Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 219 (2008) 329 – 349
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Reduction of the Gibbs phenomenon for smooth functions with jumps by the ε-algorithm☆

Bernhard Beckermann, Ana C. Matos*, Franck Wielonsky

Laboratoire Painlevé UMR 8524, UST Lille, F-59655 Villeneuve d'Ascq CEDEX, France

Received 14 September 2006; received in revised form 12 November 2007

Dedicated to our friend Claude Brezinski on the occasion of his retirement

Abstract

Recently, Brezinski has proposed to use Wynn's ε-algorithm in order to reduce the Gibbs phenomenon for partial Fourier sums of smooth functions with jumps, by displaying very convincing numerical experiments. In the present paper we derive analytic estimates for the error corresponding to a particular class of hypergeometric functions, and obtain the rate of column convergence for such functions, possibly perturbed by another sufficiently differentiable function. We also analyze the connection to Padé–Fourier and Padé–Chebyshev approximants, including those recently studied by Kaber and Maday.

MSC: 41A21; 41A20; 42A16

Keywords: Fourier series; Gibbs phenomenon; Convergence acceleration; ε-Algorithm; Padé–Fourier approximants; Padé–Chebyshev approximants

1. Introduction

Spectral methods for partial differential equations consist in approximating the exact solution by considering a discrete version of the original problem and expressing its solution as a truncated Fourier series or as an expansion in a basis of orthogonal polynomials. These methods have been mainly developed in the last decades with the advent of the fast Fourier transform and the use of tensorization for multi-dimensional problems, see [3,16] for more details. When the solution is smooth, the convergence of the approximate solution is usually geometric (also termed exponential or spectral in the literature), meaning that the error decays as $e^{-\alpha N}$, $\alpha > 0$, N the number of coefficients in the expansion. In contrast, when the solution exhibits discontinuities, as is the case, e.g., in physical problems with shocks or in image compression, the convergence is poor, due to the occurrence of the Gibbs phenomenon. Then, one needs to filter the data, that is to design more accurate solutions than the rough Fourier expansion. At this stage, if the locations and amplitudes of the jumps are unknown, one first has to compute them. Different techniques have recently been proposed, such as nonlinear optimization procedures [8,9,21,20] originally based on Prony's method for exponential

☆ This work was partly supported by INTAS network NeCCA 03-51-6637.
* Corresponding author.
E-mail addresses: bbecker@math.univ-lille1.fr (B. Beckermann), matos@math.univ-lille1.fr (A.C. Matos), wielonsk@math.univ-lille1.fr (F. Wielonsky).

0377-0427/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2007.11.011

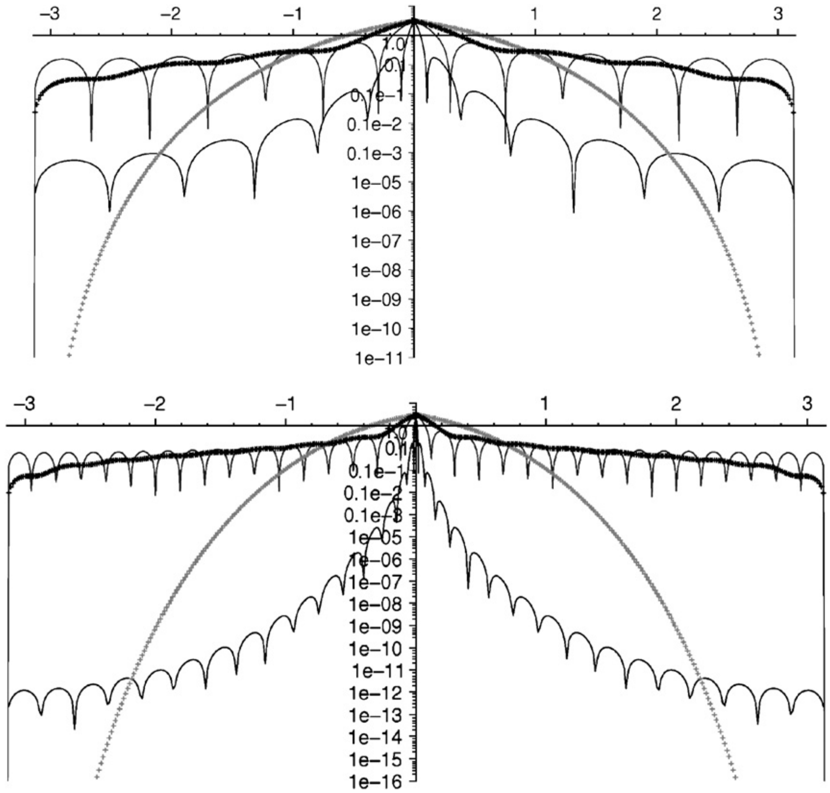

Fig. 1. The modulus of the error of approximation on a logarithmic scale of the Fourier series of the sew tooth function (2.2). On the top we use the first 7 and on the bottom the first 17 coefficients of the Fourier series.

approximation, or such as exploiting classical formulas for the jumps as limits of the conjugate or derivative of the Fourier series, and accelerating the convergence by using suitable "concentration" kernels [13,14,19]. Assuming knowledge of the jumps, Gottlieb and Shu [17] have developed a method based on projecting the Fourier expansion onto a space spanned by the Gegenbauer polynomials. These are associated with the weight $(1 - x^2)^\lambda$, where the exponent λ grows linearly like N, the number of Fourier coefficients, which has the effect of restoring geometric convergence in any subinterval that does not include a discontinuity. Another approach consists in subtracting off the singularities [10] which leads to a numerical method of high order for computing derivatives and integrals of solutions. These singularities are assumed to be either discontinuities of the function (or its derivatives) or branch points such that those of x^α, $\alpha \in \mathbb{R}$. The speed of convergence is then governed by the regularity of the remaining part of the original function.

For smoothing the Gibbs phenomenon, Brezinski [5] has recently proposed a quite simple procedure based on Padé approximation (or equivalently the ε-algorithm which consists in the better conditioned "value problem") of the analytic function $G_n(f)(z)$, $z = \mathrm{e}^{\mathrm{i}t}$, $t \in [-\pi, \pi]$, the real part of which is the truncated Fourier series of the unknown solution f. The efficiency of this technique, called the complex ε-algorithm, is shown in [5, Section 4] which presents many numerical experiments. Note that the acceleration of Fourier series via the ε-algorithm applied to the partial sums of $G_n(f)$ has been already proposed in [29], without discussing the link with the Gibbs phenomenon. Wynn gives several examples where classical linear acceleration procedures for Fourier series like Cesaro means or de la Vallée–Poussin means have convergence behavior clearly weaker than the one discussed here, see Fig. 1. Here, one should also mention [18] where the model function $f(t) = \mathrm{sign}(\cos(t))$ is discussed, and interesting acceleration properties are obtained by using linear Padé–Chebyshev approximants, whose denominators are of fixed degrees, see [18, Theorem 4.10].

The aim of this paper is to provide error estimates for Brezinski's method for some "model" functions f, possibly perturbed by a sufficiently smooth function. Actually, these "model" functions will consist in real parts of hypergeometric functions, see Section 2, and these specific functions already entails a large set of interesting examples, with jumps

Table 1
Some examples for Fourier series f, their associated power series $G(f)$ and explicit formulas for Padé approximants $[n+k|k]$ or reversed Padé denominators $q_{n+k,k}$ (up to a normalization constant)

	$f(t)$	$G(f)(z)$	Padé approximant/reversed denominator
(a)	$\check{f}(t-t_0)$	$G(\check{f})(\mathrm{e}^{-\mathrm{i}t_0}z)$	$[n+k\|k]_{G(f)}(z)=[n+k\|k]_{G(\check{f})}(\mathrm{e}^{-\mathrm{i}t_0}z)$
(b)	$\check{f}(2t)$	$G(\check{f})(z^2)$	$[2n+2k\|2k]_{G(f)}(z)=[n+k\|k]_{G(\check{f})}(z^2)$
(c)		$p(z)+cz^{n+1}G(\check{f})(z)$ $\deg\ p\leqslant n,\ c\in\mathbb{C}.$ Notice: $G(f)=z^{\ell}G^{(\alpha,\beta)}$ $\Longrightarrow G(\check{f})=G^{(\alpha+n+1-\ell,\beta)}$	$[n+k\|k]_{G(f)}(z)$ $=p(z)+cz^{n+1}[k-1\|k]_{G(\check{f})}(z)$
(d)		$G^{(\alpha,\beta)}(z)$	$q_{n+k-1,k}(z)=P_k^{(\alpha+n,\beta)}(1-2z)$
$(d)_1$	$s(t)$, see (2.1)	$-2i\log(1-z)$ $=2iz,\ G^{(0,0)}(z)$	$q_{n+k,k}(z)=P_k^{(n,0)}(1-2z)$
$(d)_2$	$\mathrm{sign}(\cos(t))=$ $(s(t-\frac{\pi}{2})-s(t+\frac{\pi}{2}))/\pi$	$\frac{4z}{\pi}G^{(-\frac{1}{2},0)}(-z^2)$	$q_{2n+2k,2k}(z)=P_k^{(n-\frac{1}{2},0)}(1+2z^2)$
$(d)_3$	$\|\sin(\frac{t}{2})\|$	$\frac{2}{\pi}-\frac{4z}{3\pi}G^{(-\frac{1}{2},1)}(z)$	$q_{n+k,k}(z)=P_k^{(n-\frac{1}{2},1)}(1-2z)$
$(d)_4$	$\|\sin(t)\|+\sin(t)$	$\frac{2}{\pi}-iz-\frac{4z^2}{3\pi}G^{(-\frac{1}{2},1)}(z^2)$	$q_{2n+2k,2k}(z)=P_k^{(n-\frac{1}{2},1)}(1-2z^2)$
(e)	$(\cos(t)-1)^{\ell}\check{f}(t),$ $a_{j+\ell}=2^{-\ell}\Delta^{2\ell}\check{a}_j,$ $b_{j+\ell}=2^{-\ell}\Delta^{2\ell}\check{b}_j$	$p(z)+\frac{(z-1)^{2\ell}}{2^{\ell}z^{\ell}}[G(\check{f})(z)-G_{\ell-1}(\check{f})(z)]$ $\deg\ p\leqslant\ell-1,$	
$(e)_1$	$(\cos(t)-1)^{\ell}\check{f}(t)$ $G(\check{f})=G^{(\alpha,\beta)}$	$p+cz^{\ell}G^{(\alpha,\beta+2\ell)}$ $\deg\ p\leqslant\ell-1,\ c=\frac{2^{-\ell}(\beta+1)_{2\ell}}{(\alpha+\beta+2)_{2\ell}}$	$q_{n+2\ell+k-1,k}(z)=P_k^{(\alpha+n,\beta+2\ell)}(1-2z)$
$(e)_2$	$(\cos(t)-1)s(t)$	$-\frac{3}{2}iz+\frac{iz^2}{3}G^{(0,2)}(z)$	$q_{n+k+1,k}(z)=P_k^{(n,2)}(1-2z)$

The quantities $c, t_0, t_1, t_2, \alpha, \beta$ occurring in the table are real numbers ($\alpha, \beta > -1$), also n, k, ℓ are nonnegative integers ($n \geqslant -k$ for cases (a) and (b)), and p are suitable polynomials, not necessarily the same for different rows of the table.

possibly in higher order derivatives, see Table 1. We will show in Theorem 5.1 that the quantity $\varepsilon_{2k}^{(n)}$ corresponding to Pade approximants with fixed denominator degree k allow for $n \to \infty$ to improve convergence by, roughly, the factor n^{-2k} as long as we stay away from the singularities of f, and by $n^{-2(1-\tau)k}$ if we approach these singularities up to a distance of order $n^{-\tau}$ for some $\tau \in [0, 1)$. Notice that the error at the singularity itself cannot be zero since, implicitly, we approach a function with singularities by a $\mathscr{C}^{\infty}$ function.

Brezinski's Padé approach has the advantage of being simple and does not require any a priori knowledge about the location of singularities. However, there are more sophisticated methods which allow for better approximation in a neighborhood of the singularities once the location of these singularities is known, see for instance the numerical experiments reported in [7, Fig. 7]. Besides the linear Eckhoff approach [10] of subtracting the "singular" part, we would like to mention in particular the recent method of Driscoll and Fornberg [7] of constructing nonlinear approximants with built-in singularities. These latter authors suggest to consider a "simple" function S (e.g., $S(z)=\log(1-z)$) having the same singularities as $G(f)$ (after a change of variables $z=\mathrm{e}^{\mathrm{i}t}$) which are supposed to be known. Subsequently they construct a Hermite–Padé form (p_1, p_2, p_3) of the triple of functions $(1, S, G(f))$ (which means that p_j are polynomials of prescribed degree such that the form $p_1 + Sp_2 + G(f)p_3$ has the highest possible order at zero), and then they solve the equation $p_1(z) + S(z)p_2(z) + G(f)(z)p_3(z) = 0$ for the unknown $G(f)(z)$ to obtain an approximation of it. Numerical experiments provided in [7] show that this Hermite–Padé approach allows to obtain a good rate of convergence also in a neighborhood of the singularities, but no theoretical error estimates have been given so far. We plan to investigate this question in a future publication.

The paper is organized as follows. In Section 2 we introduce the complex ε-algorithm in more details and define the functions to which our analysis applies. In Section 3, we exhibit a link between the Padé–Chebyshev approximants of a Chebyshev series (and more general rational approximants of Fourier series) and the ordinary Padé approximants of the corresponding Taylor series. We also relate our analysis to the results of [18]. In Section 4, we study the rate of convergence of Padé approximants (in a column of the Padé table) to some specific hypergeometric functions, the real parts of which correspond to functions with prescribed discontinuities, possibly occurring in higher order derivatives. In Section 5, we extend the previous estimates by adding functions with continuous derivatives (up to some order depending on the degree of approximation) to the previous ones. Finally, in Section 6, we present numerical results.

2. Description of the method

The procedure is as follows: see [5]. Let

$$S_n(f)(t)=\frac{a_0}{2}+\sum_{j=1}^{n}[a_j\ \cos(jt)+b_j\ \sin(jt)]$$

be the partial sum $S_n(f)$ of the Fourier series of a function $f:[-\pi,\pi]\mapsto\mathbb{R}$ with jumps, and add to it i times the conjugate part

$$\widetilde{S}_n(f)(t)=\sum_{j=1}^{n}[a_j\ \sin(jt)-b_j\ \cos(jt)]$$

in order to get

$$S_n(f)(t)+i\widetilde{S}_n(f)(t)=G_n(f)(\mathrm{e}^{\mathrm{i}t}),$$

with $G_n(f)$ the nth Taylor series of the (formal) series

$$G(f)(z)=\sum_{j=0}^{\infty}c_j(f)z^j,\quad c_0(f)=\frac{a_0}{2}\quad\text{and for } j>1,\ \ c_j(f)=a_j-ib_j.$$

Then apply the ε-algorithm to the sequence of partial sums $(G_n(f)(\mathrm{e}^{\mathrm{i}t}))_n$ for fixed t, and use the real part of the resulting quantities $\varepsilon_{2k}^{(n)}(t)$ for approaching $f(t)=\mathrm{Re}(G(f)(\mathrm{e}^{\mathrm{i}t})$. In this paper we consider functions of the form $f=f_1+f_2$ where f_1 has prescribed discontinuities and is smooth elsewhere while f_2 has sufficiently fast decreasing Fourier coefficients. For such functions, the partial Fourier series converges slowly and presents the so-called Gibbs phenomenon of oscillation close to the singularities of f (or f_1). The acceleration properties of the ε-algorithm will essentially depend on f_1.

A typical jump function f_1 considered in this paper is given by (a multiple of) the 2π periodic saw tooth function

$$s(t)=\pi+t\ \text{ for } t\in(-\pi,0],\quad s(t)=-\pi+t\ \text{ for } t\in(0,\pi],\tag{2.1}$$

having one jump of absolute value 2π at $t=0$ in $[-\pi,\pi)$. We have for the saw tooth function the Fourier expansion

$$S_n(s)(t)=-2\sum_{j=1}^{n}\frac{\sin(jt)}{j}\quad\text{and thus } G(s)(z)=2i\sum_{j=1}^{\infty}\frac{z^j}{j}=-2i\ \log(1-z).\tag{2.2}$$

The error obtained by approximating $s(t)$ via partial sums and via the ε-algorithm is displayed in Fig. 1: we remark that the error for the partial sums (in solid gray) is strongly oscillating, and, according to the Gibbs phenomenon, remains quite large (about 1/10) even for higher order Fourier sums. The error for Cesaro means (in dotted black) is smoother, but about of the same size, even for arguments far from 0, the singularity of our function, whereas the de la Vallée–Poussin mean (in dotted gray) gives better approximants only far from 0. In contrast, the errors for $\varepsilon_6^{(0)}$ and $\varepsilon_{16}^{(0)}$ (in solid black) are much smaller, even for arguments closer to the singularity.

In the present paper we will consider for $G(f_1)$ more general hypergeometric functions of the form

$$G^{(\alpha,\beta)}(z)=\left(\begin{matrix}\alpha+1,1\\ \alpha+\beta+2\end{matrix}\middle|\,z\right)\quad\text{where}\quad\left(\begin{matrix}a,b\\ c\end{matrix}\middle|\,z\right)=\sum_{j=0}^{\infty}\frac{(a)_j(b)_j}{(c)_j\,j!}z^j,\tag{2.3}$$

$\alpha, \beta > -1$, and $(a)_0 = 1$, $(a)_j = a(a+1)\cdots(a+j-1)$ for $j > 0$, is the usual Pochhammer symbol. Throughout, we denote by $P_k^{(\alpha,\beta)}$ the Jacobi polynomial of degree k, orthogonal with respect to the measure $(1-x)^\alpha(1+x)^\beta\,\mathrm{d}x$, such that

$$\int_{-1}^{1} (P_k^{(\alpha,\beta)}(x))^2(1-x)^\alpha(1+x)^\beta\,\mathrm{d}x = \frac{2^{\alpha+\beta+1}\Gamma(k+\alpha+1)\Gamma(k+\beta+1)}{(2k+\alpha+\beta+1)\Gamma(k+\alpha+\beta+1)k!}, \tag{2.4}$$

see for instance [6, Chapter 5, Eq. (2.18)]. As seen in Table 1 (d)$_1$–(d)$_4$, (e)$_2$, (e), and verified by elementary computations, such a class of hypergeometric series allows us to include other functions $f_1(t)$ with a particular jump behavior, like $\mathrm{sign}(\cos(t))$ (compare with [18]), having two discontinuities, or like $|\sin(t/2)|$, and $(1-\cos(t))\,s(t)$, respectively, with first (and second) order derivative having a discontinuity at 0. Moreover, combined with Table 1(a)–(c), (e)$_1$, we may easily construct other examples where the argument is shifted, or where $f_1 \in \mathscr{C}^{\ell-1}([-\pi,\pi])$, with its ℓth derivative having one discontinuity.

Our main tool in deriving error estimates for the complex ε-algorithm will be the connection to Padé approximation of perturbations of the (shifted) logarithm $z \mapsto \log(1-z)$ (or more generally of Stieltjes functions with respect to measures related to the Jacobi orthogonal polynomials) on the unit circle. Indeed, as already mentioned in [5], it is well known that

$$\varepsilon_{2k}^{(n)}(t) = [n+k|k]_{G(f)}(\mathrm{e}^{\mathrm{i}t}) \tag{2.5}$$

and hence we will have to estimate the modulus of

$$f(t) - \mathrm{Re}(\varepsilon_{2k}^{(n)}(t)) = \mathrm{Re}(G(f)(\mathrm{e}^{\mathrm{i}t}) - [n+k|k]_{G(f)}(\mathrm{e}^{\mathrm{i}t})). \tag{2.6}$$

In particular, for $n \to \infty$ and k fixed, we will have to find a Montessus de Ballore type convergence theorem for perturbed Stieltjes functions. This is given in Theorem 5.1 where we show that, roughly, we gain a factor n^{-2k} for t away from jumps, and a slightly weaker rate if we approach the jump up to a factor $n^{-\tau}$ for some $\tau \in (0,1)$.

For even f (and hence $b_j = 0$ for all j), $S_n(f)(\arccos(x))$ is the partial Chebyshev series of $x \mapsto F(x) := f(\arccos(x))$. Here, according to the well-known formula $T_j(x) = \cos(j\ \arccos(x))$ for the Chebyshev polynomials, it is not difficult to see that $\mathrm{Re}(\varepsilon_{2k}^{(n)}(\arccos(x)))$ is a rational function in x, the so-called Padé–Chebyshev approximant of type $(n+k,k)$ of g due to Gragg [2, pp. 383–387]. Following the nomenclature of Baker and Graves-Morris [2, Section 7.4], there are other approaches to rational approximation of Chebyshev series, and these latter methods can also be adapted to Fourier series (see Section 3) or to series of general orthogonal polynomials (see, e.g., [15]): for the so-called *linear* Padé–Chebyshev approximant $R_{m,n}^{\mathrm{LC}} = P_{m,n}^{\mathrm{LC}}/Q_{m,n}^{\mathrm{LC}}$, and the *nonlinear* Padé–Chebyshev approximant $R_{m,n}^{\mathrm{NC}} = P_{m,n}^{\mathrm{NC}}/Q_{m,n}^{\mathrm{NC}}$, respectively, we look for polynomials $P_{m,n}^{\mathrm{LC}}$, $P_{m,n}^{\mathrm{NC}}$ of degree $\leqslant m$ and $Q_{m,n}^{\mathrm{LC}}$, $Q_{m,n}^{\mathrm{NC}}$ of degree $\leqslant n$ such that either the linearized error $FQ_{m,n}^{\mathrm{LC}} - P_{m,n}^{\mathrm{LC}}$ or the error $F - R_{m,n}^{\mathrm{NC}}$ itself is orthogonal to $T_0, T_1, \ldots, T_{m+n}$. Linear Padé–Chebyshev approximants are easy to compute (a solution of a linear system of equations with matrix of coefficients being Toeplitz plus Hankel) but require more coefficients of the Chebyshev series; the acceleration properties of $R_{m,n}^{\mathrm{LC}}(x)$ for the sign function and n fixed, $m \to \infty$ have been discussed in [18]. It was observed numerically by Fleischer [11] and proved rigorously Gonchar et al. in [15] for Markov functions that nonlinear Padé–Chebyshev diagonal approximants ($m = n \to \infty$) have better approximation properties than the linear ones. However, the nonlinear approximants are in general difficult to compute, which limits their impact in practical applications. We will show in Theorem 3.1 that, provided that f is even and the Padé approximant $[n+k|k]_{G(f)}$ of $G(f)$ has no poles in the closed unit disk $|z| \leqslant 1$, the nonlinear Padé–Chebyshev approximant $R_{n+k,k}^{\mathrm{NC}}$ of $F = f \circ \arccos$ is given by

$$R_{n+k,k}^{\mathrm{NC}}(\cos(t)) = \mathrm{Re}(\varepsilon_{2k}^{(n)}(t)) = \mathrm{Re}([n+k|k]_{G(f)}(\mathrm{e}^{\mathrm{i}t})). \tag{2.7}$$

This observation, which to our knowledge is original, may help to understand the convergence properties of $\mathrm{Re}(\varepsilon_{2k}^{(n)}(t))$. A formal link between the denominators of both rational approximants has been given before by Paszkowski [25], without mentioning the necessary hypothesis on the Padé approximant.

The information given in Table 1 requires additional explanations and proofs. Concerning Table 1(a), we notice that a translation of the argument t of f is equivalent to a multiplication of the variable z of $G(f)$ with a constant of modulus one

$$\breve{f}(t) = f(t-t_0) \quad\Longrightarrow\quad S_n(\breve{f})(t) = S_n(f)(t-t_0) \quad\Longrightarrow\quad G(\breve{f})(z) = G(f)(\mathrm{e}^{-\mathrm{i}t_0}z), \tag{2.8}$$

since indeed $\breve{a}_j = a_j \cos(jt_0) - b_j \sin(jt_0)$, $\breve{b}_j = a_j \sin(jt_0) + b_j \cos(jt_0)$, and thus $c_j(\breve{f}) = \breve{a}_j - i\breve{b}_j = \mathrm{e}^{-\mathrm{i}jt_0} c_j(f)$. In the last column of rows (a)–(c) we recall some well-known properties of Padé approximation [2]. One may express the Padé approximants of $G^{(\alpha,\beta)}$ by means of the Gauss continued fraction (see, e.g., [1, Chapter 5]), in particular, there exist explicit formulas for the Padé denominator [1, Eq. (5.11)] which will enable us to estimate quite precisely the Padé error on the unit circle of such functions. For the sake of completeness, this connection between Jacobi orthogonal polynomials and the (reversed) Padé denominators of $G^{(\alpha,\beta)}$ claimed in the last column of Table 1(d) will be shown in Lemma 4.1 in Section 3. The claims in rows $(\mathrm{d})_1$–$(\mathrm{d})_4$ are obtained by combining (d) with the statements of (a)–(c), we leave the details for the reader. The claims in rows (e)–$(\mathrm{e})_2$ are again not too difficult to verify and left to the reader. From the information in Table 1, we see that, in order to study the error (2.6) for f being equal to one of the functions $s(t)$, $s(t-t_1) - s(t-t_2)$,sign(cos(t)), $|\sin(t/2)|$, or $(1-\cos(t))s(t)$, it is sufficient to estimate the modulus of

$$G^{(\alpha,\beta)}(\mathrm{e}^{\mathrm{i}t}) - [n+k|k]_{G^{(\alpha,\beta)}}(\mathrm{e}^{\mathrm{i}t})$$

(up to some explicitly known constant not depending on n, k) in terms of the distance of $\mathrm{e}^{\mathrm{i}t}$ to the singularities of f. This will be done in Section 4 below. We will also show in Theorem 5.1 that similar bounds hold true for $f = f_1 + f_2$ with f_1 as before and f_2 sufficiently smooth.

3. Rational approximants of Fourier series

The most efficient way of evaluating the value at $z = \mathrm{e}^{\mathrm{i}t}$ of a Padé approximant is known to be Wynn's ε-algorithm, as described in Table 2. In this section, we relate the approximant (2.5) to other rational approximants such as linear or nonlinear Padé–Chebyshev and Padé–Fourier approximants.

In what follows we denote as usual by $L_2([-\pi,\pi])$ the set of square integrable functions on $[-\pi,\pi]$, with norm

$$\|f\|_2 := \left(\frac{1}{\pi}\int_{-\pi}^{\pi} |f(t)|^2\,\mathrm{d}t\right)^{1/2}.$$

For a Fourier series

$$f(t) = \frac{a_0}{2} + \sum_{j=1}^{\infty}(a_j \cos(jt) + b_j \sin(jt)),$$

recall that, by Parceval's formula, $\|f\|_2^2 = |a_0|^2/2 + \sum_{j=1}^{\infty}(|a_j|^2 + |b_j|^2)$. We also deal with the Hardy space H_2 of functions G being analytic in the unit disk $\mathbb{D}$, with

$$\|G\|_2 := \lim_{r\to 1_-}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |G(r\mathrm{e}^{\mathrm{i}t})|^2\,\mathrm{d}t\right)^{1/2} < \infty.$$

One may show, see, e.g., [24, Chapter 3.3], that for $G(z) = \sum_{j=0}^{\infty} G_j z^j$, one has

$$\|G\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |G(\mathrm{e}^{\mathrm{i}t})|^2\,\mathrm{d}t = \sum_{j=0}^{\infty} |G_j|^2.$$

Table 2
Evaluating the Padé approximant $[n+k|k]_G(z)$ via the ε-algorithm

Input:	integers $n, k \geqslant 0$, partial sum $G_{n+2k}(z) = \sum_{j=0}^{n+2k} g_j z^j$, a fixed argument $z = \mathrm{e}^{\mathrm{i}t}$
Initialization:	for $\ell = 0, \ldots, n+2k$: $\varepsilon_0^{(\ell)} = G_\ell(z)$, $\varepsilon_{-1}^{(\ell)} = 0$.
Recurrence:	for $j = 0, 2, \ldots, 2k-1$, for $\ell = 0, \ldots, n+2k-j-1$: $\varepsilon_{j+1}^{(\ell)} = \varepsilon_{j-1}^{(\ell+1)} + (\varepsilon_j^{(\ell+1)} - \varepsilon_j^{(\ell)})^{-1}$
Output:	$[n+k|k]_G(z) = \varepsilon_{2k}^{(n)}$

In particular we may represent such functions by means of the Cauchy formula with contour being the unit circle $\mathbb{T}$. As a consequence, we have for our real-valued Fourier series f that $f \in L_2([-\pi, \pi])$ if and only if $G(f) \in H_2$, with

$$\|G(f)\|_2^2 = \|f\|_2^2 - \frac{|a_0|^2}{4}. \tag{3.1}$$

We will show in the proof of Theorem 3.1 below that $\mathrm{Re}(\varepsilon_{2k}^{(n)}(t))$ is a trigonometric rational function in $t \in [-\pi, \pi]$, with the numerator of degree $n+k$, and the denominator of degree k. Following the nomenclature of Baker and Graves-Morris [2, Section 7.4], the quantity $\mathrm{Re}(\varepsilon_{2k}^{(n)}(t))$ equals the Padé–Fourier approximant of type $(n+k, k)$ of f. There are other approaches to trigonometric rational approximation of Fourier series, and these latter methods can be also adapted to series of general orthogonal polynomials, see for instance Section 1 for Chebyshev series. For the so-called *linear* Padé–Fourier approximant $R_{m,n}^{\mathrm{LF}} = P_{m,n}^{\mathrm{LF}}/Q_{m,n}^{\mathrm{LF}}$, and the *nonlinear* Padé–Fourier approximant $R_{m,n}^{\mathrm{NF}} = P_{m,n}^{\mathrm{NF}}/Q_{m,n}^{\mathrm{NF}}$, respectively, we look for trigonometric polynomials $P_{m,n}^{\mathrm{LF}}, P_{m,n}^{\mathrm{NF}}$ of degree $\leqslant m$ and $Q_{m,n}^{\mathrm{LF}}, Q_{m,n}^{\mathrm{NF}}$ of degree $\leqslant n$ such that either the linearized error $fQ_{m,n}^{\mathrm{LF}} - P_{m,n}^{\mathrm{LF}}$ or the error $f - R_{m,n}^{\mathrm{NF}}$ itself is orthogonal to the functions $\sin(jt)$ and $\cos(jt)$ for $j = 0, 1, \ldots, m+n$.

We have the following link between these rational approximants.

Theorem 3.1. *Let $n, k \geqslant 0$, and consider the real-valued Fourier series $f(t) := a_0/2 + \sum_{j=1}^{\infty}[a_j \cos(jt) + b_j \sin(jt)] \in L_2([-\pi, \pi])$ together with the associated series $G(z) = G(f)(z) = a_0/2 + \sum_{j=1}^{\infty}[a_j - ib_j]z^j$. Suppose that the linear Padé approximant $[n+k|k]_G = P/Q$ of G has no poles in the closed unit disk, i.e.,*

$$\deg P \leqslant n+k, \quad \deg Q \leqslant k, \quad (G(z)Q(z) - P(z))/z^{n+2k+1} \textit{is analytic around } 0, \quad \forall |z| \leqslant 1 : Q(z) \neq 0.$$

Then the nonlinear Padé–Fourier approximant $R_{n+k,k}^{\mathrm{NF}}$ of f exists, and

$$R_{n+k,k}^{\mathrm{NF}}(t) = \mathrm{Re}(\varepsilon_{2k}^{(n)}(t)) = \mathrm{Re}([n+k|k]_G(\mathrm{e}^{\mathrm{i}t})), \quad t \in [-\pi, \pi].$$

If moreover $f(t) = F(\cos(t))$ is even (and thus $b_j = 0$ for all j), then the nonlinear Padé–Chebyshev approximant $R_{n+k,k}^{\mathrm{NC}}$ of the Chebyshev series $F(x) = a_0/2 + \sum_{j=1}^{\infty} a_j T_j(x)$ exists, and

$$R_{n+k,k}^{\mathrm{NC}}(\cos(t)) = \mathrm{Re}(\varepsilon_{2k}^{(n)}(t)) = \mathrm{Re}([n+k|k]_G(\mathrm{e}^{\mathrm{i}t})), \quad t \in [-\pi, \pi].$$

Proof. Denote by P^* and Q^*, respectively, the polynomials obtained by taking the complex conjugate of the coefficients of P and Q, then

$$\mathrm{Re}([n+k|k]_G)(\mathrm{e}^{\mathrm{i}t}) = \frac{P(\mathrm{e}^{\mathrm{i}t})Q^*(\mathrm{e}^{-\mathrm{i}t}) + P^*(\mathrm{e}^{-\mathrm{i}t})Q(\mathrm{e}^{\mathrm{i}t})}{2Q(\mathrm{e}^{\mathrm{i}t})Q^*(\mathrm{e}^{-\mathrm{i}t})} =: R(t).$$

Here, the numerator and the denominator are trigonometric polynomials in t of degree $n+k$ and k, respectively, showing that R is indeed a candidate for the nonlinear Padé–Fourier approximant of type $(n+k, k)$ of f. If in addition f is even, then, with the coefficients of G, also the coefficients of P and Q can be chosen to be real. In this latter case, $P = P^*$ and $Q = Q^*$, implying that both numerator and denominator of $R(t)$ are even, and thus cosine polynomials. Using the relation $T_j(x) = \cos(j \arccos(x))$ it follows that $R(\arccos(x))$ is indeed a rational function in x, and thus a candidate for the nonlinear Padé–Chebyshev approximant of type $(n+k, k)$ of $F = f \circ \arccos$.

In order to conclude, we only need to show that the real-valued function $f - R$ is orthogonal to the functions $\cos(jt)$ and $\sin(jt)$ for $j = 0, 1, \ldots, n+2k$. We have for $j \in \{0, 1, \ldots, n+2k\}$

$$\begin{aligned} \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}jt}(f(t) - R(t))\,\mathrm{d}t &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{e}^{-\mathrm{i}jt}\left[G(\mathrm{e}^{\mathrm{i}t}) - \frac{P(\mathrm{e}^{\mathrm{i}t})}{Q(\mathrm{e}^{\mathrm{i}t})} + \overline{G(\mathrm{e}^{\mathrm{i}t}) - \frac{P(\mathrm{e}^{\mathrm{i}t})}{Q(\mathrm{e}^{\mathrm{i}t})}}\right]\mathrm{d}t \\ &= \frac{1}{2\pi i}\int_{|\zeta|=1}\left[G(\zeta) - \frac{P(\zeta)}{Q(\zeta)}\right]\frac{\mathrm{d}\zeta}{\zeta^{j+1}} + \overline{\frac{1}{2\pi i}\int_{|\zeta|=1}\left[G(\zeta) - \frac{P(\zeta)}{Q(\zeta)}\right]\zeta^j\frac{\mathrm{d}\zeta}{\zeta}}. \end{aligned}$$

By construction, G is an element of the Hardy space H_2, and so is $G - P/Q$ by assumption on Q. In particular, $G - P/Q$ is analytic in the unit disk and vanishes at zero, and therefore the second integral on the right-hand side vanishes. The

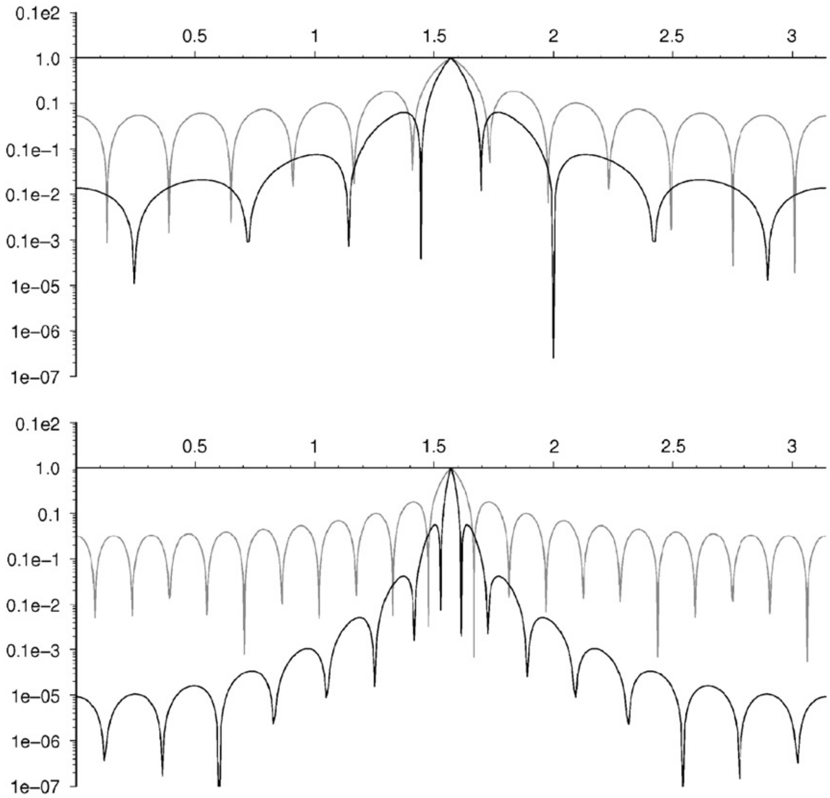

Fig. 2. The modulus of the error of approximation on a logarithmic scale at $x = \cos(t)$ of the linear (in gray) and the nonlinear Padé–Chebyshev approximant (in black) of $F(x) = \mathrm{sign}(x)$. We used the degrees $(n, k) = (4, 0)$ on the top and $(n, k) = (8, 4)$ on the bottom figure.

first integral equals the $(j+1)$th coefficient of the Taylor expansion of $G - P/Q$ at zero. By assumption, we have that $1/Q$ is analytic in the unit disk, and that the first $n + 2k + 1$ coefficients in the Taylor expansion of $GQ - P$ do vanish. Hence also the first integral equals zero, and the above claim follows by taking real and imaginary parts. $\square$

Let us illustrate the previous result with the sign function $F(x) = \mathrm{sign}(x)$. We have

$$f(t) = \mathrm{sign}(\cos(t)) = \frac{1}{\pi}\left(s\left(t - \frac{\pi}{2}\right) - s\left(t + \frac{\pi}{2}\right)\right),$$

and, according to (2.8), we get the Chebyshev series expansion

$$F(x) = \frac{4}{\pi}\sum_{j=0}^{\infty}\frac{(-1)^j}{2j+1}T_{2j+1}(x).$$

The convergence properties of the linear Padé–Chebyshev approximants $R^{\mathrm{LC}}_{n+k,k}$ of the sign function $F(x) = \mathrm{sign}(x)$ for fixed k and $n \to \infty$ has been discussed in detail in [18]. Paszkowski [25] gave an explicit expression for the nonlinear Padé–Chebyshev approximants $R^{\mathrm{NC}}_{n+k,k}$ of the sign function. Let us recover its denominator via Theorem 3.1: we have to compute the Padé approximant of $G = G(f)$, an odd function. In this case, it is well known and easy to verify that the Padé table of $G(f)$ has a 2×2 block structure

$$[2n + 2k - 1|2k]_G = [2n + 2k|2k]_G = [2n + 2k - 1|2k + 1]_G = [2n + 2k|2k + 1]_G. \tag{3.2}$$

In particular, the denominator of the linear Padé approximant of G of degree $[2n + 2k|2k + 1]_G$ are vanishing at zero, and the hypothesis of Theorem 3.1 fails to hold. However, for the other three members of the block (3.2), the denominator is the same, and its reversed counterpart has been given in Table 1(d)$_2$. In particular, all zeros of the denominator lie in $(-\mathrm{i}\infty, -i) \cup (i, \mathrm{i}\infty)$. Thus Theorem 3.1 gives us the following formula for the denominator of the nonlinear Padé–Chebyshev approximant of the sign function

$$Q^{\mathrm{NC}}_{2n+2k-1,2k}(\cos(t)) = |P_k^{(n-1/2,0)}(1+2\mathrm{e}^{2\mathrm{i}t})|^2$$

in terms of a Jacobi orthogonal polynomial.

In contrast, the approximants of index $(\mathcal{N}, \mathcal{M})$ for the sign function used in [18] for $\mathcal{N} \geqslant \mathcal{M} - 1$ are rational functions of numerator degree $2\mathcal{N}+1$ and denominator degree $2\mathcal{M}$, which coincide with the linear Padé–Chebyshev approximants $R^{\mathrm{LC}}_{2n+2k-1,2k} = R^{\mathrm{LC}}_{2n+2k,2k}$ for $n \geqslant 0$. The authors in [18] use an explicit formula for the denominator given by Németh and Páris in [23]. Our numerical experiments reported in Fig. 2 for $k=2$ and $n \in \{0,4\}$ seem to indicate that the nonlinear Padé–Chebyshev approximants have better approximation properties.

4. Error estimates for hypergeometric functions of type $G^{(\alpha,\beta)}$

In this section we study the Padé approximants to the functions

$$G^{(\alpha,\beta)}(z) = \left(\begin{array}{c} \alpha+1, 1 \\ \alpha+\beta+2 \end{array} \middle| z \right) = \int_0^1 \frac{\mathrm{d}\mu^{(\alpha,\beta)}(y)}{1-yz}, \tag{4.1}$$

where $\alpha, \beta > -1$ and the measure $\mathrm{d}\mu^{(\alpha,\beta)}$ has the support $[0,1]$ and the density

$$\mathrm{d}\mu^{(\alpha,\beta)}(y) = \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)} y^\alpha (1-y)^\beta \,\mathrm{d}y.$$

We notice that the diagonal Padé approximants $[k|k]_{G^{(\alpha,\beta)}}$ (and, by Table 1(c), also the offdiagonal Padé approximants) are the even convergents of the Gauss continued fraction [1, Chapter 5], from which one may conclude uniform convergence for $k \to \infty$ on compact subsets of $\mathbb{C} \backslash [1, +\infty)$. In the present context we are more interested in convergence on the unit circle including in particular points close to $z=1$. Also, we are interested in convergence of columns $[n+k|k]_{G^{(\alpha,\beta)}}$ for fixed k and $n \to \infty$.

Here it is useful to recall the well-known explicit representation of the Padé denominator of $G^{(\alpha,\beta)}$ in terms of Jacobi polynomials [1, Eq. (5.11)]. Indeed, we just need to use the representation of $G^{(\alpha,\beta)}$ as a Stieltjes function, see (4.1), to relate the reversed Padé denominators to the polynomials orthonormal with respect to the measure $\mathrm{d}\mu^{(\alpha,\beta)}$ on $[0,1]$, see [1, Eq. (7.7)].

Lemma 4.1. *Let $\alpha, \beta > -1$, and $k \geqslant 0, n \geqslant -1$ be two integers. The $(n+k-1, k)$ Padé denominator $Q_{n+k-1,k}$ of $G^{(\alpha,\beta)}$ is unique (up to multiplication with a scalar). More precisely, we may normalize such that for the reversed denominator we get the formula*

$$\widetilde{Q}_{n+k-1,k}(z) = z^k Q_{n+k-1,k}(1/z) = \gamma_{k,n}^{(\alpha,\beta)} P_k^{(\alpha+n,\beta)}(1-2z),$$

where $P_k^{(\alpha+n,\beta)}$ denotes the classical Jacobi polynomial, see (2.4), and

$$\gamma_{k,n}^{(\alpha,\beta)} = \sqrt{\frac{(2k+\alpha+n+\beta+1)k!(\alpha+\beta+2)_{n+k-1}}{(\alpha+1)_{n+k}(\beta+1)_k}}. \tag{4.2}$$

Moreover, for $|z|=1$, the following upper bound holds true:

$$|Q_{n+k-1,k}(z)^2(G^{(\alpha,\beta)}(z) - [n+k-1|k]_{G^{(\alpha,\beta)}}(z))| \leqslant \frac{1}{\mathrm{dist}(z,[0,1])}. \tag{4.3}$$

Proof. We have

$$G^{(\alpha,\beta)}(z) = \int \left[1 + yz + \cdots + (yz)^{n-1} + \frac{(yz)^n}{1-yz}\right] \mathrm{d}\mu^{(\alpha,\beta)}(y) = c_{n-1}(z) + z^n G_n^{(\alpha,\beta)}(z),$$

where $c_{n-1}(z)$ is a polynomial of degree $n-1$ in z and

$$G_n^{(\alpha,\beta)}(z) = \int_0^1 \frac{y^n \, \mathrm{d}\mu^{(\alpha,\beta)}(y)}{1-yz}.$$

The Padé approximant $[n+k-1|k]_{G^{(\alpha,\beta)}}$ is obtained from the Padé approximant $[k-1|k]_{G_n^{(\alpha,\beta)}}$ in the following way:

$$[n+k-1|k]_{G^{(\alpha,\beta)}}(z) = c_{n-1}(z) + z^n [k-1|k]_{G_n^{(\alpha,\beta)}}(z).$$

Since $G_n^{(\alpha,\beta)}$ is a Stieltjes function, the reversed Padé denominator $\widetilde{Q}_{n+k-1,k}(z)$ equals the orthonormal polynomial of degree k with respect to the measure $y^n \, \mathrm{d}\mu^{(\alpha,\beta)}(y)$ supported on the interval [0, 1] (up to normalization with a scalar), see, e.g., [1, Eq. (7.7)]. For the Jacobi polynomials of indices $(\alpha+n, \beta)$, we have

$$\int_0^1 P_i^{(\alpha+n,\beta)}(1-2y) P_j^{(\alpha+n,\beta)}(1-2y) y^n \, \mathrm{d}\mu^{(\alpha,\beta)}(y) = 0 \quad \text{for } i \neq j, \tag{4.4}$$

and it is easily checked from (2.4) that

$$\int_0^1 (P_k^{(\alpha+n,\beta)}(1-2y))^2 y^n \, \mathrm{d}\mu^{(\alpha,\beta)}(y) = \frac{(\alpha+1)_{n+k}(\beta+1)_k}{(2k+\alpha+n+\beta+1)k!(\alpha+\beta+2)_{n+k-1}} = (\gamma_{k,n}^{(\alpha,\beta)})^{-2}. \tag{4.5}$$

Hence, $\widetilde{Q}_{n+k-1,k}(z) = \gamma_{k,n}^{(\alpha,\beta)} P_k^{(\alpha+n,\beta)}(1-2z)$, and these polynomials are orthonormal with respect to the measure $y^n \, \mathrm{d}\mu^{(\alpha,\beta)}(y)$.

It is well known, see, e.g., [4, Chapter 1, Theorem 1.17], that the error of the Padé approximant $[k-1|k]$ of a function $f(z) = c(1/(1-xz))$, where c is a linear form acting on the variable x, is given by

$$f(z) - [k-1|k]_f(z) = \frac{z^{2k}}{\widetilde{P}_k(z)^2} c\left(\frac{P_k(x)^2}{1-xz}\right), \tag{4.6}$$

where P_k is the orthogonal polynomial of degree k with respect to the linear form c. Therefore, the linearized error of the Padé approximant $[k-1|k]_{G_n^{(\alpha,\beta)}}$ of the Stieltjes function $G_n^{(\alpha,\beta)}$ is given by

$$Q_{n+k-1,k}(z)(G_n^{(\alpha,\beta)}(z) - [k-1|k]_{G_n^{(\alpha,\beta)}}(z)) = \frac{z^{2k}}{Q_{n+k-1,k}(z)} \int \frac{\widetilde{Q}_{n+k-1,k}(y)^2}{1-zy} y^n \, \mathrm{d}\mu^{(\alpha,\beta)}(y),$$

which leads for $G^{(\alpha,\beta)}$ to the linearized error

$$Q_{n+k-1,k}(z) G^{(\alpha,\beta)} - P_{n+k-1,k}(z) = \frac{z^{2k+n}}{Q_{n+k-1,k}(z)} \int \frac{\widetilde{Q}_{n+k-1,k}(y)^2}{1-zy} y^n \, \mathrm{d}\mu^{(\alpha,\beta)}(y). \tag{4.7}$$

From (4.7) and the orthonormality properties of $\widetilde{Q}_{n+k-1,k}(z)$, we obtain for $|z|=1$, $z \neq 1$ the following upper bound

$$|Q_{n+k-1,k}(z)(Q_{n+k-1,k}(z) G^{(\alpha,\beta)} - P_{n+k-1,k}(z))| \leqslant \frac{1}{\mathrm{dist}(z, [0,1])}. \qquad \square$$

The next lemma gives estimates on the modulus of $Q_{n+k-1,k}(\mathrm{e}^{\mathrm{i}t})$.

Lemma 4.2. *Set*

$$v_{k,n} = \left|Q_{n+k-1,k}(-1)\right| = \gamma_{k,n}^{(\alpha,\beta)} P_k^{(\alpha+n,\beta)}(3). \tag{4.8}$$

Then,

$$|Q_{n+k-1,k}(\mathrm{e}^{\mathrm{i}t})|^2 \leqslant v_{k,n}^2 \leqslant \frac{2^{2k}}{k!}\frac{(\alpha+\beta+2)_{n+2k}(n+k+1+\alpha+\beta)_k}{(\alpha+1)_{n+k}(\beta+1)_k}, \tag{4.9}$$

and for $0<\delta\leqslant|t|\leqslant\pi$, *we have*

$$|Q_{n+k-1,k}(\mathrm{e}^{\mathrm{i}t})|^2 \geqslant |Q_{n+k-1,k}(\mathrm{e}^{\mathrm{i}\delta})|^2 \geqslant \frac{2^k}{k!}\frac{(\alpha+\beta+2)_{n+2k}}{(\alpha+1)_n(\beta+1)_k}(1-\cos\,\delta)^k. \tag{4.10}$$

Proof. We recall the following representation of the Jacobi polynomials, see [28, Chapter 4, Eq. 4.21.2, p. 62],

$$P_k^{(\alpha,\beta)}(x) = \frac{(\alpha+1)_k}{k!}\,{}_2F_1\left(\begin{matrix}-k,\,k+\alpha+\beta+1\\ \alpha+1\end{matrix}\,\middle|\,\frac{1-x}{2}\right). \tag{4.11}$$

Thus we get for the reversed denominators

$$\widetilde{Q}_{n+k-1,k}(z) = \gamma_{k,n}^{(\alpha,\beta)}\frac{(\alpha+n+1)_k}{k!}\,{}_2F_1\left(\begin{matrix}-k,\,\alpha+n+1+\beta+k\\ \alpha+n+1\end{matrix}\,\middle|\,z\right). \tag{4.12}$$

For $\alpha,\beta>-1$, we denote by a_i the coefficient of x^i of the ${}_2F_1$ polynomial. Then

$$\mathrm{sign}(a_i) = \mathrm{sign}((-k)(-k+1)\cdots(-k+i-1)) = (-1)^i,$$

and the coefficients of $\widetilde{Q}_{n+k-1,k}(z)$ have alternating signs, which implies that

$$\max_{|z|=1}\left|\widetilde{Q}_{n+k-1,k}(z)\right| = \left|\widetilde{Q}_{n+k-1,k}(-1)\right|.$$

Denote by $x_{j,k}\in(0,1)$, $j=1,\ldots,k$, the zeros of $\widetilde{Q}_{n+k-1,k}$. We have

$$\begin{aligned}\left|\widetilde{Q}_{n+k-1,k}(-1)\right| &= \gamma_{k,n}^{(\alpha,\beta)}\frac{(\alpha+n+1)_k}{k!}\frac{|(-k)_k|}{k!}\frac{(\alpha+n+\beta+k+1)_k}{(\alpha+n+1)_k}\left|\prod_{j=1}^{k}(-1-x_{jk})\right|\\ &\leqslant 2^k\gamma_{k,n}^{(\alpha,\beta)}\frac{(\alpha+n+\beta+k+1)_k}{k!}.\end{aligned}$$

Plugging the expression (4.2) of $\gamma_{k,n}^{(\alpha,\beta)}$ into the square of the last upper bound leads, after some computations to (4.9). Let us now obtain the lower bound. We can write

$$\widetilde{Q}_{n+k-1,k}(z) = \widetilde{Q}_{n+k-1,k}(0)\prod_{j=1}^{k}\left(1-\frac{z}{x_{jk}}\right).$$

Then we obtain

$$|\widetilde{Q}_{n+k-1,k}(\mathrm{e}^{\mathrm{i}t})|^2 = |\widetilde{Q}_{n+k-1,k}(0)|^2\prod_{j=1}^{k}\frac{2}{x_{jk}}\left(\frac{x_{jk}+1/x_{jk}}{2}-\cos(t)\right).$$

This shows that $\left|\widetilde{Q}_{n+k-1,k}(\mathrm{e}^{\mathrm{i}t})\right|^2$ is increasing with t and so, for $0<\delta\leqslant|t|<\pi$,

$$\begin{aligned}|\widetilde{Q}_{n+k-1,k}(\mathrm{e}^{\mathrm{i}t})|^2 &\geqslant |\widetilde{Q}_{n+k-1,k}(\mathrm{e}^{\mathrm{i}\delta})|^2\\ &\geqslant \left|\widetilde{Q}_{n+k-1,k}(0)\right|^2\prod_{j=1}^{k}\left[\frac{2}{x_{jk}}(1-\cos(\delta))\right] = \frac{2^k\left|\widetilde{Q}_{n+k-1,k}(0)\right|^2}{\prod_{j=1}^{k}x_{jk}}(1-\cos(\delta))^k.\end{aligned} \tag{4.13}$$

The quotient $\widetilde{Q}_{n+k-1,k}(0)/\prod_{j=1}^{k}(-x_{jk})$ equals the leading coefficient of $\widetilde{Q}_{n+k-1,k}(x)$, which, in view of (4.12), equals

$$\gamma_{k,n}^{(\alpha,\beta)}\frac{(\alpha+n+1)_k}{k!}\frac{(-k)_k(\alpha+n+\beta+k+1)_k}{(\alpha+n+1)_k k!}=(-1)^k\gamma_{k,n}^{(\alpha,\beta)}\frac{(\alpha+n+\beta+k+1)_k}{k!}.$$

On the other hand,

$$\widetilde{Q}_{n+k-1,k}(0)=\gamma_{k,n}^{(\alpha,\beta)}P_k^{(\alpha+n,\beta)}(1)=\gamma_{k,n}^{(\alpha,\beta)}\frac{(\alpha+n+1)_k}{k!}. \qquad \square$$

Making use of these two expressions in (4.13) leads to (4.10).

Corollary 4.3. *Assume f is such that* $G(f)=G^{(\alpha,\beta)}$. *Then, for* $0<\delta\leqslant|t|\leqslant\pi$ *and for all integers* $k\geqslant 0$, $n\geqslant -1$, *we have that*

$$|f(t)-\mathrm{Re}(\varepsilon_{2k}^{(n)}(t))|\leqslant\frac{k!(\beta+1)_k(\alpha+1)_{n+1}}{2^k\sin\delta(1-\cos\delta)^k(\alpha+\beta+2)_{n+1+2k}}, \tag{4.14}$$

which implies that, for any $0<\delta\leqslant\pi$, $0\leqslant\tau\leqslant 1$ *and* $k\geqslant 0$,

$$\max_{\delta/n^{\tau}\leqslant|t|\leqslant\pi}|f(t)-\mathrm{Re}(\varepsilon_{2k}^{(n)}(t))|=\mathcal{O}(n^{-(1-\tau)(2k+1)-\beta})\quad\text{as } n\to\infty. \tag{4.15}$$

Proof. We know that

$$f(t)-\mathrm{Re}(\varepsilon_{2k}^{(n)}(t))=\mathrm{Re}(G(f)-[n+k|k]_{G(f)})(\mathrm{e}^{\mathrm{i}t}),$$

and, by (4.3), the modulus of the last expression is less than

$$(\mathrm{dist}(\mathrm{e}^{\mathrm{i}t},[0,1])|Q_{n+k,k}(\mathrm{e}^{\mathrm{i}t})|^2)^{-1}\leqslant\frac{k!(\alpha+1)_{n+1}(\beta+1)_k}{2^k\sin\delta(1-\cos\delta)^k(\alpha+\beta+2)_{n+1+2k}},$$

for $0<\delta\leqslant|t|\leqslant\pi$, where in the last inequality we have used (4.10). This proves (4.14) from which (4.15) is immediate. $\square$

This result shows the quite impressive convergence properties of the columns of the Padé table for the Stieltjes function $G^{(\alpha,\beta)}$. If we fix the parameters $\delta>0$, $\tau=0$ and the column k of the table, then the error corresponding to the sequence of approximants $[n+k|k](\mathrm{e}^{\mathrm{i}t})$ is of order $\mathcal{O}(n^{-2k-1-\beta})$ as n tends to infinity. This fact explains for the fast convergence of the approximants observed when applying the ε-algorithm to the Fourier series of functions like $G^{(\alpha,\beta)}$.

5. Error estimates for the sum of a $G^{(\alpha,\beta)}$ function and a smooth function

In this section, we show that the results of Corollary 4.3 remain valid when adding a smooth perturbation to a function *f* as in Section 4.

Theorem 5.1. *Let* $f=f_1+f_2$ *with* $G(f_1)=G^{(\alpha,\beta)}$, $\alpha,\beta>-1$, *and* $f_2\in\mathscr{C}^{m-1}(\mathbb{R})$ *a* 2π*-periodic function such that* $f_2^{(m)}$ *exists almost everywhere, with* $f_2^{(m)}\in L_1([0,2\pi])$. *Let* $0<\delta\leqslant\pi$, $0\leqslant\tau<1$, *and* $k\geqslant 0$ *an integer such that* $m\geqslant 2k+5/2+\beta-\tau$. *Then the same estimate as in Corollary* 4.3 *holds true, namely,*

$$\max_{\delta/n^{\tau}\leqslant|t|\leqslant\pi}|f(t)-\mathrm{Re}(\varepsilon_{2k}^{(n)}(t))|=\mathcal{O}(n^{-(1-\tau)(2k+1)-\beta})\quad \textit{as } n\to\infty.$$

In the sequel, we set

$$p_{k,n}(x)=\gamma_{k,n}^{(\alpha,\beta)}P_k^{(\alpha+n,\beta)}(1-2x),$$

that is, $p_{k,n}$ is the orthonormal polynomial of degree k with respect to the measure $y^n\,\mathrm{d}\mu^{(\alpha,\beta)}(y)$ (and thus the reversed $(n+k-1,k)$ Padé denominator of $G(f_1)$, see the preceding section). We also set

$$E_n(f) = f - S_{n-1}(f)$$

for the remainder of order n in the Fourier expansion of f.

Before proving Theorem 5.1, we establish three preliminary lemmas.

Lemma 5.2. *Let $f = f_1 + f_2$ be a function on $[-\pi,\pi]$ with $G(f_1) = G^{(\alpha,\beta)}$ and assume that the Fourier coefficients of f_2 decrease sufficiently fast. Namely, we suppose that there exists an $N_k \geqslant 0$ such that, for all $n \geqslant N_k$,*

$$\|E_n(f_2)\|_2 \left(\sum_{j=0}^{k} v_{j,n}^2\right) < \frac{1}{2}, \tag{5.1}$$

where the numbers $v_{j,n} = |p_{j,n}(-1)|$ have been defined in (4.8)*. Then, for all $n \geqslant N_k$, the $(n+k-1,k)$ Padé denominator $Q_{n+k-1,k}$ of $G(f)$ is unique* (*up to multiplication with a scalar*), *and its reversed counterpart admits the decomposition*

$$\widetilde{Q}_{n+k-1,k}(z) = p_{k,n}(z) + \sum_{j=0}^{k-1} a_{j,k,n}\, p_{j,n}(z), \tag{5.2}$$

with coefficients $a_{j,k,n}$ satisfying

$$|a_{j,k,n}| \leqslant 2 v_{j,n} v_{k,n} \|E_n(f_2)\|_2 < 1, \quad j = 0,\ldots,k-1. \tag{5.3}$$

Proof. Let $c^{(n)}$ be the linear form acting on the space of polynomials such that $c^{(n)}(z^j)$ is the coefficient of z^{n+j} in the power series of $G(f)$. In view of the integral representation (4.1) of $G^{(\alpha,\beta)}$ and the Cauchy formula for $G(f_2) \in H_2$,

$$G(f_2)(z) = \frac{1}{2\pi i}\int_{|\xi|=1} \frac{G(f_2)(\xi)}{\xi - z}\,\mathrm{d}\xi,$$

we have that for any polynomial P,

$$c^{(n)}(P) = \int y^n P(y)\,\mathrm{d}\mu^{(\alpha,\beta)}(y) + \frac{1}{2\pi i}\int_{|\xi|=1} \xi^{-n-1} P\left(\frac{1}{\xi}\right) G(f_2)(\xi)\,\mathrm{d}\xi. \tag{5.4}$$

The second integral equals

$$\frac{1}{2\pi i}\int_{|\xi|=1} \xi^{n-1} P(\xi) G(f_2)(1/\xi)\,\mathrm{d}\xi = \frac{1}{2\pi i}\int_{|\xi|=1} \xi^{n-1} P(\xi) g_n(1/\xi)\,\mathrm{d}\xi,$$

where in the last inequality we have set $g_n(\xi) = G(E_n(f_2))(\xi)$, and have used the fact that $\xi^{n-1}(f_2(1/\xi) - E_n(f_2)(1/\xi)) = \xi^{n-1} S_{n-1}(f_2)(1/\xi)$ is analytic in the closed unit disk.

Let us suppose for a moment that there is a reversed denominator $\widetilde{Q}_{n+k-1,k}$ of the Padé approximant $[n+k-1|k]_{G(f)}$ which is of degree exactly k. Then, after possibly multiplying with a scalar, there exists coefficients $a_{j,k,n}$ such that (5.2) holds. The order condition for the linearized Padé error yields the orthogonality conditions

$$c^{(n)}(\widetilde{Q}_{n+k-1,k}\, p_{j,n}) = 0, \quad j = 0,\ldots,k-1. \tag{5.5}$$

Setting

$$A_{l,j} = -\frac{1}{2\pi i}\int_{|\xi|=1} \xi^{n-1} p_{j,n}(\xi) p_{l,n}(\xi) g_n\left(\frac{1}{\xi}\right)\mathrm{d}\xi, \quad l,j = 0,\ldots,k-1,$$

these orthogonality relations rewrite as

$$(I - A)a = b,$$

where A denotes the matrix $(A_{l,j})_{l,j=0,\dots,k-1}$, $a=(a_{0,k,n},\dots,a_{k-1,k,n})^{\mathrm{T}}$ and $b=(A_{0,k},\dots,A_{k-1,k})^{\mathrm{T}}$. From the Cauchy–Schwarz inequality together with (3.1) and (4.9) we obtain that

$$\left|A_{l,j}\right|^2 \leqslant \left(\frac{1}{2\pi}\int_{|\xi|=1}\left|p_{j,n}(\xi)p_{l,n}(\xi)\right|^2|\mathrm{d}\xi|\right)\left(\frac{1}{2\pi}\int_{|\xi|=1}|g_n(1/\xi)|^2|\mathrm{d}\xi|\right) \leqslant v_{j,n}^2 v_{l,n}^2\|E_n(f_2)\|_2^2, \tag{5.6}$$

so that, in view of (5.1), we have the following upper bound for the Frobenius norm of A,

$$\|A\|_F^2=\sum_{l,j=0}^{k-1}A_{j,l}^2 \leqslant \|E_n(f_2)\|_2^2\left(\sum_{j=0}^{k-1}v_{j,n}^2\right)^2<\frac{1}{4}.$$

As a consequence, the matrix $(I-A)$ is invertible and the vector $a=(a_{0,k,n},\dots,a_{k-1,k,n})^{\mathrm{T}}$ is given by

$$a=(I-A)^{-1}b=\sum_{m=0}^{\infty}A^m b.$$

Let us show by induction on m that

$$\forall m\geqslant 0,\quad |(A^m b)_j|\leqslant \frac{v_{j,n}v_{k,n}}{2^m}\|E_n(f_2)\|_2,\quad j=0,\dots,k-1. \tag{5.7}$$

When $m=0$, this is true, see (5.6). Assume we have

$$|(A^{m-1}b)_j|\leqslant \frac{v_{j,n}v_{k,n}}{2^{m-1}}\|E_n(f_2)\|_2,\quad j=0,\dots,k-1.$$

Then,

$$\begin{aligned}|(A^m b)_j|^2 &\leqslant \left(\sum_{l=0}^{k-1}A_{j,l}^2\right)\sum_{l=0}^{k-1}|(A^{m-1}b)_l|^2\\ &\leqslant \left(v_{j,n}^2\left(\sum_{l=0}^{k-1}v_{l,n}^2\right)\|E_n(f_2)\|_2^2\right)\left(\frac{v_{k,n}^2}{4^{m-1}}\left(\sum_{l=0}^{k-1}v_{l,n}^2\right)\|E_n(f_2)\|_2^2\right)\\ &\leqslant 4^{-m}v_{k,n}^2v_{j,n}^2\|E_n(f_2)\|_2^2,\end{aligned}$$

where in the last inequality we have used (5.1). Hence, (5.7) holds true for any $m\in\mathbb{N}$. It implies that

$$\left|a_{j,k,n}\right| \leqslant \sum_{m=0}^{\infty}\left|(A^m b)_j\right| \leqslant 2v_{j,n}v_{k,n}\|E_n(f_2)\|_2,$$

where the last upper bound is less than 1 in view of (5.1).

It finally remains to show that $\widetilde{Q}_{n+k-1,k}$ is necessarily of degree k (which also implies that this reversed Padé denominator is unique up to multiplication with a constant). By contradiction, suppose that $\kappa:=\deg\widetilde{Q}_{n+k-1,k}<k$. Then, after possibly multiplying with a scalar, we may write

$$\widetilde{Q}_{n+k-1,k}(z)=p_{\kappa,n}(z)+\sum_{j=0}^{\kappa-1}a_{j,\kappa,n}p_{j,n}(z),$$

and get the same estimates for the coefficients $A_{l,j}$ and $a_{j,\kappa,n}$. In particular, relation (5.5) for $j=\kappa$ leads to

$$
\begin{aligned}
0 = |c^{(n)}(p_{\kappa,n}\widetilde{Q}_{n+k-1,k})| &= \left|1 - A_{\kappa,\kappa} - \sum_{j=0}^{\kappa-1} a_{j,\kappa,n}A_{\kappa,j}\right| \\
&\geqslant 1 - |A_{\kappa,\kappa}| - \sum_{j=0}^{\kappa-1} |a_{j,\kappa,n}|\,|A_{\kappa,j}| \geqslant 1 - v_{\kappa,n}^2\,\|E_n(f_2)\| - \frac{1}{2}v_{\kappa,n}^2\,\|E_n(f_2)\| \geqslant \frac{1}{4},
\end{aligned}
$$

a contradiction. Hence $\deg \widetilde{Q}_{n+k-1,k} = k$. $\square$

We now give a lower bound for the modulus of $\widetilde{Q}_{n+k-1,k}(z)$.

Lemma 5.3. *Let $f = f_1 + f_2$ satisfy the assumptions of Lemma 5.2. Then, for all $n \geqslant N_k$, we have*

$$
|\widetilde{Q}_{n+k-1,k}(\mathrm{e}^{\mathrm{i}t})| \geqslant \frac{1}{2}|p_{k,n}(\mathrm{e}^{\mathrm{i}t})| \quad \textit{provided that } \frac{\pi}{3}\frac{\sqrt{k(\beta+k)}}{n-1} \leqslant |t| \leqslant \pi. \tag{5.8}
$$

Proof. By the interlacing property of the (simple) zeros $x_{l,j} \in (0,1)$, $l = 1, \ldots, j$, of the orthonormal polynomials $p_{j,n}$, we can write

$$
\frac{p_{j-1,n}(z)}{p_{j,n}(z)} = \sum_{l=1}^{j} \frac{\beta_{lj}}{z - x_{l,j}}, \qquad j = 1, \ldots, k,
$$

with $\beta_{lj} > 0$. Then,

$$
\begin{aligned}
\left|\frac{p_{j-1,n}(z)}{p_{j,n}(z)}\right| &\leqslant \sum_{l=1}^{j} \frac{|x_{l,j}|}{|z - x_{l,j}|}\frac{\beta_{lj}}{|x_{l,j}|} \\
&\leqslant \max_{1\leqslant l\leqslant j}\left|\frac{x_{l,j}}{z - x_{l,j}}\right| \left|\frac{p_{j-1,n}(0)}{p_{j,n}(0)}\right| \leqslant \frac{1}{\mathrm{dist}(z,[0,1])}\left|\frac{p_{j-1,n}(0)}{p_{j,n}(0)}\right|.
\end{aligned}
$$

Since $p_{j,n}(0) = \gamma_{j,n}^{(\alpha,\beta)} P_j^{(n+\alpha,\beta)}(1)$, we obtain using (4.11) that, for $j = 1, \ldots, k$,

$$
\begin{aligned}
\left|\frac{p_{j-1,n}(0)}{p_{j,n}(0)}\right| &= \frac{\gamma_{j-1,n}^{(\alpha,\beta)}}{\gamma_{j,n}^{(\alpha,\beta)}}\left|\frac{P_{j-1}^{(\alpha+n,\beta)}(1)}{P_j^{(\alpha+n,\beta)}(1)}\right| \\
&= \sqrt{\frac{2j-1+\alpha+\beta+n}{2j+1+\alpha+\beta+n}\,\frac{j(\beta+j)}{(\alpha+\beta+n+j)(\alpha+j+n)}} \leqslant \frac{\sqrt{k(\beta+k)}}{n-1}.
\end{aligned}
$$

Thus, by assumption on t,

$$
\left|\frac{p_{j-1,n}(\mathrm{e}^{\mathrm{i}t})}{p_{j,n}(\mathrm{e}^{\mathrm{i}t})}\right| \leqslant \frac{1}{\mathrm{dist}(\mathrm{e}^{\mathrm{i}t},[0,1])}\frac{\sqrt{k(\beta+k)}}{n-1} \leqslant \frac{1}{\sin|t/2|}\frac{\sqrt{k(\beta+k)}}{n-1} \leqslant \frac{1}{3}.
$$

Since from (5.3) we know that $|a_{j,k,n}| < 1$, we obtain

$$
\left|\frac{\widetilde{Q}_{n+k-1,k}(\mathrm{e}^{\mathrm{i}t})}{p_{k,n}(\mathrm{e}^{\mathrm{i}t})}\right| \geqslant 1 - \left|\sum_{j=0}^{k-1} a_{j,k,n}\frac{p_{j,n}(\mathrm{e}^{\mathrm{i}t})}{p_{k,n}(\mathrm{e}^{\mathrm{i}t})}\right| \geqslant 1 - \sum_{j=0}^{k-1}\left|\frac{p_{j,n}(\mathrm{e}^{\mathrm{i}t})}{p_{k,n}(\mathrm{e}^{\mathrm{i}t})}\right| \geqslant 1 - \sum_{j=0}^{k-1}\frac{1}{3^{k-j}} = \frac{1}{2}. \qquad \square
$$

Lemma 5.4. *Let $f = f_1 + f_2$ satisfy the assumptions of Lemma* 5.2, *and let*

$$e_{n,k}(z) = G(f)(z) - [n+k-1|k]_{G(f)}(z)$$

be the error corresponding to the Padé approximant $[n+k-1|k]_{G(f)}$. Then for all $n \geqslant N_k$ and $|z| = 1$ we have

$$|Q_{n+k-1,k}(z)^2 e_{n,k}(z)| \leqslant \frac{2}{\operatorname{dist}(z,[0,1])} + 4v_{k,n}^2 \sum_{j=0}^{\infty} \|E_{n+j}(f_2)\|_2. \tag{5.9}$$

Proof. By adapting the reasoning leading to (4.7), the error $e_{n,k}(z)$ can be written in the following way:

$$e_{n,k}(z) = \frac{z^{n+2k}}{Q_{n+k-1,k}(z)^2} c^{(n)} \left(\frac{\widetilde{Q}_{n+k-1,k}(x)^2}{1-xz} \right),$$

where $c^{(n)}$ has been defined at the beginning of the proof of Lemma 5.2. Replacing $c^{(n)}$ by the expression obtained there, we get

$$\begin{aligned} &Q_{n+k-1,k}(z)^2 e_{n,k}(z) \\ &\quad = z^{n+2k} \int \frac{\widetilde{Q}_{n+k-1,k}(y)^2}{1-yz} y^n \, \mathrm{d}\mu^{(\alpha,\beta)}(y) + \frac{z^{n+2k}}{2\pi i} \int_{|\xi|=1} \frac{\xi^{n-1} \widetilde{Q}_{n+k-1,k}(\xi)^2}{1-\xi z} g_n(1/\xi) \, \mathrm{d}\xi. \end{aligned}$$

Let us denote by I_1 and I_2 the two terms in the previous sum. We first bound the modulus of I_1. Using the decomposition (5.2), we have

$$\widetilde{Q}_{n+k-1,k}(y)^2 = \sum_{j,l=0}^{k} a_{j,k,n} a_{l,k,n} p_{j,n}(y) p_{l,n}(y),$$

where $a_{k,k,n} = 1$. From the orthonormality of the $p_{j,n}$ with respect to the measure $y^n \mathrm{d}\mu^{(\alpha,\beta)}(y)$ and the fact that $|z| = 1$, we obtain

$$|I_1| \leqslant \operatorname{dist}(z,[0,1])^{-1} \int \widetilde{Q}_{n+k-1,k}(y)^2 y^n \mathrm{d}\mu^{(\alpha,\beta)}(y) = \operatorname{dist}(\mathrm{z},[0,1])^{-1} \sum_{j=0}^{k} a_{j,k,n}^2. \tag{5.10}$$

Moreover, from (5.3) and assumption (5.1), we derive that

$$\sum_{j=0}^{k-1} a_{j,k,n}^2 \leqslant 4v_{k,n}^2 \|E_n(f_2)\|_2^2 \left(\sum_{j=0}^{k-1} v_{j,n}^2 \right) \leqslant 1,$$

which, together with (5.10) and the fact that $a_{k,k,n} = 1$, shows that

$$|I_1| \leqslant 2/\operatorname{dist}(z,[0,1]). \tag{5.11}$$

For the second term, we have

$$\begin{aligned} I_2 &= \frac{z^{n+2k}}{2\pi i} \sum_{j=0}^{\infty} z^j \int_{|\xi|=1} \xi^{n-1+j} \widetilde{Q}_{n+k,k}(\xi)^2 g_n(1/\xi) \, \mathrm{d}\xi \\ &= \frac{z^{n+2k}}{2\pi i} \sum_{j=0}^{\infty} z^j \int_{|\xi|=1} \xi^{n-1+j} \widetilde{Q}_{n+k-1,k}(\xi)^2 g_{n+j}(1/\xi) \, \mathrm{d}\xi, \end{aligned}$$

where in the second equality, we have used the fact that $\xi^{n+j-1}(g_{n+j}-g_n)(1/\xi)$ is analytic in the unit disk. Then, by applying the Cauchy–Schwarz inequality to the integrals, we obtain that the modulus of I_2 satisfies

$$|I_2| \leqslant \max_{|z|=1} |\widetilde{Q}_{n+k-1,k}(z)^2| \sum_{j=0}^{\infty} \left[\frac{1}{2\pi} \int_{|\xi|=1} \left|g_{n+j}(\xi)\right|^2 |\mathrm{d}\xi| \right]^{1/2}$$

$$= \max_{|z|=1} |\widetilde{Q}_{n+k-1,k}(z)^2| \sum_{j=0}^{\infty} \|E_{n+j}(f_2)\|_2. \tag{5.12}$$

Using (5.1)–(5.3), and the first inequality in (4.9), we obtain, for $n \geqslant N_k$ and $|z|=1$,

$$\left|\widetilde{Q}_{n+k-1,k}(z)\right|^2 = \left| \sum_{j=0}^{k} a_{j,k,n} p_{j,n}(z) \right|^2 \leqslant 2 \left(v_{k,n}^2 + \left| \sum_{j=0}^{k-1} a_{j,k,n} p_{j,n}(z) \right|^2 \right)$$

$$\leqslant 2v_{k,n}^2 + 8v_{k,n}^2 \|E_n(f_2)\|_2^2 \left(\sum_{j=0}^{k-1} v_{j,n}^2 \right)^2 \leqslant 4v_{k,n}^2. \qquad \square$$

Hence, inequality (5.9) follows from (5.11) and (5.12).

Proof of Theorem 5.1. Since $f_2^{(m)} \in L_1$, we know from the Riemann–Lebesgue Lemma [27, (5.14)] that its Fourier coefficients satisfy

$$|a_j(f_2^{(m)})| + |b_j(f_2^{(m)})| = \mathrm{o}(1) \quad \text{as } j \to \infty.$$

Taking into account that $|a_j(f_2^{(m)})|+|b_j(f_2^{(m)})| = j^m(|a_j(f_2)|+|b_j(f_2)|)$ for $j \geqslant 1$, we may conclude that, as $n \to \infty$,

$$\|E_n(f_2)\|_2 = \mathrm{o}(n^{1/2-m}), \qquad \sum_{j=0}^{\infty} \|E_{n+j}(f_2)\|_2 = \mathrm{o}(n^{3/2-m}).$$

On the other hand, we know from (4.9) that $v_{k,n}^2 = \mathcal{O}(n^{2k+1+\beta})$ as $n \to \infty$ for fixed α, β, k, and $1/2 - m + 2k + 1 + \beta \leqslant \tau - 1 \leqslant 0$, by assumption on f_2 and τ. Hence, as $n \to \infty$,

$$\|E_n(f_2)\|_2 \left(\sum_{j=0}^{k} v_{j,n}^2 \right) = \mathrm{o}(1), \qquad v_{k,n}^2 \sum_{j=0}^{\infty} \|E_{n+j}(f_2)\|_2 = \mathrm{o}(n^{\tau}),$$

and the assumption of Lemma 5.2 is true for sufficiently large N_k by the choice of m. We also observe that

$$\left[\frac{\delta}{n^{\tau}}, \pi \right] \subset \left[\frac{\pi}{3} \frac{\sqrt{k(\beta+k)}}{n-1}, \pi \right],$$

for sufficiently large n by assumption on δ, τ.

Using Lemmas 5.3 and 5.4 we get for the Padé error that

$$\max_{\delta/n^{\tau} \leqslant |t| \leqslant \pi} \left| e_{n,k}(\mathrm{e}^{\mathrm{i}t}) \right| \leqslant \max_{\delta/n^{\tau} \leqslant |t| \leqslant \pi} \frac{4}{\left|p_{n,k}(\mathrm{e}^{\mathrm{i}t})\right|^2} \left[\frac{2}{\operatorname{dist}(\mathrm{e}^{\mathrm{i}\delta/n^{\tau}}, [0,1])} + \mathrm{o}(n^{\tau})_{n\to\infty} \right],$$

and hence Theorem 5.1 follows from (4.10). $\square$

Notice that, according to the explicit form of Lemmas 5.3 and 5.4, it would be possible to give more explicit bounds for the Padé error in case where explicit expressions for $\|E_n(f_2)\|_2$ are available.

6. Numerical results

We have seen already in Fig. 1 that for the saw tooth function s of (2.2) and hence

$$G(s)(z) = -2i\log(1-z) = 2izG^{(0,0)}(z)$$

(compare with Table 1(d$_1$)) we have an impressive acceleration of convergence via the ε-algorithm even for low order. Indeed, as shown in Corollary 4.3, the error is dominated by the Padé error on the unit circle of the diagonal approximants $[3|3]_{G(s)}(z)=2iz[2|3]_{G^{(0,0)}}(z)$ and $[8|8]_{G(s)}(z)=2iz[7|8]_{G^{(0,0)}}(z)$, which is quite small: for the second approximant we obtain for $z=\mathrm{e}^{\mathrm{i}t}$, $\delta=\pi/4\leqslant|t|\leqslant\pi$ from Corollary 4.3 the upper bound $3.57\times10^{-8}/[\sin\delta(1-\cos\delta)^8]=8.08\times10^{-7}$, which is approximately attained for $t=\delta$.

Let us compare in Fig. 3 these findings with a 2π-periodic function being $\mathscr{C}^0$ but having a derivative with a discontinuity at zero, namely

$$f(t) = \left|\sin\left(\frac{t}{2}\right)\right| = \frac{2}{\pi} - \frac{4\cos(t)}{3\times5\times\pi} - \frac{4\cos(2t)}{5\times7\times\pi} - \frac{4\cos(3t)}{7\times9\times\pi} - \cdots \Longrightarrow G(f)(z) = \frac{2}{\pi} - \frac{4z}{3\pi}G^{(-1/2,1)}(z)$$

(compare with Table 1 (d$_1$)). We again observe that the error for the partial sums is strongly oscillating, and remains quite large even for higher order Fourier sums, namely about $\frac{1}{100}$ for order 7, and 1/1000 for order 17 (smaller as for the saw tooth function in Fig. 1, since of cause the latter function is less regular). In this example we see that both linear acceleration procedures, namely the Cesaro means and the de la Vallée–Poussin mean, give very disappointing

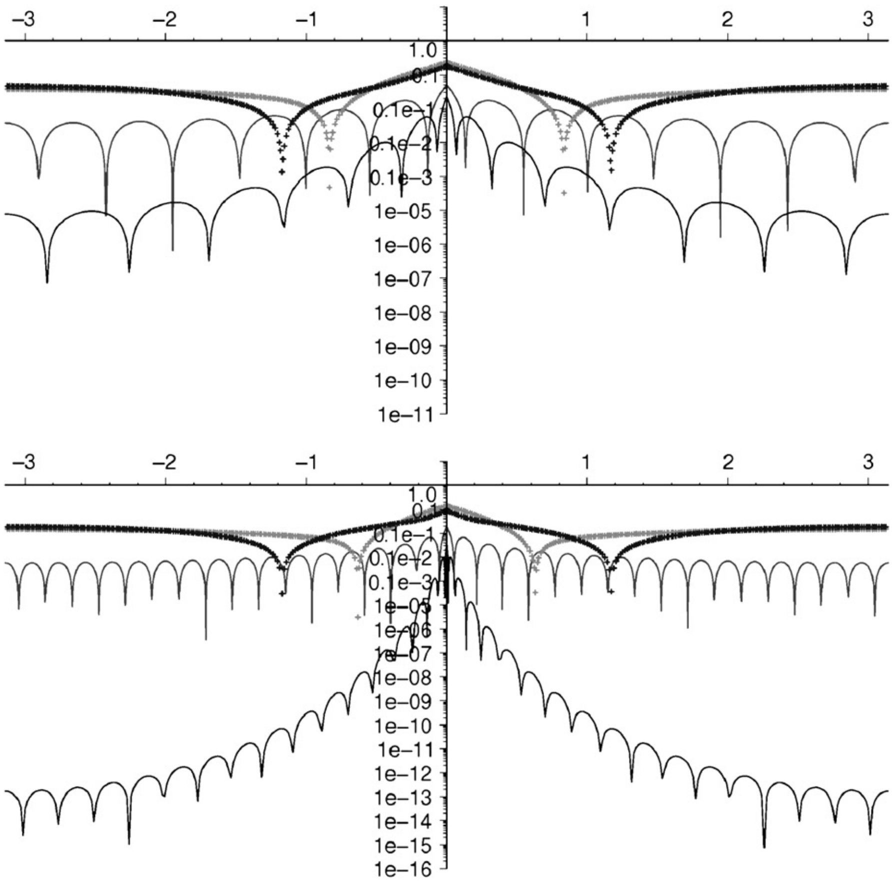

Fig. 3. The modulus of the error of approximation on a logarithmic scale of the Fourier series of $f(t)=|\sin(t/2)|$. We have drawn the error of the partial sums (in solid gray), of their Cesaro means (in crossed black), the de la Vallée–Poussin mean (in crossed gray) and Padé error (in solid black), where on the top we use the first 7 and on the bottom figure the first 17 coefficients of the Fourier series.

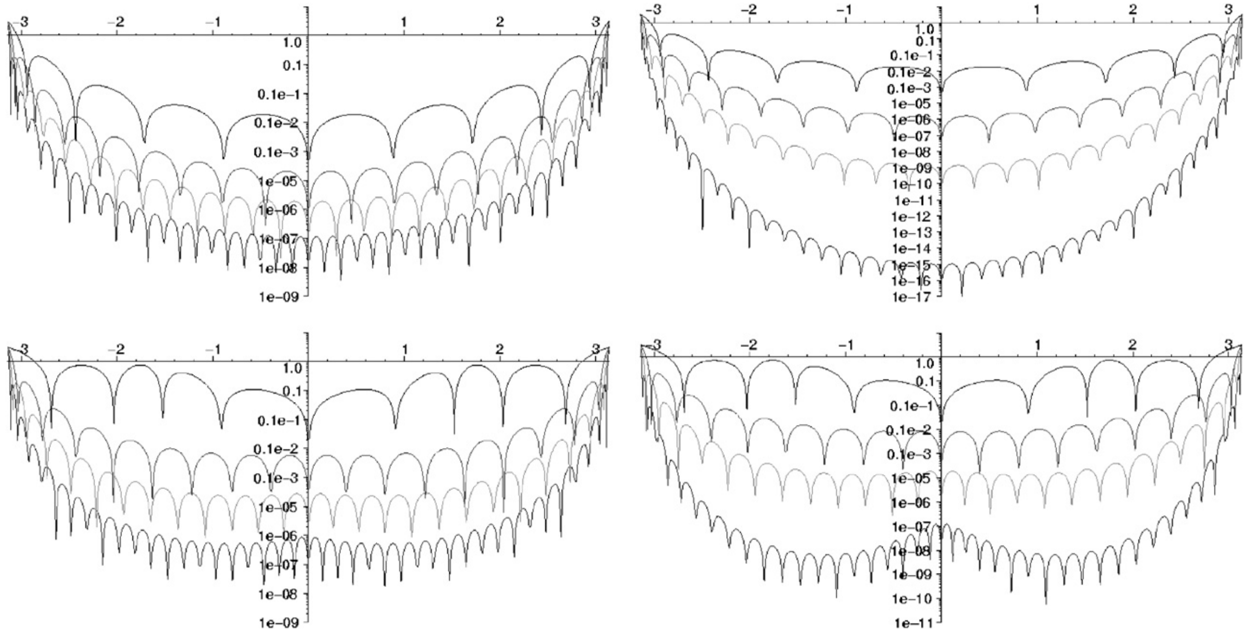

Fig. 4. The modulus of the error of approximation on a logarithmic scale of the Fourier series of $f_1(t) = s(t-\pi)$ (on the top) and of $f(t) = f_1(t) + (1-\cos(t))^3 s(t)$ (on the bottom). On the left we display the error for $\varepsilon_4^{(2\ell)}$ and on the right for $\varepsilon_{4+2\ell}^{(0)}$, $\ell = 0, 2, 4, 8$ from the top to the bottom.

results (in what follows we will no longer display them). According to Table 1(c), the error obtained by the real part of $\varepsilon_{2k}^{(n)}(t)$ (here for $k=3$ and 8) is dominated by $4/3\pi$ times the error on the unit circle of $[k-1|k]_{G^{(-1/2,1)}}$, the latter being estimated in Corollary 4.3. Again, even for arguments close to the singularity $t=0$ we have a quite impressive convergence improvement. We should mention that all numerical experiments have been performed using Maple with sufficiently high precision such that rounding errors can be neglected.

In our last example in Fig. 4 we consider the functions

$$f_1(t) = s(t-\pi) \notin \mathscr{C}^0, \quad f_2(t) = (1-\cos(t))^3 s(t) \in \mathscr{C}^5 \backslash \mathscr{C}^6$$

and hence $G(f_1)(z) = -2izG^{(0,0)}(-z)$. We compare the improvements obtained for columns of the Padé table (here $k=2$ and $n=2\ell$ for $\ell \in \{0,2,4,8\}$) and diagonals (here $n=0$ and $k=2+\ell$ for $\ell \in \{0,2,4,8\}$). Notice that the number of terms of the Fourier series required for $\varepsilon_4^{(2\ell)}$ and for $\varepsilon_{4+2\ell}^{(0)}$ is the same (namely $5+2\ell$). We observe that in both cases there is improvement of convergence for increasing ℓ; however, the rate is much more interesting for our diagonal sequence $\varepsilon_{2+2\ell}^{(0)}$, in particular for $f(t) = f_1(t)$.

The error for $\varepsilon_2^{(2\ell)}$ on the bottom of Fig. 4 (that is, $f(t) = f_1(t) + f_2(t)$) and sufficiently large ℓ has been discussed (implicitly) in Theorem 5.1: since $G(f_1)(z) = -2izG^{(0,0)}(-z)$, we replace $z = e^{it}$ by $-z = e^{i(\pi-t)}$, and set $k=2$, $\alpha=\beta=0$. Also, $f_2 \in \mathscr{C}^5$, and $f_2^{(6)} \in L_1$ (with one jump), such that $m=6$, showing that the asymptotic rate $\mathcal{O}(n^{-5(1-\tau)})$ of Theorem 5.1 is valid for all $0 \leqslant \tau < 1$. In order to be more precise for finite n, we have to compute explicitly the quantity N_2 (or even N_k for diagonal sequences) in the hypothesis of Lemma 5.2. Observe that, by (4.8) and (4.11),

$$\sum_{j=0}^{2} v_{j,n}^2 = (n+1) + (n+3)(2n+3)^2 + (n+5)(n+2)^2(2n+5)^2 \leqslant 5(n+5)(n+2)^4,$$

and from Table 1(e)$_1$ for $n \geqslant 3$

$$\|f_2 - S_{n-1}(f_2)\|_{L_2([-\pi,\pi])}^2 \leqslant 4\left(\frac{2^{-3}(1)_6}{(2)_6}\right)^2 \sum_{j=n-3}^{\infty}\left(\frac{(1)_j}{(8)_j}\right)^2 \leqslant \left(\frac{1}{28}\right)^2 \sum_{j=n-2}^{\infty}\left(\frac{(1)_7}{(j)_7}\right)^2 \leqslant \frac{180^2}{13}(n-3)^{-13}.$$

Thus, with the very rough choice $N_2 \geqslant 83$, the hypothesis of Lemma 5.2 is true, and a combination of Lemmas 5.3 and 5.4 and (4.10) enables us to establish more explicit bounds for $n \geqslant N_2$.

Finally, we should comment on the peak of the error on the lower right plot of Fig. 4 around $t = 0$: indeed, the influence of f_2 on $[n + k|k]_{G(f_1+f_2)}$ is negligible for fixed k and $n \to \infty$, but this is no longer true for fixed n and $k \to \infty$: here the zeros of the Padé denominator also detect the singularities of $G(f_2)$.

7. Concluding remarks and open questions

In the present paper we have established a link between the complex ε-algorithm applied to partial Fourier sums, and the nonlinear Padé–Chebyshev and Padé–Fourier approximants. We were able to show by deriving explicit error estimates for a class of hypergeometric functions that the complex ε-algorithm allows one to accelerate convergence of partial Fourier sums. In particular, as observed numerically by Brezinski [5], this technique allows one to smooth the Gibbs phenomenon for functions which either themselves or their higher order derivatives have a jump. Finally, we have shown that the rate of convergence for columns is preserved even after smooth perturbations of the underlying function.

There are several remarks or open questions about this field of research:

- We have seen in our numerical experiments that, for ray sequences of the form $k = [\lambda n]$ with $\lambda > 0$, we get a rate of convergence better than that of columns (k fixed). Here one could derive a result similar to Theorem 5.1 by using the strong asymptotics of Jacobi polynomials with varying parameters as derived in [12,22]. However, then for the hypothesis (5.1) we would require very smooth f_2 with exponentially decaying $\|E_n(f_2)\|_2$, obtained, for instance, for rational $G(f_2)$. It would also be interesting to combine our findings with those of Rakhmanov [26] who discusses the error of diagonal Padé approximants where $G(f_1)$ is a Stieltjes function and $G(f_2)$ is rational.
- A nice test function s_m not included in our class of hypergeometric functions would be the mth primitive of the saw tooth function s of (2.1), with jth derivative being continuous for all $j \neq m$, and having one jump for $j = m$. Notice that, by (2.2),
$$G(s_m)(z) = 2i^{4m+1} \sum_{j=1}^{\infty} \frac{z^j}{j^{m+1}} = \frac{2i^{4m+1}}{m!} z \int_0^1 \frac{(\log(1/y))^m}{1 - yz}\, \mathrm{d}y,$$
that is, we essentially get a Stieltjes function. Therefore, it would be interesting to extend Corollary 4.3 and Theorem 5.1 to general Stieltjes functions.
- It would be nice to understand the convergence behavior for functions f having several jumps, like $t \mapsto f_0(t) = s(t - t_0) - s(t + t_0)$, having two jumps at $\pm t_0$, and reducing to a multiple of $t \mapsto \mathrm{sign}(\cos(t))$ for $t_0 = \pi/2$. One may derive an explicit formula for the diagonal Padé denominators of $G(f_0)$, showing that the poles stay outside the unit disk, but are no longer on the real axis but now on a circle orthogonal to the unit circle and intersecting the unit circle at $\mathrm{e}^{\pm i t_0}$.
- Sometimes the data available in spectral methods for PDEs are partial sums of Legendre series. We suspect that by exploiting the link with Baker–Gammel approximants [2, Section 7.2] we should get similar convergence results.
- From the numerical experiments, it seems very plausible that the Gibbs overshoot at the discontinuity is reduced when applying the complex ε-algorithm. It would be interesting to theoretically answer this conjecture.

Acknowledgments

The authors would like to thank the anonymous referees whose remarks have helped them in improving the presentation of the paper.

References

[1] G.A. Baker, Essentials of Padé Approximants, Academic Press, London, 1975.
[2] G.A. Baker, P. Graves-Morris, Padé Approximants, second ed., Cambridge University Press, Cambridge, 1996.
[3] Ch. Bernardi, Y. Maday, Spectral methods, Handbook of Numerical Analysis, vol. V, North-Holland, Amsterdam, 1997.
[4] C. Brezinski, Padé-Type Approximation and General Orthogonal Polynomials, Birkhauser, Basel, 1980.

[5] C. Brezinski, Extrapolation algorithms for filtering series of functions, and treating the Gibbs phenomenon, Numer. Algorithms 36 (2004) 309–329.
[6] T. Chihara, An introduction to orthogonal polynomials, Mathematics and its Applications, vol. 13, Gordon and Breach Science Publishers New York-London-Paris, 1978.
[7] T. Driscoll, B. Fornberg, A Padé-based algorithm for overcoming the Gibbs phenomenon, Numer. Algorithms 26 (2001) 77–92.
[8] K.S. Eckhoff, Accurate and efficient reconstruction of discontinuous functions from truncated series expansions, Math. Comput. 61 (1993) 745–763.
[9] K.S. Eckhoff, Accurate reconstructions of functions of finite regularity from truncated Fourier series expansions, Math. Comput. 64 (1995) 671–690.
[10] K.S. Eckhoff, On a high order numerical method for functions with singularities, Math. Comput. 67 (1998) 1063–1087.
[11] J. Fleischer, Nonlinear Padé approximants for Legendre series, J. Math. Phys. 14 (1973) 246–248.
[12] W. Gawronski, B. Shawyer, Strong asymptotics and the limit distribution of the zeros of Jacobi polynomials $P_n^{(an+\alpha,bn+\beta)}$, in: P. Nevai, A. Pinkus (Eds.), Progress in Approximation Theory, Academic Press, New York, 1991, pp. 379–404.
[13] A. Gelb, E. Tadmor, Detection of edges in spectral data, Appl. Comput. Harmon. Anal. 7 (1999) 101–135.
[14] A. Gelb, E. Tadmor, Detection of edges in spectral data II. Nonlinear enhancement, SIAM J. Numer. Anal. 38 (2000) 1389–1408.
[15] A.A. Gonchar, E.A. Rakhmanov, S.P. Suetin, On the rate of convergence of Padé approximants of orthogonal expansions, in: A.A. Gonchar, E.B. Saff (Eds.), Progress in Approximation Theory, Springer, Berlin, 1992, pp. 169–190.
[16] D. Gottlieb, S. Orszag, Numerical analysis of spectral methods: theory and applications, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 26, SIAM, Philadelphia, 1977.
[17] D. Gottlieb, C.-W. Shu, On the Gibbs phenomenon and its resolution, SIAM Rev. 39 (1997) 644–668.
[18] S.M. Kaber, Y. Maday, Analysis of some Padé–Chebyshev approximants, SIAM J. Numer. Anal. 43 (2005) 437–454.
[19] G. Kvernadze, Determination of the jumps of a bounded function by its Fourier series, J. Approx. Theory 92 (1998) 167–190.
[20] G. Kvernadze, Approximating the jump discontinuities of a function by its Fourier–Jacobi coefficients, Math. Comput. 73 (2004) 731–751.
[21] G. Kvernadze, T. Hagstrom, H. Shapiro, Locating discontinuities of a bounded function by the partial sums of its Fourier series, J. Sci. Comput. 4 (1999) 301–327.
[22] A. Martinez-Finkelshtein, R. Orive, Riemann–Hilbert analysis for Jacobi polynomials orthogonal on a single contour, J. Approx. Theory 134 (2005) 137–170.
[23] G. Németh, G. Páris, The Gibbs phenomenon in generalized Padé approximation, J. Math. Phys. 26 (1985) 1175–1178.
[24] E.M. Nikishin, V.N. Sorokin, Rational Approximations and Orthogonality, Translations of Mathematical Monographs, vol. 92, American Mathematical Society, Providence, RI, 1991.
[25] S. Paszkowski, Approximation uniforme des fonctions continues par des fonctions rationnelles, Zastosow. Math. 6 (1963) 441–458.
[26] E.A. Rakhmanov, Convergence of diagonal Padé approximants, Math. Sbornik 33 (1977) 243–260.
[27] W. Rudin, Real and Complex Analysis, McGraw-Hill, Singapore, 1987.
[28] G. Szegő, Orthogonal Polynomials, vol. XXIII, American Mathematical Society, Colloquium Publications, 1975.
[29] P. Wynn, Transformations to accelerate the convergence of Fourier series, in: Gertrude Blanch Anniversary Volume, Wright Patterson Air Force Base, Aerospace Research Laboratories, Office of Aerospace Research, United States Air Force, 1967, pp. 339–379.