# Padé-Jacobi Approximants

Jan Hesthaven, Sidi Mahmoud Kaber

# Padé-Jacobi Approximants

J.S. Hesthaven[†] and S.M. Kaber[‡]

December 14, 2016

**Abstract**

We discuss projection based Padé-Jacobi approximants in general and present in particular an exact rational approximation to the Sign function. This serves as vehicle to analyze the behavior of Padé-Jacobi approximants for discontinuous functions. The analysis shows that the Padé-Jacobi approximant is superior in several ways to classic polynomial approximations of discontinuous functions, provided the parameters in the approximations are chosen carefully. Guidelines for this is obtained through the analysis.

## 1 Introduction

The nonuniform pointwise convergence, known as the Gibbs phenomenon, of polynomial approximations to discontinuous function is a well known and much studied phenomenon, see e.g. [10] and references therein. Among the consequences of the Gibbs phenomenon is the lack of convergence at the jump with an overshoot of approximately 9% of the jump size, a global $\mathcal{O}(N^{-1})$ convergence rate in mean, and a steepness of the approximation right at the jump being proportional to the length, $N$, of the polynomial expansion.

The literature is rich with methods trying to reduce or even eliminate these problems. The perhaps simplest approach is that of modal filtering, essentially relying on forcing the expansion to converge more rapidly [20, 10, 13]. An alternative approach is physical space filtering using mollifiers [11, 19], yielding similar behavior. Both methods, however, do not overcome the lack of convergence at the point of discontinuity. To achieve this, information about the shock location is needed. With this, the Gibbs phenomenon can be completely resolved [10], albeit this approach has considerable practical problems.

In this work we shall discuss the use of rational functions, Padé-Jacobi approximants, for the representation of discontinuous functions. As rational functions are richer than simple polynomial expansions, one can hope that the impact of the discontinuity will be less severe and, further, that one could use this as a postprocessing tool to reduce the impact of the Gibbs phenomena in polynomial expansions.

To study the fundamental behavior of Padé-Jacobi approximants of discontinuous functions, we present a family of exact rational approximations to the Sign function, considered as a prototype of discontinuous

functions. This enables a complete analysis of the behavior of this approximation as characterized by the maximum size of the overshoot and the achievable steepness at the point of discontinuity. As we shall show, the use of a rational approximation allows one to dramatically reduce the overshoot and increase the steepness while recovering high order accuracy away from the jump.

There has been some recent activity in the exploration of Padë-forms for the reconstruction of Gibbs oscillations. In particular, work for the Fourier case can be found in [8, 6, 4], for the Chebyshev in [16], and for the Legendre approximations in [5, 12]. However, much of this has been of a qualitative character and for special polynomial families only.

In Sec. 2, we recall some properties of the Jacobi polynomials and the Padé-Jacobi problem. Section 3 is devoted to the derivation of an exact solution of the Padé-Jacobi approximation problem for the Sign function. In Sec. 4, we consider the optimization of the Padé-Jacobi solution by varying several of the free parameters. Section 5 contains a few remarks.

# 2 Jacobi Polynomials and Padé-Jacobi Approximations

In the following we shall recall various definitions and properties of Jacobi polynomials and expansions, as well as define exactly what we mean by Padé-Jacobi approximations in this work.

## 2.1 Jacobi Polynomials and Expansions

For $\alpha > -1$, the symmetric Jacobi polynomials, $P_n^{(\alpha)}(x)$, also known as the ultraspherical polynomials, are defined as the polynomial eigenfunctions to the singular Strum-Liouville problem

$$\mathcal{A}_\alpha P_n^{(\alpha)}(x) = \lambda_n^\alpha P_n^{(\alpha)}(x) \quad , \quad x \in [-1,1] \quad , \tag{1}$$

where

$$\mathcal{A}_\alpha \varphi = -\frac{1}{\omega_\alpha}\left(\omega_{\alpha+1}\varphi'\right)' ,$$

with the weight function

$$\omega_\alpha = (1-x^2)^\alpha \quad ,$$

and the eigenvalue

$$\lambda_n^\alpha = n(n+2\alpha+1) \quad .$$

One easily proves that the Jacobi polynomials are the unique polynomial solution [18] to Eq.(1), once a normalization is chosen. The standard choice, also used here, is

$$P_n^{(\alpha)}(1) = \frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)},$$

where $\Gamma(x), x \geq 0$ represents the classic Euler Gamma function. Recall that $\Gamma(n+1) = n\Gamma(n) = n!$.

We introduce the Pochhammer symbol

$$(z)_n = \frac{\Gamma(z+n)}{\Gamma(z)} = (z+1)(z+2)\dots(z+n-1) \ ,$$

and note that $(1)_n = n!$ and $(z)_1 = z$. Recall also that for $k \in \mathsf{N}$

$$\forall n \geq k+1 \ : \ (-k)_n = 0 \ . \tag{2}$$

An important property of the Pochhammer symbol is expressed in the Saalchütz's formula [7]

$${}_3F_2(-n, a, b; d, 1+a+b-d-n; 1) = \frac{(d-a)_n (d-b)_n}{(d)_n (d-a-b)_n}, \tag{3}$$

with ${}_3F_2(-n, a, b; d, 1+a+b-d-n; 1)$ being the hypergeometric function defined as

$${}_3F_2(a, b, c; d, e; z) := \sum_{k=0}^{\infty} \frac{(a)_k (b)_k (c)_k}{(d)_k (e)_k} \frac{1}{k!} z^k \ .$$

Well known examples of ultraspherical polynomials are the Chebyshev polynomials ($\alpha = -1/2$) and the Legendre polynomials ($\alpha = 0$). The ultraspherical polynomials have a number of important properties which we shall exploit. In particular, all the polynomials are mutually orthogonal in the inner product

$$(u, v)_\alpha = \int_{-1}^{1} u(x) v(x) \omega_\alpha \, dx, \tag{4}$$

with the associated weighted $L^2_\alpha$ norm

$$\|u\|^2_\alpha = (u, u)_\alpha \ .$$

The normalization is given by

$$\gamma^\alpha_n = \left( P^{(\alpha)}_n, P^{(\alpha)}_n \right)_\alpha = \frac{2^{2\alpha+1}}{2n+2\alpha+1} \frac{\Gamma(n+\alpha+1)^2}{\Gamma(n+1)\Gamma(n+2\alpha+1)} \ . \tag{5}$$

Another important property of the ultraspherical polynomials is their even-odd characteristics

$$P^{(\alpha)}_n(x) = (-1)^n P^{(\alpha)}_n(-x) \ . \tag{6}$$

We also recall the special value at $x = 0$ as

$$P^{(\alpha)}_{2n}(0) = (-1)^n 2^{-2n} \begin{pmatrix} 2n+\alpha \\ n \end{pmatrix} \tag{7}$$

and zero otherwise due to Eq.(6).

Finally we shall need the relation [18]

$$\frac{d}{dx} P^{(\alpha)}_n(x) = \frac{1}{2} (n+2\alpha+1) P^{(\alpha+1)}_{n-1}(x) \ . \tag{8}$$

If we now consider functions, $u(x) \in L^2_\alpha$, i.e., for which $\|u\|_\alpha < \infty$, we can seek polynomial approximations as

$$u(x) = \sum_{n=0}^{\infty} \hat{u}_n P_n^{(\alpha)}(x) \quad , \quad \hat{u}_n = \frac{1}{\gamma_n^\alpha} \left( u, P_n^{(\alpha)} \right)_\alpha \quad ,$$

by orthogonality.

Let us consider the truncated expansion

$$u_N^{(\alpha)}(x) = \sum_{n=0}^{N} \hat{u}_n P_n^{(\alpha)}(x), \tag{9}$$

i.e., $u_N^{(\alpha)} \in \mathsf{P}_N$ where $\mathsf{P}_N$ is the space of algebraic polynomials of degree less than or equal to $N$. The orthogonality of the Jacobi polynomials implies

$$\forall p \in \mathsf{P}_N \quad : \quad \left( u - u_N^{(\alpha)}, p \right)_\alpha = 0.$$

It is well known that the polynomial expansion is convergent in the mean but not uniformly. In particular, if the smoothness is measured in the Sobolev space $H_\alpha^p$ of functions $u$ and their derivatives up to order $p$ in $L_\alpha^2$, there exists a constant $c_p$ such that

$$\| u - u_N^{(\alpha)} \|_\alpha \leq c_p N^{-p} \| u \|_{H_\alpha^p}.$$

For a smooth function, i.e., $p$ large, this provides an accurate approximation and the approximation error decreases rapidly to zero as $N$ goes to infinity. This is one of the main motivations for using spectral methods for solving partial differential equations with regular solutions. We refer the reader to [1].

However, for problems with discontinuous solutions, the expansion exhibits non-uniform convergence and a phenomenon known as the Gibbs phenomenon [14] as illustrated in Fig. 1 where the truncated ($N = 20$ and $N = 100$) Legendre expansion, i.e., for $\alpha = 0$, of the Sign function is displayed. One observes the oscillations, especially near the discontinuity. If the parameter $N$ is increasing, the size of the oscillations decrease everywhere except near the discontinuity where $\mathcal{O}(1)$ oscillations (overshoot/undershoot) remain. Furthermore, the global nature of the oscillations limits the pointwise accuracy to first order away from $x = 0$.

One of the objectives of this work is to consider Padé-Jacobi approximations of the Sign function and attempt to answer the question of which one among this family would be best suited to approximate the Sign function. As measures of success we shall consider

- the overshoot/undershoot of the approximation at the point of discontinuity as characterized by the Gibbs constant of the expansion.
- the ability to reproduce the discontinuity characterized by the steepness of the approximation.

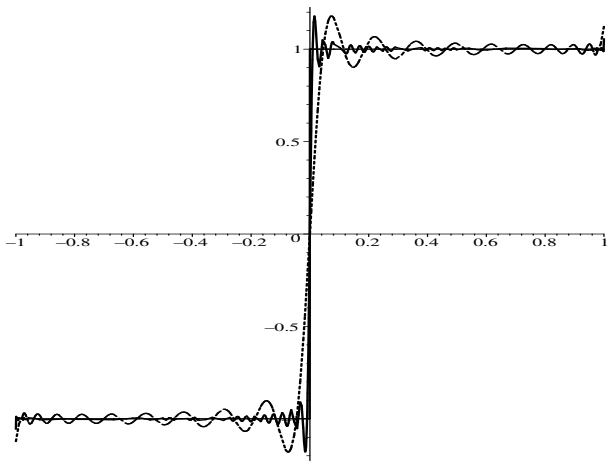


Figure 1: Legendre approximations of the Sign function: $N = 20$ (dashed) and $N = 100$ (solid).

## 2.2 Padé-Jacobi Approximants

We shall consider the Padé-Jacobi approximation to $u$ in a Galerkin sense, i.e., find $\mathcal{P} \in \mathsf{P}_N$ and $\mathcal{Q} \in \mathsf{P}_M$ such that

$$(\mathcal{Q}u - \mathcal{P}, p)_\alpha = 0, \qquad \forall p \in \mathsf{P}_K \ , \tag{10}$$

with $K \leq M + N$. This shall be used to define the linear Galerkin type Padé-Jacobi approximation of order $(N, M)$ to $u$ as the rational function

$$\mathcal{R}_{N,M}(x) = \frac{\mathcal{P}(x)}{\mathcal{Q}(x)} \ ,$$

where $(\mathcal{P}, \mathcal{Q})$ satisfies Eq.(10).

**Remark 2.1** *For $M = 0$ and $K = N$, the pair $(\mathcal{Q} \equiv 1, \mathcal{P} = u_N^{(\alpha)})$ defined in (9) is a solution of the Padé-Jacobi approximation problem.*

It is important for practical purposes that the complete knowledge of $u$ is not needed to solve the problem (10), only $u_{N+M}^{(\alpha)}$ is required. In Eq. (10) we take $p = P_k^{(\alpha)}$ with $k = N + 1, \ldots K$, to get

$$\widehat{(\mathcal{Q}u)}_k^{(\alpha)} = 0, \qquad \forall k = N + 1, \cdots, K.$$

This is a linear system of $K - N$ equations and $M + 1$ unknowns (the coefficients of $\mathcal{Q}$ in a basis of $\mathsf{P}_M$). Once a non trivial solution of this system is found (such a solution always exists if $K \leq N + M$), the numerator $\mathcal{P} \in \mathsf{P}_N$ is simply computed by

$$\widehat{\mathcal{P}}_k^{(\alpha)} = \widehat{(\mathcal{Q}u)}_k^{(\alpha)}, \qquad \forall k = 0, \cdots, N.$$

Hence the main problem is the determination of the denominator.

# 3 Analysis of the Sign-Function

Let us now return to a more thorough analysis of the behavior of the Padé-Jacobi approximation for the most basic discontinuous function – the Sign function

$$u(x) = \begin{cases} -1 & x < 0 \\ 1 & x > 0 \end{cases} \quad . \tag{11}$$

## 3.1 An Exact Solution

For the purpose of analysis, we shall seek the approximation to the Sign function using

$$\mathcal{P}(x) = \sum_{n=0}^{N} \hat{p}^{\alpha}_{2n+1} P^{(\alpha)}_{2n+1}(x) \in \mathsf{P}_{2N+1} \quad ,$$

and

$$\mathcal{Q}(x) = \sum_{m=0}^{M} \hat{q}^{\alpha}_{2m} x^{2m} \in \mathsf{P}_{2M} \quad ,$$

where we have used the parity of the problem and Eq.(6) to reduce the complexity of the problem.

We must now seek $\mathcal{P} \in \mathsf{P}_{2N+1}$ and $\mathcal{Q} \in \mathsf{P}_{2M}$ to satisfy Eq.(10), i.e.,

$$\left(\mathcal{Q}u - \mathcal{P}, P^{(\alpha)}_{2k+1}\right)_{\alpha} = 0 \quad , \quad \forall k \leq K \quad ,$$

where we have again utilized the parity of the problem to reduce the complexity.

Orthogonality of the Jacobi basis immediately yields

$$\left(\mathcal{P}u, P^{(\alpha)}_{2k+1}\right)_{\alpha} = \left(\mathcal{Q}u, P^{(\alpha)}_{2k+1}\right)_{\alpha} = \gamma^{\alpha}_{2k+1} \widehat{(\mathcal{Q}u)}^{\alpha}_{2k+1} = 0 \quad N < k \leq K \quad ,$$

which is a linear system of $(K - N) \times (M + 1)$, with the unknowns being the coefficients of the denominator $\mathcal{Q}$, i.e., $\hat{q}^{\alpha}_{2m}$. Clearly for $K \leq N + M$ this linear system will always have at least one nontrivial solution. In what remains we shall restrict ourselves to the special case

$$K = N + M \quad .$$

To compute the numerator, we observe that

$$\hat{p}^{\alpha}_{2k+1} - \widehat{(\mathcal{Q}u)}^{\alpha}_{2k+1} = 0 \quad , \quad 0 \leq k \leq N, \tag{12}$$

i.e., once $\mathcal{Q}$, and hence, $\widehat{\mathcal{Q}u}^{\alpha}_{2k+1}$ is computed, the numerator follows immediately.

Let us thus focus on the computation of the denominator, or rather its coefficients, satisfying

$$\widehat{(\mathcal{Q}u)}^{\alpha}_{2k+1} = \sum_{m=0}^{M} \hat{q}^{\alpha}_{2m} \frac{1}{\gamma^{\alpha}_{2k+1}} \left(x^{2m} u, P^{(\alpha)}_{2k+1}\right)_{\alpha} = 0 \quad .$$

We shall need the following results. Define for the integers $k$ and $m$

$$\mathcal{I}_{m,k}^{\alpha} = \int_0^1 x^{2m} P_{2k+1}^{(\alpha)} \omega_\alpha \, dx \quad \text{and} \quad \mathcal{J}_{m,k}^{\alpha} = \int_0^1 x^{2m+1} P_{2k+1}^{(\alpha)} \omega_\alpha \, dx, \quad (13)$$

**Lemma 3.1** *For $m \geq 1$ and $k \geq 0$*

$$\mathcal{I}_{m,k}^{\alpha} = \mathcal{I}_{0,k}^{\alpha} \frac{m!(1/2)_m}{(-k+1/2)_m (k+\alpha+2)_m}$$

*with*

$$\mathcal{I}_{0,k}^{\alpha} = \frac{k+\alpha+1}{\lambda_{2k+1}^{\alpha}} P_{2k}^{(\alpha+1)}(0) = \frac{(-1)^k}{2k+1} \frac{1}{2^{2k+1}} \frac{1}{k!} \frac{\Gamma(2k+\alpha+2)}{\Gamma(k+\alpha+2)}.$$

*For $m < k$, $\mathcal{J}_{m,k}^{\alpha} = 0$ and*

$$\forall j \in \mathsf{N} \quad : \quad \mathcal{J}_{k+j,k}^{\alpha} = \frac{(k+1)_j}{j!} \frac{(k+3/2)_j}{(2k+\alpha+5/2)_j} \mathcal{J}_{k,k}^{\alpha} \quad (14)$$

*with*

$$\mathcal{J}_{k,k}^{\alpha} = \frac{1}{2} \frac{\gamma_{2k+1}^{(\alpha)}}{\theta_{2k+1}^{(\alpha)}}$$

*and $\theta_j^{\alpha}$ the coefficient of $x^j$ in $P_j^{(\alpha)}(x)$.*

Proof: Using Eq.(1) we have

$$\lambda_{2k+1}^{\alpha} \mathcal{I}_{m,k}^{\alpha} = \int_0^1 x^{2m} \mathcal{A}_\alpha P_{2k+1}^{(\alpha)} \omega_\alpha \, dx = -\int_0^1 x^{2m} \left( \omega_{\alpha+1} \left( P_{2k+1}^{(\alpha)} \right)' \right)' \, dx.$$

Recalling the singular nature of $\omega_\alpha$, integration by parts twice yields

$$\begin{aligned} \lambda_{2k+1}^{\alpha} \mathcal{I}_{m,k}^{\alpha} &= 2m \int_0^1 x^{2m-1} \omega_{\alpha+1} \left( P_{2k+1}^{(\alpha)} \right)' (x) \, dx \\ &= -2m \int_0^1 \left( \omega_{\alpha+1} x^{2m-1} \right)' P_{2k+1}^{(\alpha)} \, dx \\ &= -2m(2m-1) \mathcal{I}_{m-1,k}^{\alpha} + 2m(2m+2\alpha+1) \mathcal{I}_{m,k}^{\alpha}. \end{aligned}$$

From this, we recover the recurrence

$$\mathcal{I}_{m,k}^{\alpha} = \frac{-2m(2m-1)}{\lambda_{2k+1}^{\alpha} - 2m(2m+2\alpha+1)} \mathcal{I}_{m-1,k}^{\alpha} = \frac{-m(2m-1)}{(2k-2m+1)(k+m+\alpha+1)} \mathcal{I}_{m-1,k}^{\alpha}.$$

We finally note that

$$\begin{aligned} \lambda_{2k+1}^{\alpha} \mathcal{I}_{0,k}^{\alpha} &= -\int_0^1 \left( \omega_{\alpha+1} \left( P_{2k+1}^{(\alpha)} \right)' \right)' \, dx = -\left[ \omega_{\alpha+1} \left( P_{2k+1}^{(\alpha)} \right)' \right]_0^1 \\ &= \left( P_{2k+1}^{(\alpha)} \right)' (0) = (k+\alpha+1) \, P_{2k}^{\alpha+1}(0). \end{aligned}$$

Combining Eq.(7) with the above result yields

$$\begin{aligned} \mathcal{I}_{m,k}^{\alpha} &= \frac{m(m-1/2)}{(-k+m-1/2)(k+m+\alpha+1)} \mathcal{I}_{m-1,k}^{\alpha} \\ &= \mathcal{I}_{0,k}^{\alpha} \frac{m!(1/2)_m}{(-k+1/2)_m (k+\alpha+2)_m}. \end{aligned}$$

Concerning $\mathcal{J}^{\alpha}_{m,k}$, we observe that $P^{(\alpha)}_j(x) = \theta^{(\alpha)}_j x^j + q^{(\alpha)}_j$ with $q^{(\alpha)}_j \in \mathsf{P}_{j-1}$ :

$$\mathcal{J}^{\alpha}_{k,k} = \frac{1}{2}\int_{-1}^{1} \frac{P^{(\alpha)}_{2k+1} - q^{(\alpha)}_{2k+1}}{\theta^{(\alpha)}_{2k+1}} P^{(\alpha)}_{2k+1}\omega_\alpha \, dx = \frac{1}{2}\frac{\gamma^{(\alpha)}_{2k+1}}{\theta^{(\alpha)}_{2k+1}} \quad , \quad \theta^{\alpha}_j = \frac{1}{2^j}\frac{1}{j!}\frac{\Gamma(2j+2\alpha+1)}{\Gamma(j+2\alpha+1)} \quad .$$

The proof of (14) follows the same lines. □ Thus, to find a Padé-Jacobi approximant to the Sign function, we must seek a solution to

$$\sum_{m=0}^{M} \hat{q}^{\alpha}_{2m} \frac{m!(1/2)_m}{(-k+1/2)_m(k+\alpha+2)_m} = 0. \tag{15}$$

One non-unique solution is given in the following

**Proposition 3.2** *The coefficients, $\hat{q}^{\alpha}_{2m}$, defined for $m \in [0, M]$ as*

$$\hat{q}^{\alpha}_{2m} = \frac{(-M)_m(A)_m(-A+M+\alpha+3/2)_m}{(m!)^2(1/2)_m},$$

*is a solution to Eq.(15) with*

$$A = -(N+1/2) \quad .$$

Proof: Inserting the above result into Eq.(15) yields

$$\sum_{m=0}^{M} \frac{(-M)_m(A)_m(-A+M+\alpha+3/2)_m}{m!(-k+1/2)_m(k+\alpha+2)_m} \quad .$$

Using Eq.(3) this can be written as

$$\frac{(k+\alpha+2-A)_M(k-M-1/2+A)_M}{(k+\alpha+2)_M(k-M-1/2)_M} \quad .$$

Recalling Eq.(2) we immediately get two solutions to Eq.(15) from each of the two terms in the numerator

$$A = M+N+\alpha+2 \quad , \quad A = -N-1/2 \quad .$$

In both cases, we have $(k-M-N)_M$ which vanishes for all $k \in (N+1, M+N)$. □ Thus, the denominator takes the form

$$\begin{aligned} \mathcal{Q}(x) &= \sum_{m=0}^{M} \frac{(-M)_m(-N-1/2)_m(N+M+\alpha+2)_m}{(m!)^2(1/2)_m} x^{2m} \\ &= {}_3F_2(-M, -N-1/2, N+M+\alpha+2; 1, 1/2; x^2) \quad . \end{aligned} \tag{16}$$

Before we continue with the development of the Padé-Jacobi approximation, let us consider a few properties of $Q(x)$.

**Lemma 3.3** *Provided $N + 3/2 > M$ we have*

$$\forall m \in [0, M] \ : \hat{q}_{2m} > 0,$$

*and, hence,*

$$Q(x) \geq Q(0) = \hat{q}_0 = 1 \quad .$$

Proof: Consider

$$\hat{q}_{2m} = \frac{(-M)_m(-N-1/2)_m(N+M+\alpha+2)_m}{(m!)^2(1/2)_m} \quad .$$

Clearly,

$$(-M)_m = -M(-M+1)(-M+2)\dots(-M+m-1) = (-1)^m \frac{M!}{(M-m)!} \quad ,$$

and

$$\begin{aligned}(-N-1/2)_m &= (-N-1/2)(-N+1/2)(-N+3/2)\dots(-N-3/2+m) \\ &= (-1)^m \frac{\Gamma(N+3/2)}{\Gamma(N+3/2-m)} \quad .\end{aligned}$$

As $m \in [0, M]$, $(-M)_m(-N-1/2)_m > 0$ provided only that $N+3/2-M > 0$, hence completing the proof. □ Thus all

roots of $Q(x)$ are complex, ensuring that the Padé-Jacobi approximation to the Sign-function always exists.

The location of the roots can be specified a bit more

**Lemma 3.4** *Assume that $N \gg M$, $\alpha$ fixed, and $z \in \mathsf{C}$ be a root of $Q(x)$. Then*

$$\frac{1}{2M} \leq N^2|z|^2 \leq M^3.$$

For a proof of this Lemma, see Proposition 4.7 of [16]. As we shall see shortly, this results also gives some indications of how well one can expect the to approximate the Sign function since there is a direct relation between the position of the poles and the ability of the approximation to reproduce the discontinuity.

The sharpness in $N$ can be realized by considering the limit of large $N$ in which case

$$(-N-1/2)_m(N+M+\alpha+2)_m \simeq (-1)^m N^{2m} \quad ,$$

such that

$$\mathcal{Q}(x) = {}_3F_2(-M, -N-1/2, N+M+\alpha+2; 1, 1/2; x^2) \simeq {}_1F_2(-M; 1, 1/2; -N^2x^2) \quad .$$

However, since the ${}_1F_2(a1; b1, b2; z)$ is independent of $N$, the roots of $\mathcal{Q}(x)$ can not decay faster than $N^{-1}$. It is worth emphasizing that this result assumes that $M$ is fixed, i.e., making $M \propto N$ and/or $\alpha \propto N$ may yield qualitative differences in the approximation as we shall indeed see shortly.

Let us now return to the determination of the numerator,

$$P(x) = \sum_{n=0}^{N} \hat{p}_{2n+1} P_{2n+1}^{(\alpha)}(x). \tag{17}$$

The coefficients of this polynomial are given in the following

**Lemma 3.5** *The coefficients $\hat{p}^{\alpha}_{2n+1}$ in (17) are defined for $n \in [0, N]$ as*

$$\hat{p}^{\alpha}_{2n+1} = 2\frac{\mathcal{I}^{\alpha}_{0,n}}{\gamma^{\alpha}_{2n+1}} \frac{(N+n+\alpha+5/2)_M (n-N-M)_M}{(n+\alpha+2)_M (n+1/2-M)_M} \quad ,$$

*where $\mathcal{I}^{\alpha}_{0,n}$ is given in Lemma 3.1.*

Proof: From the definition of the Padé-Jacobi approximation and the orthogonality of the Jacobi polynomials, we immediately recover from (12)

$$\begin{aligned}
\gamma^{\alpha}_{2n+1}\hat{p}^{\alpha}_{2n+1} &= 2\sum_{m=0}^{M} \hat{q}_{2m}\mathcal{I}^{\alpha}_{m} \\
&= 2\mathcal{I}^{\alpha}_{0,n} \sum_{m=0}^{M} \frac{(-M)_m(-N-1/2)_m(N+M+\alpha+2)_m}{m!(1/2-n)_m(n+\alpha+2)_m} \\
&= 2\mathcal{I}^{\alpha}_{0,n} \frac{(N+n+\alpha+5/2)_M (n-N-M)_M}{(n+\alpha+2)_M (n+1/2-M)_M} \quad ,
\end{aligned}$$

where the last reduction follows from the Saalchütz's formula (3). □

Using the identity $(n-N-M)_M = (-1)^M (N-n+M)!/(N-n)!$, one can express the coefficients $\hat{p}^{\alpha}_{2n+1}$ in the form

$$\hat{p}^{\alpha}_{2n+1} = 2(-1)^M \frac{\mathcal{I}^{\alpha}_{0,n}}{\gamma^{\alpha}_{2n+1}} \frac{(N+n+\alpha+5/2)_M (N-n+M)!}{(n+\alpha+2)_M (n+1/2-M)_M (N-n)!}. \quad (18)$$

In (16), the denominator was written in a geometric form. Now we seek the numerator in the form : $\mathcal{P}(x) = xS^{\alpha}_{N,M}\ {}_3F_2(.,.,.;.,.;x^2)$, with $S^{\alpha}_{N,M}$ being the steepness, i.e., the value at $x = 0$ of the derivative of the rational approximation $\mathcal{R}(x) = \mathcal{P}(x)/\mathcal{Q}(x)$. The following Proposition allows us to recover the numerator on a hypergeometric form.

**Proposition 3.6** *The numerator, $\mathcal{P}(x)$, for the Padé-Jacobi approximation to the Sign function, Eq.(11), takes the form*

$$\mathcal{P}(x) = xS^{\alpha}_{N,M}\,{}_3F_2(-N, -M+1/2, N+M+\alpha+5/2; 3/2, 3/2; x^2).$$

*where the steepness, $S^{\alpha}_{N,M}$ is given as*

$$S^{\alpha}_{N,M} = \frac{4}{\sqrt{\pi}} \frac{M!}{N!} \frac{\Gamma(N+3/2)\Gamma(N+M+\alpha+5/2)}{\Gamma(M+1/2)\Gamma(N+M+\alpha+2)} \quad .$$

Proof: Let us compute the Jacobi coefficients of $\mathcal{T} \in \mathsf{P}_{2N+1}$ defined by

$$\mathcal{T}(x) = xS^{\alpha}_{N,M}\,{}_3F_2(-N, -M+1/2, N+M+\alpha+5/2; 3/2, 3/2; x^2).$$

For $n = 0 \cdots, N$

$$\begin{aligned}
\frac{\gamma^{\alpha}_{2n+1}}{S^{\alpha}_{N,M}}\hat{t}^{\alpha}_{2n+1} &= \int_{-1}^{1} x\,{}_3F_2(-N, -M+1/2, N+M+\alpha+5/2; 3/2, 3/2; x^2) P^{(\alpha)}_{2n+1}(x)\omega_{\alpha}(x)dx \\
&= 2\sum_{k=0}^{N} \frac{(-N)_k(-M+1/2)_k(N+M+\alpha+5/2)_k}{(3/2)_k(3/2)_k} \frac{1}{k!}\mathcal{J}_{k,n}
\end{aligned}$$

with $\mathcal{J}^{\alpha}_{k,n}$ defined in (13). Using Lemma 3.1, we get

$$\frac{\gamma^{\alpha}_{2n+1}}{S^{\alpha}_{N,M}}\hat{t}^{\alpha}_{2n+1} = 2\sum_{p=0}^{N-n}\frac{(-N)_{n+p}(-M+1/2)_{n+p}(N+M+\alpha+5/2)_{n+p}}{(3/2)_{n+p}(3/2)_{n+p}(1)_{n+p}}\mathcal{J}_{n+p,n}.$$

By the identity $(z)_{n+p} = (z)_n\,(z+n)_p$, we get

$$\hat{t}^{\alpha}_{2n+1} = 2X^{\alpha}_{N,M,n}\sum_{p=0}^{N-n}\frac{(-N+n)_p(-M+1/2+n)_p(N+M+\alpha+5/2+n)_p}{(3/2+n)_p(3/2+n)_p(1+n)_p}\mathcal{J}_{n+p,n}$$

with

$$X^{\alpha}_{N,M,n} = \frac{S^{\alpha}_{N,M}}{\gamma^{\alpha}_{2n+1}}\frac{(-N)_n(-M+1/2)_n(N+M+\alpha+5/2)_n}{(3/2)_n(3/2)_n(1)_n}.$$

Using (14), we get

$$\hat{t}^{\alpha}_{2n+1} = Y^{\alpha}_{N,M,n}\sum_{p=0}^{N-n}\frac{(-N+n)_p(-M+1/2+n)_p(N+M+\alpha+5/2+n)_p}{(3/2+n)_p(2n+\alpha+5/2)_p}\frac{1}{p!}$$

with $Y^{\alpha}_{N,M,n} = 2X^{\alpha}_{N,M,n}\mathcal{J}^{\alpha}_{n,n}$. By use of the Saalchütz's formula (3), we obtain

$$\begin{aligned}\hat{t}^{\alpha}_{2n+1} &= Y^{\alpha}_{N,M,n}\frac{(M+1)_{N-n}(-N-M-\alpha-1)_{N-n}}{(3/2+n)_{N-n}(-N-n-\alpha-3/2)_{N-n}}\\ &= Y^{\alpha}_{N,M,n}\frac{\Gamma(n+3/2)\Gamma(2n+\alpha+5/2)}{\Gamma(N+3/2)\Gamma(N+n+\alpha+5/2)}\frac{\Gamma(N+M-n+1)\Gamma(N+M+\alpha+2)}{\Gamma(M+1)\Gamma(n+M+\alpha+2)}.\end{aligned}$$

with

$$\begin{aligned}Y^{\alpha}_{N,M,n} &= \frac{\pi}{4\theta^{\alpha}_{2n+1}}\frac{(-N)_n(-M+1/2)_n(N+M+\alpha+5/2)_n}{n!\Gamma(3/2+n)^2}S^{\alpha}_{N,M}\\ &= \sqrt{\pi}\frac{(-1)^n}{\theta^{\alpha}_{2n+1}}\frac{\Gamma(N+3/2)}{n!\Gamma(n+3/2)^2\,(N-n)!}\\ &= \times\frac{\Gamma(M+1)\Gamma(-M+1/2+n)}{\Gamma(M+1/2)\Gamma(-M+1/2)}\frac{\Gamma(N+M+n+\alpha+5/2)}{\Gamma(N+M+\alpha+2)}\\ &= \frac{1}{\sqrt{\pi}}\frac{(-1)^n}{\theta^{\alpha}_{2n+1}}\frac{\Gamma(N+3/2)}{n!\Gamma(n+3/2)^2\,(N-n)!}\\ &\quad\times(-1)^M\frac{\Gamma(M+1)\Gamma(-M+1/2+n)}{\Gamma(N+M+\alpha+2)}\Gamma(N+M+n+\alpha+5/2).\end{aligned}$$

The last simplification follows from the reflection formula the reflection formula

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}\quad.$$

Finally

$$\begin{aligned}\hat{t}^{\alpha}_{2n+1} &= \frac{1}{\sqrt{\pi}}\frac{(-1)^n}{\theta^{\alpha}_{2n+1}}\frac{\Gamma(2n+\alpha+5/2)}{n!\Gamma(n+3/2)(N-n)!\Gamma(N+n+\alpha+5/2)}\\ &\quad\times\frac{(-1)^M\Gamma(-M+1/2+n)\Gamma(N+M-n+1)\Gamma(N+M+n+\alpha+5/2)}{\Gamma(n+M+\alpha+2)}\end{aligned}$$

This is to be compared with the coefficients $\hat{p}^{\alpha}_{2n+1}$ given by Lemma 3.5 (see also (18)):

$$\begin{aligned}\hat{p}^{\alpha}_{2n+1} &= 2\frac{\mathcal{I}^{\alpha}_{0,n}}{\gamma^{\alpha}_{2n+1}}\frac{\Gamma(n+\alpha+2)}{(N-n)!\Gamma(n+1/2)\Gamma(N+n+\alpha+5/2)} \\ &\quad\times(-1)^M\frac{\Gamma(N+M+n+\alpha+5/2)(N+M-n)!\Gamma(n+1/2-M)}{\Gamma(n+M+\alpha+2)}\end{aligned}$$

The ratio of the two coefficients is

$$\begin{aligned}\frac{\hat{t}^{\alpha}_{2n+1}}{\hat{p}^{\alpha}_{2n+1}} &= \frac{1}{2\sqrt{\pi}}\frac{(-1)^n\gamma^{\alpha}_{2n+1}}{\mathcal{I}^{\alpha}_{0,n}\theta^{\alpha}_{2n+1}}\frac{\Gamma(2n+\alpha+5/2)}{n!\Gamma(n+3/2)}\frac{\Gamma(n+1/2)}{\Gamma(n+\alpha+2)} \\ &= \frac{1}{2\sqrt{\pi}}\frac{(-1)^n\gamma^{\alpha}_{2n+1}}{\mathcal{I}^{\alpha}_{0,n}\theta^{\alpha}_{2n+1}}\frac{\Gamma(2n+\alpha+5/2)}{n!(n+1/2)\Gamma(n+\alpha+2)}.\end{aligned}$$

Straightforward computations give

$$\frac{\gamma^{\alpha}_{2n+1}}{\mathcal{I}^{\alpha}_{0,n}\theta^{\alpha}_{2n+1}} = (-1)^n 2^{4n+2\alpha+3} n!(2n+1)\Gamma(n+\alpha+2)\frac{\Gamma(2n+\alpha+2)}{\Gamma(4n+2\alpha+4)}$$

By the Legendre duplication formula[1], we have

$$\frac{\Gamma(2n+\alpha+2)}{\Gamma(4n+2\alpha+4)} = \frac{\sqrt{\pi}}{2^{4n+2\alpha+3}}\frac{1}{\Gamma(2n+\alpha+5/2)}$$

and

$$\frac{\gamma^{\alpha}_{2n+1}}{\mathcal{I}^{\alpha}_{0,n}\theta^{\alpha}_{2n+1}} = (-1)^n\sqrt{\pi}n!(2n+1)\frac{\Gamma(n+\alpha+2)}{\Gamma(2n+\alpha+5/2)}.$$

Hence $\hat{t}^{\alpha}_{2n+1} = \hat{p}^{\alpha}_{2n+1}$, which means the equality of the two odd polynomials, $\mathcal{T}$ and $\mathcal{P}$. □ Lemma 3.6 and Eq. (16) gives the main result in the form of a hypergeometric representation of the Padé-Jacobi approximation of Eq.(11)

**Theorem 3.7** *For all integers $N$ and $M$,*

$$\mathcal{R}^{\alpha}_{N,M}(x) = S^{\alpha}_{N,M}\,x\,\frac{{}_3F_2(-N,-M+1/2,N+M+\alpha+5/2;3/2,3/2;x^2)}{{}_3F_2(-M,-N-1/2,N+M+\alpha+2;1,1/2;x^2)},$$

*with the steepness $S^{\alpha}_{N,M}$ defined in Proposition 3.6, is a Padé-Jacobi approximation of order $(N,M)$ to the Sign function, Eq.(11).*

Note in particular that for the special case of $M = 0$, this includes the Jacobi polynomial approximation of a step function and, thus, enables the general analysis of the Gibbs phenomenon for this case also.

Let us first consider two extremal cases: $M = 0$ (polynomial approximation) and $N = 0$ (reciprocal polynomial approximation).

1. Polynomial (Jacobi) approximation.

$$\mathcal{R}^{\alpha}_{N,0} = S^{\alpha}_{N,0}x\,{}_3F_2(-N,1/2,N+\alpha+5/2;3/2,3/2;x^2)$$

$\mathcal{R}^{\alpha}_{N,0}$ is nothing but the orthogonal projection (with respect to the inner product (4)) of the Sign function onto $\mathsf{P}_{2N+1}$ defined in (9).

[1]Legendre duplication formula: $\sqrt{\pi}\Gamma(2z) = 2^{2z-1}\Gamma(z)\Gamma(z+1/2)$

It has been shown in [15] that for all Jacobi approximants $S^{\alpha}_{N,0} \simeq 4N/\pi$. We give here the precise value of the steepness

$$S^{\alpha}_{N,0} = \frac{4}{\pi}\frac{\Gamma(N+3/2)}{N!}\frac{\Gamma(N+\alpha+5/2)}{\Gamma(N+\alpha+2)} \simeq \frac{4}{\pi}N.$$

2. Reciprocal polynomial approximation.

$$\mathcal{R}^{\alpha}_{0,M}(x) = S^{\alpha}_{0,M}x\,\frac{1}{{}_3F_2(-M,-1/2,M+\alpha+2;1,1/2;x^2)}$$

with the steepness

$$S^{\alpha}_{0,M} = 2\frac{M!}{\Gamma(M+1/2)}\frac{\Gamma(M+\alpha+5/2)}{\Gamma(M+\alpha+2} \simeq 2M.$$

# 4 Optimized Approximations

Let us consider in a bit more detail the question of the Gibbs phenomena and attempt to understand whether it is better behaved in the Padé-Jacobi approximation as compared to the classical polynomial approximations above.

We shall in particular consider the questions of steepness of the Padé-Jacobi approximation and the size of the overshoot as measured by the Gibbs constant, $G^{\alpha}_{N,M}$.

We consider five cases

1. **Case 1.** The classic polynomial case, $M = 0$, with the Gibbs constant denoted $G^{\alpha}_{\cdot,0}$.
2. **Case 2.** The reciprocal polynomial case, $N = 0$, with the Gibbs constant denoted $G^{\alpha}_{0,\cdot}$.
3. **Case 3.** The case of $M$ going to infinity with $N$, however constrained such that $M = cN^s$ and fixed $c > 0$ and $s > 0$. In this case we shall denote the Gibbs constant as $G^{\alpha}_{\cdot,c,s}$.
4. **Case 4.** The case of $M$ fixed ($\neq 0$) with the Gibbs constant denoted $G^{\alpha}_{\cdot,M}$.
5. **Case 5.** The case of $N$ fixed ($\neq 0$) with the Gibbs constant denoted $G^{\alpha}_{N,\cdot}$.

## 4.1 Optimize the steepness

If we first consider the steepness of the Padé-Jacobi approximation, then this is defined as

$$\frac{d}{dx}\mathcal{R}_{N,M}(0) = \mathcal{P}'_N(0) = S^{\alpha}_{N,M} \ ,$$

due to the symmetry of the problem, i.e., the steepness measures the ability to reproduce the discontinuity.

**Cases 1, 4.** From Lemma 3.6 we immediately get for a fixed $M$

$$\mathcal{P}'_N(0) = S^{\alpha}_{N,M} \simeq \frac{4}{\sqrt{\pi}}\frac{\Gamma(M+1)}{\Gamma(M+1/2)}\sqrt{N(N+M+\alpha)} \simeq \frac{4}{\sqrt{\pi}}\frac{\Gamma(M+1)}{\Gamma(M+1/2)}N \ ,$$

for large $N$ and fixed $M$ and $\alpha$. Hence, in this case there is no qualitative difference between the pure polynomial case ($M = 0$) and a fixed value of $M$. All polynomials behave, asymptotically, as the Legendre case of $\alpha = 0$.

**Cases 2, 5.** For a fixed $N$ (and $M \to +\infty$), the steepness grows like

$$\frac{4}{\sqrt{\pi}} \frac{\Gamma(N + 3/2)}{N!} M.$$

**Case 3.** If $M \propto N$, we get a more interesting result. For $M = cN$, with positive constant $c$

$$S^{\alpha}_{N,cN} \simeq \frac{4}{\sqrt{\pi}} \sqrt{c(c+1)} N^{3/2}$$

which improve drastically the steepness in comparison with the polynomial case. It is important to note that $M = N$ is not needed to recover this improved steepness, simply that $M \propto N$. In the case $M = cN^s$, the steepness is

$$S^{\alpha}_{N,cN^s} \simeq \frac{4}{\sqrt{\pi}} c_s N^{s+1/2},$$

with $c_s = c$ except for $s = 1$: $c_1 = \sqrt{c(c+1)}$.

We remind the reader of Lemma 3.4 which reflects these steepness results in a sligthly different way, i.e., for $s = 1$ we can not hope for better than $S^{\alpha}_{N,cN} \simeq N^{-3/2}$ as reflected in how quickly the poles approach zero.

## 4.2 Optimize the Gibbs constant

Let us now also consider the Gibbs constant, defined as the maximum overshoot of the Padé-Jacobi approximation. We proceed as in [17] for the Chebyshev approximation and seek an $\eta > 0$ such that the error function $x \in [0, 1] \mapsto \mathcal{R}_{N,M}(x) - 1$ takes its maximum at the point $x = \eta/N^{\beta}$ as $N$ goes to infinity ($\beta$ is a fixed real number to be made precise shortly). We shall call this limit the Gibbs constant, $G^{\alpha}_{N,M}$.

**Case 1.** ($M = 0$). We seek $\eta > 0$ such that the error function takes its maximum at the point $x = \eta/N$ as $N$ goes to infinity. In this case, we have

$$\begin{aligned} 1 + G^{\alpha}_{\cdot,0} &= \lim_{N \to +\infty} \mathcal{R}_{N,0}(\frac{\eta}{N}) \\ &= \frac{4}{\pi} \eta \lim_{N \to +\infty} {}_3F_2(-N, 1/2, N + \alpha + 5/2; 3/2, 3/2; (\frac{\eta}{N})^2) \\ &= \frac{4}{\pi} \eta \, {}_1F_2(1/2; 3/2, 3/2; -\eta^2). \end{aligned}$$

$\eta$ is determined as the smallest positive solution of

$$\frac{d}{d\eta} \left[ \eta \sum_{k \geq 0} \frac{(1/2)_k}{(3/2)_k (3/2)_k} \frac{1}{k!} (-\eta^2)^k \right] = 0.$$

Using the identities $(1/2)_k = \frac{1}{2^{2k}}\frac{(2k)!}{k!}$ and $(3/2)_k = (2k+1)(1/2)_k$, we obtain

$$\begin{aligned} 0 &= \frac{d}{d\eta}\left[\eta\sum_{k\geq 0}\frac{2^{2k}}{(2k+1)^2}\frac{(-\eta^2)^k}{(2k)!}\right] = \frac{1}{2}\frac{d}{d\eta}\left[\sum_{k\geq 0}\frac{(-1)^k}{2k+1}\frac{(2\eta)^{2k+1}}{(2k+1)!}\right] \\ &= \frac{1}{2\eta}\sum_{k\geq 0}(-1)^k\frac{(2\eta)^{2k+1}}{(2k+1)!} = \frac{\sin(2\eta)}{2\eta}. \end{aligned}$$

Hence $\eta = \pi/2$ and the Gibbs constant is

$$G^{\alpha}_{\cdot,0} = -1 + \frac{4}{\pi}\int_0^{\pi/2}\frac{\sin(2t)}{2t}dt = -1 + \frac{2}{\pi}Si(\pi) \simeq 0.178\,979\,744$$

with $\mathrm{Si}(z) = \int_0^z \frac{\sin s}{s}ds$ being the Sine integral. This is a classic result in Fourier and Chebyshev approximations (see e.g. [9]). It has also been shown for general Jacobi approximations in [15] using properties of orthogonal polynomials.

**Case 2.** (Case $N = 0$). The right scaling is $x = \eta/M$ as $M$ goes to infinity. Using the same arguments as before, we obtain

$$1 + G^{\alpha}_{0,\cdot} = \lim_{M\to+\infty}\mathcal{R}_{M,0}(\frac{\eta}{M}) = 2\eta\frac{1}{{}_1F_2(-1/2;1,1/2;-\eta^2)}.$$

$\eta$ is determined as the smallest positive solution of the equation

$$\frac{d}{d\eta}\left[\frac{1}{f(\eta)}\right] = 0, \qquad f(\eta) = \frac{{}_1F_2(-1/2;1,1/2;-\eta^2)}{\eta}.$$

We observe that $f'(\eta) = J_0(2\eta)/\eta^2$ with $J_0(t)$ being the Bessel function of the first kind of order 0. Thus, $2\eta$ equals $j_{0,1}$, the first zero of $J_0$, as

$$2\eta = j_{0,1} \simeq 2.404\,825\,557\,8,$$

and the Gibbs constant

$$G^{\alpha}_{0,\cdot} = -1 + j_{0,1}\frac{1}{{}_1F_2(-1/2;1,1/2;-(j_{0,1}/2)^2)} \simeq 0.051\,356\,067.$$

Here again all the Jacobi approximants give the same Gibbs constant, the one given in [17] for the Chebyshev case, $\alpha = -1/2$. An example of reciprocal polynomial approximation in Fig.2.

**Case 3.** ($M = cN^s$). Consider now the case $M = cN^s$ with fixed $c > 0$ and $s > 0$. In this case the steepness grows like $\frac{4}{\sqrt{\pi}}c_s N^{s+1/2}$ and we seek $\eta > 0$ that maximizes the error function $x = \eta/N^{s+1/2}$ as $N$ goes to infinity. Defining $G^{\alpha}_{N,\,c,s} = G^{\alpha}_{N,\,cN^s}$ we have

$$\begin{aligned} 1 + G^{\alpha}_{\cdot,\,c,s} &= \frac{4}{\sqrt{\pi}}c_s\eta \\ &\quad \lim_{N\to+\infty}\frac{{}_3F_2(-N,-cN+1/2,(1+c)N+\alpha+5/2;3/2,3/2;(\frac{\eta}{N^{s+1/2}})^2)}{{}_3F_2(-cM,-N-1/2,(1+c)N+\alpha+2;1,1/2;(\frac{\eta}{N^{s+1/2}})^2)} \\ &= f^*(c_s\eta), \end{aligned}$$

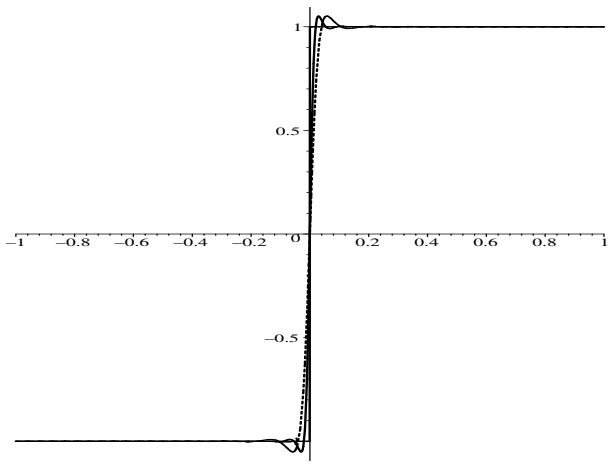


Figure 2: Padé-Jacobi approximations: $N = 0$, $\alpha = -1/2$, $M = 20$ (dashed), $M = 40$ (solid).

with $f^*(z) = \dfrac{4}{\sqrt{\pi}} z \, \dfrac{{}_0F_2(;3/2,3/2;z^2)}{{}_0F_2(;1,1/2;z^2)}$.

The unknown $\eta_{c,s}$ is determined as the smallest positive solution of the equation $\frac{d}{d\eta}\left[f^*(c_s\eta)\right] = 0$. This equation shows that $\eta_{c,s}$ does not depend on the Jacobi parameter $\alpha$ and

$$\eta_{c,s} = \frac{\eta^*}{c_s},$$

with $\eta^*$ the first positive zero of $f^*$. The Gibbs constant $G^{\alpha}_{\cdot,\,c,s} = -1 + f^*(\eta^*)$ is independent of $\alpha$, $c > 0$ and $s$.

Numerical experiments, finding the location of the first maximum of the analytic expression, yields the approximations

$$\eta_{\tilde{c}} \simeq 1.344\,947, \quad G^{\alpha}_{\cdot,\,c,s} \simeq 0.008\,149 \ .$$

This is the value given in [17] for the special case $\alpha = 0$, $c = 1$ and $s = 1$. Examples of this type of approximation in shown in Fig.3 and Fig.4

**Case 4**. $(M \neq 0)$. We seek $\eta > 0$ such that the error function takes its maximum at the point $x = \eta/N$ as $N$ goes to infinity. In this case, we have

$$\begin{aligned} 1 + G^{\alpha}_{\cdot,M} &= \frac{4}{\sqrt{\pi}} \frac{\Gamma(M+1)}{\Gamma(M+1/2)} \, \eta \\ &\quad \times \lim_{N\to+\infty} \frac{{}_3F_2(-N,-M+1/2,N+M+\alpha+5/2;3/2,3/2;(\frac{\eta}{N})^2)}{{}_3F_2(-M,-N-1/2,N+M+\alpha+2;1,1/2;(\frac{\eta}{N})^2)} \\ &= \frac{4}{\sqrt{\pi}} \frac{\Gamma(M+1)}{\Gamma(M+1/2)} \, \eta \frac{{}_1F_2(-M+1/2;3/2,3/2;-\eta^2)}{{}_1F_2(-M;1,1/2;-\eta^2)}. \end{aligned}$$

$\eta_M$ (independent of $\alpha$) is determined as the smallest positive solution of the equation $\frac{d}{dz}\left[f_M(z)\right] = 0$ with

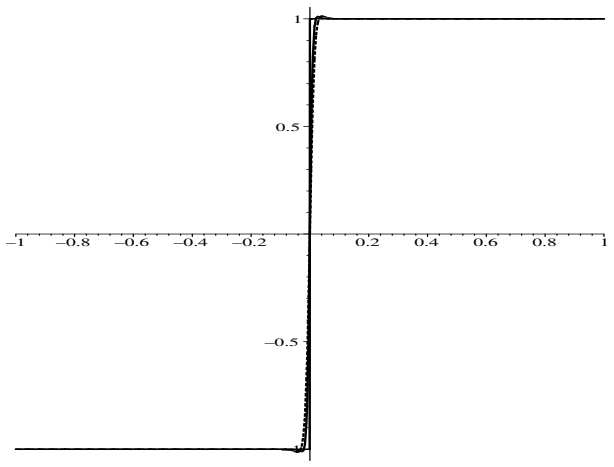


Figure 3: Padé-Jacobi approximations: $N = 10$, $\alpha = 0$, $M = 5$ (dashed), $M = 10$ (solid).

| $M$ | 0 | 1 | 5 | 10 | 20 | 50 | 100 | 200 |
|---|---|---|---|---|---|---|---|---|
| $\eta_M$ | 1.5708 | 1.0144 | 0.56 | 0.4098 | 0.2951 | 0.1888 | 0.134 | 0.095 |
| $G^{\alpha}_{\cdot,M}$ | 0.1789 | 0.0302 | 0.012 | 0.01 | 0.009 | 0.00851 | 0.0083 | 0.00823 |

Table 1: Computational evidence for the scaling of the Gibbs constant for values of fixed $M$

.

$$f_M(z) = \frac{4}{\sqrt{\pi}} \frac{\Gamma(M+1)}{\Gamma(M+1/2)}\, z\, \frac{{}_1F_2(-M+1/2; 3/2, 3/2; -z^2)}{{}_1F_2(-M; 1, 1/2; -z^2)}.$$

The Gibbs constant $G^{\alpha}_{\cdot,M} = -1 + f_M(\eta_M)$ is likewise independent of $\alpha$.

We have not been able to complete the analysis of this function. In Table 1 we show results for numerically finding, by seeking the position of the first maximum of the analytic expression, $\eta_M$ and $G^{\alpha}_{\cdot,M}$ for fixed values of $M$.

Based on these computations we conjecture that

$$\lim_{M\to+\infty} \sqrt{M}\, \eta_M = \eta^* = 1.344947$$

and

$$\lim_{M\to+\infty} G^{\alpha}_{\cdot,M} = -1 + \lim_{M\to+\infty} f_M(\eta_M) = -1 + f^*(\eta^*) = G^{\alpha}_{\cdot,M} \simeq 0.008\,149$$

The effect of changing $N$ for fixed $M$ can be seen by comparing Fig.3 and Fig.4.

**Case 5.** ($N \neq 0$). In this case the steepness grows like $\frac{4}{\sqrt{\pi}} \frac{\Gamma(N+3/2)}{N!} M$ and

$$1 + G^{\alpha}_{N,\cdot} = \frac{4}{\sqrt{\pi}} \frac{\Gamma(N+3/2)}{N!}\, \eta \frac{{}_1F_2(-N; 3/2, 3/2; -\eta^2)}{{}_1F_2(-N-1/2; 1, 1/2; -\eta^2)}.$$

$\eta_N$ (independent of $\alpha$) is determined as the smallest positive solution of the equation $\frac{d}{dz}\left[g_N(z)\right] = 0$ with

$$g_N(z) = \frac{4}{\sqrt{\pi}} \frac{\Gamma(N+3/2)}{N!}\, z\, \frac{{}_1F_2(-N;3/2,3/2;-z^2)}{{}_1F_2(-N-1/2;1,1/2;-z^2)}.$$

The Gibbs constant $G^{\alpha}_{N,.} = -1 + g_N(\eta_N)$ is also independent of $\alpha$ and

$$\lim_{N\to+\infty} g_N(\eta/\sqrt{N}) = f^*(\eta^*) \ ,$$

based on computational experimentation as for Case 4.

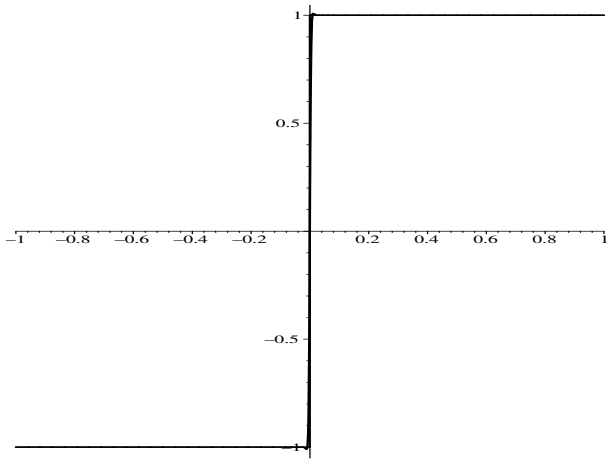


Figure 4: Padé-Jacobi approximations: $\alpha = 0$, $N = 30$, $M = 10$.

# 5 Concluding Remarks

We have derived an exact rational Galerkin approximation $\mathcal{R}^{\alpha}_{N,M}$ of the Sign function based on Jacobi expansions and investigated its ability to reduce the Gibbs phenomenon. The analysis contains the cases for $M$ or $N$ fixed, including the classic polynomial results, the case where both go to infinity with $M = cN^s$. The latter case is superior in terms of the Gibbs constant and steepness.

- The steepness of the approximation grows, as $N \to +\infty$, like $\frac{4}{\sqrt{\pi}} c_s N^{s+1/2}$ (in the case $c > 0$ and $s > 0$). Recall that in the polynomial case, the steepness is $\simeq \frac{4}{\pi} N$.
- The Gibbs constant is about 22 times less than that of a polynomial approximation.

One case we have not considered in detail is the one where $\alpha$ is a function of $N$ and/or $M$. Although the analysis indicates that this could be interesting, it is less practical and unlikely to behave well numerically for high

values of $\alpha$ as also observed for Gibbs reconstruction methods based on Gegenbauer expansions [10, 2]. Nevertheless, we intend to consider this in more detail later. Furthermore, the rate of convergence of $\mathcal{R}^{\alpha}_{N,M}$ to the Sign function (the acceleration of convergence problem) remains open, yet is important to understand the value of Padé-Jacobi approximations for postprocessing. Some partial results on this can be found in [16] for the Padé-Chebyshev case and we hope to generalize these in the near future.

# Acknowledgment

The work of JSH was partly supported by NSF Career Award DMS0132967, by an NSF International Award NSF-INT 0307475, and by the Alfred P. Sloan Foundation through a Sloan Research Fellowship.

# References

[1] C. Bernardi, Y. Maday, *Spectral methods, Handbook of numerical analysis*, Vol. V, North-Holland , Amsterdam, 1997.

[2] J.P. Boyd, *Trouble with Gegenbauer reconstruction for defeating Gibbs phenomenon: Runge phenomenon in the diagonal limit of Gegenbauer polynomial approximations*, J. Comput. Phys. **204**(2005), pp. 253-264

[3] C.W. Clenshaw and K. Lord, *Rational approximations from Chebyshev series*, in *Studies in Numerical Analysis*, B.K.P. Scaife ed., Academic Press, London, 1974, pp. 95-113.

[4] , W.S. Don, S.M. Kaber, M.S. Sun, *Fourier-Padé Approximations and Filtering for the Spectral Simulations of Incompressible Boussinesq Convection Problem*, accepted for publication in Math of comp.

[5] L. Emmel, S.M. Kaber, Y. Maday, *Padé-Jacobi filtering for spectral approximations of discontinuous solutions*. Numer. Algo. **33**(2003),pp. 251-264.

[6] T. Driscoll and B. Fornberg, *A Padé-based algoritheorem for overcoming the Gibbs phenomenon*, Numer. Algo. **26**(2001), pp. 77-92.

[7] A. Erdélyi, W. Magnus, F. Oberhettinger, and F.G. Tricomi, *Higher Transcendental Functions I+II+III*, Robert E. Krieger Publishing, Malabar, FL, 1981.

[8] J.F. Geer, *Rational Trigonometric Approximations using Fourier Series Partial Sums*, J. Sci. Comput. **10**(1995), pp. 325-356.

[9] D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*. CBMS-NSF **26**. SIAM, Philadelphia, 1978.

[10] D. Gottlieb and C.W. Shu, *On the Gibbs Phenomenon and its Resolution*, SIAM Review **39**(1997), pp. 644-668.

[11] D. Gottlieb and E. Tadmor, *Recovering Pointwise Values of Discontinuous Data with Spectral Accuracy*. In Progress and Supercomputing in Computational Fluid Dynamics. Birkhäuser, Boston, 1984. pp. 357-375.

[12] , J. S. Hesthaven, S.M. Kaber, and L. Lurati, *The Padé-Legendre Interpolation Method*, In preparation.

[13] J.S. Hesthaven and M. Kirby, *Filtering in Legendre Spectral Methods*, Math. Comp. 2005 – submitted.

[14] A. J. Jerri, *The Gibbs phenomenon in Fourier analysis, splines and wavelet approximations*, Kluwer Academic Publishers, Dordrecht, 1998.

[15] S.M. Kaber, *The Gibbs phenomenon for Jacobi expansions*, Technical rep. Université Pierre & Marie Curie, 2004. Available at `http://www.ann.jussieu.fr/publications/2005/R05003.html`.

[16] S.M. Kaber and Y. Maday, *Analysis of some Padé-Chebyshev approximants*. Accepted for publication in SIAM Journal on Numerical Analysis, 2004.

[17] G. Németh, and G. Paris, *The Gibbs phenomenon in generalized Padé approximation*. J. Math. Phys. 26, 1985.

[18] G. Szegö, *Orthogonal Polynomials*, American Mathematical Society, 1930.

[19] J. Tanner and E. Tadmor, *Adaptive Mollifiers - High Resolution Recover of Piecewise Smooth Data from its Spectral Information*, Found. Comput. Math. **2**(2002), pp. 155-189.

[20] H. Vandeven, *Family of Spectral Filters for Discontinuous Problems*, J. Sci. Comput. **8**(1991), pp. 159-192.