SHARP GROWTH ESTIMATES FOR SUBDIAGONAL RATIONAL PADÉ APPROXIMATIONS (FIRST DRAFT)

FRANK NEUBRANDER, KORAY ÖZER, AND LEE WINDSPERGER

Abstract. The computational powers of computer algebra systems like Maple or Mathematica are used to prove polynomial identities that are essential to obtain sharp growth estimates for subdiagonal rational Padé approximations of the exponential function.

1. Introduction

Let \( r = \frac{P}{Q} \) be an \( \mathcal{A} \)-stable rational approximation to the exponential function of order \( m \); i.e., \( P \) and \( Q \) are polynomials with \( p := \deg(P) \leq \deg(Q) =: q \), and

(i) \(|r(z) - e^z| \leq C_m |z|^{m+1} \) for \(|z|\) sufficiently small, and

(ii) \(|r(z)| \leq 1 \) for \( \Re(z) \leq 0 \).

It is a well-known result of Padé [2] that \( m \leq p + q \) for all rational approximations to the exponential function. The rational approximations of maximal order \( m = p + q \) are called Padé approximations. They are of the form \( r = \frac{P}{Q} \), where

\[
P(z) = \sum_{j=0}^{p} b[j,p] z^j, \quad \text{where} \quad b[j,p] = \frac{(m-j)!p!}{m!j!(p-j)!} \]

and

\[
Q(z) = \sum_{j=0}^{q} a[j,q] (-z)^j, \quad \text{where} \quad a[j,q] = \frac{(m-j)!q!}{m!j!(q-j)!}.
\]

Moreover, for every Padé approximation \( r(z) = \frac{P(z)}{Q(z)} \) of the exponential of order \( m = p + q \) we have that

\[
r(z) - e^z = \frac{(-1)^{q+1}}{Q(z)} \frac{1}{m!} z^{m+1} e^z \int_0^1 s^p (1-s)^q e^{-sz} ds
\]

(see, for example, [4], Section 75 (Die Exponentialfunktion), or [5]). As shown in [1], Padé approximations are \( \mathcal{A} \)-stable if and only if \( q - 2 \leq p \leq q \). A rational Padé approximation \( r(z) = \frac{P(z)}{Q(z)} \) is called subdiagonal if \( p = q - 1 \). In particular, a subdiagonal Padé approximation is always \( \mathcal{A} \)-stable and of odd approximation order \( m = 2q - 1 \).

In applications (see, for example [?], [3], [6]), the following identities and estimates are useful.

---

Date: September 10, 2013.

1991 Mathematics Subject Classification. 65R10, 44A10, 41A20, 41A25, 47D06.

Key words and phrases. Laplace transform inversion, A-stable rational functions, rational approximation of semigroups, time discretization.

Acknowledgments. The first author would like to thank Moritz Egert from the University of Darmstadt for inspiring discussions during his visit at LSU.
Proposition 1. Let \( r(z) = \frac{P(z)}{Q(z)} \) be a subdiagonal Padé approximation, then

\[
|Q(is)|^2 = Q(is) * Q(-is) = \sum_{j=0}^{q} d[2j, q] s^{2j},
\]

where \( d[0, q] = 1, d[2q, q] = \left(\frac{(q-1)!}{(2q-1)!}\right)^2 \), and

\[
d[2j, q] = \frac{1}{2j!} \prod_{k=1}^{j} \frac{(q-k+1)}{(2q-k)(2q-2k+1)} \text{ if } 0 < j < q.
\]

In particular,

(a) \( \sup_{s \in \mathbb{R}} \frac{1}{|Q(is)|} = 1 \).

(b) \( \sup_{s \in \mathbb{R}} \left| \frac{q^n}{Q(is)} \right| = \frac{(2q-1)!}{(q-1)!} \frac{m!}{n!} \).

(c) \( \sup_{s \in \mathbb{R}} \left| \frac{q^n}{Q(is)} \right| \leq \sqrt{\frac{1}{d[2n, q]}} \) for all \( 0 \leq n \leq q \).

Proof. Clearly,

\[
|Q(is)|^2 = Q(is) * Q(-is) = \left( \sum_{j=0}^{n} (-1)^j j! a[j, q] s^j \right) * \left( \sum_{j=0}^{n} j! a[j, q] s^j \right) = \sum_{n=0}^{2q} d[n, q] s^n,
\]

where \( d[n, q] \) is given by the Cauchy product

\[
d[n, q] = \sum_{j=0}^{n} (-1)^j j! a[j, q] * i^{n-j} a[n-j, q] = i^n \sum_{j=0}^{n} (-1)^j a[j, q] * a[n-j, q]
\]

and where we set \( a[j, q] := 0 \) if \( j > q \). It can be easily seen that \( d[n, q] = 0 \) if \( n \) is odd and it follows immediately that\(^1\)

\[
d[0, q] = a[0, q]^2 = 1 \quad \text{and} \quad d[2q, q] = a[q, q]^2 = \left( \frac{p!}{m!} \right)^2 = \left( \frac{(q-1)!}{(2q-1)!} \right)^2.
\]

For \( 0 < j < q \) and \( n = 2j \) the identity

\[
d[n, q] = d[2j, q] = (-1)^j \sum_{k=0}^{2j} (-1)^k a[k, q] * a[2j-k, q] = \frac{1}{2j!} \prod_{k=1}^{j} \frac{(q-k+1)}{(2q-k)(2q-2k+1)}
\]

can be proven with the following Mathematica code.

\[
\text{a}[j_, q_] := ((2q - 1 - j)!q!)/((2q - 1)!j!(q - j)!);^2
\]

\[
d[n_, q_] := (-1)^{n/2} \text{Sum}[(-1)^j a[j, q] a[n-j, q], \{j, 0, n\}];
\]

\[
\text{Ans}[n_, q_] := 1/((2^{n/2})(n/2)!);^2 \text{Product}[(q - k + 1)/((2^q q - k)(2q - 2k + 1)), \{k, 1, n/2\}];
\]

\[
\text{FullSimplify}[d[2m, q] - \text{Ans}[2m, q]]
\]

This proves (1.4). Since \( d[0, q] = 1 \) and \( d[2j, q] \geq 0 \) for all \( j \in \mathbb{N}_0 \) it follows from (1.4) that \( \min_{s \in \mathbb{R}} |Q(is)| = 1 \). Moreover, since

\[
\frac{|Q(is)|^2}{s^{2n}} = \sum_{j=0}^{q} d[2j, q] s^{2j-2n} \geq d[2n, q],
\]

\(^1\)Other identities are \( d[2, q] = \frac{q}{2m}, d[2q - 2, q] = \left(\frac{q}{m}\right)^2 \) and \( d[2q - 4, q] = \left(\frac{q}{m-1}\right)^2 \).

\(^2\)Notice that \( a[j, q] \) is well defined as long as \( 0 \leq j < 2q \) and that Mathematica will evaluate \( a[j, q] \) to be zero if \( q < j < 2q \) since in this case \( 1/(q-j)! = 0. \)
it follows that
\[
|s_n^{\prime}\left|Q(is)\right| \leq \frac{1}{\sqrt{d[2n, q]}}
\]
for all \(s \in \mathbb{R}\) and that \(\sup_{s \in \mathbb{R}} \left|\frac{s_n^{\prime}}{Q(is)}\right| = d[2q, q].\)

**Proposition 2.** Let \(r(z) = \frac{P(z)}{Q(z)}\) be a subdiagonal Padé approximation, then

(a) \(\left|\frac{Q'(is)}{Q(is)}\right| \leq 1\) for all \(s \in \mathbb{R}\).
(b) \(\left|\frac{P'(is)}{Q(is)}\right| \leq 1\) for all \(s \in \mathbb{R}\).
(c) \(\left|r'(is)\right| \leq 2\) for all \(s \in \mathbb{R}\).\(^3\)

**Proof.** It follows from
\[
Q'(z) = \sum_{j=1}^{q} a[j, q] j(-1)^j z^{j-1} = \sum_{j=0}^{q-1} \tilde{a}[j, q] (-1)^{j+1} z^j,
\]
where \(\tilde{a}[j, q] = (j + 1)a[j + 1, q]\), that
\[
|Q'(is)|^2 = Q'(is) * Q'(-is) = \left( \sum_{j=0}^{q} (-1)^j (-1)^j a[j, q] s^j \right) \left( \sum_{j=0}^{q} (-1)^j (-1)^j a[j, q] s^j \right) = \sum_{n=0}^{2q-2} \tilde{d}[n, q] s^n,
\]
where
\[
\tilde{d}[n, q] = \sum_{j=0}^{n} (-1)^j \tilde{a}[j, q] \ast i^{n-j} \tilde{a}[n-j, q] = i^n \sum_{j=0}^{n} (-1)^j \tilde{a}[j, q] \ast \tilde{a}[n-j, q]
\]
and where we set \(\tilde{a}[j, q] := 0\) if \(j > q - 1\). It can be easily seen that \(\tilde{d}[n, q] = 0\) if \(n\) is odd. For \(0 \leq j < q\) and \(n = 2j\) the identity
\[
\tilde{d}[n, q] = \tilde{d}[2j, q] = \frac{q(q-j)}{(2q-j-1)(2q-2j-1)} d[2j, q]
\]
can be proven with the following Mathematica code.

Now, \(\left|\frac{Q'(is)}{Q(is)}\right| \leq 1\) for all \(s \in \mathbb{R}\) if and only
\[
0 \leq Q(is)Q(-is) - Q'(is)Q'(-is) = d[2q, q] s^{2q} + \sum_{j=0}^{q-1} \left( d[2j, q] - \tilde{d}[2j, q] \right) s^{2j}
\]
for all \(s \in \mathbb{R}\). However, the last statement holds since \(d[2j, q] - \tilde{d}[2j, q] \geq 0\) for all \(0 \leq j \leq q - 1\). This proves statement (a). To prove (b) observe that
\[
P(z) = \sum_{j=0}^{q-1} b[j, q] z^j, \text{ where } b[j, q] = \frac{(2q-1-j)! (q-1)!}{(2q-1)! j!(q-1-j)!} = \frac{q-j}{q} a[j, q]
\]
\(^3\)In fact, the sharper estimate \(\max_{s \in \mathbb{R}} |r'(is)| = 1\) holds. However, we do not jet have a working Mathematica code to prove this estimate.
for \(0 \leq j\). Then

\[ |P(is)|^2 = P(is) * P(-is) = \left( \sum_{j=0}^{q-1} i^j b[j, q] s^j \right) * \left( \sum_{j=0}^{q-1} (-1)^j b[j, q] s^j \right) = \sum_{n=0}^{2q-2} e[n, q] s^n, \]

where

\[ e[n, q] = \sum_{j=0}^{n} i^j b[j, q] * i^{n-j}(-1)^{n-j} b[n - j, q] = (-i)^n \sum_{j=0}^{n} (-1)^j b[j, q] * b[n - j, q] \]

and where we set \(b[j, q] := 0\) if \(j > q - 1\). It can be easily seen that \(e[n, q] = 0\) if \(n\) is odd. For \(0 \leq j < q\) and \(n = 2j\) the identity

\[ e[n, q] = d[n, q] \]

can be proven with the following Mathematica code.

```mathematica
bb[j_, q_] := ((q - j)/q)a[j, q];
e[n_, q_] := (-1)^n Sum[(-1)^j b[j, q] b[n - j, q], {j, 0, n}];
FullSimplify[d[2r, x]/Ans[2r, x]]
```

This shows that

\[ (1.5) \quad |Q(is)|^2 - |P(is)|^2 = d[2q, q]s^{2q} = \left( \frac{(q - 1)!}{(2q - 1)!} \right)^2 s^{2q} \]

(see also Theorem 3.3 in [1]). It follows from

\[ P'(z) = \sum_{j=1}^{q} b[j, q] j z^{j-1} = \sum_{j=0}^{q-2} b[j, q] z^j, \]

where \(b[j, q] = (j + 1)b[j + 1, q]\), that

\[ |P'(is)|^2 = P'(is) * P'(-is) = \left( \sum_{j=0}^{q-1} i^j \tilde{b}[j, q] s^j \right) * \left( \sum_{j=0}^{q-1} (-i)^j \tilde{b}[j, q] s^j \right) = \sum_{n=0}^{2q-2} \tilde{e}[n, q] s^n, \]

where

\[ \tilde{e}[n, q] = \sum_{j=0}^{n} i^j \tilde{b}[j, q] * (-i)^{n-j} \tilde{b}[n - j, q] = (-i)^n \sum_{j=0}^{n} (-1)^j \tilde{b}[j, q] * \tilde{b}[n - j, q] \]

and where we set \(\tilde{b}[j, q] := 0\) if \(j > q - 1\). It can be easily seen that \(\tilde{e}[n, q] = 0\) if \(n\) is odd. For \(0 \leq j < q\) and \(n = 2j\) the identity

\[ \tilde{e}[n, q] = \tilde{e}[2j, q] = \frac{(q - 1)(q - j - 1)}{(2q - j - 1)(2q - 2j - 1)} d[2j, q] = \frac{q - 1}{q} \frac{q - j - 1}{q - j} d[2j, q] \]

can be proven with the following Mathematica code.

```mathematica
bb[j_, q_] := (j + 1)b[j + 1, q];
e[n_, q_] := (-1)^n Sum[(-1)^j b[j, q] b[n - j, q], {j, 0, n}];
EEAns[n_, q_] := ((q - 1)(q - n/2 - 1)/((2 q - n/2 - 1)(2 q - n - 1))) Ans[n, q];
FullSimplify[ee[2r, x]/EEAns[2r, x]]
```

It follows from (1.5) that

\[ |P(is)|^2 - |P'(is)|^2 = |Q(is)|^2 - d[2q, q]s^{2q} - |P'(is)|^2 = \sum_{j=0}^{2q-2} (d[2j, q] - \tilde{e}[2j, q]) s^{2j} \geq 0 \]
where
\[ d[2j, q] - e[2j, q] = (1 - \frac{(q - 1)(q - j - 1)}{(2q - j - 1)(2q - 2j - 1)})d[2j, q] = \frac{(3q - 2j - 2)(q - j)}{(2q - j - 1)(2q - 2j - 1)}d[2j, q] \geq 0 \]
for all \( 0 \leq j \leq q - 1 \). This shows statement (b). Finally, statement (c) follows from
\[
 r'(z) = \frac{P'(z)Q(z) - P(z)Q'(z)}{Q(z)^2} = r(z) \left( \frac{P'(z)}{P(z)} - \frac{Q'(z)}{Q(z)} \right)
\]
and the fact that \( |r(z)| \leq 1 \) for \( Re(z) \leq 0 \) (A-stability). \( \square \)

For subdiagonal Padé approximation with \( p = q - 1 \), the Perron representation (1.3) can be written as
\[
 r(z) - e^z = (-1)^{q+1} \frac{z^q}{Q(z)} \frac{1}{(2q - 1)!} \int_0^1 f(t, q) z^q e^{-tz} dt,
\]
where
\[
 f(t, q) := t^{q-1}(1-t)^q = \sum_{j=0}^{q} (-1)^j \frac{q!}{j!(q-j)!} t^{q-j-1}.
\]
Then
\[
 f^{(k)}(t, q) := \sum_{j=0}^{q} (-1)^j a[j, q, k] t^{q-j-1-k} = (-1)^{q-k-1} \sum_{i=q-k-1}^{2q-k-1} (-1)^{i} \tilde{a}[i, q, k] t^i,
\]
where \( a[j, q, k] = \frac{q!}{j!(q-j)!} \frac{(q+j-1)!}{(q+j-1-k)!} \) and
\[
 \tilde{a}[i, q, k] = a[i - q + k + 1, q, k] = \frac{q!}{(i-q+1+k)! (2q - i - 1 - k)!} \frac{(i+k)!}{i!}.
\]
Notice that and that \( \tilde{a}[i, q, k] = 0 \) if \( i < q - k - 1 \) or \( i > 2q - k - 1 \). Let \( e_n(t) := t^n \). Then
\[
 f(t, q) = e_{q-1}(t)e_q(1-t) \text{ and therefore } f^{(k)}(t, q) = \sum_{i=0}^{k} \binom{k}{i} e_{q-1}^{(i)}(t)e_q^{(k-i)}(1-t).
\]
It follows that
\[
 f^{(k)}(0, q) = f^{(k)}(1, q) = 0
\]
for all \( 0 \leq k \leq q - 2 \). Thus, using integration by parts,
\[
 r(z) - e^z = (-1)^{q+1} \frac{z^q}{Q(z)} \frac{1}{(2q - 1)!} \int_0^1 f^{(k)}(t, q) z^{q-k} e^{-tz} dt
\]
for all \( 0 \leq k \leq q - 1 \). In applications one has to estimate the integrant in (1.6) for \( z = is \). To this end the following \( L_2 \)-estimates are useful. It follows from the above that
\[
 |f^{(k)}(t, q)|^2 = \left( \sum_{i=q-k-1}^{2q-k-1} (-1)^i \tilde{a}[i, q, k] t^i \right)^2 = \sum_{j=2(q-k-1)}^{2(2q-k-1)} b[j, q, k] t^j,
\]
where
\[
 b[j, q, k] = (-1)^j \sum_{i=0}^{j} \tilde{a}[i, q, k] \tilde{a}[j - i, q, k].
\]
For $0 \leq k \leq q - 1$ let

$$G(t, q, k) = \int |f^{(k)}(t, q)|^2 \, dt = \sum_{j=2}^{2(q-k-1)} b[j, q, k] \frac{1}{j + 1} t^{j+1}.$$ 

Then, since $G(0, q, k) = 0$ for $0 \leq k \leq q - 1$,

$$\int_0^1 |f^{(k)}(t, q)|^2 \, dt = G(1, q, k),$$

where

$$G(1, q, k) = \sum_{j=2(q-k-1)}^{2(2q-k-1)} b[j, q, k] \frac{1}{j + 1} = 2k \frac{(2q - 2)! (2q)!}{(4q - 2j + 1) (q - j)} \prod_{j=1}^{k} \frac{(4q - 2j - 1)(2j - 1)}{(2q - 2j - 1)(2j - 1)}.$$  

The last equality can be validated by Mathematica for $0 \leq k < q \leq 100$. Unfortunately, it seems that a symbolic proof of the identity requires some extra tools (to be completed shortly).

References


Frank Neubrander, Louisiana State University, neubrand@math.lsu.edu
Koray Özer, Roger Williams University, kozer@rwu.edu
Lee Windsperger, Winona State University, lwindsperger@winona.edu