

# A Brief Introduction to Order Stars

张陈成 519071910019

## Introduction

In this assignment, we aim to introduce a technique called *order stars*, which provides us with a better understanding of stability theories. Padé approximations and multiple real pole approximation is mainly discussed in this article.

## Stability Functions

### Introduction of stability function

The idea of stability functions is initially raised by Dahlquist in 1963. One simple example is presented in the application of implicit Euler method to Dahlquist's equation  $y' = \lambda y$ . In detail, the local linearisation on  $y = y_0$  with stepsize  $h$  gives the linear equation  $y_1 = y_0 + h\lambda y_1$ . We notice that  $y_1 = R(h\lambda)y_0$  converges for each  $y_0$  whence  $|R(z)| = \frac{1}{|1-z|} < 1$ .

### Stability function of Runge-Kutta method

The  $r$ -stage RK-method is given in form of

$$K_j = f(t_n + c_j h, y_n + h \sum_{i=1}^r a_{ij} K_i)$$
$$y_{n+1} = y_n + h \sum_{j=1}^r b_j K_j.$$

For  $y' = \lambda y$ , it yields that  $K = zAK + \lambda y_n$ , where

$$K = (K_1, K_2, \dots, K_r) \in \mathbb{R}^r.$$

Hence

$$\begin{pmatrix} I - zA & 0 \\ -zb^T & 1 \end{pmatrix} \cdot \begin{pmatrix} K/\lambda \\ y_{n+1} \end{pmatrix} = y_n \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

By taking the inverse, we obtain

$$\begin{pmatrix} K/\lambda \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} (I - zA)^{-1} & 0 \\ zb^T(I - zA)^{-1} & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot y_n.$$

We deduce that  $y_{n+1} = (zb^T(I - zA)^{-1}\mathbf{1} + 1)y_n$ . Therefore

$$R(z) = (zb^T(I - zA)^{-1}\mathbf{1} + 1) = \frac{\det(z\mathbf{1}b^T + I - zA)}{\det(I - zA)}.$$

Consider the exact solution  $y_1 = e^z \cdot y_0$ , we say the RK method is of order  $p$  whence

$$e^z - R(z) = Cz^{p+1} + O(|z|^{p+2}), \quad |z| \ll 1, C \neq 0.$$

## Absolute stability

The stability domain is defined as  $\mathcal{D} := \{z : |R(z)| < 1\}$ . For arbitrary  $y_0$  and each  $z \in \mathcal{D}$  in a Dahlquist equation, it yields that  $y_n = (R(z))^n y_0 \rightarrow 0$  as  $n \rightarrow \infty$ .

For ODEs, a method is absolute stable (A-stable in short) whence  $S \supset \mathbb{C}^- = \{z : \operatorname{Re}(z) < 0\}$ . A RK method with  $R(z) = \frac{P(z)}{Q(z)}$  is A-stable whenever the following conditions holds

- $|R(iy)| \leq 1$  for each  $y \in \mathbb{R}$ .
- $R(z)$  is analytic on  $\mathbb{C}^-$ .

## Padé Approximations

In comparison of the numerical estimate  $y_1 = R(z)y_0$  and the explicit solution  $y_1 = e^z y_0$ , we say the RK method is of order  $s$  whence  $e^z Q(z) = P(z) + C_1 z^{s+1} + C_2 z^{s+2} + \dots$ . Here  $P(z)$  and  $Q(z)$  are polynomials with degree  $\leq s$ ,  $P(0) = Q(0) = 1$ . It is clear that  $C_i$ 's are uniquely determined by  $Q(z)$ .

For  $Q(z) = \sum_{i=0}^s q_i z^i$  ( $q_0 = 1$ ), taking Taylor expansion yields that

$$\begin{aligned} P(z) &= \sum_{k=0}^s \left( \sum_{i=0}^k \frac{q_i}{(k-i)!} \right) z^k \\ &= q_0 + \left( \frac{q_1}{0!} + \frac{q_0}{1!} \right) z + \dots \\ &\quad + \left( \frac{q_0}{s!} + \frac{q_1}{(s-1)!} + \dots + \frac{q_s}{1!} \right) z^s. \end{aligned}$$

We also deduce the error constants  $C_j = \sum_{i=0}^s \frac{q_i}{(s+j-i)!}$ . In order to symmetrise  $P(z)$  and  $Q(z)$ , we introduce the  $M$ -polynomial

$$M(x) := \sum_{k=0}^s \frac{x^k}{k!} q_{s-k}.$$

Hence  $Q(z) = \sum_{i=0}^s M^{(s)}(0) z^{s-i}$ ,  $P(z) = \sum_{i=0}^s M^{(s)}(1) z^{s-i}$ . The error constants are given by integrals, i.e.,  $C_1 = \int_0^1 M(x) dx$ ,  $C_2 = \int_0^1 (1-x)M(x) dx$ .

Nørsett and Wanner (1979) shows  $M(x) = u(x)$  for collocation method based on points  $\{c_i\}_{i=1}^s$ , provided that

$$u(x) = (s!)^{-1} \prod_{i=1}^s (x - c_i).$$

The main idea of Padé's work is to estimates  $e^z$  with rational function  $\frac{P(z)}{Q(z)}$  to the greatest extent. We shall summerise his construction as follows:

One notices that  $M(x) = \frac{x^k(x-1)^j}{(k+j)!}$  satisfies  $M^{(i)}(0) = 0$  for each  $i = 0, 1, \dots, k-1$ ,  $M^{(i)}(1) = 0$  for each  $i = 0, 1, \dots, j-1$ .

Hence

$$P_{kj} = \sum_{i=0}^k \frac{\binom{k}{i}}{\binom{j+k}{i}} z^i,$$

$$Q_{kj} = \sum_{i=0}^j \frac{\binom{j}{i}}{\binom{j+k}{i}} (-z)^i = P_{jk}(-z).$$

Let  $R_{kj}(z) = \frac{P_{kj}(z)}{Q_{kj}(z)}$ . Then

$$e^z - R_{kj}(z) = (-1)^j \frac{j!k!}{(j+k)!(j+k+1)!} z^{j+k+1} + O(z^{j+k+2}).$$

The uniqueness is due to  $M^{(k)}(0) = M^{(j)}(1) = 0$ .

Here  $M(x)$  is called  $C$ -polynomial, which we shall discuss in the following sections.

## Order Stars

### Ideas of order stars

The idea of ordered stars is led to better understand whether  $R_{jk}$  fails to satisfy A-stability. It is known that  $R_{11}$ , the midpoint method, is A-stable. Whereas,

$$R_{15}(z) = \frac{1 + \frac{1}{6}z}{1 - \frac{5}{6}z + \frac{1}{3}z^2 - \frac{1}{12}z^3 + \frac{1}{72}z^4 - \frac{1}{720}z^5}$$

violates A-stability, in fact,  $|R_{15}(i)| > 1$ . In general,  $R(z)$  is A-stable whence

- $|R(it)| \leq 1$  for each  $t \in \mathbb{R}$ .
- All poles of  $R(z)$  lie in  $\mathbb{C}^+$ .

The command `approx=PadeApproximant[E^z,{z,0,{1,2}}]` in mathematica can be utilised to generate  $E_{12}$ .

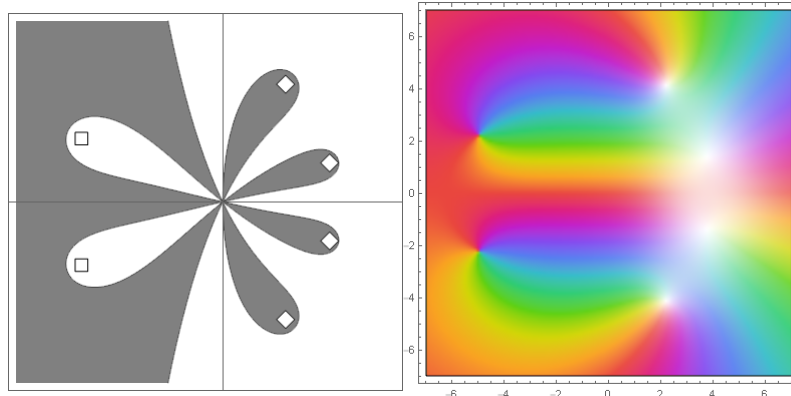
The definition of order star is given by

$$A = \{z \in \mathbb{C} : |q(z)| > 1\}.$$

Here  $q(z) := \frac{R(z)}{e^z}$  is called the order star of  $R$ . It is clear that the order star is symmetric w.r.t. real axis. As a immediate corollary,  $R(z)$  is A-stable whence  $A \cap i\mathbb{R} = \emptyset$  and all poles of  $q(z)$  lies in  $\mathbb{C}^+$ .

### Star Structure

The star structure of  $R_{15}$  can be visualised by the following codes (in mathematica)



```
Needs["FunctionApproximations`"];
approx = PadeApproximant[E^z, {z, 0, {2, 4}}];
orderStarPlot[approx, E^z]
```

```
ComplexPlot[approx, {z, -7 - 7 I, 7 + 7 I}]
```

The `ComplexPlot` command fails to plots the star structure of  $A$ , due to the lack of accuracy near the origin.

The star structure has the following behaviours.

### The even partition near the origin

One notices that the order stars gives the even partition in the vicinity of the origin. This is due to the error formula

$$q(z) - 1 = Cz^{s+1} + O(z^{s+2})$$

induces an  $s + 1$ -covering map near the origin.

### The asymptomatic behaviour as $r \rightarrow \infty$

One notices that

- for fixed  $\theta \in (-\pi/2, \pi/2)$ , there exists  $r(\theta) > 0$  such that  $re^{i\theta} \notin A$  as  $r > r(\theta)$ .
- for fixed  $\theta \in (\pi/2, 3\pi/2)$ , there exists  $r(\theta) > 0$  such that  $re^{i\theta} \in A$  as  $r > r(\theta)$ .

This is due to  $|e^z|$  is much stronger than each given polynomial as  $\operatorname{Re}(z) > 0$ ; much weaker if  $\operatorname{Re}(z) < 0$ .

Moreover, the border  $\partial A$  possesses only two branches which go to infinity. If  $R(z) = Kz^l + O(z^{l-1})$  for  $z \rightarrow \infty$ , then the branches asymptotically approach  $x = \log |K| + l \log |y|$ .

**Proof.** Let

$$\begin{aligned}\varphi_1(\theta) &= |e^z|^2 = e^{2r \cos \theta}, \\ \varphi_2(\theta) &= |R(z)|^2 = R(re^{i\theta})R(re^{-i\theta}).\end{aligned}$$

Then  $(\log \varphi_1)_\theta = -2r \sin \theta$ ,  $(\log \varphi_2)_\theta = 2r \operatorname{Re} \left( ie^{i\theta} \frac{R'(re^{i\theta})}{R(re^{i\theta})} \right)$ .

Since  $\frac{R'(z)}{R(z)} \rightarrow 0$  as  $z \rightarrow \infty$ , there exists  $\varepsilon > 0$  such that

$$(\log \varphi_1)_\theta < (\log \varphi_2)_\theta, \quad \theta \in [\varepsilon, \pi - \varepsilon].$$

There exists exact one  $\theta$  such that  $\varphi_1(\theta) = \varphi_2(\theta)$ , thus  $\theta$  can be regarded as a function of  $r$  (as well as some fixed  $\varepsilon$  for each  $r$ ). As a consequence, the intersection of  $|R(z)|$  and  $|e^z|$  lying in

$$\bigcup_{r \in \mathbb{R}} \left( r \times \{[\varepsilon, \pi - \varepsilon]\} \right) \subset \mathbb{H}$$

is the curve  $\theta(r)$ . Since  $|K| \cdot r^l = e^x + O(r^{l-1})$ , we deduce that

$$\log |K| + l \log |y|(1 + E_1(z)) = x \cdot (1 + E_2(z)).$$

Here

- $E_1(z) \rightarrow 0$  as  $r \rightarrow \infty$ , since  $\frac{y}{x} \rightarrow 0$  as  $r \rightarrow \infty$ .
- $\lim_{r \rightarrow \infty} p(r) = E_2(r) \rightarrow 0$  for each polynomial  $p(r)$ .

Thus we deduce the desired approximation

$$x = \log |K| + l \log |y|.$$

### The poles in each white fingers

We claim that each bounded white sector at the origin contains at least a pole. Moreover, the number of exponential fitting points  $\#\{z \in \partial A : R(z) = e^z\}$  equals the the number of poles (counting multiplicities).

**Proof.** For a sector  $\Omega$  with  $\partial\Omega \subset \partial A$ . Let  $\gamma : [s_0, s_1] \rightarrow \mathbb{C}$  be a parametrisation of  $\partial\Omega$  such that  $|\gamma'| \equiv 1$ . Then  $\vec{s} = (x(s)', y(s)')$  is a tangent vector along  $\partial\Omega$ ,  $\vec{n} = (y(s)', -x(s)')$  is a (exterior) normal vector. Write  $q(z) = q(x + iy)$  as  $r(x, y) \cdot e^{i\varphi(x, y)}$ .

First,  $\nabla_{\vec{n}}(\log r) \leq 0$  since  $\log r$  increases in  $\Omega$ . The Cauchy-Riemann equation gives

$$\nabla_{\vec{t}}(\varphi) = \nabla_x \varphi + \nabla_y \varphi = -\nabla_y r + \nabla_x r = \nabla_{\vec{n}}(\varphi) \leq 0.$$

Consider  $\vec{t} \cdot q'(\gamma(s)) = i \cdot q(\gamma(s)) \cdot \nabla_{\vec{t}}(\varphi)$ . It reveals that  $\nabla_{\vec{t}}(\varphi) \neq 0$  except finite many points (that is, exponential fitting points). Thus the vector  $q(z)$  along the contour curve  $\gamma(t)$  revolution at least  $2\pi$  every when finishing a loop in  $\partial\Omega$ . Since the total change of arguments are at least  $m \cdot 2\pi$ , the the  $m$  sectors contains at least  $m$  poles (counting multiplicities).

We claim that the number of exponential fitting points  $\#\{z \in \partial A : R(z) = e^z\}$  equals the the number of poles (counting multiplicities). Indeed, the number of exponential fitting points are as many as the loops in  $\partial\Omega$ , due to  $\nabla_{\vec{t}}(\varphi)$  is monotonous.

### Order stars for rational approximations

Let  $R(z)$  be a rational approximation of order  $s$ , with  $m_0$  zeros and  $m_\infty$  poles. Then we have the following theorems:

**Theorem 1.** When  $R(z)$  is A-stable, then the number of different zeros in  $\mathbb{C}^-$  is no less than  $\frac{s-2}{2}$ .

**Proof.** It reveals that  $A$  has at least  $k_1$  white sectors in  $\mathbb{C}^-$ . Otherwise, some of the arguments of the poles exceed  $[-\pi/2, \pi/2]$ .

**Theorem 2.** Suppose  $s \geq 2m_\infty - 1$ ,  $|R(\infty)| \leq 1$ , then  $R(z)$  is A-stable.

**Proof.**  $|R(\infty)| \leq 1$  implies  $m_0 \leq m_\infty$ . One notices that at least  $\lfloor (s+1)/2 \rfloor$  sectors start in  $\mathbb{C}^+$  and cannot cross the imaginary axis, thus must be bounded. As a result, the number of poles in  $\mathbb{C}^+$  is no less than  $\lfloor (s+1)/2 \rfloor$ .

We claim that all poles are simple, since the second equality holds in the following inequality,

$$\lfloor (s+1)/2 \rfloor \leq \#\{\text{poles in } \mathbb{C}^+\} \leq m_\infty \leq \frac{s+1}{2}.$$

Therefore, all poles are in  $\mathbb{C}^+$ . It (with  $|R(\infty)| \leq 1$ ) yields that  $R(z)$  is A-stable.

**Theorem 3.** Suppose  $s \geq 2m_\infty - 3$ ,  $|R(\infty)| \leq 1$ , and the coefficients of  $Q(z)$  have alternating signs, then  $R(z)$  is A-stable.

**Proof.** Similarly, we deduce that

$$\lfloor (s+1)/2 \rfloor \leq \#\{\text{poles in } \mathbb{C}^+\} \leq m_\infty \leq \frac{s+3}{2}$$

Then we have either 0 or 1 pole in  $\mathbb{C}^-$ . Assume that the latter holds, the symmetry implies that the pole must be on  $\mathbb{R}^-$ . One notice that  $Q(x) \neq 0$  for  $x \in \mathbb{R}^-$ , which contradicts our assumption.

**Theorem 4.** Let  $m'_0$  and  $m'_\infty$  denotes the number of distinct zeros and poles. Then  $m'_0 + m'_\infty \geq s$ .

**Proof.** One shall regard the order star as a CW-complex in  $\overline{\mathbb{C}}$  (the  $\infty$ -compactification of  $\mathbb{R}$ ). Via

$$\begin{aligned} 2 &= \chi(\overline{\mathbb{C}}) = \sum_{\ell=0}^2 (-1)^\ell \cdot s_\ell \\ &\leq 1 - (p+1) + (m'_0 + m'_\infty + 2), \end{aligned}$$

one obtains  $m'_0 + m'_\infty \geq s$ .

Here  $s_1 \geq p+1$ ,  $s_2 \leq m'_0 + m'_\infty + 2$ .

**Theorem 5.** (Theorem and Conjecture of Ehle)  $R_{kj}$  is A-stable whence  $k \leq j \leq k+2$ . All poles and zeros are simple.

**Proof.**  $\Leftarrow$ :  $|R(\infty)| \leq i$  equals  $j \geq k$ . When all poles are in  $\mathbb{C}^+$ ,  $j \leq k+2$ .

$\Rightarrow$ : If  $j \leq k+2$ , then  $s \geq 2j-2$ , thus  $j \leq k+2$ . Moreover,  $|R(\infty)| \leq i$  equals  $j \geq k$ .

The inequality

$$m'_0 + m'_\infty \geq s = j + k$$

reveals all poles and zeros are simple.

## A-Stability of restricted Padé approximations.

### Intorduction to Laguerre polynomials

The Lagguerre polynomial is utilised for *Gaussian quadrature* to numerically compute integrals of the form

$$\int_0^\infty f(x) e^{-x} dx.$$

Since  $\left(\frac{d}{dx} - 1\right)e^x = 0$ , we can conclude that

$$L_n(x) = \frac{1}{n!} \left(\frac{d}{dx} - 1\right)^n x^n.$$

The inner product is defined as

$$\langle f, g \rangle := \int_0^\infty f(x) g(x) e^{-x} dx.$$

For alrge values, we have Hilb's asymptotic formula

$$L_k(x) = \frac{e^{x/2}}{\sqrt{\pi\sqrt{xk}}} \cdot (\cos[2\sqrt{xk} - \pi/4] + \sqrt{xk} \cdot O(1)).$$

# Multiple real pole approximation

## The motivation

The stability function of real pole approximation has real poles only, i.e., all poles of  $Q(z)$  are on  $\mathbb{R}$ . By symmertry, each white finger in  $A$  is enclosed by a black finger in  $\mathbb{C} - A$ . As a result, the remaining (finite)  $s - 1$  sectors contains at least  $s - 1$  zeros.

We conclude that the order of real pole approximation is at most  $\deg P + 1$ .

In order to estimate  $R(A)$  more efficiently, we hope  $Q(A) = (I - \gamma A)^p$ . A multiple real pole approximation of  $e^{-x}$  is of the form

$$R_p^m(z) = \frac{\sum_{i=0}^m a_i x^i}{(1 + zx)^p}.$$

In fact, there exists  $m$  different choice of  $z$  such that the resulting order is  $m + 1$  (optimal order).

## C polynomial

The  $M(x)$  defined in the previous section is given by

$$M(x) = \frac{x^k(x-1)^j}{(k+j)!}.$$

As a consequence, we deduce that

$$R(x) = \frac{P(x)}{Q(x)} = \frac{\sum_{i=0}^s M^{(s)}(1)x^{s-i}}{\sum_{i=0}^s M^{(s)}(0)x^{s-i}}.$$

As for multiple real pole approximation, we have

$$S_p^m(x) = \frac{\sum_{j=0}^m M_p^{(p-j)}(z; 0)x^j}{\sum_{j=0}^m M_p^{(p-j)}(z; 1)x^j}.$$

Here  $S_p^m(x)$  denotes some  $R_p^m(x)$  with order at least  $m$ . Since

$$M_p^{(p-j)}(z; 1) = (-1)^{p+j} z^j L_p^{(p-j)}(0) = z^j \binom{p}{j},$$

we obtian that  $\sum_{j=0}^m M_p^{(p-j)}(z; 1)x^j = (1 + zx)^p$ .

As a consequence,

$$\begin{aligned} S_p^m(x) &= \frac{\sum_{j=0}^m (-1)^{p+j} L_p^{(p-j)}(z^{-1})(zx)^j}{(1 + zx)^p} \\ &\stackrel{\gamma=-z}{=} \frac{\sum_{j=0}^m (-1)^j L_p^{(p-j)}(\gamma^{-1})(\gamma x)^j}{(1 - \gamma x)^p}. \end{aligned}$$

The error constant is  $C = (-1)^{j+1} \frac{\gamma^k}{k+1} L'_{k+1}(\gamma^{-1})$ . As a result, the optimal order attains whenever  $L'_{k+1}(\gamma^{-1})' = 0$ . We write  $\gamma_1^{-1} < \gamma_2^{-1} < \dots < \gamma_k^{-1}$ .

For such  $\gamma_\nu$ 's, the order star for restricted Padé approximation contains one bounded finger with multiplicity  $k - \nu + 1$ .

As a corollary, the restrict Padé approximation with optimal order is A-stable if

$$\lfloor k/2 \rfloor \leq \nu \leq \lfloor (k+1)/2 \rfloor.$$

**Proof.** There are  $k - \nu + 1$  fingers begins in  $\mathbb{C}^+$ ,  $\nu - 1$  fingers begin in  $\mathbb{C}_-$ . We claim at most  $k + 2$  fingers starts in  $\mathbb{C}^+$  or in  $\mathbb{C}_-$ . If not, then some fingers must cross the imaginary axis, which contradicts the A-stability.

## The optimal order of A-stable restrict Padé approximation

The restrict Padé approximation is A-stable whenever

$$L_k(\gamma_\nu^{-1}) = \lim_{z \rightarrow \infty} |R(z)| \leq 1.$$

By Hilb's asymptotic formula, we obtains the  $\nu$ -th extremum point of  $L_{k+1}$  is approximately

$$x_\nu = \frac{\pi^2(\nu + 1/4)^2}{4(k+1)}.$$

By  $|L(x_\nu)| > 1$ , we obtains the following proposition

$$k \geq \begin{cases} 1 & \nu = 1, \\ 5 & \nu = 2, \\ 9 & \nu = 3, \\ 6\nu - 10 & \nu \geq 4. \end{cases}$$

Indeed, it is verifacated by Norsett that the restrict Padé approximation with optimal order is A-stable whenever

$k =$	$\nu =$
1	1
2	1
3	1
5	2

## References

[Nor74] Syvert P. Nørsett. One-step methods of hermite type for numerical integration of stiff systems. *BIT Numerical Mathematics*, 14(1): 63–77, Mar 1974.

[Sof96] M. Sofroniou. Order stars and linear stability theory. *Journal of Symbolic Computation*, 21(1): 101–131, 1996.

[Wan78] G. Wanner, E. Hairer, and S. P. Norsett. Order stars and stability theorems. *BIT Numerical Mathematics*, 18(4): 475–489, Dec 1978.