The stability of Padé and generalized Padé approximations

John Butcher

The University of Auckland
New Zealand

International Conference
Combinatorics and Control
Miraflores de la Sierra

24 June 2010
1 Motivation

- Stability of Runge–Kutta methods
- Padé approximations as stability functions
- A-stable methods
1 Motivation
- Stability of Runge–Kutta methods
- Padé approximations as stability functions
- A-stable methods

2 Introduction to Padé approximations
- Padé approximations
- Applications to stability of numerical methods
- Formula for \((d, n)\) approximation
- Example

3 Stability of numerical methods
- A-functions
- Order stars and order arrows
- Order arrow proof for \(d > n + 2\)
1 Motivation
   - Stability of Runge–Kutta methods
   - Padé approximations as stability functions
   - A-stable methods

2 Introduction to Padé approximations
   - Padé approximations
   - Applications to stability of numerical methods
   - Formula for \((d, n)\) approximation
   - Example

3 Stability of numerical methods
   - A-functions
   - Order stars and order arrows
   - Order arrow proof for \(d > n + 2\)

4 Generalized Padé approximations

5 Developments of order arrows
Motivation

Stability of Runge–Kutta methods
Padé approximations as stability functions
A-stable methods
Stability of Runge–Kutta methods

When solving stiff differential equations by a Runge–Kutta method we need to pay attention to stability of the method. For a problem
\[ y'(x) = f(y(x)), \]
the sensitivity of numerical approximations to small perturbations, depends can be modelled by the variational equation
\[ Y'(x) = \frac{\partial f}{\partial y} Y \]
and the behaviour of this equation depends on the eigenvalues of the Jacobian matrix \( \partial f / \partial y \).
The simplest case is when there is a constant, and possibly complex, eigenvalue \( q \).
Let \( z = hq \), where \( h \) is the stepsize.
Stability of Runge–Kutta methods

When solving stiff differential equations by a Runge–Kutta method we need to pay attention to stability of the method. For a problem

\[ y'(x) = f(y(x)), \]

the sensitivity of numerical approximations to small perturbations, depends can be modelled by the variational equation

\[ Y'(x) = \frac{\partial f}{\partial y} Y \]

and the behaviour of this equation depends on the eigenvalues of the Jacobian matrix \( \frac{\partial f}{\partial y} \).

The simplest case is when there is a constant, and possibly complex, eigenvalue \( q \).

Let \( z = hq \), where \( h \) is the stepsize.
Stability of Runge–Kutta methods

When solving stiff differential equations by a Runge–Kutta method we need to pay attention to stability of the method. For a problem

$$y'(x) = f(y(x)),$$

the sensitivity of numerical approximations to small perturbations, depends can be modelled by the variational equation

$$Y'(x) = \frac{\partial f}{\partial y} Y$$

and the behaviour of this equation depends on the eigenvalues of the Jacobian matrix $\partial f / \partial y$.

The simplest case is when there is a constant, and possibly complex, eigenvalue $q$.

Let $z = hq$, where $h$ is the stepsize.
Stability of Runge–Kutta methods

When solving stiff differential equations by a Runge–Kutta method we need to pay attention to stability of the method. For a problem

\[ y'(x) = f(y(x)), \]

the sensitivity of numerical approximations to small perturbations, depends can be modelled by the variational equation

\[ Y'(x) = \frac{\partial f}{\partial y} Y \]

and the behaviour of this equation depends on the eigenvalues of the Jacobian matrix \( \partial f / \partial y \).

The simplest case is when there is a constant, and possibly complex, eigenvalue \( q \).

Let \( z = hq \), where \( h \) is the stepsize.
For a Runge–Kutta method with stages and output defined by

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_s \\ y_1 \end{bmatrix} = y_0 \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + h \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} \end{bmatrix} \begin{bmatrix} f(Y_1) \\ f(Y_2) \\ \vdots \\ f(Y_s) \end{bmatrix}$$

the output is defined, in the case $hf(Y) = zY$ by $y_1 = R(z)y_0$, where

$$R(z) = 1 + zb^T(I - zA)^{-1}1$$

The “Stability Region” is the set of points in the complex plane for which

$$|R(z)| \leq 1$$

We will now look at three examples.
For a Runge–Kutta method with stages and output defined by

$$
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_s \\
y_1
\end{bmatrix}
= y_0
\begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix}
+ h
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1s} \\
a_{21} & a_{22} & \cdots & a_{2s} \\
\vdots & \vdots & \ddots & \vdots \\
a_{s1} & a_{s2} & \cdots & a_{ss}
\end{bmatrix}
\begin{bmatrix}
f(Y_1) \\
f(Y_2) \\
\vdots \\
f(Y_s)
\end{bmatrix}
$$

the output is defined, in the case $hf(Y) = zY$ by $y_1 = R(z)y_0$, where

$$
R(z) = 1 + zb^T(I - zA)^{-1}1
$$

The “Stability Region” is the set of points in the complex plane for which

$$
|R(z)| \leq 1
$$

We will now look at three examples
For a Runge–Kutta method with stages and output defined by

$$
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_s \\
y_1
\end{bmatrix} = y_0 
\begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix} + h
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1s} \\
a_{21} & a_{22} & \cdots & a_{2s} \\
\vdots & \vdots & \ddots & \vdots \\
a_{s1} & a_{s2} & \cdots & a_{ss}
\end{bmatrix}
\begin{bmatrix}
f(Y_1) \\
f(Y_2) \\
\vdots \\
f(Y_s)
\end{bmatrix}
$$

the output is defined, in the case $hf(Y) = zY$ by $y_1 = R(z)y_0$, where

$$R(z) = 1 + zb^T(I - zA)^{-1}1$$

The “Stability Region” is the set of points in the complex plane for which

$$|R(z)| \leq 1$$

We will now look at three examples
Example 1: 4th order method of Kutta

\[
\begin{bmatrix}
A \\
b^T
\end{bmatrix} =
\begin{bmatrix}
0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}, \quad R(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4
\]

Example 2: 4th order Gauss method

\[
\begin{bmatrix}
A \\
b^T
\end{bmatrix} =
\begin{bmatrix}
\frac{1}{4} & \frac{1}{4} - \frac{1}{6}\sqrt{3} \\
\frac{1}{4} + \frac{1}{6}\sqrt{3} & \frac{1}{4}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2}
\]

Example 3: 3rd order Radau IIA method

\[
\begin{bmatrix}
A \\
b^T
\end{bmatrix} =
\begin{bmatrix}
\frac{5}{12} & \frac{1}{12} \\
\frac{3}{4} & \frac{1}{4}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}
\]
Example 1: 4th order method of Kutta

\[
\begin{bmatrix}
\frac{A}{b^T}
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 1 & 0 \\
\frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}
\end{bmatrix}, \quad R(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4
\]

Example 2: 4th order Gauss method

\[
\begin{bmatrix}
\frac{A}{b^T}
\end{bmatrix} = \begin{bmatrix}
\frac{1}{4} & \frac{1}{4} - \frac{1}{6}\sqrt{3} \\
\frac{1}{4} + \frac{1}{6}\sqrt{3} & \frac{1}{4} \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2}
\]

Example 3: 3rd order Radau IIA method

\[
\begin{bmatrix}
\frac{A}{b^T}
\end{bmatrix} = \begin{bmatrix}
\frac{5}{12} & -\frac{1}{12} \\
\frac{3}{4} & \frac{1}{4} \\
\frac{3}{4} & \frac{1}{4}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}
\]
Motivation
Stability of Runge–Kutta methods
Padé approximations as stability functions
A-stable methods

Example 1: 4th order method of Kutta

\[
\frac{A}{b^T} = \begin{bmatrix}
0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 1 & 0 \\
\frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}
\end{bmatrix}, \quad R(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4
\]

Example 2: 4th order Gauss method

\[
\frac{A}{b^T} = \begin{bmatrix}
\frac{1}{4} \\
\frac{1}{4} + \frac{1}{6}\sqrt{3} \\
\frac{1}{4} - \frac{1}{6}\sqrt{3} \\
\frac{1}{2}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2}
\]

Example 3: 3rd order Radau IIA method

\[
\frac{A}{b^T} = \begin{bmatrix}
\frac{5}{12} \\
\frac{3}{4} \\
\frac{1}{4}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}
\]
Example 1: 4th order method of Kutta

\[ \frac{A}{b^T} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad R(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4 \]

Example 2: 4th order Gauss method

\[ \frac{A}{b^T} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} - \frac{1}{6}\sqrt{3} \\ \frac{1}{4} + \frac{1}{6}\sqrt{3} & \frac{1}{4} \end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2} \]

Example 3: 3rd order Radau IIA method

\[ \frac{A}{b^T} = \begin{bmatrix} \frac{5}{12} & \frac{1}{12} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2} \]
Motivation

Stability of Runge–Kutta methods

Padé approximations as stability functions

A-stable methods

Example 1: 4th order method of Kutta

\[
\begin{bmatrix}
\frac{A}{b^T}
\end{bmatrix} = \begin{bmatrix}
0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 1 & 0 \\
\frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}
\end{bmatrix}, \quad R(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4
\]

Example 2: 4th order Gauss method

\[
\begin{bmatrix}
\frac{A}{b^T}
\end{bmatrix} = \begin{bmatrix}
\frac{1}{4} & \frac{1}{4} - \frac{1}{6}\sqrt{3} \\
1 + \frac{1}{6}\sqrt{3} & \frac{1}{4} \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2}
\]

Example 3: 3rd order Radau IIA method

\[
\begin{bmatrix}
\frac{A}{b^T}
\end{bmatrix} = \begin{bmatrix}
\frac{5}{12} & -\frac{1}{12} \\
\frac{3}{4} & \frac{1}{4} \\
\frac{3}{4} & \frac{1}{4}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}
\]
Example 1: 4th order method of Kutta

\[
\frac{A}{b^T} = \begin{bmatrix}
0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}, \quad R(z) = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4
\]

Example 2: 4th order Gauss method

\[
\frac{A}{b^T} = \begin{bmatrix}
\frac{1}{4} & \frac{1}{4} - \frac{1}{6}\sqrt{3} \\
\frac{1}{4} + \frac{1}{6}\sqrt{3} & \frac{1}{4} \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2}
\]

Example 3: 3rd order Radau IIA method

\[
\frac{A}{b^T} = \begin{bmatrix}
\frac{5}{12} & -\frac{1}{12} \\
\frac{3}{4} & \frac{1}{4} \\
\frac{3}{4} & \frac{1}{4}
\end{bmatrix}, \quad R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}
\]
The three stability functions, we have referred to, are Padé approximations to the exponential function.

Example 1: \(1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4\) is the \([0, 4]\) approximation.

Example 2: \(\frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2}\) is the \([2, 2]\) approximation.

Example 3: \(\frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}\) is the \([2, 1]\) approximation.
The three stability functions, we have referred to, are Padé approximations to the exponential function

Example 1: \( 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4 \) is the [0, 4] approximation

Example 2: \( \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2} \) is the [2, 2] approximation

Example 3: \( \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2} \) is the [2, 1] approximation
The three stability functions, we have referred to, are Padé approximations to the exponential function

Example 1: \( 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4 \) is the \([0, 4]\) approximation

Example 2: \( \frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2} \) is the \([2, 2]\) approximation

Example 3: \( \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2} \) is the \([2, 1]\) approximation
The three stability functions, we have referred to, are Padé approximations to the exponential function

Example 1: $1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4$ is the $[0, 4]$ approximation

Example 2: $\frac{1 + \frac{1}{2}z + \frac{1}{12}z^2}{1 - \frac{1}{2}z + \frac{1}{12}z^2}$ is the $[2, 2]$ approximation

Example 3: $\frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}$ is the $[2, 1]$ approximation
Motivation
Stability of Runge–Kutta methods
Padé approximations as stability functions
A-stable methods

Padé approximations as stability functions

Example 1
Motivation

Stability of Runge–Kutta methods
Padé approximations as stability functions
A-stable methods

Padé approximations as stability functions

Example 1

-3

0

Stable

Example 2

0

Stable
Padé approximations as stability functions

Example 1

Example 2

Example 3
If the stability region includes all complex numbers with non-positive real part, the method is “A-stable”.

Example 1, the [0, 4] Padé approximation, is not A-stable.

But the other two examples, which are the [2, 1] and the [2, 2] Padé approximations, are A-stable.
If the stability region includes all complex numbers with non-positive real part, the method is “A-stable”.

Example 1, the \([0, 4]\) Padé approximation, is not A-stable.

But the other two examples, which are the \([2, 1]\) and the \([2, 2]\) Padé approximations, are A-stable.
A-stable methods

If the stability region includes all complex numbers with non-positive real part, the method is “A-stable”.

Example 1, the [0, 4] Padé approximation, is not A-stable.

But the other two examples, which are the [2, 1] and the [2, 2] Padé approximations, are A-stable.
Introduction to Padé approximations

Applications to stability of numerical methods

Formula for $(d, n)$ approximation

Example
For a given function with a Taylor expansion about zero

\[ f(z) = c_0 + c_1 z + c_2 z^2 + \cdots \]

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 \( n \) and \( d \).
Padé approximations

For a given function with a Taylor expansion about zero

$$f(z) = c_0 + c_1 z + c_2 z^2 + \cdots$$

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 $n$ and $d$. 
### Padé approximations to \( \exp(z) \)

The first few entries in the Padé table for \( \exp \) are as follows:

<table>
<thead>
<tr>
<th>( d )</th>
<th>( n )</th>
<th>( 0 )</th>
<th>( 1 )</th>
<th>( 2 )</th>
<th>( 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>( 1+z )</td>
<td>( 1+z+\frac{1}{2}z^2 )</td>
<td>( 1+z+\frac{1}{2}z^2+\frac{1}{6}z^3 )</td>
</tr>
<tr>
<td>1</td>
<td>( \frac{1}{1-z} )</td>
<td>( \frac{1+\frac{1}{2}z}{1-\frac{1}{2}z} )</td>
<td>( \frac{1+\frac{2}{3}z+\frac{1}{6}z^2}{1-\frac{1}{3}z} )</td>
<td>( \frac{1+\frac{3}{4}z+\frac{1}{4}z^2+\frac{1}{24}z^3}{1-\frac{1}{4}z} )</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{1-z+\frac{1}{2}z^2} )</td>
<td>( \frac{1+\frac{1}{3}z}{1-\frac{2}{3}z+\frac{1}{6}z^2} )</td>
<td>( \frac{1+\frac{2}{3}z+\frac{1}{12}z^2}{1-\frac{2}{5}z+\frac{1}{20}z^2} )</td>
<td>( \frac{1+\frac{3}{5}z+\frac{3}{20}z^2+\frac{1}{60}z^3}{1-\frac{3}{5}z+\frac{1}{10}z^2+\frac{1}{120}z^3} )</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>( \frac{1}{1-z+\frac{1}{2}z^2-\frac{1}{6}z^3} )</td>
<td>( \frac{1+\frac{1}{4}z}{1-\frac{3}{4}z+\frac{1}{4}z^2-\frac{1}{24}z^3} )</td>
<td>( \frac{1+\frac{3}{4}z+\frac{3}{20}z^2-\frac{1}{60}z^3}{1-\frac{5}{2}z+\frac{1}{10}z^2-\frac{1}{120}z^3} )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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

\[ P_0(z)y_n + P_1(z)y_{n-1} + \cdots + P_k(z)y_{n-k} = 0, \quad (*) \]

where \( P_0, P_1, \ldots, P_k \) are polynomials.

In the case of “one-step” methods, such as Runge–Kutta methods, (*) reduces to

\[ P_0(z)y_n + P_1(z)y_{n-1} = 0 \quad (†) \]

and the stability of the difference equation is determined by the value of the rational function

\[ R(z) = -\frac{P_1(z)}{P_0(z)} \]

It is of considerable significance to know for which \( z \) values, the difference equation (†) has stable solutions.
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

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

where \( P_0, P_1, \ldots, P_k \) are polynomials.

In the case of “one-step” methods, such as Runge–Kutta methods, (*) reduces to

\[ P_0(z)y_n + P_1(z)y_{n-1} = 0 \tag{†} \]

and the stability of the difference equation is determined by the value of the rational function

\[ R(z) = -\frac{P_1(z)}{P_0(z)} \]

It is of considerable significance to know for which \( z \) values, the difference equation (†) has stable solutions.
Applications to stability of numerical methods

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

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

where \( P_0, P_1, \ldots, P_k \) are polynomials.

In the case of “one-step” methods, such as Runge–Kutta methods, (*) reduces to

\[ P_0(z)y_n + P_1(z)y_{n-1} = 0 \]  

and the stability of the difference equation is determined by the value of the rational function

\[ R(z) = -\frac{P_1(z)}{P_0(z)} \]

It is of considerable significance to know for which \( z \) values, the difference equation (†) has stable solutions.
Formula for \((d, n)\) approximation

**Theorem (Confluent divided differences)**

There exist non-zero coefficients \(a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_d\) such that

\[
b_0 f(1) + b_1 f'(1) + b_2 f''(1) + \cdots + b_d f^{(d)}(1)
- a_0 f(0) - a_1 f'(0) - a_2 f''(0) - \cdots - a_n f^{(n)}(0) = 0 \quad (*)
\]

whenever \(f\) is a polynomial of degree no more than \(n + d\).

**Proof.**

The integral

\[
\frac{1}{2\pi i} \oint \frac{f(z)dz}{(z-1)^{d+1}z^{n+1}}, \quad (\dagger)
\]

where \(C\) is a large circle with radius \(R\), is \(O(R^{-1})\) and is therefore zero. Write \((\dagger)\) in partial fractions and evaluate using the Cauchy integral formula to obtain \((*)\).
Formula for \((d, n)\) approximation

**Theorem (Confluent divided differences)**

There exist non-zero coefficients \(a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_d\) such that

\[
b_0f(1) + b_1f'(1) + b_2f''(1) + \cdots + b_df^{(d)}(1) \nonumber \\
- a_0f(0) - a_1f'(0) - a_2f''(0) - \cdots - a_nf^{(n)}(0) = 0 \quad (*)
\]

whenever \(f\) is a polynomial of degree no more than \(n + d\).

**Proof.**

The integral

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

where \(C\) is a large circle with radius \(R\), is \(O(R^{-1})\) and is therefore zero. Write \((†)\) in partial fractions and evaluate using the Cauchy integral formula to obtain \((*)\).
Theorem (Confluent divided differences)

There exist non-zero coefficients $a_0, a_1, \ldots, a_n, b_0, b_1, \ldots, b_d$ such that

$$b_0 f(1) + b_1 f'(1) + b_2 f''(1) + \cdots + b_d f^{(d)}(1)$$

$$- a_0 f(0) - a_1 f'(0) - a_2 f''(0) - \cdots - a_n f^{(n)}(0) = 0 \quad (*)$$

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

Proof.

The integral

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

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

\[
\exp(z)(b_0 + b_1 z + b_2 z^2 + \cdots + b_d z^d)
\]

\[
- (a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n)
\]
Theorem (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

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

\[
\exp(z)(b_0 + b_1 z + b_2 z^2 + \cdots + b_d z^d) - (a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n)
\]
In the case $n = 1, \ d = 2$, we find the contour integral $(\dagger)$ to be

\[
\frac{1}{2\pi i} \int_{C} \frac{f(z) \, dz}{(z - 1)^3 z^2} = \frac{1}{2\pi i} \int_{C} f(z) \left( \frac{3}{z - 1} - \frac{2}{(z - 1)^2} + \frac{1}{(z - 1)^3} - \frac{3}{z} - \frac{1}{z^2} \right) \, dz
\]

\[
= 3f(1) - 2f'(1) + \frac{1}{2}f''(1) - 3f(0) - f'(0)
\]

leading to the Padé approximation

\[
\frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}
\]
Stability of numerical methods
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 (A function)**

A rational function $R$ is an A-function if

$$|R(z)| \leq 1$$

whenever

$$\text{Re}(z) \leq 0.$$
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 (A function)**

A rational function \( R \) is an A-function if

\[
|R(z)| \leq 1
\]

whenever

\[
\text{Re}(z) \leq 0.
\]
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 (A function)**

A rational function \( R \) is an A-function if

\[
|R(z)| \leq 1
\]

whenever

\[
\text{Re}(z) \leq 0.
\]
Which Padé approximations are A-functions?

Theorem

A \((d, n)\) Padé approximation to \(\exp\) is an A-function if and only if

\[ 2 \geq d - n \geq 0. \]
Which Padé approximations are A-functions?

**Theorem**

A \((d, n)\) Padé approximation to \(\exp\) is an A-function if and only if

\[ 2 \geq d - n \geq 0. \]

<table>
<thead>
<tr>
<th>(d/n)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1+z</td>
<td>1+z+\frac{1}{2}z^2</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
</tr>
<tr>
<td>1</td>
<td>\frac{1}{1-z}</td>
<td>\frac{1+\frac{1}{2}z}{1-\frac{1}{2}z}</td>
<td>\frac{1+\frac{2}{3}z+\frac{1}{6}z^2}{1-\frac{1}{3}z}</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
</tr>
<tr>
<td>2</td>
<td>\frac{1}{1-z+\frac{1}{2}z^2}</td>
<td>\frac{1+\frac{1}{3}z}{1-\frac{2}{3}z+\frac{1}{6}z^2}</td>
<td>\frac{1+\frac{2}{3}z+\frac{1}{12}z^2}{1-\frac{1}{3}z+\frac{1}{12}z^2}</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
</tr>
<tr>
<td>3</td>
<td>\frac{1}{1-z+\frac{1}{2}z^2-\frac{1}{6}z^3}</td>
<td>\frac{1+\frac{1}{4}z}{1-\frac{3}{4}z+\frac{1}{3}z^2-\frac{1}{24}z^3}</td>
<td>\frac{1+\frac{2}{5}z+\frac{1}{20}z^2}{1-\frac{2}{5}z+\frac{1}{20}z^2}</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
</tr>
<tr>
<td>4</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
</tr>
<tr>
<td>5</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
<td>\cdots</td>
</tr>
</tbody>
</table>
The case $n > d$.

Because the numerator of $R$ has greater degree than the denominator,

$$\lim_{z \to -\infty} |R(z)| = \infty$$
The case \( n + 2 \geq d \geq n \).

Between the \((d, n)\) and the \((d + 1, n)\) Padé approximations we write \((d + t, n)\) for the approximation

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

and similarly for the \((d, n + t)\) approximation.

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 \(C_0 \geq 0, C_1 \geq 0\) and, by a continuity argument, all zeros of \(D\) remain in the right half-plane as we carry out the homotopies shown in the next slide.
The case \( n + 2 \geq d \geq n \).

Between the \((d, n)\) and the \((d + 1, n)\) Padé approximations we write \((d + t, n)\) for the approximation

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

and similarly for the \((d, n + t)\) approximation.

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 \(C_0 \geq 0, C_1 \geq 0\) and, by a continuity argument, all zeros of \(D\) remain in the right half-plane as we carry out the homotopies shown in the next slide.
The case \( n + 2 \geq d \geq n \).

Between the \((d, n)\) and the \((d + 1, n)\) Padé approximations we write \((d + t, n)\) for the approximation

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

and similarly for the \((d, n + t)\) approximation.

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 \(C_0 \geq 0\), \(C_1 \geq 0\) and, by a continuity argument, all zeros of \(D\) remain in the right half-plane as we carry out the homotopies shown in the next slide.
Introduction to Padé approximations
Stability of numerical methods

A-functions
Order stars and order arrows
Order arrow proof for $d > n + 2$
The case $d > n + 2$ was formerly known as the Ehle conjecture.

As a prelude to looking at this case, look at criteria for a stability function being an A-function.

**Theorem (Basic criterion)**

A rational function $R(z)$ is an A-function if and only if

- All poles are in the right half-plane
- $|R(iy)| \leq 1$ 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.
The case $d > n + 2$ was formerly known as the Ehle conjecture.

As a prelude to looking at this case, look at criteria for a stability function being an A-function.

**Theorem (Basic criterion)**

A rational function $R(z)$ is an A-function if and only if

- All poles are in the right half-plane
- $|R(iy)| \leq 1$ 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.
The case \( d > n + 2 \) was formerly known as the Ehle conjecture.

As a prelude to looking at this case, look at criteria for a stability function being an A-function.

**Theorem (Basic criterion)**

A rational function \( R(z) \) is an A-function if and only if

- *All poles are in the right half-plane*
- \( |R(iy)| \leq 1 \) 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.
The case $d > n + 2$ was formerly known as the Ehle conjecture.

As a prelude to looking at this case, look at criteria for a stability function being an A-function.

**Theorem (Basic criterion)**

A rational function $R(z)$ is an A-function if and only if

- All poles are in the right half-plane
- $|R(iy)| \leq 1$ 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.
The case $d > n + 2$ was formerly known as the Ehle conjecture.

As a prelude to looking at this case, look at criteria for a stability function being an A-function.

**Theorem (Basic criterion)**

A rational function $R(z)$ is an A-function if and only if

- All poles are in the right half-plane
- $|R(iy)| \leq 1$ 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 behaviour 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.

It is also the starting point of the theory of order arrows which we will now discuss.
Now consider the behaviour 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.

It is also the starting point of the theory of order arrows which we will now discuss.
Now consider the behaviour 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.

It is also the starting point of the theory of order arrows which we will now discuss.
Order stars and order arrows

Now consider the behaviour 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.

It is also the starting point of the theory of order arrows which we will now discuss.
Now consider the behaviour 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.

It is also the starting point of the theory of order arrows which we will now discuss.
Definition (Order stars)

The order star of $R$ is the set of points in the complex plane for which $|\tilde{R}(z)| > 1$.

Definition (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 for which

$$R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}.$$ 

The figures on the next slide shows the relationship between order star and order arrows:
Definition (Order stars)
The order star of $R$ is the set of points in the complex plane for which $|\tilde{R}(z)| > 1$.

Definition (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 for which

$$R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}.$$ 

The figures on the next slide shows the relationship between order star and order arrows:
Definition (Order stars)
The order star of $R$ is the set of points in the complex plane for which $|\tilde{R}(z)| > 1$.

Definition (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 for which

$$R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}.$$ 

The figures on the next slide shows the relationship between order star and order arrows:
Definition (Order stars)
The order star of \( R \) is the set of points in the complex plane for which \( |\tilde{R}(z)| > 1 \).

Definition (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 for which
\[
R(z) = \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2}.
\]

The figures on the next slide shows the relationship between order star and order arrows:
\[ w(z) = \exp(-z) \frac{1 + \frac{1}{3}z}{1 - \frac{2}{3}z + \frac{1}{6}z^2} \]

\[ |w(z)| > 1 : \]

\[ |w(z)| < 1 : \]
\[
\frac{\exp(-z)(1+\frac{1}{3}z)}{1-\frac{2}{3}z+\frac{1}{6}z^2} = 1 - \frac{1}{72}z^4 + O(|z|^5), \quad |w(z)| = 1 - \frac{1}{72} r^4 \cos(4\theta) + O(r^5)
\]
Bounded fingers contain poles
Bounded dual fingers contain zeros
Add additional contour lines, in addition to $|w(z)| = 1$
Emphasise the contour lines for $|w(z)|$
Orthogonal to these are contour lines for $\arg(w(z))$
Retain only lines of constant argument
Fade out all lines except $w(z)$ real and positive
This leaves only the order arrows
As for order stars, behaviour at zero determines order
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.
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.
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.
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.
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

\[ \frac{2\pi}{p + 1} \]

and \( d \) of them terminate at poles, the total angle subtended is at least

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

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.
Because adjacent up-arrows subtend an angle
\[
\frac{2\pi}{p + 1}
\]
and \(d\) of them terminate at poles, the total angle subtended is at least
\[
\frac{2(d - 1)}{p + 1}\pi \geq \pi \quad \text{if} \quad 2d - p > 2.
\]
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.
Because adjacent up-arrows subtend an angle
\[ \frac{2\pi}{p + 1} \]
and \( d \) of them terminate at poles, the total angle subtended is at least
\[ \frac{2(d - 1)}{p + 1} \pi \geq \pi \quad \text{if} \quad 2d - p > 2. \]

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.
4 Generalized Padé approximations
- Existence of approximations
- An example
- A-functions
- The Daniel-Moore and Butcher-Chipman “conjectures”

5 Developments of order arrows
- Modified order arrows
- Proof of the Butcher-Chipman conjecture
Generalized Padé approximations
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 \( r = 1 \).

These are outlined on the next slide.
Existence of approximations

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 \( r = 1 \).

These are outlined on the next slide.
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 $r = 1$.

These are outlined on the next slide.
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 \( f(z) = \exp_p(tz) \)
From

\[
\frac{1}{2\pi i} \oint_C \frac{f(z)\,dz}{(z - 2)^3(z - 1)z^2}
\]

\[
= \frac{1}{2\pi i} \oint_C f(z) \left( \frac{11}{16} \frac{1}{z - 2} - \frac{1}{2} \frac{1}{(z - 2)^2} + \frac{1}{4} \frac{1}{(z - 2)^3} 
- \frac{1}{z - 1} + \frac{5}{16} \frac{1}{z} + \frac{1}{8} \frac{1}{z^2} \right)\,dz
\]

we find the approximation

\[
\Phi(w, z) = w^2 \left( 1 - \frac{8}{11} z + \frac{2}{11} z^2 \right) - w \left( \frac{16}{11} \right) + \left( \frac{5}{11} + \frac{2}{11} z \right)
\]
A-functions

An A-function in the multivalue case is a function $\Phi$, 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 an order arrow proof. The necessity of 2 is known as the Butcher-Chipman conjecture; it has now been proved.
A-functions

An A-function in the multivalue case is a function \( \Phi \), 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
\[
\begin{align*}
1 & \quad 2d_0 - p \geq 0 \\
2 & \quad 2d_0 - p \leq 2
\end{align*}
\]

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 an order arrow proof. The necessity of \( 2 \) is known as the Butcher-Chipman conjecture; it has now been proved.
A-functions

An A-function in the multivalue case is a function \( \Phi \), 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 an order arrow proof. The necessity of 2 is known as the Butcher-Chipman conjecture; it has now been proved.
A-functions

An A-function in the multivalue case is a function $\Phi$, 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 an order arrow proof. The necessity of 2 is known as the Butcher-Chipman conjecture; it has now been proved.
A-functions

An A-function in the multivalue case is a function $\Phi$, 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 an order arrow proof. The necessity of 2 is known as the Butcher-Chipman conjecture; it has now been proved.
A-functions

An A-function in the multivalue case is a function $\Phi$, 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 an order arrow proof. The necessity of 2 is known as the Butcher-Chipman conjecture; it has now been proved.
The Daniel-Moore “conjecture”

**Theorem**

\[ 2d_0 - p \geq 0 \] 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 Daniel-Moore “conjecture”

Theorem

\[2d_0 - p \geq 0 \text{ for an } A\text{-approximation.}\]

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 Daniel-Moore “conjecture”

**Theorem**

\[ 2d_0 - p \geq 0 \text{ for an } A\text{-approximation.} \]

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 red lines are tangent to the arrows and are spaced at angles of \(\pi/(p + 1) = \pi/6\).
The Daniel-Moore “conjecture”

**Theorem**

\[ 2d_0 - p \geq 0 \text{ for an } A\text{-approximation}. \]

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 red 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.
The Butcher-Chipman “conjecture”

Theorem

\[ 2d_0 - p \leq 2 \text{ for an A-approximation.} \]

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.

In the B-C 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 ends of some up-arrows from zero.
The Butcher-Chipman “conjecture”

Theorem

\[ 2d_0 - p \leq 2 \text{ for an } A\text{-approximation}. \]

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.

In the B-C 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 ends of some up-arrows from zero.
The Butcher-Chipman “conjecture”

**Theorem**

\[ 2d_0 - p \leq 2 \text{ for an } A\text{-approximation.} \]

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.

In the B-C 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 ends of some up-arrows from zero.
The Butcher-Chipman "conjecture"

**Theorem**

\[ 2d_0 - p \leq 2 \text{ for an A-approximation.} \]

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.

In the B-C 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 ends of some up-arrows from zero.
Developments of order arrows
Modified order arrows

We want to simplify what happens when an arrow interacts with a stagnation point, a branch point, a pole or a zero.
Modified order arrows

We want to simplify what happens when an arrow interacts with a stagnation point, a branch point, a pole or a zero.

We will adopt a “pass on the right” convention by moving each arrow, drawn in the increasing $w$ sense, by an infinitesimal amount to the right.
We want to simplify what happens when an arrow interacts with a stagnation point, a branch point, a pole or a zero.

We will adopt a “pass on the right” convention by moving each arrow, drawn in the increasing $w$ sense, by an infinitesimal amount to the right.

We will remove all poles by replacing a polynomial sequence $(P_0, P_1, \ldots, P_r)$ by $(-t, P_0, P_1, \ldots, P_r)$ and take the limit as $t \to 0$. Although the limit does not exist on the Riemann surface, its projection onto the $Z$ plane does.
We want to simplify what happens when an arrow interacts with a stagnation point, a branch point, a pole or a zero.

We will adopt a “pass on the right” convention by moving each arrow, drawn in the increasing $w$ sense, by an infinitesimal amount to the right.

We will remove all poles by replacing a polynomial sequence $(P_0, P_1, \ldots, P_r)$ by $(-t, P_0, P_1, \ldots, P_r)$ and take the limit as $t \to 0$. Although the limit does not exist on the Riemann surface, its projection onto the $Z$ plane does.

Do the same with zeros.
We will illustrate these ideas with the [2, 0, 1] approximation
Use right-oriented arrows
Replace poles and zeros using extra sheets
Now a generic diagram for $n_0 = 3$, $p = 5$:

It could be $[3, 2]$, $[3, 1, 0]$, $[3, 0, 1]$, $[3, 0, -1, 1]$ etc
Proof of the Butcher-Chipman conjecture

Once we have proved that $n_0$ of the up-arrow from 0 terminate at poles, the proof is just the same as for the Ehle theorem.
Proof of the Butcher-Chipman conjecture

Once we have proved that $n_0$ of the up-arrow from 0 terminate at poles, the proof is just the same as for the Ehle theorem. Hence we concentrate on this intermediate result.
Proof of the Butcher-Chipman conjecture

Once we have proved that $n_0$ of the up-arrow from 0 terminate at poles, the proof is just the same as for the Ehle theorem. Hence we concentrate on this intermediate result.

Step 1: $n_0 = p$
Proof of the Butcher-Chipman conjecture

Once we have proved that $n_0$ of the up-arrow from 0 terminate at poles, the proof is just the same as for the Ehle theorem. Hence we concentrate on this intermediate result.

Step 1: $n_0 = p$
Step 2: Induction on $p - n_0$
Once we have proved that $n_0$ of the up-arrow from 0 terminate at poles, the proof is just the same as for the Ehle theorem. Hence we concentrate on this intermediate result.

Step 1: $n_0 = p$

Step 2: Induction on $p - n_0$

We will illustrate step 2, in the case $n_0 = 4$, $p = 5$
Once we have proved that $n_0$ of the up-arrow from 0 terminate at poles, the proof is just the same as for the Ehle theorem. Hence we concentrate on this intermediate result.

Step 1: $n_0 = p$

Step 2: Induction on $p - n_0$

We will illustrate step 2, in the case $n_0 = 4$, $p = 5$

We use homotopy: as $t$ moves from 0 to 1 we move from an approximation with $p - n_0 = 0$ to $p - n_0 = 1$
Proof of the Butcher-Chipman conjecture

Once we have proved that $n_0$ of the up-arrow from 0 terminate at poles, the proof is just the same as for the Ehle theorem. Hence we concentrate on this intermediate result.
Step 1: $n_0 = p$
Step 2: Induction on $p - n_0$
We will illustrate step 2, in the case $n_0 = 4$, $p = 5$
We use homotopy: as $t$ moves from 0 to 1 we move from an approximation with $p - n_0 = 0$ to $p - n_0 = 1$
First see how the order increases as $t$ approaches 1
Generalized Padé approximations
Developments of order arrows
Modified order arrows
Proof of the Butcher-Chipman conjecture
Generalized Padé approximations
Developments of order arrows
Modified order arrows
Proof of the Butcher-Chipman conjecture

$t$ close to 0
Generalized Padé approximations
Developments of order arrows
Modified order arrows
Proof of the Butcher-Chipman conjecture

$t$ close to 1
Is it possible that during the homotopy, one of the arrows which terminates on the top sheet gets detached from 0?
Is it possible that during the homotopy, one of the arrows which terminates on the top sheet gets detached from 0? If so, an arrow from a lower sheet must connect to 0 at the same time to retain order $p - 1$. 
Is it possible that during the homotopy, one of the arrows which terminates on the top sheet gets detached from 0? If so, an arrow from a lower sheet must connect to 0 at the same time to retain order $p - 1$. This means that for some $t \in (0, 1)$, the order becomes $p$. 
Is it possible that during the homotopy, one of the arrows which terminates on the top sheet gets detached from 0?
If so, an arrow from a lower sheet must connect to 0 at the same time to retain order $p - 1$.
This means that for some $t \in (0, 1)$, the order becomes $p$.
This is impossible because of the uniqueness of generalized Padé approximations.
Is it possible that during the homotopy, one of the arrows which terminates on the top sheet gets detached from 0? If so, an arrow from a lower sheet must connect to 0 at the same time to retain order $p - 1$. This means that for some $t \in (0, 1)$, the order becomes $p$. This is impossible because of the uniqueness of generalized Padé approximations. This completes the proof.