mathematics
University of Padova, Italy

Abstract

In this paper we consider the numerical approximation of A^α by contour integral. We are mainly interested to the case of A representing the discretization of the first derivative by means of a BDF formula, and $0 < \alpha < 1$. The computation of the contour integral yields a rational approximation to A^α which can be used to define k-step formulas for the numerical integration of Fractional Differential Equations.

1 Introduction

This paper deals with the numerical approximation of A^α, where $A \in \mathbb{R}^{n\times n}$ and $0 < \alpha < 1$ (see [12] Chapter 8 and the references therein for a background and a review on the most effective methods). While some of the arguments of the paper can be applied to matrices whose spectrum $\sigma(A)$ is such that $\sigma(A) \subset \mathbb{C}\setminus(-\infty, 0]$, we are mainly interested to the case of A representing a discretization of the first order derivative operator. In particular, denoting by $a_0, a_1, ..., a_p$ the $p+1$ coefficients of a Backward Differentiation Formula (BDF) of order p, with $1 \le p \le 6$, which discretizes the derivative operator (see [10] Chapter III.1 for a background), we consider lower triangular banded Toeplitz matrices of the type

$$A_p = \begin{pmatrix} a_0 & 0 & & & 0 \\ \vdots & a_0 & 0 & & \\ a_p & \vdots & \ddots & 0 & \\ 0 & \ddots & & \ddots & 0 \\ & 0 & a_p & \cdots & a_0 \end{pmatrix} \in \mathbb{R}^{n\times n}, \tag{1}$$

whose spectrum $\sigma(A_p)$ consists of the point $a_0 > 0$.

In this situation, $A_p^\alpha e_1$, $e_1 = (1, 0, ..., 0)^T$, contains the whole set of coefficients of the corresponding Fractional BDF (FBDF) formula for solving Fractional Differential Equations (FDEs, see e.g. [17] for an exhaustive overview) of

the type

$$ {}_{t_0}D_t^{\alpha} y(t) = g(t, y(t)), \quad t \in [t_0, T], \tag{2} $$

where ${}_{t_0}D_t^{\alpha}$ denotes the fractional derivative operator, and where we assume to consider a uniform discretization $t_0, t_1, ..., t_n = T$ of the time domain. FBDF formulas of order $p \geq 2$ have been introduced in [14], and extend the Grunwald-Letnikov discretization of the fractional derivative (see again [17]). We remark that the j-th entry of $A_p^{\alpha} e_1$ is just the j-th coefficient of the Taylor expansion of the generating function of the method

$$ \omega_p^{(\alpha)}(\zeta) = (a_0 + a_1\zeta + ... + a_p\zeta^p)^{\alpha}, \tag{3} $$

around $\zeta = 0$.

For a given analytic function f and a general square matrix A, we know that $f(A)$ can be represented by the contour integral

$$ f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(zI - A)^{-1} dz, \tag{4} $$

where Γ is a closed contour lying in the region of analyticity of f and enclosing the spectrum in its interior. Among the existing methods for the computation of $f(A)$ (end even $f(A)v$, $v \in \mathbb{R}^n$, see [12] for an overview), the approximation of (4) by a quadrature rule represents a strategy not much used so far. Recently, however, it has been successfully used in [11] and [22] to approximate the matrix functions here considered and the matrix functions involved in the Exponential Integrators for IVPs (the so-called phi-functions) respectively.

It is just the paper [11] which has given us the basic input for the computation of A^{α}, when A is of type (1), by means of the approximation of the contour integral. For matrices of this kind, a polynomial approximation of $A_p^{\alpha} e_1$ may lead to very poor results, especially when n is large and α is not close to 1. Indeed, in a situation like this, A_p^{α} is full lower triangular with entries that annihilate quite slowly departing from the diagonal. At the same time, since A_p has a narrow band ($p \leq 6$), any matrix polynomial $p_m(A_p)$ ($p_m \in \Pi_m$, where Π_m denotes the set of polynomials of degree at most m) is still lower triangular but the filling is attained only for $m \approx n/p$. Therefore any polynomial method for $A_p^{\alpha} e_1$ would require a number of iteration proportional to the dimension. This consideration, together with the fact that a linear system with A of type (1) can be solved with $O(n)$ floating point operations, leads us to consider rational approximations for A_p^{α}, based on the contour integral approximation. For completeness, we need to quote here the recent paper [8], where Padé type rational approximations to A^{α} are considered.

In this paper we partially follow the ideas presented in [11] for the special case of $\alpha = 1/2$. After a suitable change of variable, the contour integral representing A_p^{α} will be computed by the m-point Gauss-Legendre formula, leading to partial fraction approximations of the type

$$ A_p^{\alpha} \approx R_m(A_p) = A_p \sum\nolimits_{j=1}^{m} \gamma_j (\eta_j I + A_p)^{-1}, \tag{5} $$

$R_m(z) = p_m(z)/q_m(z)$, $p_m, q_m(z) \in \Pi_m$. The above formula can be regarded to as a new technique to evaluate the coefficients of a FBDF method, which in general is not a very simple task when n is large and $p > 1$. The most common approach seems to be the one based on the FFT ([17] Chapter 7.5), already employed in [9] in the case of $\alpha = 1/2$. More recently, in [21] the authors introduce a new approach based on the computation of the Laplace transform representing the formal power series of the unknown coefficients, using a quadrature rule on Talbot contours (see [24]) and hyperbolas.

Without the explicit computation of $R_m(A_p)e_1$, the definition of the coefficients γ_j, η_j, $j = 1, ..., m$, in (5), allows to construct the polynomials p_m and q_m of degree m such that $R_m(z) = p_m(z)/q_m(z)$ and then

$$A_p^\alpha \approx [q_m(A_p)]^{-1} p_m(A_p). \tag{6}$$

Note that $R_m(a_0 + a_1\zeta + ... + a_p\zeta^p)$ represents a rational approximation to the generating function (3). Writing a FBDF method for (2) in the matrix form

$$A_p^\alpha \mathbf{y} = h^\alpha \mathbf{g}, \tag{7}$$

where $\mathbf{y} = (y_0, ..., y_n)^T$, $\mathbf{g} = (h^{-\alpha}y_0, g_1..., g_n)^T$, being y_j an approximation of $y(t_j)$, and $g_j = g(t_j, y_j)$, the rational approximation (6) allows to define the implicit formula

$$p_m(A_p)\mathbf{y} = h^\alpha q_m(A_p)\mathbf{g}, \tag{8}$$

which asymptotically represents an mp-term finite difference equation. Computationally, the use of (8) represents a not negligible advantage since (7) is a full recursion because A_p^α is full lower triangular. While the order of the underlying FBDF formula is lost, as we shall see the Gauss-Legendre approximation of the contour integral which leads to (8), ensures a good simulation of the convergence and the linear stability properties of the FBDF method.

The paper is organized as follows. In Section 2 we study the contour integral representation of A_p^α. In Section 3 we present our methods for the numerical approximation of $A_p^\alpha e_1$, showing also some numerical examples. A theoretical error analysis (with some further experiments) is given in Section 4. In Section 5 we apply our techniques for the construction of k-step formulas for FDEs, and a numerical example on a linear problem is presented in Section 6.

2 The integral representation of A^α

As pointed out by many authors whenever the contour Γ of (4) is wide, it is more convenient to consider the integral representation of $A(A^{-1}f(A))$ if A is nonsingular. Therefore, A^α can be written as

$$A^\alpha = \frac{A}{2\pi i}\int_\Gamma z^{\alpha-1}(zI - A)^{-1}dz, \tag{9}$$

where Γ is a suitable contour. The following known result (see e.g. [1]), expresses A^α in terms of a real integral.

Proposition 1 *Let $A \in \mathbb{R}^{n\times n}$ be such that $\sigma(A) \subset \mathbb{C}\setminus(-\infty, 0]$. For $0 < \alpha < 1$ the following representation holds*

$$A^{\alpha} = \frac{A\sin(\alpha\pi)}{\alpha\pi}\int_0^{\infty}(\rho^{1/\alpha}I + A)^{-1}d\rho. \tag{10}$$

Before presenting our methods for the numerical approximation of (10) we give the following result, which can be proved by direct computation.

Proposition 2 *Let $A_p \in \mathbb{R}^{n\times n}$ be a matrix of type (1), and let $\overline{A}_p = \frac{1}{a_0}A_p$. Then the components of $(\tau I + \overline{A}_p)^{-1}e_1$, $\tau \neq -1$, are given by*

$$\begin{aligned} v_1^{(p)}(\tau) &= \frac{1}{\tau+1}, \\ v_j^{(p)}(\tau) &= \frac{c_{2,j}^{(p)}}{(\tau+1)^2} + ... + \frac{c_{j,j}^{(p)}}{(\tau+1)^j}, \quad 2 \le j \le n. \end{aligned}$$

where the coefficients $c_{i,j}^{(p)}$ depend on the order p. For $p = 1$ we simply have $\{a_0, a_1\} = \{1, -1\}$, and hence

$$v_j^{(1)}(\tau) = \frac{1}{(\tau+1)^j}, \quad 1 \le j \le n.$$

2.1 First approach, for $0 < \alpha \le 1/2$

For the computation of (10) we may consider the same approach presented in [11], where, for the special case of $\alpha = 1/2$, the authors consider a change of variable involving Jacobi elliptic functions. In our case, and for each $0 < \alpha \le 1/2$, this transformation reads

$$\rho = a_0^{\alpha}\tan\theta, \tag{11}$$

so that, by Proposition 1,

$$A_p^{\alpha} = \frac{\overline{A}_p a_0^{\alpha}\sin(\alpha\pi)}{\alpha\pi}\int_0^{\pi/2}(\tan^{1/\alpha}\theta I + \overline{A}_p)^{-1}\left(\tan^2\theta + 1\right)d\theta, \tag{12}$$

where $\overline{A}_p$ is defined as in Proposition 2.

In order to understand the reliability of the change of variable (11), and remembering that we are just interested in computing $A_p^{\alpha}e_1$, we need to study the regularity of the components of the vector

$$\begin{aligned} v^{(p)}(\theta) &= \left(\tan^2\theta + 1\right)(\tan^{1/\alpha}\theta I + \overline{A}_p)^{-1}e_1, \\ &= \left(\tan^2\theta + 1\right)\left(v_1^{(p)}(\tan^{1/\alpha}\theta), ..., v_n^{(p)}(\tan^{1/\alpha}\theta)\right)^T, \end{aligned}$$

in $[0, \pi/2]$. By Proposition 2 the components of $v^{(p)}(\theta)$ contains functions of the type

$$f_j(\theta) = \frac{\tan^2\theta + 1}{\left(\tan^{1/\alpha}\theta + 1\right)^j}, \quad j = 1, ..., n, \quad \theta \in [0, \pi/2].$$

For $0 < \alpha \leq 1/2$, $f_j(\theta)$ is bounded in $[0, \pi/2]$ for each $j \geq 1$, since $f_j(\theta) \to 1$ as $\theta \to 0$ and

$$f_j(\theta) \sim (\cos\theta)^{j/\alpha - 2} \to 0 \quad \text{as} \quad \theta \to \pi/2. \tag{13}$$

Note that the functions $f_j(\theta)$ are analytic in some open subset containing $[0, \pi/2]$ for $\alpha = 1/q$, $q = 2, 3, ...$ (i.e., when we consider the matrix q-root), and only continuous in $[0, \pi/2]$ for other values. On the other side, by (13), for $1/2 < \alpha < 1$, $f_1(\theta) \to \infty$ as $\theta \to \pi/2$, and hence the substitution (11) does not appear to be reliable anymore.

2.2 Second approach, for $1/2 \leq \alpha < 1$

For $1/2 \leq \alpha < 1$, we use the slightly different substitution

$$\rho = a_0^{\alpha} \tan^{\beta}\theta, \tag{14}$$

where we want to define $\beta > 0$ such that the arising integrand in (10) remains bounded in $[0, \pi/2]$. With this substitution we obtain

$$A_p^{\alpha} = \beta \frac{\overline{A}_p a_0^{\alpha} \sin(\alpha\pi)}{\alpha\pi} \int_0^{\pi/2} (\tan^{\beta/\alpha}\theta I + \overline{A}_p)^{-1} \left(\tan^2\theta + 1\right) \tan^{\beta-1}\theta d\theta. \tag{15}$$

Proceeding as before we have now to examine the behavior of the functions

$$f_j(\theta) = \frac{\left(\tan^2\theta + 1\right)\tan^{\beta-1}\theta}{\left(\tan^{\beta/\alpha}\theta + 1\right)^j}.$$

We have that

$$f_j(\theta) \sim (\cos\theta)^{j\beta/\alpha - 2 - (\beta-1)} \quad \text{as} \quad \theta \to \pi/2,$$

and hence defining

$$\beta = \frac{\alpha}{1-\alpha},$$

which is the solution of

$$\frac{\beta}{\alpha} - 2 - (\beta - 1) = 0,$$

we have that $f_1(\theta) \to 1$ and $f_j(\theta) \to 0$, for $j \geq 2$, as $\theta \to \pi/2$. Since $1 \leq \beta < \infty$ for $1/2 \leq \alpha < 1$, we have that $f_j(\theta) \to 0$ for $\theta \to 0$, for each $j \geq 1$. It is easy to see that the functions $f_j(\theta)$ are analytic in some open subset containing $[0, \pi/2]$ for $\alpha = 1 - 1/q$, $q = 2, 3, ...$, and only continuous for other values. Observe that the method does not apply to the case $0 < \alpha < 1/2$, since we would have $\beta - 1 < 0$ and hence $f_j(\theta) \to \infty$ as $\theta \to 0$. Observe moreover that for $\alpha = 1/2$ we have $\beta = 1$, and the method reduces to the previous one.

2.3 Third approach, for $0 < \alpha < 1$

In order to define a substitution that allows to have a bounded integrand function in (10) for each $0 < \alpha < 1$, we consider the change of variable

$$\rho = a_0^\alpha \cos^{-\beta}\theta \sin\theta. \tag{16}$$

With this substitution in (10) we obtain

$$A_p^\alpha = \frac{\overline{A}_p a_0^\alpha \sin(\alpha\pi)}{\alpha\pi} \int_0^{\pi/2} (\cos^{-\beta/\alpha}\theta \sin^{1/\alpha}\theta I + \overline{A}_p)^{-1} \frac{\beta \sin^2\theta + \cos^2\theta}{\cos^{\beta+1}\theta} d\theta. \tag{17}$$

Defining as before the functions

$$f_j(\theta) = \frac{\beta \sin^2\theta + \cos^2\theta}{(\cos^{-\beta/\alpha}\theta \sin^{1/\alpha}\theta + 1)^j \cos^{\beta+1}\theta},$$

we have that $f_j(\theta) \to 1$ for $\theta \to 0$ and for each $j \geq 1$. Moreover

$$f_j(\theta) \sim \beta\,(\cos\theta)^{j\beta/\alpha-\beta-1} \quad \text{as} \quad \theta \to \pi/2,$$

and hence, defining as before $\beta = \frac{\alpha}{1-\alpha}$, we have $f_1(\theta) \to \beta$ and $f_j(\theta) \to 0$, for $j \geq 2$, as $\theta \to \pi/2$, for each $0 < \alpha < 1$. Note that even in this case the method reduces to the first one for $\alpha = 1/2$. In this situation the functions $f_j(\theta)$ are of class C^1 in $[0, \pi/2]$, and, of course, analytic for $\alpha = 1/2$.

3 The Gauss-Legendre approximations

For the three substitutions presented in the previous section, we consider the Gauss-Legendre approximation of the corresponding integrals (12), (15), (17). For the case $\alpha = 1/2$, in [11] the authors consider the trapezoidal rule, reproducing in that way the rational approximation of the square root function proposed by Zolotarev in 1877, which has shown to be extremely efficient for the approximation of this function on wide real intervals. Here, the situation is quite different, since the spectrum of A_p reduces to one point.

Denoting by x_k and w_k, $k = 1, ..., m$, the nodes and the weights of the m-point Gauss-Legendre rule, and defining

$$\theta_k = \frac{\pi}{4}(x_k + 1), \quad k = 1, ..., m, \tag{18}$$

from (12), (15) and (17) we obtain the following three methods, which coincide for $\alpha = 1/2$.

Method 1, for $0 < \alpha \leq 1/2$

$$A_p^\alpha \approx \frac{\overline{A}_p a_0^\alpha \sin(\alpha\pi)}{4\alpha} \sum_{k=1}^{m} w_k (\tan^{1/\alpha}\theta_k I + \overline{A}_p)^{-1} \left(\tan^2\theta_k + 1\right) \tag{19}$$

Method 2, for $1/2 \le \alpha < 1$

$$A_p^\alpha \approx \frac{\overline{A}_p a_0^\alpha \sin(\alpha\pi)}{4\,(1-\alpha)} \times \sum_{k=1}^{m} w_k (\tan^{1/(1-\alpha)}\theta_k I + \overline{A}_p)^{-1} \left(\tan^2\theta_k + 1\right) \tan^{(2\alpha-1)/(1-\alpha)}\theta_k \quad (20)$$

Method 3, for $0 < \alpha < 1$

$$A_p^\alpha \approx \frac{\overline{A}_p a_0^\alpha \sin(\alpha\pi)}{4\alpha} \times \sum_{k=1}^{m} w_k (\cos^{1/(\alpha-1)}\theta_k \sin^{1/\alpha}\theta_k I + \overline{A}_p)^{-1} \frac{\frac{\alpha}{1-\alpha}\sin^2\theta_k + \cos^2\theta_k}{\cos^{1/(1-\alpha)}\theta_k} \quad (21)$$

In order to appreciate the potential of the methods just described, below we present some numerical experiments. In each examples we consider the matrices of type (1) arising from the BDF formulas of order $p = 1, 2, 3, 4$, whose sets of coefficients $\{a_0, a_1, ..., a_p\}$ are given by

$$\begin{aligned} p &= 1: \quad \{1, -1\} \\ p &= 2: \quad \{3/2, -2, 1/2\} \\ p &= 3: \quad \{11/6, -3, 3/2, -1/3\} \\ p &= 4: \quad \{25/12, -4, 3, -4/3, 1/4\} \end{aligned}$$

For each experiment the dimension of A_p is $n = 500$, and the number of nodes m of the quadrature formulas is between 2 and 16. All the computation are performed in Matlab. Denoting by $R_m(A_p)$ the rational approximation of A_p^α arising by one of the three methods, in each picture we plot the relative error

$$\frac{\left\|R_m(A_p)e_1 - A_p^\alpha e_1\right\|_2}{\left\|A_p^\alpha e_1\right\|_2},$$

where the reference solution $A_p^\alpha e_1$ is computed with the Matlab functions `expm` and `logm`.

Example 1. We take $\alpha = 1/2$ and compare our Method 1 with the trapezoidal rule applied to (12). For this example we also consider the diagonal (m, m) Padé approximations of the function $\sqrt{1+z}$, used in [26] and [8] for the computation of the matrix square root,

$$\sqrt{1+z} \approx 1 + \sum\nolimits_{k=1}^{m} \frac{a_k^{(m)} z}{1 + b_k^{(m)} z}, \quad (22)$$

where

$$a_k^{(m)} = \frac{2}{2m+1}\sin^2\frac{k\pi}{2m+1}, \quad b_k^{(m)} = \cos^2\frac{k\pi}{2m+1}.$$

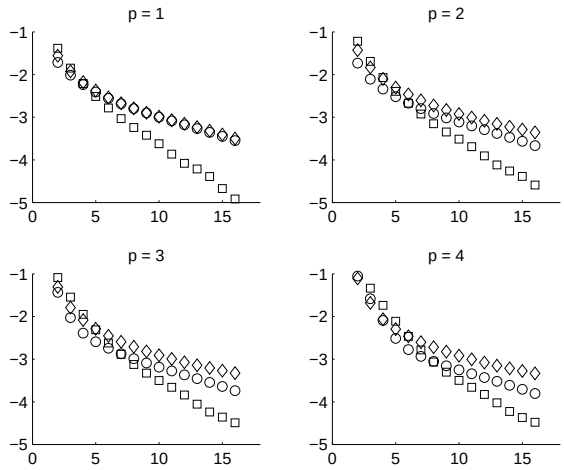

Figure 1: Relative error (in logarithmic scale) with respect to m of the trapezoidal rule applied to (12) (diamonds), diagonal Padé approximation (circles) and Method 1 (squares), for $\alpha = 1/2$.

The results are reported in Figure 1.

Example 2. In this example we consider the case of $\alpha < 1/2$. In particular we consider the values $\alpha = 1/5$ and $\alpha = 1/3$ comparing Method 1 and 3. Figures 2 and 3 display the results, that also include the behavior of the trapezoidal rule applied to (12).

Comparing Figure 2 and 3, we can observe that for the case $\alpha = 1/3$ there is little difference between the methods, but this is not true in general for smaller values of α. Indeed for $\alpha = 1/5$, Method 3 is much more accurate, and the experiments reveal that this difference increases for α close to 0.

Example 3. In this third example we consider the case of $\alpha > 1/2$ taking $\alpha = 2/3$ and $\alpha = 4/5$. We compare Method 2 and 3. Figures 4 and 5 display the results. Symmetrically (with respect to $\alpha = 1/2$) to the case of $\alpha < 1/2$, the difference between the methods becomes well marked for α close to 1. In this example we also report the behavior of the trapezoidal rule applied to (15).

It is clear from the above examples, that Method 3 seems to be able to approximate the discrete fractional derivative with an error around 10^{-4} in the range $\alpha \in [1/5, 4/5]$, with $m \leq 16$ quadrature points. If the goal is to design alternative methods for FDEs, this approach seems to be promising. On the other side, if one would rather compute with high accuracy the FBDF coefficients $A_p^\alpha e_1$, then larger values of m should be considered. It is also necessary to point out that for α small or close to 1 (Figure 2 and 5) in some cases the trapezoidal rule allows faster convergence. On the other side, for α near $1/2$, which is very important in the field of FDEs, the Gauss-Legendre approximation guarantees better results (Figure 1, 3 and 4).

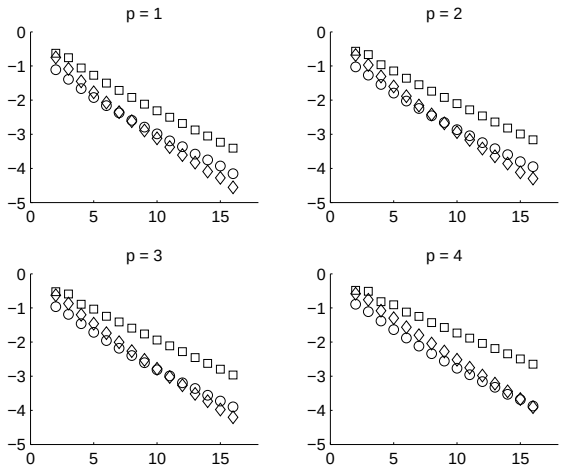

Figure 2: Relative error (in logarithmic scale) with respect to m of Method 1 (squares), Method 3 (circles) and trapezoidal rule applied to (12) (diamonds) for $\alpha = 1/5$.

We remark moreover that, independently of the quadrature rule adopted, our methods with $m = 16$ require about 0.1 seconds (for each p) to perform the computation on a simple PC, as the one we have used. On the contrary, the computation of $A_p^\alpha e_1$ by means of the Matlab instruction `expm(alpha*logm(A'))'*e1` takes about 8 seconds for $n = 500$ (the transpose of A_p is used to avoid the Schur decomposition). We remark however that this Matlab instruction cannot be considered a meaningful benchmark, since the structure of A_p cannot be fully exploited as in a rational approximation.

4 Error analysis

In this section we provide a theoretical error analysis for Method 1, in the case of $\alpha = 1/k$, $k = 2, 3, \ldots$. With some effort the result should be extendible to Method 2. Let $\chi(z)$ be a function analytic in $[-1, 1]$. For $r > 1$, let

$$\Phi_r = \left\{ z \in \mathbb{C} : z = \frac{1}{2}\left(re^{i\vartheta} + \frac{1}{re^{i\vartheta}} \right), \vartheta \in [0, 2\pi] \right\},$$

be an ellipse of the complex plane with foci in ± 1. Let moreover $R > 1$ be the smallest real number such that Φ_R contains a singularity of the function $\chi(z)$. Denoting by $E_m(\chi)$ the error of the m-point Gauss-Legendre quadrature rule, that is,

$$E_m(\chi) = \int_{-1}^{1} \chi(x)dx - \sum_{k=1}^{m} w_k \chi(x_k),$$

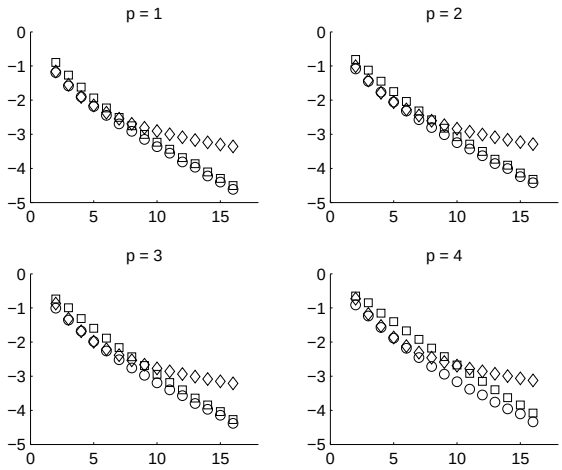

Figure 3: Relative error (in logarithmic scale) with respect to m of Method 1 (squares), Method 3 (circles) and trapezoidal rule applied to (12) (diamonds), for $\alpha = 1/3$.

from a series of results dating back to the late 60' ([3], [4], [13], [19], [23]), we know that the error can be bounded as follows

$$|E_m(\chi)| \leq C \frac{M(r)}{(r^2-1)\, r^{2m}}, \quad 1 < r < R, \tag{23}$$

where the constant C can be taken independent of m and r and

$$M(r) = \max_{\Phi_r} |\chi(z)| = \max_{[0,2\pi]} \left| \chi\left(\frac{1}{2}\left(re^{i\vartheta} + \frac{1}{re^{i\vartheta}} \right) \right) \right|.$$

For a recent review about Gauss-Legendre formula, we refer to [25] and the references therein.

Formula (23) may be difficult to use in practice, and in general it may be also rather pessimistic. Anyway, it can help us to have a better insight about the behavior of our methods with respect to some values of α. Indeed, whenever α is such that the functions $f_j(\theta)$ arising from our substitutions (11) and (14) are analytic, the location of the poles of these functions plays a crucial role since it defines R in (23). In the z-plane, looking at the corresponding $f_j\left(\frac{\pi}{4}(z+1)\right)$ (cf. (18)), for the substitution (11) (that is, for Method 1) the poles that define R, are given explicitly by

$$\begin{aligned} z &= \infty, \quad \text{for} \quad \alpha = 1/2, \\ z &= \frac{4}{\pi} \arctan\left(e^{i\alpha\pi}\right) - 1, \quad \text{for} \quad \alpha = 1/3, 1/4, ... \end{aligned} \tag{24}$$

For the substitution (14) the poles are the same: $z = \frac{4}{\pi} \arctan\left(e^{i(1-\alpha)\pi}\right) - 1$, for $\alpha = 1 - 1/q$, $q = 3, 4, ...$ The analysis cannot be applied to the substitution (16)

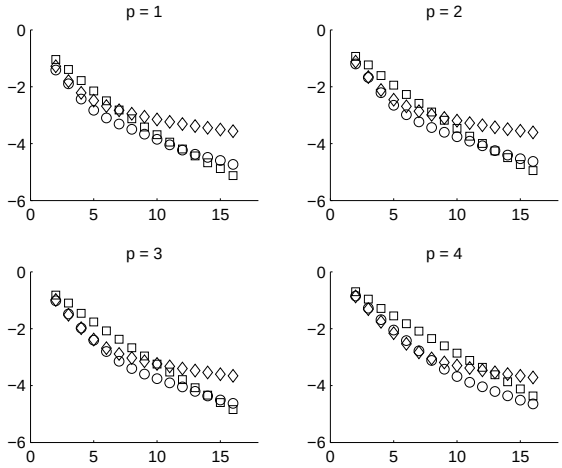

Figure 4: Relative error (in logarithmic scale) with respect to m of Method 2 (squares), Method 3 (circles) and trapezoidal rule applied to (15) (diamonds), for $\alpha = 2/3$.

(Method 3), since the corresponding functions are analytic only for $\alpha = 1/2$. For each method, the functions $f_j(\theta)$ are thus entire for $\alpha = 1/2$. In Table 1 we report the poles of the Method 1 and 2 for some values of q, together with the corresponding value of $R = R(\alpha)$ that is obtained by solving

$$\frac{1}{2}\left(Re^{i\vartheta} + \frac{1}{Re^{i\vartheta}}\right) = z,$$

that is,

$$R = b + \sqrt{b^2 + 1}, \quad \text{where} \quad b = \mathrm{Im}(\frac{4}{\pi}\arctan\left(e^{i\alpha\pi}\right)).$$

q	2	3	4	5	6	7	8	9	10
z	∞	$0.838i$	$0.561i$	$0.429i$	$0.350i$	$0.296i$	$0.257i$	$0.227i$	$0.203i$
$R(\alpha)$	∞	2.143	1.708	1.517	1.409	1.339	1.289	1.252	1.224

Table 1 - Poles in the z-plane of Method 1 and 2, and corresponding values of $R(\alpha)$.

As we can see, for α near 0 or 1, the poles are close to the interval $[-1, 1]$, so that R is close to 1 and the methods are expected to be slow (with respect to m) by (23). On the contrary, for α close to $1/2$, R can be taken larger, and the methods will be faster. This considerations are confirmed by the experiments of Figures 1-5.

Now, for $j = 1, ..., n$, let

$$E_{m,j}^{(p)} = \left| e_j^T R_m(A_p) e_1 - e_j^T A_p^\alpha e_1 \right|, \tag{25}$$

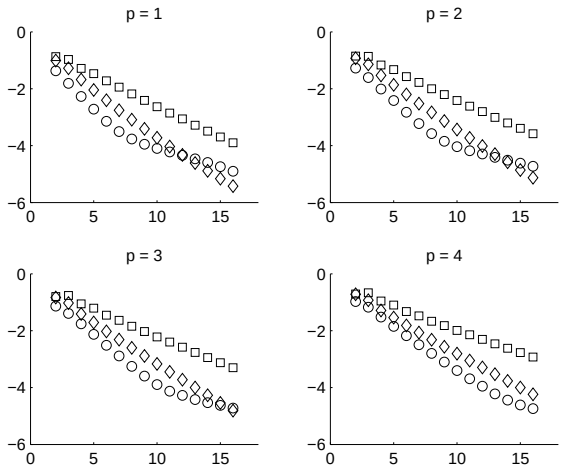

Figure 5: Relative error (in logarithmic scale) with respect to m of Method 2 (squares), Method 3 (circles) and trapezoidal rule applied to (15) (diamonds), for $\alpha = 4/5$.

be the error corresponding to the j-th component of $A_p^\alpha e_1$ obtained using one of our methods. In what follows we focus the analysis on Method 1.

In order to use (23) to bound (25), we need to work in the z-plane with the functions (cf. (12) and (18))

$$\psi_j^{(p)}(z) := e_j^T \overline{A}_p \left(\tan^{1/\alpha} \left(\frac{\pi}{4}(z+1) \right) I + \overline{A}_p \right)^{-1} e_1 \left(\tan^2 \left(\frac{\pi}{4}(z+1) \right) + 1 \right),$$

for $\alpha = 1/q$, $q = 2, 3, ...$, so that

$$E_{m,j}^{(p)} \leq C \frac{a_0^\alpha \sin(\alpha\pi)}{4\alpha} \frac{\max_{\Phi_r} \left| \psi_j^{(p)}(z) \right|}{(r^2-1)\, r^{2m}}, \quad 1 < r < R. \tag{26}$$

Let us define the functions

$$d_\alpha(\xi) = \tan^{1/\alpha} \left(\frac{\pi}{4} \left(\frac{1}{2} \left(\xi + \frac{1}{\xi} \right) + 1 \right) \right),$$

for $\xi = re^{i\theta}$, $1 < r < R$. Remember that $d_\alpha(\xi) + 1 = 0$ for $\xi = iR$.

Proposition 3 *For $z = \frac{1}{2}\left(\xi + \frac{1}{\xi}\right)$ we have*

$$\psi_j^{(p)}(z) = \frac{s_j^{(p)}(d_\alpha(\xi))(d_{1/2}(\xi)+1)}{(d_\alpha(\xi)+1)^j},$$

where $s_j^{(p)}$ is a polynomial of degree less than j, whose coefficients depend on $a_0, a_1, ..., a_p$. For $p = 1$, we simply have $s_1^{(p)}(x) \equiv 1$ and $s_j^{(p)}(x) = -x$.

Proof. The result follows straightfully from Proposition 2. ■

Before proceeding, we need the following two lemmas.

Lemma 4 *Let* $0 < \alpha \leq 1/2$ *and* $\xi = re^{i\theta}$. *Then for* $0 \leq \theta \leq 2\pi$

$$\frac{1}{|d_\alpha(\xi)+1|} \leq K(r)\frac{1}{|d_\alpha(ir)+1|}, \quad \textit{for} \quad 1 < r < R, \tag{27}$$

with $K(r) \leq 2$. *Moreover, for* $0 < \alpha < 1/2$, $K(r) \leq 3/2$ *for* $(R+1)/2 \leq r < R$.

Proof. Consider the auxiliary function

$$\varphi(\vartheta, r) := \left| \frac{d_\alpha(ir)+1}{d_\alpha(re^{i\theta})+1} \right|, \quad 1 \leq r \leq R,$$

that is 2π-periodic and such that $\varphi(\vartheta, r) = \varphi(2\pi - \vartheta, r)$. Hence we can restrict our analysis to $\vartheta \in [0, \pi]$. For each $1 < r < R$, one shows that $\varphi(\vartheta, r) \geq \varphi(0, r)$ and $\varphi(\vartheta, r)$ is monotonically increasing in $[0, \pi/2)$. Thus, the maximum with respect to ϑ is attained in $[\pi/2, \pi]$. Moreover, for each $\vartheta \in [\pi/2, \pi]$, $\varphi(\vartheta, r)$ is monotonically decreasing with respect to r. Thus we have

$$\begin{aligned} \max_{\vartheta\in[0,\pi], r\in[1,R]} \varphi(\vartheta, r) &= \max_{\vartheta\in[\pi/2,\pi], r\in[1,R]} \varphi(\vartheta, r) \\ &\leq \max_{\vartheta\in[\pi/2,\pi]} \varphi(\vartheta, 1) \\ &= \varphi(\pi, 1) \\ &= 2 \end{aligned}$$

The second part of the lemma is obtained evaluating

$$\max_{\vartheta\in[\pi/2,\pi]} \varphi(\vartheta, (R+1)/2).$$

■

Lemma 5 *Let* $\alpha^* \approx 0.271$ *be the solution of*

$$\frac{\pi}{8\alpha}\left(1 + \frac{1}{R^2}\right)\left|1 + e^{2i\alpha\pi}\right| = \frac{2}{R-1}.$$

Then for $1 \leq r \leq R$

$$|d_\alpha(ir)+1| \geq \mu_\alpha(R-r),$$

where

$$\mu_\alpha = \begin{cases} \frac{2}{R-1} & \textit{for} \quad 0 < \alpha \leq \alpha^* \\ \frac{\pi}{8\alpha}\left(1 + \frac{1}{R^2}\right)\left|1 + e^{2i\alpha\pi}\right| & \textit{for} \quad \alpha^* \leq \alpha < 1/2 \end{cases}. \tag{28}$$

Proof. We observe that $|d_\alpha(ir)+1| \to 2$ for $r \to 1$ and $|d_\alpha(ir)+1| \to 0$ for $r \to R$. Moreover one proves that

$$\begin{aligned}
\frac{d}{dr}\,|d_\alpha(ir)+1| &< 0 \quad \text{for} \quad 1<r<R,\\
\frac{d^2}{dr^2}\,|d_\alpha(ir)+1| &< 0 \quad \text{for} \quad r=1,\\
\frac{d^2}{dr^2}\,|d_\alpha(ir)+1| &> 0 \quad \text{for} \quad r=R,
\end{aligned}$$

and that the second derivative is equal to 0 only once in $1<r<R$. Therefore, since

$$\frac{d}{dr}\,|d_\alpha(ir)+1|\Big|_{r=R} = -\frac{\pi}{8\alpha}\left(1+\frac{1}{R^2}\right)\left|1+e^{2i\alpha\pi}\right|$$

one easily obtains the result. ■

For Method 1, we can state the following result, that holds for $p=1$.

Theorem 6 *Let $\alpha = 1/q$, $q = 3, 4, \ldots$. We have*

$$E^{(1)}_{m,j} \le \frac{K}{R-1} C_\alpha^j \left(\frac{2}{R+1}\right)^{2m}, \quad \textit{for} \quad m \ge 1, \tag{29}$$

where

$$C_\alpha = \frac{3}{\mu_\alpha (R-1)},$$

and

$$E^{(1)}_{m,j} \le \frac{K}{R-1}\left(\frac{4me}{j\mu_\alpha}\right)^j \exp\left(\frac{j^2}{2m}\right)\frac{1}{R^{2m+j}}, \quad \textit{for} \quad m \ge \frac{j(R+1)}{2(R-1)}, \tag{30}$$

where K is a constant depending on α but not on j and m.

Proof. Let us define

$$M_j(r) := \max_{\Phi_r} \left|\psi_j^{(1)}(z)\right|, \tag{31}$$

so that by (26) we have

$$E^{(1)}_{m,j} \le C\frac{\sin(\alpha\pi)}{4\alpha}\frac{M_j(r)}{(r^2-1)\,r^{2m}}. \tag{32}$$

By Proposition 3, for $z = \frac{1}{2}\left(\xi + \frac{1}{\xi}\right)$

$$\psi_j^{(1)}(z) = \frac{-d_\alpha(\xi)(d_{1/2}(\xi)+1)}{(d_\alpha(\xi)+1)^j}, \quad j>1.$$

Now, since the function $\left|d_\alpha(\xi)(d_{1/2}(\xi)+1)\right|$ is bounded in the annulus

$$\left\{\xi = re^{i\theta} : 1 < \overline{r} \leq r \leq R, 0 \leq \theta \leq 2\pi\right\},$$

we have that

$$M_j(r) \leq \overline{K}(r) \max_{|\xi|=r} \left| \frac{1}{(d_\alpha(\xi)+1)^j} \right|, \quad \text{for} \quad 1 < r < R, \tag{33}$$

where $\overline{K}(r) < \infty$ is a function of r, which can be uniformly bounded in $[\overline{r}, R]$. By (33) we can restrict our analysis on the function $d_\alpha(\xi)$. For $j = 1$ (33) still holds since $\psi_1^{(1)}(z) = (d_{1/2}(\xi)+1)\,(d_\alpha(\xi)+1)^{-1}$, by Proposition 3.

By Lemmas 4 and 5 we have that

$$M_j(r) \leq \overline{K}(r) \left(\frac{K(r)}{\mu_\alpha(R-r)} \right)^j, \quad \text{for} \quad 1 < \overline{r} \leq r \leq R, \tag{34}$$

where $K(r)$ is the function appearing in (27). Taking $r = (R+1)/2$, and using again the above lemmas we obtain

$$M_j(r) \leq K_\alpha \left(\frac{3}{\mu_\alpha(R-1)} \right)^j,$$

where

$$K_\alpha := \max_{\frac{R+1}{2} \leq r \leq R} \overline{K}(r).$$

Using the inequality

$$\frac{1}{\left(\frac{R+1}{2}\right)^2 - 1} \leq \frac{1}{R-1}, \tag{35}$$

in (32), we obtain (29), in which $K = CK_\alpha \frac{\sin(\alpha\pi)}{4\alpha}$.

In order to demonstrate (30), observe that for each $0 < \alpha < 1/2$ the minimum of

$$\frac{1}{(R-r)^j r^{2m}}, \tag{36}$$

with respect to r (cf. (32) and (34)), is obtained for

$$r = \frac{2m}{j+2m} R. \tag{37}$$

Since we need $r > 1$, we can use the above expression only for $m > \frac{j}{2(R-1)}$. In particular, taking m such that

$$\frac{2m}{j+2m} R \geq \frac{R+1}{2},$$

that is,

$$m \geq \frac{j(R+1)}{2(R-1)},$$

using (37) we obtain

$$\frac{M_j(r)}{r^{2m}} \leq K_\alpha \left(\frac{2}{Rc_\alpha}\right)^j \left(\frac{2m+j}{j}\right)^j \left(\frac{2m+j}{2m}\right)^{2m} \frac{1}{R^{2m}}.$$

Now, using the bound

$$\left(\frac{2m+j}{j}\right)^j \leq \left(\frac{2m}{j}\right)^j \exp\left(\frac{j^2}{2m}\right),$$

and (35) for the term $1/(r^2-1)$ in (32), we obtain (30). ■

Remark 7 *By (28), in Theorem 6 we have $C_\alpha = 3/2$ for $\alpha = 1/4, 1/5, ...,$ and $C_\alpha \approx 1.83$ for $\alpha = 1/3$.*

Theorem 8 *Let $\alpha = 1/2$. We have*

$$E^{(1)}_{m,j} \leq K \frac{\cosh^{j-1}\left(\frac{2m}{j}\right)}{\left(\left(\frac{8m}{j\pi}\right)^2 - 1\right)} \left(\frac{j\pi}{8m}\right)^{2m}, \quad \textit{for} \quad m > \max\left(\frac{j\pi}{8}, 1\right), \tag{38}$$

where K is a constant independent of j and m.

Proof. For $\alpha = 1/2$, we have that for $j \geq 2$ and $z = \frac{1}{2}\left(\xi + \frac{1}{\xi}\right)$

$$\psi^{(1)}_j(z) = -\frac{d_\alpha(\xi)}{\left(d_{1/2}(\xi)+1\right)^{j-1}}. \tag{39}$$

Defining as before

$$M_j(r) := \max_{\Phi_r} \left|\psi^{(1)}_j(z)\right| = \max_{|\xi|=r} \left|\frac{d_{1/2}(\xi)}{\left(d_{1/2}(\xi)+1\right)^{j-1}}\right|,$$

by (23) and (19) we have

$$E^{(1)}_{m,j} \leq \frac{C}{2} \frac{M_j(r)}{(r^2-1)\, r^{2m}}.$$

Following the proof of Theorem 6 we also have

$$M_j(r) \leq \overline{\overline{K}}(r) \max_{C_r} \left|\frac{1}{\left(d_{1/2}(\xi)+1\right)^{j-1}}\right|, \quad \text{for} \quad 1 < r < \infty, \tag{40}$$

where $\overline{\overline{K}}(r) < \infty$ is a function of r, which can be uniformly bounded in $[1, \infty)$. For $j = 1$ (40) still holds since $\psi^{(1)}_j(z) = 1$.

Now using Lemma 4 and well known trigonometric relations, we have

$$\begin{aligned}
\frac{1}{\left|d_{1/2}(\xi)+1\right|} &\leq \frac{2}{\left|d_{1/2}(ir)+1\right|} \\
&= 2\left|\cos^2\left(\frac{\pi}{4}\left(\frac{i}{2}\left(r-\frac{1}{r}\right)+1\right)\right)\right| \\
&\leq 2\left|\cos^2\left(\frac{\pi}{4}\left(\frac{ir}{2}+1\right)\right)\right| \\
&= 2\left(\cosh^2\left(\frac{\pi}{8}r\right)-\frac{1}{2}\right) \\
&= \cosh\left(\frac{\pi}{4}r\right),
\end{aligned}$$

so that

$$E_{m,j}^{(1)} \leq \frac{C}{2}K_{1/2}\frac{\cosh^{j-1}\left(\frac{\pi}{4}r\right)}{\left(r^2-1\right)r^{2m}}. \tag{41}$$

where $K_{1/2} := \max_{[1,\infty)} \overline{\overline{K}}(r)$. In order to define a suitable value for r we consider the approximation

$$\cosh\left(\frac{\pi}{4}r\right) \approx \frac{\exp\left(\frac{\pi}{4}r\right)}{2},$$

so that we can easily minimize with respect to r the function

$$\frac{\left(\exp\left(\frac{\pi}{4}r\right)\right)^j}{r^{2m}},$$

which yields $r = \frac{8m}{j\pi}$. Substituting this value in (41) yields (38). ■

While restricted to the case of $p = 1$, the results of Theorems 6 and 8 are representative of what happens also for $p > 1$ since by Proposition 2 and 3 the rate of convergence is always related with the behavior of the function $(d_\alpha(\xi)+1)^{-1}$. Anyway, for $p > 1$, it is quite complicated to control the coefficients $c_{i,j}^{(p)}$ of Proposition 2 (or the quantities $s_j^{(p)}(d_\alpha(\xi))$ in Proposition 3) and hence the constants given in the proofs.

Moreover, it is important to observe that the results explain the asymptotic behavior ($m \to \infty$) in the computation of a given component of $A_1^\alpha e_1$. As expected, the theorems show that the asymptotic rate of convergence is equal to $1/R^2$ (and superlinear for $\alpha = 1/2$), but also partially reveal the difficulty of the problem whenever the dimension grows, that is, when $j \to \infty$, for a fixed number of points of integration. Indeed, for $j \to \infty$ the functions $(d_\alpha(\xi)+1)^{-j}$ are no longer controllable unless we take $r \approx 1$, because j represents the multiplicity of the poles of the function $\psi_j^{(p)}(z)$. In this situation, experimentally one can see that $M_j(r)$ is close to 1 and its value attained near $\vartheta = \pi$. Theoretically we can only say that $E_{m,j} \leq const$ for m fixed and $j \to \infty$. A rigorous analysis of what happen in this situation is not very simple, because it is difficult to

provide a sharp bound for $M_j(r)$ when j is large, and consequently to define a suitable value for r, similarly to what we have made in the proof of Theorems 6 and 8. Note that the conditioning of the matrices A_p, $\kappa(A_p)$, is proportional to the dimension, so that the theorems also express the behavior of the method with respect to $\kappa(A_p)$.

Since in our situation $j \leq n$, where n is the number of points of the discretization of the derivative, we can restrict our consideration to values of n of interest to this aim. For this reason, in Figure 6, for $\alpha = 1/5, 1/3$ and $\alpha = 1/2$, we observe the error behavior with respect to the dimension, in the range $[200, 1200]$, in the computation of $A_1^\alpha e_1$ using Methods 1 and 3 with $m = 16$ points. It seems that the error stagnates with respect to the dimension.

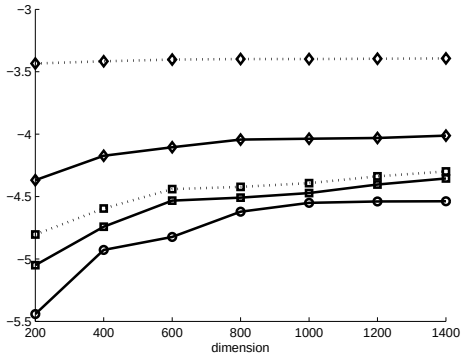

Figure 6: Relative error (in logarithmic scale) of Method 1 (dotted lines) and Method 3 (solid lines) with $m = 16$, for $\alpha = 1/5$ (diamonds), $\alpha = 1/3$ (squares) and $\alpha = 1/2$ (circles), changing the dimension of the problem. The order is $p = 1$.

In Figure 7 we also consider the relative error componentwise, that is, for each $1 \leq j \leq n$, we plot the relative error of the approximation of the j-th components of $A_1^\alpha e_1$, for some values of α and with $n = 500$.

We remark that for Method 2 the analysis given in Theorem 6 is very similar. The poles are the same, so that we can still work with the functions $(d_\alpha(\xi) + 1)^{-j}$ with $\xi = ir$. We have only differences in the constants defined during the proof.

For what concern Method 3, an analysis based on formula (23) is no longer possible because the functions involved are not analytic (the same consideration holds for Method 1 when $\alpha \neq 1/q$, $q = 2, 3, ...$). For each $0 < \alpha < 1$, $\alpha \neq 1/2$, the functions $f_j(\theta)$ arising from the substitution (16) of Method 3 are only of class C^1 at one end-point of the interval of integration, so that the asymptotic error analysis would require other tools (see e.g [25] and the reference therein), and this is beyond the purpose of the present paper. We also quote here [16] and [20] for some classical results about the error of the Gaussian rule for non-smooth functions. In any case, as demonstrated in Figures 1-5, Method 3 shows

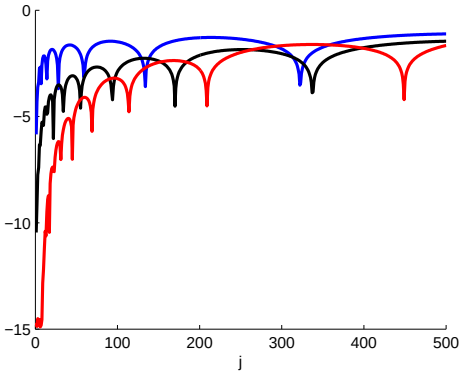

Figure 7: Componentwise relative error (in logarithmic scale) of Method 1 for $\alpha = 1/5$ (blue line), $\alpha = 1/3$ (black line), and $\alpha = 1/2$ (red line). The order is $p = 1$.

a good initial convergence (generally superior of Method 1 and 2), so that we can consider this approach of practical interest for our accuracy requirement (3-4 digits).

5 Fractional Differential Equations

A Caputo's type FDE is an equation of the form

$$ {}_{t_0}D_t^{\alpha} y(t) = g(t, y(t)), \quad t \in [t_0, T], \tag{42} $$

where $g : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$, and ${}_{t_0}D_t^{\alpha}$ denotes the Caputo's fractional derivative operator defined as

$$ D_{t_0}^{\alpha} y(t) = \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^{t} \frac{y'(u)}{(t-u)^{\alpha}} du, \quad \text{for} \quad 0 < \alpha < 1. \tag{43} $$

Among the existing definitions of fractional derivative (see [17]), Caputo's definition (43) allows to treat the initial conditions at t_0 for FDEs in the same manner as for integer-order differential equations, so that we assume $y(t_0) = y_0$. We refer here to [2] Chapter 6, for a wide background about the numerical solution of Volterra integral equation of the second kind with weakly singular kernel, by means of which the solution of (42) can be represented as follows

$$ y(t) = y(t_0) + \frac{1}{\Gamma(\alpha)} \int_{t_0}^{t} (t-s)^{\alpha-1} g(s, y(s)) ds. \tag{44} $$

In [14] the author extends the well known BDF formulas for ODEs to the fractional case, introducing methods of the type

$$ \sum\nolimits_{j=0}^{n} \omega_{n-j} y_j = h^{\alpha} g(t_n, y_n), \quad n \geq p, \tag{45} $$

where, ω_{n-j}, $0 \leq j \leq n$, are the Taylor coefficients of the generating functions

$$\omega_p^{(\alpha)}(\zeta) = (a_0 + a_1\zeta + ... + a_p\zeta^p)^\alpha, \quad \text{for} \quad 1 \leq p \leq 6, \tag{46}$$

being $\{a_0, a_1, ...a_p\}$ the coefficients of the underlying BDF formula. In [14] it is also shown the order p of the BDF formula is preserved. We remember that for this kind of equations, there is generally a intrinsic lack of regularity of the solution in a neighbor of the starting point, that is, we may have $y(t) \sim (t-t_0)^\alpha$ as $t \to t_0$. For this reason, in order to preserve the theoretical order p of the numerical method, formula (45) is generally corrected as

$$\sum\nolimits_{j=0}^{M} w_{n,j} y_j + \sum\nolimits_{j=0}^{n} \omega_{n-j} y_j = h^\alpha g(t_n, y_n), \tag{47}$$

where the sum $\sum_{j=0}^{M} w_{n,j} y_j$ is the so-called starting quadrature, whose aim is to have an exact integration of functions of the type

$$\sum\nolimits_{\substack{k,l=0 \\ k+\alpha l \leq p}} \gamma_{kl}(t-t_0)^{k+\alpha l}, \quad \gamma_{kl} \in \mathbb{R}.$$

As stated in the Introduction, an FBDF method (45) can be written in the matrix form as

$$A_p^\alpha \mathbf{y} = h^\alpha \mathbf{g}, \tag{48}$$

where $\mathbf{y} = (y_0, ..., y_n)^T$, $\mathbf{g} = (h^{-\alpha} y_0, g_1 ..., g_n)^T$, being y_j an approximation of $y(t_j)$, and $g_j = g(t_j, y_j)$. Note that the generating function (46) is just the *symbol* of the Toeplitz matrix A_p^α. Now, our approximations (19), (20), (21), leads to approximations of the type

$$A_p^\alpha \approx A_p \sum\nolimits_{j=1}^{m} \gamma_j (\eta_j I + A_p)^{-1} = [q_m(A_p)]^{-1} p_m(A_p), \tag{49}$$

where q_m and p_m are polynomials of degree m, and $p_m(0) = 0$. In this way we are able to define the recursion

$$p_m(A_p)\mathbf{y} = h^\alpha q_m(A_p)\mathbf{g}, \tag{50}$$

which yields (neglecting for a moment the starting) the implicit mp-step method

$$\sum\nolimits_{j=0}^{mp} \alpha_j y_{n-j} = h^\alpha \sum\nolimits_{j=0}^{mp} \beta_j g(t_{n-j}, y_{n-j}), \quad n \geq mp + 1. \tag{51}$$

We denote by GL(m, p) a formula of type (51), where GL remembers the use of the Gauss-Legendre rule. The starting values $y_0, ..., y_{mp}$, can be generated by the formula (45) or eventually by (47). In this sense our approach is based on the rational approximation of the generating function (46)

$$\omega_p^{(\alpha)}(\zeta) \approx \frac{\sum_{j=0}^{mp} \alpha_j \zeta^j}{\sum_{j=0}^{mp} \beta_j \zeta^j} = \frac{p_m(a_0 + a_1\zeta + ... + a_p\zeta^p)}{q_m(a_0 + a_1\zeta + ... + a_p\zeta^p)}. \tag{52}$$

For what concerns the computation of the coefficients α_j, β_j, they can be easily obtained in the following way. Let $\widetilde{A}_p \in \mathbb{R}^{(mp+1)\times(mp+1)}$ be the principal submatrix of A_p of order $mp+1$, then

$$\begin{aligned}(\alpha_0, ..., \alpha_{mp+1})^T &= p_m(\widetilde{A}_p)e_1,\\ (\beta_0, ..., \beta_{mp+1})^T &= q_m(\widetilde{A}_p)e_1.\end{aligned}$$

We remark that for the practical computation of the coefficients of the polynomials p_m and q_m, one can use the standard algorithms for converting partial fractions to polynomial quotients. For our computation we have used the Matlab function `residue`.

Unfortunately, from a theoretical point of view, we cannot expect that a GL(m,p) method could preserve the consistency of the underlying BDF formula. Indeed defining as usual the linear difference operator

$$L_h(z(t)) = \sum\nolimits_{j=0}^{mp} \alpha_j z(t-jh) - h^\alpha \sum\nolimits_{j=0}^{mp} \beta_j \left[{}_{t_0}D^\alpha_{t-jh} z(t-jh)\right],$$

where $z(t)$ is assumed to be regular as necessary, we know that the method is consistent with (42) if $\lim_{h\to 0} h^{-\alpha} L_h(z(t)) = 0$, with $t = t_0 + nh$ (cf. [7]). We can state the following result.

Proposition 9 *For each GL(m,p) method and $t = t_0 + nh$*

$$\lim_{h\to 0} h^{-\alpha} L_h(z(t)) = const. \tag{53}$$

Proof. By the expansion (eventually truncated)

$$z(t-jh) = z(t) + \sum\nolimits_{k\geq 1} \frac{(n-j)^k h^k}{k!} z^{(k)}(t_0),$$

we have

$${}_{t_0}D^\alpha_{t-jh} z(t-jh) = \sum\nolimits_{k\geq 1} \frac{(n-j)^{k-\alpha} h^{k-\alpha}}{\Gamma(k+1-\alpha)} z^{(k)}(t_0).$$

Therefore

$$L_h(z(t)) = C_0(n) z(t) + \sum\nolimits_{k\geq 1} h^k C_k(n) z^{(k)}(t), \tag{54}$$

where

$$\begin{aligned}C_0(n) &= \sum\nolimits_{j=0}^{mp} \alpha_j,\\ C_k(n) &= \frac{1}{k!}\sum\nolimits_{j=0}^{mp} \alpha_j (n-j)^k - \frac{1}{\Gamma(k+1-\alpha)} \sum\nolimits_{j=0}^{mp} \beta_j (n-j)^{k-\alpha}.\end{aligned}$$

Now

$$\sum\nolimits_{j=0}^{mp} \alpha_j = p_m\left(\sum\nolimits_{k=0}^{p} a_j\right) = p_m(0),$$

because a BDF method is consistent for $1 \leq p \leq 6$. Hence by (49) we have that $C_0(n) = 0$. For $k \geq 1$ we have

$$h^k C_k \left(\frac{t - t_0}{h} \right) = -h^\alpha \frac{(t - t_0)^{k-\alpha}}{\Gamma(k + 1 - \alpha)} \sum_{j=0}^{mp} \beta_j + O(h),$$

that proves (53) since $\sum_{j=0}^{mp} \beta_j \neq 0$. ■

While the above result express a theoretical not negligible drawback of GL(m, p) methods, numerically the situation is rather different. Indeed, already for m relatively small, a GL(m, p) method appears to be a good approximation of a method of order p, because of the quality of the approximation (52). Since a numerical method for (42) is of order p whenever $C_k(n)n^\alpha = O(h^{p-k})$ (cf. (54)), in Table 2, for $m = 12$, we report the values of the quantities $C_k(n)n^\alpha$, $k = 0, ..., p$, for $p = 1, ..., 4$, and $n = 500$. In this example $\alpha = 1/2$.

p	$C_0(n)n^\alpha$	$C_1(n)n^\alpha$	$C_2(n)n^\alpha$	$C_3(n)n^\alpha$	$C_4(n)n^\alpha$
1	-1.5e-13	-3.5e-9			
2	-9.0e-13	4.9e-8	-4.9e-6		
3	-2.1e-13	7.7e-8	2.6e-6	-1.5e-3	
4	2.1e-10	3.3e-7	5.7e-6	-3.3e-4	1.4e-1

Table 2 - Values of $C_k(n)n^\alpha$, $k = 0, ..., p$, for GL(12,p) with $p = 1, ..., 4$, and $n = 500$.

Unfortunately, for this kind of methods we cannot expect to improve the quality of the numerical solution for $h \to 0$ over a certain level determined by the approximation (52). Indeed, for $n \to \infty$, the approximation of the coefficients of the FBDF formula slowly deteriorates or, at best, stagnates (cf. Figure 6). For this reason we cannot even provide a classical analysis of the numerical order applying a GL(m, p) formula to a given problem with different values of h. Moreover, we remark that the use of the Matlab function `residue` is also responsible for additional errors, especially for $p > 1$. Indeed, as stated by Proposition 9, theoretically we should always have $C_0(n) = 0$, independently of m and p. Numerically, this theoretical property can be destroyed, as shown in Table 2.

Notwithstanding the theoretical disadvantages of a GL(m, p) formula, it is important to point out that especially for problems arising from spatial discretization, as for instance fractional diffusion equations (see e.g. [17] Section 10.10 and the references therein for some examples), the computational advantages of a mp-step formula with respect to the full recursion (45) is not negligible, especially in terms of memory saving. Moreover, as remarked in [6], in particular when $\alpha \neq 1/2$, the use of a starting quadrature as in (47), that theoretically should ensure the order of the FBDF formula, in practice may introduce substantial errors, causing unreliable numerical solutions. For high-order formulas, this is due to the severe ill-conditioning of the Vandermonde type systems one has to solve at each integration step.

We also remark, that in a typical application α, y_0 and possibly also the function g, may be only known up to a certain accuracy (see [5] for a discussion), so that one may only be interested in having a rather good approximation of the true solution. In such situations the short memory principle (consisting in the approximation of the generating function by a truncated Taylor series) is often applied, and our rational approach can be somehow regarded as a rational version of this principle.

6 A computed example

We consider the one-dimensional Nigmatullin's type equation

$$\begin{aligned} {}_0D_t^\alpha u(x,t) &= \frac{d^2u(x,t)}{dx^2}, \quad t>0, \quad x\in(0,\pi)\,, \\ u(0,t) &= u(\pi,t)=0, \\ u(x,0) &= \sin x. \end{aligned}$$

We discretize the spatial derivative using central differences on a uniform meshgrid of meshsize $\delta = \pi/(N+1)$ and Dirichlet boundary conditions. The discretization yields the N-dimensional FDE

$$ {}_0D_t^\alpha y(t) = Ly(t), \quad y(0)=y_0. \tag{55}$$

where $L=(N+1)^2\cdot tridiag(1,-2,1)$, and y_0 is the sine function evaluated at the grid points. The exact solution of (55) is given by

$$y(t) = E_\alpha(t^\alpha L)y_0,$$

where $E_\alpha(x)$ denotes the one-parameter Mittag-Leffler function (see e.g. [17] Chapter 1)

$$E_\alpha(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(k\alpha+1)}.$$

In Figure 8 some results on the approximation of $y(t)$ are reported. We compare the error at each step of the FBDF formula of order 1 and the method GL$(m,1)$ for some values of m, based on Method 3. The reference solutions have been computed using the Matlab function `funm` applied to the function `mlf` from [18] that implements the Mittag-Leffler function. The $m+1$ initial values are defined in the same manner so that they can be considered almost exact. The dimension of the problem is $N=200$, and we consider a uniform time step $h=1/n$, with $n=100$.

While the example is rather simple the results are encouraging. The cost of the GL(m,p) method is about 2/3 of the cost of the FBDF formula. Increasing n, the number of discretization points, the difference, obviously, grows. In this situation the FBDF formula produces better results near the endpoint, while the GL(m,p) methods typically stabilize around a certain error. The reason of

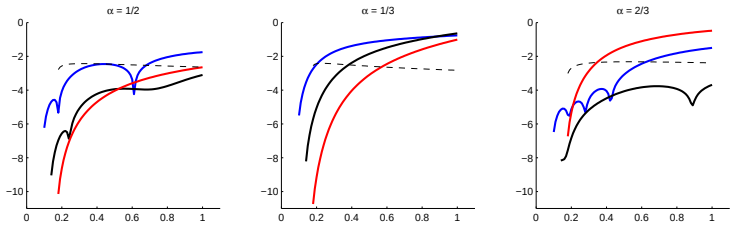

Figure 8: Step by step error (in logarithmic scale) for the FBDF method of order 1 (dashed line) and GL$(m,1)$ method (solid line), for $m = 8$ (blue), 12 (black), 16 (red). $N = 200$, $n = 100$.

this situation is that the approximation (8) is independent of h (or n), while this is not true for FBDF method, in which $\omega_p^{(\alpha)}(\zeta)$ is approximated by its truncated Taylor series. This is also the reason for which, initially, the GL(m,p) formula is much more accurate than the FBDF method.

As already mentioned, the Matlab function `residue` may introduce errors for large values of mp. This is what we have seen in other numerical experiments not reported. This of course represents a computational problem that should be fixed (cf. Figure 8 for $\alpha = 2/3$, where, unexpectedly, GL$(8,1)$ and GL$(12,1)$ behave better than GL$(16,1)$). Anyway, for a given m and p, one can of course improve the results of this Matlab function computing the corresponding coefficients once and for ever.

For what concerns the linear stability, taking $g(t, y(t) = \lambda y\,(t)$ in (42), we have that $y(t) = E_\alpha(t^\alpha \lambda) \to 0$ for

$$|\arg(\lambda) - \pi| < \left(1 - \frac{\alpha}{2}\right)\pi,$$

(see [15]). The stability region of a FBDF formula is given by

$$\mathbb{C}\backslash\left\{\omega_p^{(\alpha)}(\zeta) : |\zeta| \leq 1\right\}, \tag{56}$$

and, as expected, the GL(m,p) methods rapidly (already for m small) simulates the behavior of these formulas. In Figure 9 we show some examples involving GL(m,p) formulas, generated again by Method 3.

7 Conclusions

In this paper we have presented an alternative approach, based on the contour integral approximation of the matrix function A^α, to compute the coefficients of a FBDF formula for FDEs. Implicitly the methods presented also produce the coefficients of stationary k-step formulas for FDEs that in general are less accurate than FBDF formulas but that present important computational advantages. In our opinion the GL(m,p) methods are worth of consideration even

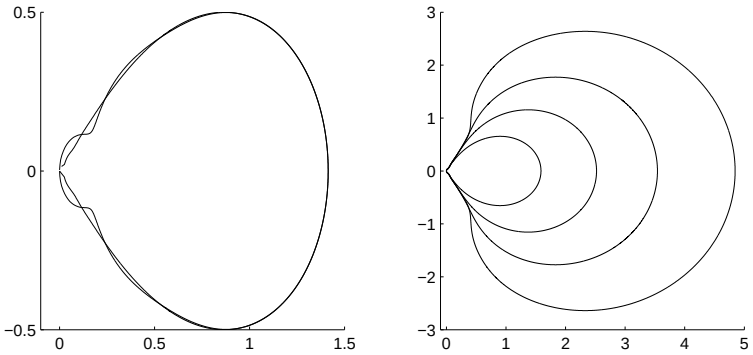

Figure 9: Left: boundary of the stability domains of GL(4,1) and GL(12,1) for $\alpha = 1/2$. Right: boundary of the stability domains of GL(8,p), for $p = 1$ (inner) to 4 (outer), for $\alpha = 2/3$.

if much work has still to be done in order to fix some computational aspect. We remark moreover that either the substitutions considered in Section 2 or the choice of the quadrature rule are responsible for this kind of formulas and the rational approximation (8). In this sense, other strategies are of course possible.

Acknowledgement 10 *The author is grateful to Igor Moret for some helpful discussions.*

References

[1] D.A. Bini, N.J. Higham, B. Meini, Algorithms for the matrix pth root, Numer. Algorithms 39 (2005), 349–378.

[2] H. Brunner, P.J. van der Houwen, The numerical solution of Volterra equations. CWI Monographs, 3. North-Holland Publishing Co., Amsterdam, 1986.

[3] M.M. Chawla, M.K. Jain, Error estimates for Gauss quadrature formulas for analytic functions, Math. Comp. 22 (1968), 82–90.

[4] M.M. Chawla, Asymptotic estimates for the error of the Gauss-Legendre quadrature formula, Comput. J. 11 (1968/1969), 339–340.

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

[6] K. Diethelm, J.M. Ford, N.J. Ford, M. Weilbeer, Pitfalls in fast numerical solvers for fractional differential equations, J. Comput. Appl. Math. 186 (2006), 482–503.

[7] L. Galeone, R. Garrappa, On multistep methods for differential equations of fractional order, Mediterr. J. Math. 3 (2006), 565–580.

[8] N.J. Higham, L. Lin, A Schur-Padé algorithm for fractional powers of a matrix, SIAM J. Matrix Anal. Appl. 32 (2011), 1056–1078.

[9] E. Hairer, C. Lubich, M. Schlichte, Fast numerical solution of weakly singular Volterra equations, J. Comput. Appl. Math. 23 (1988), 87–98.

[10] E. Hairer, S.P. Norsett, G. Wanner, Solving ordinary differential equations. I. Nonstiff problems. Second edition. Springer Series in Computational Mathematics, 8. Springer-Verlag, Berlin, 1993.

[11] N. Hale, N.J. Higham, L.N. Trefethen, Computing A^α, $log(A)$, and related matrix functions by contour integrals, SIAM J. Numer. Anal. 46 (2008), 2505–2523.

[12] N.J. Higham, Functions of matrices. Theory and computation. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[13] N.S. Kambo, Error of the Newton-Cotes and Gauss-Legendre quadrature formulas, Math. Comp. 24 (1970), 261–269.

[14] C. Lubich, Discretized fractional calculus, SlAM J. Math. Anal. 17 (1986), 704–719.

[15] C. Lubich, A stability analysis of convolution quadratures for Abel-Volterra integral equations, IMA J. Numer. Anal. 6 (1986), 87–101.

[16] D. S. Lubinsky and P. Rabinowitz, Rates of convergence of Gaussian quadrature for singular integrands, Math. Comp. 43 (1984), 219-242.

[17] I. Podlubny, Fractional differential equations. Mathematics in Science and Engineering, 198. Academic Press, Inc., San Diego, CA, 1999.

[18] I. Podlubny, Mittag–Leffler Function, http:www.mathworks.com/matlabcentral/fileexchange/8738 (2009).

[19] P. Rabinowitz, Rough and ready error estimates in Gaussian integration of analytic functions, Comm. ACM 12 (1969), 268–270.

[20] P. Rabinowitz, Rates of convergence of Gauss, Lobatto, and Radau integration rules for singular integrands, Math. Comp. 47 (1986), 625–638.

[21] A. Schädle, M. López-Fernández, C. Lubich, Fast and oblivious convolution quadrature, SIAM J. Sci. Comput. 28 (2006), 421–438.

[22] T. Schmelzer, L.N. Trefethen, Evaluating matrix functions for exponential integrators via Carathéodory-Fejér approximation and contour integrals. Electron. Trans. Numer. Anal. 29 (2007/08), 1–18.

[23] F. Stenger, Bounds on the error of Gauss-type quadratures, Numer. Math. 8 (1966), 150–160.

[24] A. Talbot, The accurate numerical inversion of Laplace transforms, J. Inst. Math. Appl. 23 (1979), 97–120.

[25] L.N. Trefethen, Is Gauss quadrature better than Clenshaw-Curtis?, SIAM Rev. 50 (2008), 67–87.

[26] Ya Yan Lu, A Padé approximation method for square roots of symmetric positive definite matrices, SIAM J. Matrix. Anal. Appl. 19 (1998), 833–845.