Adaptive point shifts in rational approximation
with optimized denominator \(^\dagger\)

Jean-Paul Berrut
Département de Mathématiques
Université de Fribourg, CH–1700 Fribourg/Pérolles, Switzerland

and

Hans D. Mittelmann
Department of Mathematics
Arizona State University, Tempe, Arizona 85287–1804, USA

Abstract: Classical rational interpolation is known to suffer from several drawbacks, such as unattainable points and randomly located poles for small numbers of nodes, as well as an erratic behavior of the error as this number grows larger. In a former article, we have suggested to obtain rational interpolants by a procedure that attaches optimally placed poles to the interpolating polynomial, using the barycentric representation of the interpolants. In order to improve upon the condition of the derivatives in the solution of differential equations, we have then experimented with a conformal point shift suggested by Kosloff and Tal-Ezer. As it turned out, such shifts can achieve a spectacular improvement in the quality of the approximation itself for functions with a large gradient in the center of the interval. This leads us to the present work which combines the pole attachment method with shifts optimally adjusted to the interpolated function. Such shifts are also constructed for functions with several shocks away from the extremities of the interval.

Key words: rational approximation, interpolation, optimal interpolation, point shifts

AMS 2000 Classification: 65D05, 41A20, 41A05.

1. Introduction

Our aim in the present work is the rational approximation of a function \(f\) on a real interval \(I\), which we may take as \([-1,1]\) without loss of generality, in such a way that the approximant interpolates \(f\) between \(N + 1\) abscissae \(x_0, \ldots, x_N\) in \(I\) which are either given in advance or images of given points under a conformal map. The reader may consult [Ber-Mit2] for an example of an application of such interpolants. We assume that \(f\) is analytic (holomorphic) within a domain \(D_1\) containing \(I\).

\(^\dagger\) This work has been supported by the Swiss National Science Foundation, grants # 21–59’272.99 and 20–6675.4.01
Let $p_N$ be the polynomial of degree at most $N$ interpolating $f$ between $x_0, \ldots, x_N$. It is well-known that, for good interpolation points (nodes) such as Chebyshev’s or Legendre’s, one has exponential convergence of the sequence $p_N$ toward $f$:

$$|p_N - f| \leq c \cdot e^{-\alpha N}$$

for some constants $c$ and $\alpha$. (This same kind of interpolation is also of interest if $f$ is merely in $C^{p-1}(I)$ for large $p$ with $f^{(p)}$ of bounded variation, in which case one reaches an order of convergence $p$:

$$|p_N - f| \leq c \cdot N^{-p}.$$  

However, the constants $c$ and $\alpha$ in the above estimate depend, explicitly or implicitly, on derivatives of $f$; when the latter are very large, the fast convergence may show up only after too large values of $N$ for practical purposes.

Rational interpolation often does much better for small values of $N$. However, classical rational interpolants $r$, in which the sum of the degrees of numerator and denominator add up to $N$, are unpredictable for small $N$. In many cases, which document more the rule than the exception, unwanted poles occur in the vicinity of $I$, or even on $I$, making the corresponding $r$ useless as approximants.

To obtain better interpolants $r$ in cases $f$ is known everywhere on $I$, we have proposed in [Ber-Mit1] to attach poles to $p_N$ in such a way as to minimize some functional, e.g., some norm $\|r - f\|$, taken as a measure of the approximation error. The decrease of the functional induced by the optimization documents the motion of the poles from their position at infinity (for $p_N$) to an optimal finite position. In sharp contrast to classical interpolants, the resulting rational interpolants $r$ are at least as good as $p_N$ and neither unwanted poles nor unattainable points [Sto] may occur.

Quite impressive results have been obtained in [Ber-Mit1] with this method, and this even with equidistant $x_k$, for which $p_N$ is notoriously useless for many $f$ when $N$ grows large (Runge’s phenomenon). For these points the computations had however to be performed in quadruple precision (about 32 digits), a manifestation of the ill-conditioning. And the higher $N$, the higher the necessary precision.

If the goal is just to approximate $f$, and if one can choose the $x_k$ – a situation we will assume here – it is clearly better to stay with Chebyshev or Legendre points. On the other hand, it is well-known by experts in spectral methods that derivatives of polynomials interpolating between such points which accumulate in the vicinity of the boundary are ill-conditioned for large $N$. Kosloff and Tal-Ezer [Kos-Tal] have therefore suggested to conformally shift these points from their, say, Chebyshev position by a conformal map $g$.
from a $y$-domain containing another copy $J$ of $I$ (with the Chebyshev points $y_k$) onto a
domain containing $D_1$, and this in such a way that $g(J) = I$ and that the new nodes $x_k =
g(y_k)$ are closer to equidistant than the $y_k$’s. These authors then suggest to approximate
$f$ by the transplant of the polynomial in $y$ of degree at most $N$ interpolating $F(y) := f(x)$
between the $y_k$. While the exponential convergence is maintained by the analyticity of $g$,
the derivatives are now better conditioned.

We have successfully applied that same idea in [Ber-Mit3] to our rational interpolant
with optimized poles of [Ber-Mit1]. As a by-product, the improvement in the approxi-
mation was particularly impressive for functions with a large gradient in the center of the
interval, a consequence of the fact that Koslof and Tal-Ezer’s shift places more nodes there.
Though not surprising in principle, this improvement impressed us by its magnitude: in
some cases it was more pronounced than that obtained from the attachment of poles.

After having studied the numerical results of [Ber-Mit3], Richard Baltensperger has
suggested that we use instead of Koslof and Tal-Ezer’s shift a much more adaptive $g$ such
as the one advocated and proven more efficient and versatile by Bayliss and Turkel in [Bay-
Tur]. The present work is devoted to the use of such adaptive point shifts for improving
upon the optimized denominator rational interpolation of [Ber-Mit1].

Section 2 reviews polynomial interpolation and the method of constructing rational
interpolants by optimally attaching poles to the interpolating polynomial, as introduced
in our earlier work. Section 3 recalls the effect of conformal point shifts on the interpolant
and its derivatives. Section 4 describes the use of Bayliss and Turkel’s point shift, which
involves two parameters, one for the location of an accumulation of points, the other for
their concentration. In Section 5 we construct a new shift that can in principle handle
an arbitrary number of shocks. The work is completed with numerical examples and
conclusions.

2. Rational interpolation with optimized denominator

This interpolation method, recently introduced in [Ber-Mit1], starts with the polynomial
interpolant. The unique polynomial $p_N$ of degree $\leq N$ that interpolates $f$ between the $x_k$
may be written in its barycentric form [Hen]

$$p_N(x) = \sum_{k=0}^{N} \frac{w_k}{x - x_k} f_k / \sum_{k=0}^{N} \frac{w_k}{x - x_k}$$ (2.1)

where

$$w_k := 1 / \prod_{i=0, i \neq k}^{N} (x_k - x_i).$$

3
For several standard point sets \( \{ x_k \}_{k=0}^{N} \), the weights \( w_k \) may be analytically computed, one of the many advantages of Lagrange interpolation [Ber]. This is in particular the case for Chebyshev points. Moreover, since the \( w_k \) appear in the numerator and in the denominator of (2.1), any common factor may be discarded. Here we will use the Chebyshev points of the second kind \( x_k = \cos k \frac{\pi}{n} \), whose simplified weights read [Sal]

\[
    w_k^* = (-1)^k \delta_k, \quad \delta_k := \begin{cases} 
        1/2, & k = 0 \text{ or } k = N, \\
        1, & \text{otherwise}. 
    \end{cases}
\]

Then (2.1) evaluates \( p_N \) at any given \( x \) in \( O(N) \) operations.

As mentioned in the introduction, the convergence \( p_N \to f \) is very rapid if \( f \) is differentiable a large number \( p \) of times, and it adapts automatically to \( p \). It is exponential for analytic functions such as those we will consider in the examples.

On the other hand, it is easy to infer from Markov’s inequality that, when functions with steep gradients (shocks) are to be approximated, the degree \( N \), and hence the number of \( x_k \), will have to grow sharply with the steepness of the shocks [Ber-Mit1]. Rational interpolation is thus to be preferred when approximating functions with large slopes. As mentioned in the introduction, classical rational interpolation is hampered by unattainable points and unpredictable poles close to or even on \( I \) (see, e.g., Figure 1b and Table 5 in [Ber-Mit1]).

For this reason, we have suggested in [Ber-Mit1] a different kind of rational interpolation in cases where \( f \) is known everywhere on \( I \) and where the computing time is not an issue. The basic idea is to start with \( p_N \), which may be viewed as a rational interpolant with all its poles at infinity, to fix \( P \), the number of poles \( z_j \) to be attached to \( p_N \), and to move the \( z_j \) from infinity toward an optimal position, i.e., one which minimizes some error functional, here \( \| r - f \|_\infty \).

Let \( \mathcal{R}_{mn} \) denote the set of all rational functions with numerator and denominator degrees at most \( m \), resp. \( n \). Our problem is therefore the following.

**Problem.** Find among all \( r \in \mathcal{R}_{NP} \) with \( r(x_k) = f(x_k) \), \( k = 0, 1, \ldots, N \), one that minimizes

\[
    \| r - f \|_\infty := \max_{x \in I} | r(x) - f(x) | .
\]

It is easy to see that this optimization problem has a solution [Ber-Mit1]. The unicity question is not settled, yet.

The resulting interpolants have many advantages over classical rational interpolants: they can have neither unattainable points nor unwanted poles, and the sequence \( \{ \| r - f \|_\infty : r \in \mathcal{R}_{NP} \} \) decreases as \( P \) increases.
The barycentric representation of \( p_N \) permits a very simple attachment of the poles: it suffices to replace the \( w_k \) in (2.1) with
\[
b_k := w_k \cdot d_k, \quad d_k := \prod_{i=1}^{P} (x_k - z_i)
\] (2.3)

(if an interpolant with the poles \( z_j \) exists, see [Ber-Mit1]). The optimization problem has been numerically solved with success in [Ber-Mit1] with standard modern algorithms. The results show that the nice properties just mentioned also arise in practice, and in particular that the resulting interpolants can indeed accommodate much more pronounced shocks than \( p_N \).

In [Ber-Mit2], such optimized rational interpolants have been computed as one part of a two-step algorithm for improving upon the polynomial pseudospectral solution of two-point boundary value problems. There, the minimized error functional was the size of the residual of the differential equation for the approximate solution \( r \). This residual involves derivatives, which are notoriously ill-conditioned for large number \( N \) of Chebyshev – or Legendre – points [Tre-Tru].

### 3. Conformal points shifts and their influence upon the derivatives

To reduce the just-mentioned ill-conditioning, Kosloff and Tal-Ezer [Kos-Tal] have suggested using in lieu of the Chebyshev points themselves their images under a conformal map that renders them closer to equidistant, an idea which has proven effective in several instances [Bal-Ber, Red-Wei-Nor, Ber-Bal, Ren-Fró]. We have applied it ourselves in [Ber-Mit3] to the above method of interpolation with optimized denominator.

There we have attached poles \( z_j \) in \( x \)-space by attaching poles \( v_j = g^{[-1]}(z_j) \) to \( P_N(y) = \sum_{k=0}^{N} \frac{w_k}{y-y_k} f_k / \sum_{k=0}^{N} \frac{w_k}{y-y_k} \) in \( y \)-space. In view of (2.3), the resulting rational interpolant reads
\[
R(y) := \frac{\sum_{k=0}^{N} \frac{w_k}{y-y_k} P \prod_{i=1}^{P} (y_k - v_i)}{\sum_{k=0}^{N} \frac{w_k}{y-y_k} P} = \frac{\sum_{k=0}^{N} \frac{w_k}{y-y_k} P \prod_{i=1}^{P} (g^{[-1]}(x_k) - g^{[-1]}(z_i))}{\sum_{k=0}^{N} \frac{w_k}{y-y_k} P \prod_{i=1}^{P} (g^{[-1]}(x_k) - g^{[-1]}(z_i))} = : r(x).
\]

and the \( v_j \) are optimized by minimizing
\[
\| R - F \|_{\infty} = \max_{y \in J} | R(y) - F(y) |,
\] (3.1)
as in (2.2).

As mentioned in the introduction, one goal of the conformal shift, besides a better approximation at steep shocks in the interior of the interval, is the improvement of the approximation of the derivatives. The derivatives of $r$ are given as functions of those of $R$ by [Ber-Mit3]

$$
 r'(x) = R'(y)y'(x) = \frac{R'(y)}{g'(y)},
$$

$$
 r''(x) = R''(y)[y'(x)]^2 + R'(y)y''(x) = \frac{R''(y)}{[g'(y)]^2} - \frac{g''(y)}{[g'(y)]^3} R'(y),
$$

(3.2)

where $R'$ and $R''$ can be expressed by the Schneider–Werner formulae

$$
 R'(y) = \begin{cases} 
 \sum_{k=0}^{N} \frac{u_k}{y - y_k} R[y, y_k] & y \neq y_i, \ i = 0(1)N, \\
 -\left( \sum_{k=0}^{N} \frac{u_k R[y_i, y_k]}{u_i} \right) & y = y_i 
\end{cases}
$$

$$
 R''(y) = \begin{cases} 
 2 \sum_{k=0}^{N} \frac{u_k}{y - y_k} R[y, y, y_k] & y \neq y_i, \ i = 0(1)N, \\
 -2 \left( \sum_{k=0}^{N} \frac{u_k R[y_i, y_i, y_k]}{u_i} \right) & y = y_i, 
\end{cases}
$$

with $R[z, z, y_k] = \frac{R(z) - R[z, y_k]}{z - y_k}$.

4. Adaptive point shifts for functions with a shock

It now remains to choose a good conformal point shift $g$. In [Ber-Mit3], we went for Koslof and Tal-Ezer’s map

$$
 g(z) = \frac{\arcsin \alpha z}{\arcsin \alpha}
$$

with a parameter $\alpha$ varying from 0 to 1. For $\alpha = 0$, the points remain the Chebyshev ones. As $\alpha \to 1$ they become equidistant. The fast convergence rates mentioned in the introduction are maintained for every fixed $\alpha < 1$.

This shift was introduced for improving the conditioning of the derivatives of the approximation in time evolution problems. In [Ber-Mit3], however, we have noticed an important by-product, namely a sharp improvement in the quality of the approximation itself for functions with a shock at 0. The aim of the present work is to take a better advantage of this effect by adapting the shift to the interpolated function. We achieve this
by considering the various proposals for shifts $g$ by Bayliss and Turkel in [Bay-Tur]. Their comparisons demonstrate the particularly remarkable efficiency of the map

$$x = g(y) = \frac{1}{\alpha} \tan[\lambda(y - \mu)] + \beta \quad (4.1)$$

where

$$\lambda = \frac{\gamma + \delta}{2}, \quad \mu = \frac{\gamma - \delta}{\gamma + \delta} \quad (4.2)$$

with

$$\gamma = \arctan[\alpha(1 + \beta)], \quad \delta = \arctan[\alpha(1 - \beta)].$$

$g$ embodies two parameters: $\beta$ determines the location of the maximal gradient, $\alpha$ its magnitude. ($g$ is constructed from $y = g^{[-1]}(x) = \mu + \frac{1}{\lambda} \arctan[\alpha(x - \beta)]$, and $\beta$ is the location of the maximal gradient of the corresponding arcus tangens; notice that $\mu = 0$ for $\beta = 0$.) Figure 1b displays $g^{[-1]}(x)$ for $\alpha = 35.35$ and $\beta = -0.5024$, the best values for approximation without the help of poles in Table 2 (Example 1).

**Figure 1**
Function and inverse change of variable in Example 1, Table 2, no poles

Bayliss and Turkel have roughly optimized the shift by computing over a grid of values of $\alpha$ and $\beta$ and selecting those parameter values for which a functional related to the interpolation error is minimized. Here we minimize the functional (3.1) with respect to the $P + 2$ variables $z_1, ..., z_P, \alpha, \beta$ using standard software. In the numerical examples discussed
in §6, we also estimate the precision of the derivatives of $r$ by evaluating $\|r' - f'\|_\infty$ and $\|r'' - f''\|_\infty$ with the formulae (3.2). A simple calculus exercise yields for the expressions containing $g$:

$$g'(y) = \alpha \cos^2 t/\lambda, \quad \frac{g''(y)}{[g'(y)]^3} = 2\alpha^2 \cos^3 t \sin t/\lambda, \quad t = \lambda(y - \mu).$$

Notice that, in contrast with the method used in [Ber-Mit1, Ber-Mit2, Ber-Mit3], the nodes $x_k$ are no longer chosen at the onset, but result from the optimization procedure. The Remes algorithm [Bra], on the other hand, yields best rational approximants in $\mathcal{R}_{N,P}$ which interpolate between (at least) $N + P + 2$ abscissae $x_k$. Our $x_k$, however, have the advantage of being images of fixed points $y_k$ under a conformal map determined by only few parameters.

5. Functions with multiple shocks

After having addressed functions with one steep gradient in the interior of $I$, it is natural to consider functions $f$ with several such shocks. When these are sufficiently close to one another, a simple change of variable with two parameters as in (4.1) will sometimes do [Bay-Tur]. In more general cases, however, two parameters will be necessary for every shock (one for location, the other for steepness).

One way is to split $I$ into subdomains and to consider another approximation on each of them [Bay-Tur]. This, though, destroys the spectral accuracy at the contact points of the various subdomains.

It seems therefore preferable to address such problems with a single conformal map involving two parameters for every shock, say $\alpha_q$ and $\beta_q$. We will now construct such a map. The basic idea arises from the observation that, as $g^{[-1]}$ in (4.1) is steep at the shock, it is relatively flat away from the latter, and the more so the steeper the shock. At such flat parts one thus can add another function to $g^{[-1]}$ without altering it too much. We therefore construct a new $g^{[-1]}$ accomodating $Q$ shocks as

$$y(x) = g^{[-1]}(x) = \mu + \frac{1}{\lambda} \sum_{q=1}^{Q} \arctan[\alpha_q (x - \beta_q)],$$

where $\lambda$ and $\mu$ are the parameters needed for ensuring that $g^{[-1]}(-1) = -1$, $g^{[-1]}(1) = 1$. A straightforward solution of this $2 \times 2$--system yields the same values of $\lambda$ and $\mu$ as in
\( (4.2) \), with
\[
\gamma := \sum_{q=1}^{Q} \arctan[\alpha_q(1 + \beta_q)]
\]
\[
\delta := \sum_{q=1}^{Q} \arctan[\alpha_q(1 - \beta_q)].
\]
To find \( g \) itself, we notice that \( g^{[-1]} \), as a sum of monotone increasing functions, is itself monotone increasing from \(-1\) to 1 and that, consequently, the equation \( g^{[-1]}(x) = y \), i.e.,
\[
\sum_{q=1}^{Q} \arctan[\alpha_q(x - \beta_q)] = \lambda(y - \mu)
\]
has a solution \( x \) for every \( y \in J \). The sum of arcus tangens may be written as a single arcus tangens, the argument of which is a rational function involving \( x^Q \). (5.2) thus will in general require a numerical solution, e.g., with the method of Dekker–Brent [Bre]. For \( Q \leq 4 \), \( y \) may be expressed according to the formulae for roots of polynomials. For simplicity, we will confine ourselves here to the case \( Q = 2 \). (5.2) then yields
\[
\frac{\alpha_1(x - \beta_1) + \alpha_2(x - \beta_2)}{1 - [\alpha_1\alpha_2(x - \beta_1)(x - \beta_2)]} = \tan[\lambda(y - \mu)] =: t
\]
or \( ax^2 + bx + c = 0 \) with
\[
a := \alpha_1\alpha_2 t, \quad b := [(\alpha_1 + \alpha_2) - a(\beta_1 + \beta_2)],
\]
\[
c := (\alpha_1\alpha_2\beta_1\beta_2 - 1)t - (\alpha_1\beta_1 + \alpha_2\beta_2).
\]
Notice that the sum formula used to transform (5.2) into (5.3) holds true merely up to multiples of \( \pi \). We have therefore chosen that solution of the quadratic equation for which \( g^{[-1]}(x) \) equals \( y \), the value to be inverted.

The derivatives, to be inserted into (3.2) when computing \( r'(x) \) and \( r''(x) \), are obviously given by
\[
y'(x) = \frac{1}{\lambda} \sum_{q=1}^{Q} \frac{\alpha_q}{1 + s_q^2},
\]
\[
y''(x) = -\frac{1}{\lambda} \sum_{q=1}^{Q} \frac{2\alpha_q^2 s_q}{(1 + s_q^2)^2}, \quad s_q := \alpha_q(x - \beta_q).
\]

6. Numerical examples

We have tested the effect of adaptive point shifts on our rational interpolation with optimized poles on two examples, one with a single shock, the other with two shocks. The
optimization problems were solved with the simulated annealing method of [Cor-Mar-Mar-Rid]. The minimized sup-norm (3.1) has thereby been estimated by considering the 1000 equally spaced points

$$\hat{y}_\ell = -\frac{5}{4} + \frac{\ell - 110}{999} \frac{4}{4}, \quad \ell = 1(1)1000,$$

on the interval $[-5/4, 5/4]$ and computing the maximal absolute value of the error at those $\hat{y}_\ell$ lying in $[-1, 1]$.

**Example 1.** The function is

$$f(x) = e^{1/(x+1.2)} + \cos \pi (x + 0.5) + \frac{\text{erf}(\delta(x + 0.5))}{\text{erf}(\delta)}, \quad \delta = \sqrt{5}e.$$

It consists of a function suggested in [Hem] as a test for solution methods for two-point boundary value problems, shifted by $-0.5$ and supplemented with a term with essential singularity at $x = -1.2$, i.e., on the real axis to the left of $I$. Figure 1a shows $f$ for $\epsilon = 10^6$ between $-0.7$ and $1$, in order for the shock to be distinguishable (at $-1$ the value of $f$ is about 147).

**Table 1**

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\alpha$</th>
<th>Poles</th>
<th>$|r - f|$</th>
<th>$|r' - f'|$</th>
<th>$|r'' - f''|$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0</td>
<td></td>
<td>1.867e-1</td>
<td>2.361e+1</td>
<td>3.454e+3</td>
</tr>
<tr>
<td>0</td>
<td>1.0</td>
<td>$(-.4902, \pm 2.011e-2)$</td>
<td>5.224e-4</td>
<td>2.158e-1</td>
<td>8.241e+1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$(-.5056, \pm 2.228e-2)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$(-.5178, \pm 5.613e-2)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-.5185</td>
<td>7.408</td>
<td></td>
<td>9.447e-9</td>
<td>5.012e-6</td>
<td>1.138e-2</td>
</tr>
<tr>
<td>-.4976</td>
<td>8.273</td>
<td>$(-1.027, \pm 3.147e-3)$</td>
<td>1.279e-11</td>
<td>6.270e-9</td>
<td>1.654e-5</td>
</tr>
<tr>
<td>-.4981</td>
<td>8.519</td>
<td>$(-1.030, \pm 3.574e-3)$</td>
<td>2.495e-12</td>
<td>1.754e-9</td>
<td>5.659e-6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$(1.062, \pm 5.346e-3)$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

First we have approximated $f$ for $\epsilon = 10'000$. The results with $N = 100$ are displayed in Table 1. $\beta$ and $\alpha$ are the parameters used in the change of variable (4.1) — $\alpha = 1$ and $\beta = 0$ thereby meaning that no change is made — the third column gives the optimized pole
pairs and the last three show the approximation errors $\|r - f\|$, $\|r' - f'\|$ and $\|r'' - f''\|$, estimated in the same way as the minimized norms (3.1) — but now in $x$–space with equidistant $\hat{x}_t := \hat{y}_t$.

**Table 2**

Effect of an optimized Bayliss–Turkel point shift on rational approximation with optimized poles in Example 1 with $\epsilon = 1'000'000$ and $N = 240$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\alpha$</th>
<th>Poles</th>
<th>$|r - f|$</th>
<th>$|r' - f'|$</th>
<th>$|r'' - f''|$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0</td>
<td>$(-0.5001, \pm 2.017e-3)$, $(-0.5041, \pm 2.080e-2)$</td>
<td>1.119</td>
<td>4.204e + 2</td>
<td>4.287e + 5</td>
</tr>
<tr>
<td>0</td>
<td>1.0</td>
<td>$(-0.5001, \pm 2.017e-3)$, $(-0.5041, \pm 2.080e-2)$</td>
<td>8.031e - 4</td>
<td>2.628e + 1</td>
<td>4.491e + 4</td>
</tr>
<tr>
<td>$-0.5041$</td>
<td>35.35</td>
<td>$(-1.007, \pm 9.506e-4)$</td>
<td>2.873e - 9</td>
<td>2.335e - 5</td>
<td>8.499e - 2</td>
</tr>
<tr>
<td>$-0.5009$</td>
<td>39.79</td>
<td>$(-1.006, \pm 6.682e-4)$</td>
<td>6.340e - 11</td>
<td>2.466e - 7</td>
<td>2.431e - 3</td>
</tr>
<tr>
<td>$-0.4968$</td>
<td>42.59</td>
<td>$(1.046, \pm 6.790e-3)$</td>
<td>4.860e - 12</td>
<td>4.062e - 8</td>
<td>2.151e - 3</td>
</tr>
</tbody>
</table>

The table documents the successive improvements in the approximation: the polynomial has an error of about $10^{-1}$, which decreases to $5 \cdot 10^{-4}$ with the attachment of 6 poles. Without poles, on the other hand, but with an optimized Bayliss–Turkel point shift (4.1), the error becomes as small as $10^{-8}$.

Shifting to the next order derivative worsens the error by about $10^2$ when the precision is low, by $10^3$–$10^4$ when $\|r - f\|$ is small. This is consistent with the theoretical factor for Chebyshev points, about $\mathcal{O}(N)$ in the interior of the interval and $\mathcal{O}(N^2)$ close to the extremities [Bre-Eve]. The factor can therefore be expected to be $\mathcal{O}(N)$ as long as the approximation error — which arises in the interior in our examples — dominates this differentiation error, $\mathcal{O}(N^2)$ when it becomes smaller. Without point shift this derivative–induced loss of accuracy at the endpoints would be even more pronounced.

For approximating interior shocks the change of variable is more powerful than the pole attachment; and if it would not be for the function $e^{1/(x+1.2)}$, attaching poles would not bring much improvement. Because of this steep function outside, but nevertheless close to the approximation interval, the poles are effective here; changes of variable such as (5.1) cannot do any good, they are even harmful since they shift the nodes away from
the boundary (this in contrast to the sinh–shift used in [Tan-Tru], which on the other hand moves the points away from the interior shocks). The last two examples of Table 1 correspondingly document the efficiency of the pole attachment, with two poles and with four: we end up with a precision of about $2 \cdot 10^{-12}$, with three digits lost for every derivative. The poles come to the vicinity of the boundary.

Similar results were obtained with $\epsilon = 10^6$, the sole difference being that, due to the larger value of $N$ necessary for a good approximation, the loss in the precision induced by the differentiation is more pronounced, as documented in Table 2 for $N = 240$.

**Example 2.** Here we started from the above function, shifted by 0.75 instead of $-0.5$, and we have added to it a tanh–term, centered at $-0.5$, and the same exponential as before:

$$f(x) = e^{1/(x+1.2)} + \cos \pi (x - 0.75) + \frac{\text{erf}(\delta(x - 0.75))}{\text{erf}(\delta)} + \tanh(\eta(x + 0.5)), \quad \delta = \sqrt{5\epsilon}.$$  

Figure 2a shows $f$ between $-0.7$ and 1 for $\epsilon = 10^4$ and $\eta = 10^2$.

**Figure 2**

Function and change of variable in Example 2, Table 3, no poles

Our results are summarized in Table 3. Without shift the poles come close to the location of the shocks, as in example 1. The point shift is again already impressively efficient without poles. Attaching the latter does not bring too much improvement here. A look at their location reveals that they take care of other difficult stretches than the
left boundary layer. Indeed, the first pair comes close to the median of the location of the two shocks. This may arise from the fact that the concentration of so many points at the shocks depletes the median region too much for a reasonable approximation with the remaining information. The second pair helps at the right shock, which comes so close to 1 (figure 2b) that the change of variable cannot move as many points toward it than it does to the left shock. The obtained precision of $10^{-9}$ is nevertheless quite remarkable for an analytic approximation using “only” 200 interpolation points. And the second derivative stays at a good $1.58 \cdot 10^{-3}$.

Table 3

Effect of an optimized two–shock point shift on the rational approximation with optimized poles in Example 2 with $\epsilon = 10^000$, $\eta = 100$ and $N = 200$

| $\beta_1$ | $\alpha_1$ | $\beta_2$ | $\alpha_2$ | Poles | $||r - f||$ | $||r' - f'||$ | $||r'' - f''||$ |
|----------|-----------|-----------|-----------|-------|------------|-----------|----------|
| 0        | 1.0       | 0.0       | 1.0       |       | 4.301e-2   | 1.043e+1  | 3.106e+3 |
| 0        | 1.0       | 0.0       | 1.0       | (.5000, ±1.572e-2) (7.293, ±7.826e-2) (7.494, ±3.717e-2) | 1.109e-5   | 3.598e-3  | 1.301    |
| -.4924   | 13.25     | .7125     | 5.114     |       | 1.728e-8   | 2.022e-5  | 2.151e-2 |
| -.4819   | 15.40     | .7141     | 5.777     | (.710, ±7.037e-2) | 1.667e-9   | 4.097e-6  | 3.066e-3 |
| -.4855   | 15.41     | .7154     | 4.888     | (.1199, ±8.128e-2) (.7084, ±1.048e-1) | 1.363e-9   | 2.191e-6  | 1.577e-3 |

7. Conclusion

The combination of adaptive conformal point shifts and optimal pole attachment, as suggested in the present work, turns out to be a very efficient means of improving upon the Chebyshev interpolating polynomial when approximating functions with shocks. We have thereby used a point shift suggested by Bayliss and Turkel before constructing one that can in principle handle a function with an arbitrary number of difficult stretches.

With a single shock, in example 1, the precision was improved from $10^{-1}$–$10^0$ to as much as $10^{-12}$, quite an achievement. With two shocks, one of them quite close to the boundary, the error which $N = 200$ worsens to $10^{-9}$, but a loss of $10^{-3}$ for each derivative is very satisfactory for an analytic approximation.

In the future we intend to employ these point shifts for improving upon the two–point
boundary value problem solver introduced in [Ber-Mit2].

References


[Ren-Frö] Renaut R., Fröhlich J., A pseudospectral Chebyshev method for the 2D wave 
equation with domain stretching and absorbing boundary conditions, *J. Comput. 

[Sal] Salzer H. E., Lagrangian interpolation at the Chebyshev points $x_{n,\nu} = \cos(\nu \pi/n)$, 
$\nu = 0(1)n$; some unnoted advantages, *The Computer J.* 15 (1972) 156–159.

1983).

[Tan-Tru] Tang T., Trummer M. R., Boundary layer resolving pseudospectral methods for 

[Tre-Tru] Trefethen L. N., Trummer M. R., An instability phenomenon in spectral methods, 