STABILITY BARRIERS FOR PADÉ AND GENERALISED PADÉ APPROXIMATIONS

John Butcher

The University of Auckland, Auckland, New Zealand
butcher@math.auckland.ac.nz

For a given function $f(z) = c_0 + c_1 z + c_2 z^2 + \cdots$, with a Taylor expansion about zero, it is sometimes possible for given integers $n, d \geq 0$, to obtain a "Padé approximation":

$$f(z) = \frac{a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n}{b_0 + b_1 z + b_2 z^2 + \cdots + b_d z^d} + O(z^{n+d+1}),$$

and sometimes these are more accurate approximations than the Taylor series up to z^{n+d} terms. In the case of $f(z) = \exp(z)$, a unique approximation exists for all choices of $n, d \in \{0,1,2,\ldots,\}$.

Numerical methods for differential equations when applied to a linear test problem generate a sequence of approximations which satisfy a difference equation of the form

$$P_0(z) y_n + P_1(z) y_{n-1} + \cdots + P_k(z) y_{n-k} = 0, \tag{1}$$

where $P_0, P_1, \ldots, P_k$ are polynomials. In the case of "one-step methods, such as Runge-Kutta methods, (1) reduces to

$$P_0(z) y_n + P_1(z) y_{n-1} = 0 \tag{2}$$

and the stability of the difference equation is determined by the value of the rational function $R(z) = -P_1(z)/P_0(z)$. It is of considerable significance to know for which z values the difference equation (2) has stable solutions.

Theorem 1. (Confluent divided differences). *There exist non-zero coefficients* $a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_d$ *such that*

$$\begin{aligned} &b_0 f(1) + b_1 f'(1) + b_2 f''(1) + \cdots + b_d f^{(d)}(1) - a_0 f(0) - a_1 f'(0) - \\ &- a_2 f''(0) - \cdots - a_n f^{(n)}(0) = 0, \end{aligned} \tag{3}$$

whenever f *is a polynomial of degree no more than* $n + d$.

Proof. The integral

$$\frac{1}{2\pi i} \oint_C \frac{f(z)dz}{(z-1)^{d+1} z^{n+1}}, \tag{4}$$

where C is a large circle with radius R, is $O(R^{-1})$ and is therefore zero. Write (4) in partial fractions and evaluate using the Cauchy integral formula obtain (3).

Theorem 2. (Padé approximation formula). *Using the notation of the confluent divided difference theorem, the unique (d,n) Padé approximation to* $\exp(z)$ *is given by*

$$\frac{a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n}{b_0 + b_1 z + b_2 z^2 + \cdots + b_d z^d}.$$

Proof. Use the function $t \mapsto \exp_{n+d}(tz)$, where $\exp_{n+d}$ denotes the exponential series truncated at the $(tz)^{n+d}/(n+d)!$ term. To within $O(z^{n+d+1})$, the result is

$$\exp(z)(b_0 + b_1 z + b_2 z^2 + \cdots + b_d z^d) - (a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n)$$

For many one-step numerical methods, the stability function is a Padé approximation to exp. The numerical method is a-stable if the corresponding approximation is an A-function.

Definition 3. (A function). *A rational function R is an A-function if* $|R(z)| \le 1,$ *whenever* $\mathrm{Re}(z) \le 0.$

Theorem 4. $A(d,n)$ *Padé approximation to exp is an A-function if and only if* $2 \ge d - n \ge 0,$

Proof.

The case $n > d.$ Because $R = O(z^{n-d}),$ as $|z| \to \infty,$ it follows that $\lim_{z\to-\infty}|R(z)| = \infty.$

The case $n + 2 \ge d \ge n.$ Between the (d,n) and the $(d+1,n)$ Padé approximations we write $(d+1,n)$ for the approximation

$$\frac{N_{d+t,n}}{D_{d+t,n}} = \frac{(1-t)N_{dn} + tN_{d+1,n}}{(1-t)D_{dn} + tD_{d+1,n}},$$

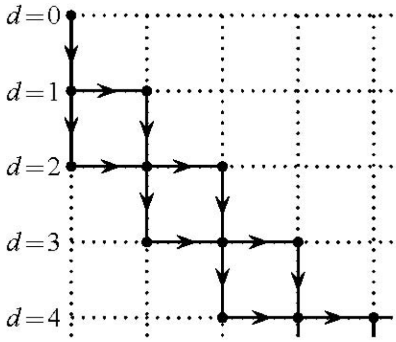
for $0 \le t \le 1,$ and similarly for the $(d,\ n+t)$ approximation. For convenience we will refer to one of these approximations as $N/D.$

For $(d+t)-n$ (or $d-(n+t)$ respectively) in [0,2], it can be verified that

$$|D(iy)|^2 - |N(iy)|^2 = C_0 y^{2d} + C_1 y^{2d+2},$$

where $C_0 \ge 0, C_1 \ge 0$ and, by a continuity argument, all zeros of D remain in the left half-plane as we carry out the homotopies indicates by arrows in the following diagram

n=0 *n*=1 *n*=2 *n*=3 *n*=4



The case $d > n + 2$ was formerly known as the Ehle conjecture. As a prelude to looking at this case, we will look at criteria for a stability function to be an A-function.

Theorem 5. (Basic criterion). *A rational function R(z) is an A-function if and only if (i) all poles are in the left half-plane and (ii)* $|R(iy)| \le 1$ *for all iy on the imaginary axis.*

Proof. Since R is analytic in the left half-plane, its maximum modulus occurs on the boundary. But the maximum modulus on the boundary does not exceed 1.

Now consider the behavior of the function $\tilde{R}$ given by $\tilde{R}(z) = \exp(-z)R(z).$ The functions R and $\tilde{R}$ have the same poles and, furthermore, $|R(iy)| = |\tilde{R}(iy)|.$ Hence the basic criterion applies equally to $\tilde{R}$ as to $R.$

The "relative stability function" $\tilde{R}$ was used as the basis for the theory of order stars, formulated by Hairer, Nørsett and Wanner [5]. It is also the starting point of the theory of order arrows [2], which we will discuss in his paper.

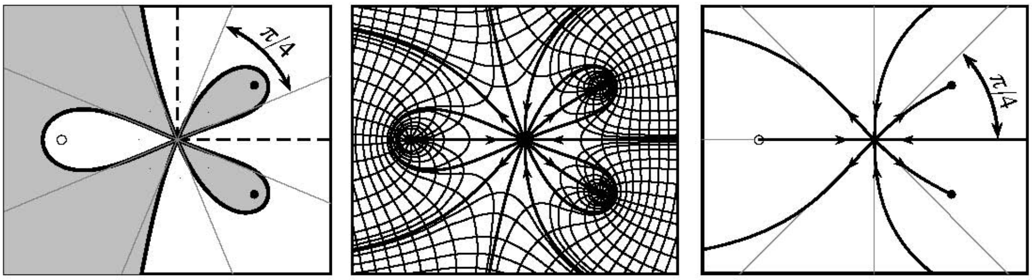
Definition 6 (Order stars). *The order star of* R *is the set of point in the complex plane for which* $|\tilde{R}(z)| > 1.$

Definition 7 (Order arrows). *The order arrows of* R *are the lines made up from points in the complex plane for which* $\tilde{R}(z)$ *is real and positive.*

We consider the example of the (2,1) Padé approximation $R(z)-\left(1+\frac{1}{3}z\right)\Big/\left(1-\frac{2}{3}z+\frac{1}{6}z^2\right)$. The following figures show, from left to right, (a) the order stars for this approximation, showing poles within bounded fingers and a zero within the bounded dual finger and showing the angles subtended by the fingers.

(b) Contour lines for constant argument, including the order arrow boundaries, superimposed on contour lines for constant magnitude, including the order star boundaries.

(c) Order arrows showing up-arrows terminating at poles or $-\infty$ and down-arrows terminating at a zero or at $+\infty$, and showing the angles subtended by the arrows.



(a) (b) (c)

We observe the following properties of order arrows for a (d, n) approximation:

. There are $p+1$ up-arrows emanating from zero alternating with the same number of down-arrows.

. The angle between each up-arrow and the next down-arrow is $\pi/(p+1.)$

. Up-arrows terminate at poles or at $-\infty$.

. Down-arrows terminate at zeros or at $+\infty$.

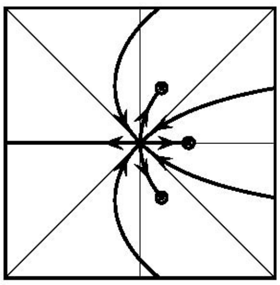
. Because up-arrows from zero cannot cross down-arrows from zero, every pole and every zero is at the end of an arrow.

Because adjacent up-arrows subtend an angle $2\pi/(p+1)$, and d of them terminate at poles, the total angle subtended is at least

$$\frac{2(d-1)}{p+1}\pi \geq \pi \text{ if } 2d-p>2.$$

Hence, either up-arrows terminating at poles are tangential to the imaginary axis or protrude into the left half-plane. In the latter case, there are poles in the left half-plane or an up-arrow crosses back across the imaginary axis.

We will illustrate this result in the [3,0] case.



We now turn our attention to generalized Padé approximation. We want to find an approximation to exp of the form

$$\Phi(w,z)=w^r P_0(z)+w^{r-1}P_1(z)+\cdots+P_r(z),$$

where the degrees of $P_0, P_1, \ldots, P_r$ have degrees $d_0, d_1, \ldots, d_r$, respectively, such that $\Phi(\exp(z), z) = O(z^{p+1})$, with p defined by $p = d_0 + d_1 + \cdots + d_r + r - 1$. We carry out steps in a similar way to the rational case $r = 1$. This method of construction was introduced in [1].

1. Evaluate the integral

$$\frac{1}{2\pi i}\oint_C \frac{f(z)dz}{(z-r)^{d_0+1}(z-r+1)^{d_1+1}\cdots z^{d_r+1}},$$

where f is a polynomial of degree no more than $p-1$.

2. The result is zero.
3. By using partial fractions obtain a version of the confluent divided difference formula.
4. Substitute $f(z) = \exp_p(tz)$.

An A-function in the multivalue case is a function Φ, given by

$$\Phi(w,z) = w^r P_0(z) + w^{r-1} P_1(z) + \cdots + P_r(z),$$

such that z in the left half-plane and w outside the unit disc do not exist for which $\Phi(w,z) = 0$.

A-functions are stability functions of A-stable numerical methods.

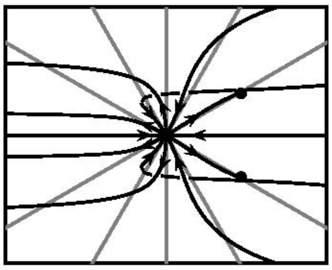
The natural generalization of the necessary and sufficient condition for a Padé approximation to be an A-function is that

1. $2d_0 - p \geq 0,$
2. $2d_0 - p \leq 2.$

The two conditions are not sufficient but each is necessary. The necessity of 1 was formerly known as the Daniel-Moore conjecture; it was proved using order stars but I will give the hint of an order arrow proof. The necessity of 2 is known as the Butcher-Chipman conjecture [4]; it has now been proved using methods discussed in this paper.

Theorem 8. $2d_0 - p \geq 0$ *for an A-approximation.*

We illustrate how this theorem is proved using the [2,0,0,0,0] approximation as an example, with $d_0 = 2$ and $p = 5$.



The thin grey lines are tangent to the arrows and are spaced at angles of $\pi/(p+1) = \pi/6$. Hence there exist up-arrows tangent to the imaginary axis.

Theorem 9. $2d_0 - p \leq 2$ *for an A-approximation.*

The proof is similar to the proof of the Ehle "conjecture". In the Daniel-Moore result there were not enough poles to match the up-arrows from zero and some of them had to cross (or be tangential to) the imaginary axis. But, in the Butcher-Chipman result, there are *too many* poles at the end of up-arrows from zero and hence some poles must be in the left half-plane or else an arrow starting in a left-oriented direction must cross back over the imaginary axis to terminate at a pole in the right half-plane. The difficult part of the proof is to show that all poles are at the end of some arrows from zero. The full proof is given in [3].

References

1. J.C. Butcher. The A-Stability of methods with Padé and generalized Padé stability functions, Numer. Algorithms 31 (2002), 47-58.
2. J.C. Butcher. Numerical Methods for Ordinary Differential Equations, Second Edition, J.Wiley, Chichester, (2008).
3. J.C. Butcher. Order and stability of generalized Padé approximations, Appl. Numer. Math. 59 (2009), 558-567.
4. J.C. Butcher and F.H. Chipman, Generalized Padé approximations to the exponential function, BIT 32 (1992), 118-130.
5. G.Wanner, E.Hairer and S.P.Nørsett, Order stars and stability theorems, BIT 18 (1978), 475-489.