AN ANALYSIS OF DISCRETISATION METHODS
FOR ORDINARY DIFFERENTIAL EQUATIONS.

Neil Pitcher,
The Queen's College,
Oxford.


"Unto the pure all things are pure."
The apostle Paul ... Titus 1:15 (A.V.)
ACKNOWLEDGEMENTS.

I should like to express my gratitude to my supervisor, Dr. Sean Mc.Kee, in the first place for suggesting this area of work, and for his helpful comments and suggestions throughout my time of research.

My thanks are also due to the Science Research Council for their financial support for my three years at The Queen's College, Oxford, which I have enjoyed very much.

Finally I should like to mention my parents, especially my late father, Mr. Rex Pitcher, to whom this thesis is dedicated, for their affection, help and encouragement over the many years of my education.

To my late father, Mr. Rex Pitcher.
Numerical methods for solving initial value problems in ordinary differential equations are studied. A notation is introduced to represent cyclic methods in terms of two matrices, $A_h$ and $B_h$, and this is developed to cover the very extensive class of $m$-block methods. Some stability results are obtained and convergence is analysed by means of a new consistency concept, namely optimal consistency. It is shown that optimal consistency allows one to give two-sided bounds on the global error, and examples are given to illustrate this. The form of the inverse of $A_h$ is studied closely to give a criterion for the order of convergence to exceed that of consistency by one. Further convergence results are obtained, the first of which gives the orders of convergence for cases in which $A_h$ and $B_h$ have a special form, and the second of which gives rise to the possibility of the order of convergence exceeding that of consistency by two or more at some stages. In addition an alternative proof is given of the superconvergence result for collocation methods. In conclusion the work covered is set in the context of that done in recent years by various authors.
CONTENTS

1. Introduction
   1.1. Preliminaries
   1.2. Linear Multistep Methods
   1.3. Cyclic Linear Multistep Methods

2. Block Semi Explicit Methods
   2.1. Definitions
   2.2. Cyclic Linear Multistep Methods
   2.3. Predictor-Corrector Methods
   2.4. Explicit Runge-Kutta Methods
   2.5. Hybrid Methods
   2.6. Extrapolation

3. Block Implicit Methods
   3.1. Definitions
   3.2. Advanced Linear Multistep Methods
   3.3. Implicit Runge-Kutta Methods
   3.4. One-Leg Methods
   3.5. Collocation Methods

4. Block Methods
   4.1. Truncation Error
   4.2. The Numerical Stability of Block Semi Explicit Methods
   4.3. Zero Stability
   4.4. Convergence Theorems
5. Two-Sided Error Bounds.
   5.1 Optimal Consistency.
   5.2 The Form of $A_h^{-1}$.
   5.3 Improved Convergence.

6. Optimal Consistency of Orders $(p_1, p_2, \ldots, p_m)$.
   6.1 Two-Sided Error Bounds.
   6.2 A Special Class of Methods.
   6.3 Two-Sided Error Bounds.
   6.4 A Further Convergence Theorem.
   6.5 Collocation.

7. Conclusions.
   7.1 The Work of Spijker.
   7.2 The Work of Albrecht.
   7.3 Comparison with some Results of Stetter.
   7.4 The Work of Skeel.
   7.5 The Work of Cooper.
1.1 Preliminaries.

Let \( T \) be some positive real number. We shall be concerned in this thesis with the numerical solution of the initial value ordinary differential equation
\[
y'(t) = f(t, y(t)) \quad (1.1.1)
\]
with \( y(0) = y_0 \).

We take \( f \) to be in the class of Lipschitz continuous functions in which for every \( f \) there exists a Lipschitz constant \( L \) such that for any \( y_1, y_2 \) and \( t \in [0, T] \),
\[
|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2| \quad (1.1.2).
\]

Let us use a basic steplength \( h > 0 \) along the interval \([0, T]\) such that \( N = T/h \) is a positive integer. For any real number \( i \in [0, N] \) we define \( t_i = ih \), so that \( t_0 = 0 \) and \( t_N = T \).

A numerical method is a computational means of obtaining approximations to \( y(t) \) at the points \( \{t_i : i \in S\} \), where \( S \) is a finite subset of \([0, N]\). For any \( i \) we denote the approximation to \( y(t_i) \) by \( y_i \).

1.2 Linear Multistep Methods.

Any linear multistep method may be written in the form
\[
\sum_{j=0}^{k} \alpha_{k-j} y_{i-j} = h \sum_{j=0}^{k} \beta_{k-j} f(t_{i-j}; y_{i-j}) \quad , \quad i = k(1)N \quad (1.2.1),
\]
where \( k, \alpha_0, \alpha_1, \ldots, \alpha_k, \beta_0, \beta_1, \ldots, \beta_k \) are real constants independent of \( h \), and \( k \) is known as the stepnumber. A single application of this method computes \( y_i \) from the known values
A linear multistep method is CONVERGENT of ORDER p \in \mathbb{N} if for \(i = 0(1)N\), \(y_i - y(t_i) = O(h^p)\) as \(h \to 0\).

A linear multistep method is ZERO STABLE if no root of the polynomial \(p(\xi) = \sum_{j=0}^{k} \alpha_j \xi^j\) has modulus greater than one, and every root with modulus one is simple.

The LOCAL TRUNCATION ERROR at the \(i\)th stage, \(i \geq k\), is

\[\Theta_i = \sum_{j=0}^{k} \alpha_{k-j} y(t_{i-j}) - h \sum_{j=0}^{k} \beta_{k-j} y'(t_{i-j}).\]

A linear multistep method is CONSISTENT of ORDER p \in \mathbb{N} if

\[\Theta_i = O(h^{p+1})\] as \(h \to 0\) for \(i \geq k\).

If \(\Theta_i = C^{(1)} h^{p+1} y^{(p+1)}(t_i) + O(h^{p+2})\) for all \(y(t) \in C^{(p+1)}[0,T]\), where \(C^{(1)}\) is some constant independent of \(h\), then \(C^{(1)}\) is the ERROR CONSTANT.

A linear multistep method is convergent of order p if and only if it is both zero stable and consistent of order p.

In the above linear multistep method (1.2.1) we may stipulate that \(\alpha_k = 1\) and thus have \(2k+1\) free parameters \(\alpha_0, \alpha_1, \ldots, \alpha_{k-1}, \beta_0, \beta_1, \ldots, \beta_k\) in constructing a method. Expanding \(\Theta_i\) as
a Taylor series about $t_i$ and assuming $y(t)$ to have sufficient continuity we observe that it will be possible to choose these constants so as to give methods with order of consistency as high as $2k$. Yet the construction of linear multistep methods with order of convergence as high as this is precluded by considerations of zero stability, as the following theorem shows.

**Theorem 1.2.6.** (6)

No zero stable linear multistep method of stepnumber $k$ can have order of consistency exceeding $2\lfloor \frac{k}{2} \rfloor + 2$, where for any real $x$, $\lfloor x \rfloor$ denotes the integral part of $x$.

### 1.3 Cyclic Linear Multistep Methods

Donelson and Hansen (9) proposed cyclic (linear multistep) methods as a means of surmounting the barrier presented by theorem 1.2.6. Such a method consists of cycles of a finite number, say $m$, of linear multistep methods, and may be written in the form

$$\sum_{j=0}^{k} \alpha_{k-j} y_{i-j} = h \sum_{j=0}^{k} \beta_{k-j} f(t_{i-j}, y_{i-j}), \quad i = k(1)N,$$

where

$$\forall \equiv i-k+1 \pmod{m}, \quad 1 \leq \forall \leq m.$$

Suppose we are given $k$ starting values $\bar{y}_i$, approximating $y(t_i)$, for $i = O(1)k^{-1}$. Then we may represent the cyclic method in the form
Let us define the vectors $y$, $f$, $g \in \mathbb{R}^{N+1}$ as follows, taking all vector and matrix suffices to begin at zero rather than one. The $i^{th}$ component of $y$ is $y_i$, $f_i = f(t_i, y_i)$ and $g_i = \begin{cases} y_i & \text{if } 0 \leq i \leq k-1, \\ 0 & \text{if } k \leq i \leq N. \end{cases}$

We may write $A_h y = hB_h f + g$, where $A_h$ and $B_h$ are lower-triangular $(N+1)\times(N+1)$ matrices with bandwidth $(k+1)$ independent of $h$.

The arrays $(1,3,1)$ recur in $A_h$ and $(1,3,2)$ recur in $B_h$ respectively.
Let us partition these arrays into $m \times m$ matrices $A_0, A_1, \ldots, B_0, B_1, \ldots, B_r$, so that

$$
(A_0, A_1, \ldots, A_r) = \begin{bmatrix}
\alpha_0^{(1)} & \cdots & \alpha_k^{(1)} \\
\vdots & \ddots & \vdots \\
\alpha_0^{(m)} & \cdots & \alpha_k^{(m)}
\end{bmatrix}
$$

and

$$
(B_0, B_1, \ldots, B_r) = \begin{bmatrix}
\beta_0^{(1)} & \cdots & \beta_k^{(1)} \\
\vdots & \ddots & \vdots \\
\beta_0^{(m)} & \cdots & \beta_k^{(m)}
\end{bmatrix},
$$

where

$$
\gamma = 1 + \left[\frac{k-1}{m}\right] \quad (2\times, p.218).
$$

So $A_h = \begin{bmatrix} 1 & \cdots & A_0 & A_1 \cdots & A_r \end{bmatrix}$ and

$$
B_h = \begin{bmatrix} 0 & \cdots & 0 & B_0 & \cdots & B_r \end{bmatrix}
$$

Example 1.3.1.

Consider the application of the following two linear multistep methods alternately.

$$
y_{i+1} - y_i = \frac{1}{2}h(f_i + 4f_{i-1} + f_{i-2}),
$$

$$
y_{i+1} - y_i = \frac{1}{12}h(5f_{i+1} + 8f_{i} - f_{i-1}).
$$

$k = m = 2$, so $\gamma = 1$. 
A cyclic method is zero stable if \( \max_{0 \leq i,j \leq N} |(A_h^{-1})_{ij}| = O(1) \) as \( h \to 0 \).

We note that any linear multistep method may be written as a cyclic method with \( m = 1 \), giving \( \gamma = k \), \( A_\theta = \alpha \), \( B_\theta = \beta \) for \( \theta = O(1)\gamma \). Definition 1.3.2 is a generalisation of definition 1.2.2 to take in the class of cyclic methods of which linear multistep methods form a subclass (18). For cyclic methods the order of convergence is defined as in definition 1.2.1. As for consistency, a cyclic method is consistent of order \( p \) if each of its constituent linear multistep methods is consistent of order at least \( p \). We then have the following convergence theorem.

Theorem 1.3.3.

Suppose we are given \( k \) starting values \( \tilde{y}_0, \ldots, \tilde{y}_{k-1} \) such that \( \tilde{y}_i - y(t_i) = O(h^p) \) as \( h \to 0 \) for \( i = O(1)k-1 \). Then any cyclic method using these starting values is convergent of order \( p \) if it is both zero stable and consistent of order \( p \).

For the proof of this see theorem 2 of (9).

Definition 1.3.4.

For \( A_h \) of the form (1.3.1) we define the companion matrices \( G_1, G_2, \ldots, G_m \) and the matrix \( G \) by
\[ G_{\nu} = \begin{bmatrix}
0 & 1 & \cdots & \cdots & 1 \\
\varepsilon_0/\alpha_k & \varepsilon_0/\alpha_k & \cdots & \cdots & \varepsilon_0/\alpha_k \\
\varepsilon_0/\alpha_k & \varepsilon_0/\alpha_k & \cdots & \cdots & \varepsilon_0/\alpha_k \\
& \cdots & \cdots & \cdots & \cdots
\end{bmatrix}, \text{ a } k \times k \text{ matrix,}
\]

\[ \nu = 1(1)m, \quad \text{and} \]

\[ G = G_m G_{m-1} \cdots G_1. \quad (1.3.3) \]

Theorem 1.3.5. (18)

\[ 0 \leq \max_{i,j} \frac{1}{h} |(A_h^{-1})_{ij}| = O(1) \text{ as } h \to 0 \text{ if no eigenvalue of } G \]

has modulus greater than one, and every eigenvalue with modulus one has linear elementary divisors.

It is possible to choose the \( m \) linear multistep methods to have order of consistency greater than the bounds imposed by theorem 1.2.6, so that each individual linear multistep method is zero unstable and yet for there to be zero stability in the cyclic method, as the following example shows.

Example 1.3.6. (9, p. 149)

\[
\begin{bmatrix}
1 \\
0 & 1 \\
0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & -57 & 24 & 33 \\
0 & 0 & 0 & 0 & 0 & -456 & -1347 & 720 \\
0 & 0 & 0 & 0 & 0 & 0 & -57 & 24 & 33 \\
\end{bmatrix}
\begin{bmatrix}
y_0 \\
y_1 \\
y_2 \\
y_3 \\
y_4 \\
y_5 \\
\vdots
\end{bmatrix}
= \begin{bmatrix}
y_0 \\
y_1 \\
y_2 \\
y_3 \\
y_4 \\
y_5 \\
\vdots
\end{bmatrix}
\]
By direct computation we obtain

\[
\begin{pmatrix}
0 & 19/11 & -8/11 \\
0 & 19/11 & -8/11 \\
0 & 19/11 & -8/11 \\
\end{pmatrix}
\]

\[\det (G - \zeta I) = \zeta^2 (\zeta - 1),\] so we have zero stability, yet for each constituent linear multistep method \( k = 3 \) and consistency is of order five.
CHAPTER TWO.

**BLOCK SEMI EXPLICIT METHODS.**

2.1 Definitions.

The matrix notation (1.3.1) can be extended to give a notation which will represent many different kinds of methods.

**Definition 2.1.1.**

For any \( \gamma \in \mathbb{R}^m \) and \( r, n \in \mathbb{N} \), where \( r \) is independent of \( h \), we define the RESTRICTION OPERATOR \( \Delta_h : C[0,T] \rightarrow \mathbb{R}^{r+nm} \) as follows. For any \( \xi \in C[0,T] \),

\[
[\Delta_h \xi(t)]_i = \begin{cases} 
\xi(ih), & i = O(1)r^{-1}, \\
\xi((r-1+s\gamma_m+\gamma_v)h), & i = r + sm + \nu - 1.
\end{cases}
\]

For any \( r, I_r \) and \( 0_r \) will represent the \( r \times r \) identity and zero matrices respectively.

**Definition 2.1.2.**

A method is \( m \)-BLOCK SEMI EXPLICIT if \( \gamma \in \mathbb{R}^m \) is given and we may represent the method in the form

\[
A_h \gamma^* = hB_h f^* + g,
\]

where \( A_h \) and \( B_h \) are lower triangular \((r+nm)\times(r+nm)\) matrices with finite bandwidth independent of \( h \) and have the form
where \( r \) is a positive integer independent of \( h \), \( A_0, A_1, \ldots, A_r, B_0, B_1, \ldots \), \( B_r \) are \( m \times m \) matrices, \( A_0, \ldots, A_{r-1}; B_0, \ldots, B_{r-1} \), \( \Theta = O(1)^{r-1} \), are matrices with \( m \) rows, and \( y_x^*, f_x^*, g \in R^{r+nm} \).

\( y_x^* \) is the computed approximation to \( y([\Delta t]_{i-1}) \), \( f_x^* = f([\Delta t]_{i-1}, y_x^*) \) and

\[
\mathcal{E}_{i} = \begin{cases} 
\mathcal{Y}_{i}, & i = O(1) \tau - 1, \\
0, & i \geq r.
\end{cases}
\]

Moreover \([\Delta t]_0 = 0 \) and \([\Delta t]_{r+nm-1} = 0 \).

\((r-1+m\gamma)h = T \) (2.1.3), so \( N = r - 1 + m\gamma_m \).

Any block semi explicit method then may be fully represented

\((*)\) The matrices \( A_0^{(\theta)}, \ldots, A_{r-1}^{(\theta)}, B_0^{(\theta)}, \ldots, B_{r-1}^{(\theta)} \) are just \( A_0, \ldots, A_{r-1}, B_0, \ldots, B_{r-1} \) with some zero columns deleted, for \( \Theta = O(1)^{r-1} \).
by the \( m \times (1 + 2(\gamma + 1)m) \) array

\[
\begin{pmatrix}
A_0, A_1, \ldots, A_\gamma | B_0, B_1, \ldots, B_\gamma | y
\end{pmatrix}
\quad (2.1.4).
\]

For any block semi explicit method \( A_\gamma \) and \( B_\gamma \) are both lower triangular, and if \( B_\gamma \) is strictly lower triangular we shall call the method \( m\)-BLOCK EXPLICIT.

The class of \( m\)-block semi explicit methods is very large, and we now give a wide range of methods covered by this definition.

2.2 Cyclic Linear Multistep Methods.

The linear multistep method (1.2.1) is a 1-cyclic method and the array (2.1.1f) takes the form \( \begin{pmatrix} \alpha_0, \alpha_1, \ldots, \alpha_\gamma \end{pmatrix} \left| \beta_0, \beta_1, \ldots, \beta_\gamma \right| y \), and the method is called explicit if \( \beta_\gamma = 0 \). For each \( i \), \( y_i^* = y_i \) and \( N = r + n - 1 \).

For example 1.3.1 we have the array

\[
\begin{pmatrix}
-1 & 0 & 1 & 0 & 1/3 & 4/3 & 1/3 & 0 & 1
0 & 0 & -1 & 1 & 0 & -1/12 & 8/12 & 5/12 & 2
\end{pmatrix},
\]

a 2-block semi explicit method, \( y_i^* = y_i \) for each \( i \) and \( N = r + 2n - 1 \).

Example 2.2.1.

Consider a cyclic method formed by alternating the explicit and implicit Euler linear multistep methods. This is a 2-block semi explicit method, and the array is

\[
\begin{pmatrix}
0 & -1 & 1 & 0 & 0 & 1 & 0 & 0 & 1
0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & 2
\end{pmatrix}.
\]

Example 2.2.2.

Consider a cyclic method represented as a 3-block semi explicit method, for which \( y_i^* = y_i \) for each \( i \) and \( N = r + 3n - 1 \).
Example 2.2.3.

Another such cyclic method is represented by

\[
\begin{pmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & -2 & 1 & 0 \\
0 & 0 & 0 & 0 & -2 & 3/2 & 1/2 & 1 \\
0 & 0 & 3 & -22/3 & 32/3 & -10/3 & 1 & 0 & 0 \\
0 & 0 & 0 & -17/12 & 15/4 & -13/4 & -1/12 & 1 & 0 \\
0 & 0 & 0 & 0 & -21/16 & 155/48 & -61/48 & 41/48 & 1
\end{pmatrix}
\]

Example 2.2.4.

A cyclic method representable as a 2-block semi explicit method:

\[
\begin{pmatrix}
0 & 0 & -1 & 1 & 0 & 0 & -1/6 & 1 & 5/6 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & -1 & 1 & 3/6 & 1 & 5/6 & 1 & 2
\end{pmatrix}
\]

Example 2.2.5.

Another such cyclic method is

\[
\begin{pmatrix}
0 & 0 & -1 & 1 & 0 & 1 & 1/6 & 1 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & -1 & 1 & -1 & -1/3 & -1 & 1/3 & -1 & 5/6 & 1 & 2
\end{pmatrix}
\]

We see then that for a cyclic method composed of m linear multistep methods \( y = (1,2, \ldots, m)^T \), each of the examples given will be referred to later.
2.3 Predictor-Corrector Methods.

Suppose we have an explicit linear multistep predictor

\[ \sum_{j=0}^{k} \alpha_{j}^{*} y_{i-j} = h \sum_{j=0}^{k-1} \beta_{j}^{*} f_{i-j} \]

followed by \( \mu \) applications of the implicit corrector

\[ \sum_{j=0}^{k} \alpha_{j}^{*} y_{i-j} = h \sum_{j=0}^{k} \beta_{j}^{*} f_{i-j} \]  \( (2.3.1) \)

Then in the array \((2.4.1)\) we have \( \tau = k \), \( m = \mu + 1 \),

\[
A_{\theta} = \begin{pmatrix} 0 & \cdots & 0 & \alpha_{\theta}^{*} \\ & \ddots & \ddots & \ddots \\ & & 0 & \cdots \\ & & & 0 \end{pmatrix}, \quad \theta = O(1)^{k-1},
\]

\[
A_{k} = \begin{pmatrix} \alpha_{1}^{*} \\ \vdots \\ \alpha_{k}^{*} \end{pmatrix}, \quad \gamma = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix},
\]

\[
B_{k} = \begin{pmatrix} 0 & \cdots & 0 & \beta_{k}^{*} \\ & \ddots & \ddots & \ddots \\ & & 0 & \cdots \\ & & & 0 \end{pmatrix}
\]

in \( P(EC)^{\mu}E \) mode \( (\text{see (15), p.}\, 87) \)

\[
B_{\theta} = \begin{pmatrix} 0 & \cdots & 0 & \beta_{\theta}^{*} \\ & \ddots & \ddots & \ddots \\ & & 0 & \cdots \\ & & & 0 \end{pmatrix}, \quad \theta = O(1)^{k-1},
\]

and in \( P(EC)^{\mu} \) mode

\[
B_{\theta} = \begin{pmatrix} 0 & \cdots & 0 & \beta_{\theta}^{*} & 0 \\ 0 & \cdots & 0 & \beta_{\theta} & 0 \\ & \ddots & \ddots & \ddots \\ & & 0 & \cdots & 0 \end{pmatrix}
\]
This then is a \((\mu+1)\)-block explicit method, and

\[
\Delta t = (0, h, \ldots, (r-1)h, rh, \ldots, rh,(r+1)h, \ldots, (r+n-1)h)
\]

and \(N = r + n - 1\).

**Example 2.3.1.**

Consider also a predictor-corrector method with a Milne modifier in PMECME mode (15, p. 94), in which we have

\[
P: y_i^0 - y_i^1 = \frac{4h}{3}[2f(t_{i-1}, y_i^1) - f(t_{i-2}, y_i^2) + 2f(t_{i-3}, y_i^3)],
\]

\[
M: y_i^0 - y_i^1 = \frac{112}{121}(y_i^1 - y_i^0) = 0,
\]

\[
C: y_i^1 - \frac{2y_i^1}{3} + \frac{1}{3}y_i^0 = \frac{3h}{8}[f(t_i, y_i^0) + 2f(t_{i-1}, y_{i-1}^1) - f(t_{i-2}, y_{i-2}^1)],
\]

\[
M: y_i^1 - y_i^1 + \frac{9}{121}(y_i^1 - y_i^0) = 0.
\]

This is a 4-block explicit method in which

\[
A_0 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
\]

\[
A_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
\]

\[
A_2 = 0_4,
\]

\[
A_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 112/121 & 0 & -112/121 & 0 \\ 0 & 0 & 0 & -9/8 \\ 0 & 0 & 0 & 0 \end{pmatrix},
\]

\[
A_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -9/121 & 0 & -112/121 & 1 \end{pmatrix},
\]

\[
B_0 = 0_4, B_1 = \begin{pmatrix} 0 & 0 & 0 & 0/3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
\]

\[
B_2 = \begin{pmatrix} 0 & 0 & 0 & -4/3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3/8 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
\]
2.4 Explicit Runge-Kutta Methods.

An $m$ stage explicit Runge-Kutta method is usually written

$$y_i - y_{i-1} = h \sum_{j=1}^{m} c_j k_j,$$

$$k_1 = f(t_{i-1}, y_{i-1}),$$

$$k_j = f(t_{i-1} + b_j h, y_{i-1} + h \sum_{q=1}^{j-1} b_{jq} k_q), \quad j = 2(1)m,$$

where $a_j = \sum_{q=1}^{j-1} b_{jq}, \quad j = 2(1)m$.

An alternative representation is

$$y_{i-1} + a_2 = y_{i-1} + h b_{21} f_{i-1},$$

$$y_{i-1} + a_j = y_{i-1} + h (b_{31} f_{i-1} + b_{32} f_{i-1} + a_2),$$

$$y_{i-1} + a_m = y_{i-1} + h (b_{m1} f_{i-1} + b_{m2} f_{i-1} + a_2 + \cdots + b_{m,m-1} f_{i-1} + a_{m-1}),$$

$$y_i = y_{i-1} + h (a_1 f_{i-1} + a_2 f_{i-1} + a_3 f_{i-1} + \cdots + a_m f_{i-1} + a_m),$$

which can be viewed as an $m$-block explicit method with the array

$$B_3 = \begin{pmatrix} 0 & 0 & 0 & 8/3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3/4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3/8 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$
Consider a semi explicit Runge-Kutta method, for which the formulation (2.4.1) has
\[ k_j = f(t_{i-1} + h a_j, y_{i-1} + h \sum_{q=1}^{m} b_{jq} k_q), \]
so that \( B_z \) is lower triangular, but not strictly so. Lambert (15, p.154) quotes the scheme

\[ y_1 - y_{i-1} = \frac{1}{6}(k_1 + 4k_2 + k_3), \]
\[ k_1 = f(t_{i-1}), \]
\[ k_2 = f(t_{i-1} + h/2, y_{i-1} + h k_1/4 + h k_2/4), \]
\[ k_3 = f(t_{i-1} + h, y_{i-1} + h k_2). \]

This is a 3-block semi explicit method with array

\[
\begin{pmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 1/2 \\
0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1/6 & 2/3 & 1/6 & 0 & 1.
\end{pmatrix}
\]

2.5 Hybrid Methods.

A k-step hybrid formula takes the form

\[ \sum_{j=0}^{k-1} \alpha_{k-j} y_{i-j} = h \sum_{j=0}^{k} \beta_{k-j} f_{i-j} + h^{p+1} f_{i-k+\mu} \quad (2.5.1), \]
where \( \mu \neq \{1, 2, \ldots, k\} \).

Definition 2.5.1.

The hybrid formula (2.5.1) has ORDER \( p \) and ERROR CONSTANT \( C \) if for all \( y(t) \in C^{(p+1)}[0, T] \),

\[ \sum_{j=0}^{k} \alpha_{k-j} y(t_{i-j}) - h \sum_{j=0}^{k} \beta_{k-j} y'(t_{i-j}) - h^{p+1} y(t_{i-k+\mu}) = C h^{p+1} y^{(p+1)}(t_i) + O(h^{p+2}). \]
Example 2.5.2.

In the case when $\beta_k = 0$ we need a predictor formula to obtain an estimate of $y_{i-k+\mu}$, and the hybrid method is formulated as a $P_{\mu_k}EP_{\mu_k}^H$ algorithm. For example, on p. 180 of (15) a $P_{2/3}EP_{1/3}^H$ method is given:

$$P_{2/3} : y_{i-1/3} = \frac{25}{27}y_{i-1} - \frac{25}{27}y_{i-2} = \frac{h}{27}(50r_{i-1} + 20r_{i-2})$$

$$P_{1/3}^H : y_{i-1} - y_{i-1/3} = \frac{h}{4}(3r_{i-1/3} + r_{i-1})$$

This is a 2-block explicit method with array

$$\begin{bmatrix}
0 & -52/27 & 0 & 25/27 & 1 & 0 & 0 & 20/27 & 0 & 50/27 & 0 & 0 & 2/3 \\
0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 1/4 & 3/4 & 0 & 1
\end{bmatrix}_h.$$

$\varepsilon = 2$ and $\Delta \, t = (0, h, 5h/3, 2h, 3h, 3h/3, 3h, \ldots)^T$.

Example 2.5.3.

When $\beta_k \neq 0$ we need in addition to a $P_{\mu_k}$ predictor a $P_k$ predictor, and the hybrid method is formulated as a $P_{\mu_k}EP_{\mu_k}^H$ algorithm. For example, on p. 175 of (15) a $P_{1/2}EP_{1/2}^H$ scheme is given:

$$P_{1/2} : y_{i-1/2} - y_{i-1} = \frac{3h}{8}(3r_{i-1} + r_{i-2})$$

$$P_k : y_{i-1}^P = 4y_{i-1} - 5y_{i-2} = h(4r_{i-1}^C + 2r_{i-2}^C)$$

$$C_{1/2} : y_{i-1}^C - y_{i-1} = \frac{h}{6}(r_{i-1}^P + r_{i-1}^C + 4r_{i-2}^C)$$

where for each $j$, $r_{i,j}^P = f(t_{i,j}, y_{i,j}^P)$ and $r_{i,j}^C = f(t_{i,j}, y_{i,j}^C)$. This is a 3-block explicit method with array

$$\begin{bmatrix}
0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 3/8 & 0 & 0 & 9/8 & 0 & 0 & 0 & 1/2 \\
0 & 0 & -5 & 0 & 0 & 4 & 0 & 1 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 2/3 & 1/6 & 0 & 1
\end{bmatrix}_h.$$

$r = 2$ and $\Delta \, t = (0, h, 3h/2, 2h, 2h, 5h/2, 3h, 3h, \ldots)^T$. 
Example 2.5.4.

We may also use a $P\mu$ mode. For the same $P_k$ and $C_H$ as in example 2.5.3 we have the same array, except that

$$B_0 = \begin{pmatrix} 0 & 3/8 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B_1 = \begin{pmatrix} 0 & 9/8 & 0 \\ 0 & 4 & 0 \\ 0 & 1/6 & 0 \end{pmatrix}.$$  

Example 2.5.5.

$P\mu$ methods may also be considered. For example Butcher (3, p.130) quotes

$$P_{1/2} : y_{i-1/2}^0 - y_{i-1}^0 = h f_{i-1/2}^0,$$

$$P_H : y_i^p - y_{i-1}^c = h (2 f_{i-1}^p - f_{i-1}^c),$$

$$C_H : y_i^c - y_{i-1}^c = \frac{h}{6} (4 f_{i-1}^p + f_i^p + f_{i-1}^c),$$

which is a 3-block explicit method with array

$$\begin{pmatrix} 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & -1 & 2 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 & 1 & 6 & 2/3 & 1/6 & 0 & 1 \end{pmatrix} \quad \text{and} \quad r = 1.$$

We note that this is just a 3-stage explicit Runge-Kutta method.

Example 2.5.6.

Kohfeld and Thompson (13) proposed another mode, $P\mu$ in order to take advantage of the fact that hybrid methods have very small error constants. For instance Lambert (15, pp.176-177) quotes the following $P_{0.7} EP_{0.5} EP\mu$ method, $\beta$ being some constant.

$$P_{0.7} : y_{i-0.3} + 5.7937275 y_{i-1} - 3.5721 y_{i-2} - 3.2216275 y_{i-3} = h (3.6869175 f_{i-1} + 6.07257 f_{i-2} + 0.9558675 f_{i-3}),$$

$$P_{0.5} : y_{i-0.5} + \left(\frac{45}{64} - \frac{357}{250} \beta\right) y_{i-1} + \left(- \frac{25}{16} + \frac{3213}{250} \beta\right) y_{i-2} + \left(- \frac{9}{64} + \frac{357}{250} \beta\right) y_{i-3}.$$
This is a 4-block explicit method with

\[
A_0 = \begin{bmatrix}
0 & 0 & 0 & -3.2216275 \\
0 & 0 & 0 & -\frac{9}{64} + \frac{357\beta}{250} \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}, \quad A_1 = \begin{bmatrix}
0 & 0 & 0 & -3.5721 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix},
\]

\[
A_2 = \begin{bmatrix}
0 & 0 & 0 & 5.7937275 \\
0 & 0 & 0 & -\frac{45}{64} - \frac{357\beta}{25} \\
0 & 0 & 0 & -1 \\
0 & 0 & 0 & -1
\end{bmatrix}, \quad A_3 = I_4,
\]

\[
B_0 = \begin{bmatrix}
0 & 0 & 0 & 0.9558675 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}, \quad B_1 = \begin{bmatrix}
0 & 0 & 0 & 6.07257 \\
0 & 0 & \frac{15}{16} - \frac{4291\beta}{500} \\
0 & 0 & -\frac{7}{714} & 0 \\
0 & 0 & 0 & 0
\end{bmatrix},
\]

\[
B_2 = \begin{bmatrix}
0 & 0 & 0 & 3.6869175 \\
0 & 0 & \frac{45}{32} - \frac{4063\beta}{500} \\
0 & 0 & \frac{221}{714} & 0 \\
0 & 0 & \frac{1}{6} & 0
\end{bmatrix}, \quad B_3 = \begin{bmatrix}
0 & 0 & 0 & 0 \\
\beta & 0 & 0 & 0 \\
\frac{500}{714} & 0 & 0 & 0 \\
0 & \frac{2}{3} & \frac{1}{6} & 0
\end{bmatrix}, \quad \gamma = \begin{bmatrix}
0.7 \\
0.5 \\
1 \\
1
\end{bmatrix}.
\]

2.6 Extrapolation.

Example 2.6.1.

In this example we apply the trapezoidal rule with steplength \( \frac{1}{2}h \) twice, followed by the same formula with steplength \( h \) once, and then
extrapolate the two values obtained.

\[ y_{i\frac{1}{2}} - y_{i-1} = \frac{h}{4} (f_{i\frac{1}{2}} + f_{i-1}) , \]
\[ y_i - y_{i\frac{1}{2}} = \frac{h}{4} (f_i + f_{i-1}) , \]
\[ y_{i\frac{1}{2}} - y_{i-1} = \frac{h}{2} (f_{i\frac{1}{2}} + f_{i-1}) , \]
\[ y_i + \frac{1}{2} y_{i\frac{1}{2}} - \frac{h}{2} y_{[0]} = 0 . \]

For each \( j \) and \( l \), \( f_j^{[l]} = f(t_j, y_j^{[l]}) \). This is a \( 4 \)-block semi explicit method with array

\[
\begin{bmatrix}
0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 1/2 \\
0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 1 \\
0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 1/2 & 0 & 0 & 1/2 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & -1/3 & 1/3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}.
\]

Example 2.6.2.

Consider a particular case of Gragg's method (15, pp.191,192).

We have \( y_{i-1/2} - y_{i-1} = \frac{h}{8} f_{i-1} \),
\[ y_{i-1} - y_{i-1} = h f_{i-1/2} , \]
\[ y_{i+1/2} - y_{i-1/2} = h f_i , \]
\[ y_i - \frac{1}{4} y_{i+1/2} - \frac{1}{2} y_i - \frac{1}{4} y_i - \frac{1}{2} y_i = 0 , \]
\[ y_i - \frac{1}{4} y_i - \frac{1}{2} y_i - \frac{1}{4} y_i - \frac{1}{2} y_i = 0 . \]

\[ y_{i-3/4} - y_{i-1} = \frac{h}{8} f_{i-1} , \]
\[ y_{i-1/2} - y_{i-1} = \frac{h}{8} f_{i-1} , \]
\[ y_{i-1/4} - y_{i-3/4} = \frac{h}{8} f_{i-1/2} , \]
\[ y_i - y_{i-1/2} = \frac{h}{8} f_{i-1/2} , \]
\[ y_{i+1/4} - y_{i-3/4} = \frac{h}{8} f_{i-1/4} , \]
\[ y_{i+1/4} - y_{i-3/4} = \frac{h}{8} f_{i-1/4} , \]
\[
\begin{align*}
\dot{y}_i - \frac{1}{4}y_{i+1/4} - \frac{1}{2}y_i - \frac{1}{4}y_{i-1/4} &= 0, \\
y_i - \frac{4}{3}y_i + \frac{1}{3}y_i &= 0.
\end{align*}
\]

This is an 11-block explicit method with

\[
A_0 = \begin{pmatrix}
0 & \cdots & 0 & -1 \\
\cdot & \cdots & -1 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & -1 \\
\cdot & \cdots & -1 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 0 \\
0 & \cdots & 0 & 0
\end{pmatrix},
B_0 = \begin{pmatrix}
0 & \cdots & 0 & 1/2 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 1/4 \\
\cdot & \cdots & 0 \\
\cdot & \cdots & 0 \\
0 & \cdots & 0 & 0
\end{pmatrix},
\]

\[
A_1 = \begin{pmatrix}
1 \\
0 & 1 \\
-1 & 0 & 1 \\
-1/4 & -1/2 & -1/4 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & -1 & 0 & 1 \\
0 & -1 & 0 & 1 \\
0 & -1 & 0 & 1 \\
0 & -1/2 & -1/4 & 1 \\
0 & 0 & 0 & 1/3 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1/3 \\
1
\end{pmatrix},
\]
\[ B_1 = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\
\end{bmatrix}, \]

\[ \gamma = \left( \frac{1}{2}, 1, \frac{3}{2}, 1, \frac{1}{4}, 1/2, \frac{3}{4}, 1, \frac{5}{4}, 1, 1 \right)^T. \]
3.1 Definitions.

So far we have considered methods of the form $A_h y^* = hB_h f^* + g$ where $A_h$ and $B_h$ are both lower triangular. There are also classes of methods however which may be represented in this way, but with non-zero elements above the diagonal in these matrices.

Definition 3.1.1.

A method is $m$-BLOCK IMPLICIT if $\gamma \in \mathbb{R}^m$ is given and it may be represented by the array $(A_0, A_1, \ldots, A_{\tau} | B_0, B_1, \ldots, B_{\tau} | \gamma)$ as (2.1.4) with at least one of $A_{\tau}$ and $B_{\tau}$ not lower triangular.

Again we shall have $A_h y^* = hB_h f^* + g$ and if $A_{\tau}$ is not lower triangular we may premultiply through by

\[
\begin{bmatrix}
I_r & A_{\tau}^{-1} & A_{\tau}^{-1} & \cdots & A_{\tau}^{-1} \\
A_{\tau} & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\end{bmatrix}
\]

(3.1.1) to reduce $A_h$ to lower triangular form, leaving $B_h$ not lower triangular. Thus for any $m$-block implicit method we may consider only $B_h$ to have non-zero elements above the diagonal, without loss of generality.

3.2 Advanced Linear Multistep Methods.

This class of methods has been studied by de Hoog and Williams (26) and more recently by McKee and Pitcher (17).

Example 3.2.1.

Consider Clippinger and Dimsdale's method (14):
\[ -\frac{1}{2}y_{i+1} + y_i - \frac{1}{2}y_{i-1} = \frac{h}{4}(f_{i-1} - f_{i+1}), \]
\[ y_{i+1} - y_i - \frac{h}{3}(\bar{f}_{i-1} + 4\bar{f}_i + \bar{f}_{i+1}). \]

Having computed \( y_{i-1} \) we have implicit equations to solve for \( y_i \) and \( y_{i+1} \) together. This is a 2-block implicit method with array

\[
\begin{pmatrix}
0 & -1/2 & 1 & -1/2 \\
0 & -1 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
0 & 1/4 & 0 & -1/4 & 1 \\
0 & 1/3 & 4/3 & 1/3 & 2
\end{pmatrix}
A_1^{-1} = \begin{pmatrix}
1 & 1/2 \\
0 & 1
\end{pmatrix},
\]

so premultiplying through by the matrix \((3.1.1)\) we have also the array

\[
\begin{pmatrix}
0 & -1 & 1 & 0 & 0 & 5/12 & 2/3 & -1/12 & 1 \\
0 & -1 & 0 & 1 & 0 & 1/3 & 4/3 & 1/3 & 2
\end{pmatrix}
(3.2.1).
\]

3.3 Implicit Runge-Kutta Methods.

An implicit \( \mu \)-stage Runge-Kutta method is defined by

\[
y_i - y_{i-1} = h \sum_{j=1}^{\mu} b_j k_j,
\]
\[
k_j = f(t_{i-1} + h a_j, y_{i-1} + h \sum_{q=1}^{\mu} b_{jq} k_q), j = 1(1) \mu,
\]

where \( a_j = \sum_{q=1}^{\mu} b_{jq}, j = 1(1) \mu. \)

This is a \((\mu+1)\)-block implicit method with array

\[
\begin{pmatrix}
\ldots & \ldots & 0 & -1 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & 0 & -1 \\
0 & \ldots & 0 & 0 & -1
\end{pmatrix} I_{\mu+1}
\begin{pmatrix}
\ldots & \ldots & \ldots & \ldots & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \ldots & 0 & 0 & 0 & a_1 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & \ldots & 0 & 0 & 0 & 0 & a_{\mu} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & \ldots & 0 & 0 & 0 & 0 & 0 & a_{\mu+1} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & \ldots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
(3.3.2).
\]

Example 3.3.1.

Lambert (15, p.153) quotes the following 2-stage method:

\[
y_i - y_{i-1} = \frac{1}{2}h(k_1 + k_2),
\]
\[
k_1 = f(t_{i-1} + \left(\frac{1}{2} + \frac{\sqrt{3}}{6}\right)h, y_{i-1} + \frac{h}{4} k_1 + (\frac{1}{4} + \frac{\sqrt{3}}{6})h k_2),
\]

\[ k_2 = f(t_{i-1} + \frac{\sqrt{3}}{2}h, y_{i-1} + \frac{h}{4} k_1 + (\frac{1}{4} + \frac{\sqrt{3}}{6})h k_2). \]
\[ k_2 = f(t_{i-1} + (\frac{1}{2} - \frac{\beta}{6})h, y_{i-1} + (\frac{1}{4} - \frac{\beta}{6})hk_1 + \frac{h}{4}k_2), \]

for which the array is

\[
\begin{bmatrix}
0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 & \frac{1}{4} & (\frac{1}{4} + \frac{\beta}{6}) & 0 & 1 & \frac{1}{2} + \frac{\beta}{6} \\
0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & (\frac{1}{4} - \frac{\beta}{6}) & \frac{1}{4} & 0 & 1 & \frac{1}{2} - \frac{\beta}{6} \\
0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 1
\end{bmatrix}.
\]

3.4 One Leg Methods.

If we consider the equation \( y' = f(y) \), a one leg method is

\[ \sum_{j=0}^{k} \alpha_{k-j}y_{i-j} = hf(\sum_{j=0}^{k} \alpha_{k-j}y_{i-j}), \]

which may be written as

\[ \hat{y}_{i-k} = \sum_{j=0}^{k} \alpha_{k-j}y_{i-j}, \]

\[ \sum_{j=0}^{k} \alpha_{k-j}y_{i-j} = hf(\hat{y}_{i-k}), \]

array

\[
\begin{bmatrix}
-\alpha_0 & -\alpha_1 & \cdots & -\alpha_{k-1} & \alpha_k & 0 & \cdots & 0 & 0 & 0 & 1 \\
\alpha_0 & \alpha_1 & \cdots & \alpha_{k-1} & 0 & \cdots & 0 & 1 & 0 & 1
\end{bmatrix}.
\]

Example 3.4.1.

Dahlquist (8) quotes the method \( y_1 - y_{i-1} = hf(\frac{1}{2}y_1 + \frac{1}{2}y_{i-1}), \)

which has the array \( \begin{bmatrix} 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \end{bmatrix} \).

3.5 Collocation Method.

A collocation method consists of finding \( u(t) \in \mathbb{P}_\mu \) such that

\[ u(0) = y_0 \quad \text{and} \quad u'(a_jh) = f(a_jh, u(a_jh)), \]

\( v = 1(1)\mu \), and then putting

\[ y_1 = u(h) \quad (3.5.1) \].

We then proceed in the same way to obtain \( y_2 \),

and so on. Nørsett (20) has shown that this is equivalent to the

implicit Runge-Kutta method (3.3.2) with

\[ b_{\nu q} = \int_{0}^{a_\nu} q(\eta)d\eta \]

and
\[ a_{1}, a_{2}, \ldots, a_{\mu} \] are the nodes of the collocation and we have a \((\mu+1)\)-block implicit method. The equivalence of the collocation and the Runge-Kutta method is in the sense that the Runge-Kutta method gives a discrete set of solutions which coincide exactly with the values taken by the continuous collocation polynomial at the points \(0, a_{1} h, a_{2} h, \ldots, a_{\mu} h, h\).

Consider now a Galerkin method, which consists of finding \(u(t) \in \mathcal{P}_{\mu}\) such that \(u(0) = y_{0}\) and

\[
\int_{0}^{h} [u'(\eta) - f(\eta, u(\eta))] \alpha_{\nu}(\eta) \, d\eta = 0
\]

for \(\nu = 0(1)\mu\ (3.5.2)\), where the \(\alpha_{\nu}(\eta)\) are \(\nu^{th}\) degree polynomial basis functions and \(\mu \geq 1\). We then put \(y_{1} = u(h)\) and then obtain \(y_{2}\) by the same procedure, and so on. The integral in (3.5.2) is replaced by a quadrature formula,

\[
\int_{0}^{h} \varphi(\eta) \, d\eta \approx h \sum_{\nu=1}^{\mu} w_{\nu} \varphi(a_{\nu} h),
\]

where \(0 < a_{1} < a_{2} < \ldots < a_{\mu} < 1\) and \(w_{\nu} \neq 0\) for each \(\nu\).

Hulme (12) showed that this Galerkin method is equivalent to the collocation given above, and hence to an implicit Runge-Kutta method. Hence Galerkin methods may also be considered within the framework of \(m\)-block implicit methods.

In an interpolatory quadrature method (12) we find \(u(t) \in \mathcal{P}_{\mu}\) such that \(u(0) = y_{0}\) and

\[
u(0) + b \sum_{j=1}^{\mu} b_{\nu q} f(a_{q} h, u(a_{q} h)) \text{ for } \nu = 1(1)\mu,
\]

where

\[
b_{\nu q} = \int_{0}^{a_{\nu}} l_{q}(\eta) \, d\eta \text{, and then put } y_{1} = u(h), \text{ and then proceed in the same way to find } y_{2} \text{ and so on.}
\]
Hulme (12) showed that this method too is equivalent to the collocation method given, so that interpolatory quadrature methods may also be considered in our formulation.

Consider now a perturbed collocation method (20), for which we have a pre-assigned matrix

\[
\hat{P} = \begin{pmatrix}
P_{01} & \cdots & P_{0,\mu} \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0
\end{pmatrix},
\]

\[
N_j(t) = \frac{1}{j!} \sum_{i=0}^{\mu} (p_{ij} - \delta_{ij}) t^i, \quad p_{\mu j} = 0 \text{ for } j = 1(1)\mu, \quad \text{where } \delta_{ij} \text{ is the Kronecker } \delta,
\]

and for any \( u(t) \in \mathcal{P}_{\mu} \),

\[
(Pu)(t) = u(t) + \sum_{j=1}^{\mu} N_j(t) h^j u(j)(0) \in \mathcal{P}_{\mu}.
\]

A general perturbed collocation (PECO) method consists of finding \( u(t) \in \mathcal{P}_{\mu} \) such that \( u(0) = y_0 \) and \( u'(a \nu h) = f(a \nu h, (Pu)(a \nu h)) \) for \( \nu = 1(1)\mu \), and then putting \( y_1 = u(h) \). We proceed in the same way to obtain \( y_2 \) and so on.

Nørsett (20) showed that PECO is equivalent to the implicit Runge-Kutta method (3.3.1) with

\[
b_{\nu q} = \int_0^a l_q(\eta) d\eta + \sum_{\alpha=1}^{\infty} N_\alpha(a \nu l_{q}^{(\alpha-1)}(0) \text{ and } c_\nu = \int_0^1 l_{\nu}(\eta) d\eta.
\]
CHAPTER FOUR

BLOCK METHODS

Having seen the wide variety of methods which fall under the block semi explicit and block implicit umbrellas, we may define the class of block methods as the union of the block semi explicit and block implicit methods. Each block method then is characterised by an array \((A_0, A_1, \ldots, A_T \mid B_0, B_1, \ldots, B_T \mid \gamma)\).

4.1 Truncation Error

From now on the symbols \(C\) and \(M\), with or without subscripts or superscripts, will be assumed to be real, strictly positive constants, independent of \(h\), unless otherwise stated.

Definition 4.1.1

The LOCAL TRUNCATION ERROR VECTOR \(\Theta^* \in \mathbb{R}^{r+nm}\) is defined for an \(m\)-block method by \(\Theta^* = A_n \Delta_n y(t) - hB_n \Delta_n y'(t) - g\).

Definition 4.1.2

The \(r\) starting values \(\tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_{r-1}\) are ACCURATE of order \(p \in \mathbb{N}\) if there exist \(C'_0, C'_1, \ldots, C'_{r-1}\) such that \(\Theta^*_i = C'h^p + O(h^{p+1})\) for \(i = O(1) r-1\).

Definition 4.1.3

An \(m\)-block method is CONSISTENT of order \(p \in \mathbb{N}\) if there exist \(c_1, c_2, \ldots, c_m\), some of which may be zero, such that for every \(y(t) \in C^{(p+1)}[0, T]\), \(\Theta^*_{r+sm+v-1} = c_v (\Delta_n y^{(p+1)}(t))_{r+sm} h^p + O(h^{p+2})\) as \(h \to 0\), for \(v = 1(1)m\) and \(s = O(1)n-1\).

For block methods zero stability is defined as for cyclic methods.

Definition 4.1.4

An \(m\)-block method is CONVERGENT of order \(p \in \mathbb{N}\) if there exists \(C_U\) such that \(\|e\|_\infty \leq C_U h^p\), where the GLOBAL ERROR VECTOR \(e\) is \(y^* - \Delta_n y(t)\).
For each $\nu = 1(1)m$ we refer to the $\nu^{th}$ stage of an $m$-block method as corresponding to the components $y^*_{r+s m+\nu-1}$, $s = 0(1)n^{-1}$.

4.2 The Numerical Stability of Block Semi Explicit Methods.

For numerical stability we consider the differential equation $y' = \lambda y$ for some $\lambda < 0$. A method is absolutely stable for $h = h\lambda$ if $y^*_{r+s m-1}$ is decreasing as $n \to \infty$. Now for an acceptable method it is desirable to have $A$-stability, which is the term given to denote the situation where we have absolute stability for all complex values of $h$ with negative real part.

Any block semi explicit method may be written $A_h y^* = hB_h f^* + g = h\lambda B_h y^* + g$, so $y^* = (A_h - \bar{h}B_h)^{-1} g$. Hence the solution is decreasing if there exists $M^*$ such that $\| (A_h - \bar{h}B_h)^{-1} \|_\infty \leq M^*$.

Now $A_h$ and $B_h$ have recurring arrays like those for cyclic methods (1.3.2), so we may define the companion matrices $G_1, G_2, \ldots, G_m$ and the matrix $G$ for $A_h - \bar{h}B_h$ as in definition 1.3.4. We now need the following theorem.

Theorem 4.2.1.

If the elements of $(A_h - \bar{h}B_h)$ are independent of $n$, then there exists $M^*$ such that $\| (A_h - \bar{h}B_h)^{-1} \|_\infty \leq M^*$ if and only if $G$ has all its eigenvalues strictly inside the unit circle.

This result can be found in (18). Hence we obtain the following theorem.

Theorem 4.2.2.

An $m$-block semi explicit method is absolutely stable for some $\bar{h}$ if the eigenvalues of $G$ all lie strictly within the unit circle.
Definition 4.2.3.

For the linear multistep method \((1.2.1)\) the FIRST and SECOND CHARACTERISTIC POLYNOMIALS are respectively

\[ \phi(\xi) = \sum_{j=0}^{k} \alpha_{k-j} \xi^{k-j} \quad \text{and} \quad \sigma(\xi) = \sum_{j=0}^{k} \beta_{k-j} \xi^{k-j}. \]

Definition 4.2.4.

For the linear multistep method \((1.2.1)\) the STABILITY POLYNOMIAL is \(\pi(\xi, \bar{h}) = \phi(\xi) - \bar{h} \sigma(\xi).\)

Definition 4.2.5.

For an \(m\)-block semi explicit method the stability polynomial is

\[ \pi(\xi, \bar{h}) = \det(G - \xi I), \]

where \(G\) is as in theorem 4.2.2.

For a given \(\bar{h}\) a method is absolutely stable if all the roots of the equation \(\pi(\xi, \bar{h}) = 0\) are less than one in modulus. In this section we shall consider the absolute stability of predictor-corrector methods, obtaining the stability polynomials given by Lambert (15, pp. 97, 98), and shall then examine the stability of certain extrapolation methods.

Proposition 4.2.6.

For the predictor-corrector scheme \((2.3.1)\) in \(P(EC)^{\mu_E}\) mode the stability polynomial is

\[ \pi_{P(EC)^{\mu_E}}(\xi, \bar{h}) = \phi(\xi) - \bar{h} \sigma(\xi) + \mu(\bar{h})[\phi^*(\xi) - \bar{h} \sigma^*(\xi)], \]

where

\[ \mu(h) = \frac{(h \beta)^\mu(1 - h \beta^\mu)}{1 - (h \beta)^\mu} \]

and \(\phi^*(\xi)\) and \(\sigma^*(\xi)\) are the stability polynomials of the predictor and corrector respectively.

Proof.

\(A_{n} - \bar{h} B_{n}\) is composed of the repeated array
where without loss of generality we take \( \alpha_k = \alpha_k = 1 \).

\[
G_1 = \begin{pmatrix}
0 & 1 & 1 & 1 \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & -\alpha + \bar{h}_0 \\
\mu & \mu & \mu & \mu
\end{pmatrix}
\]

\[
G_\nu = \begin{pmatrix}
0 & 1 & 1 & 1 \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & -\alpha + \bar{h}_0 \\
\mu - \nu + 1 & \mu & \mu & \mu
\end{pmatrix}
\]

\( \nu = 2(1)^{\mu+1} \). By direct computation we obtain \( G = G_{\mu+1} \mu \ldots G_2 G_1 = \ldots \).
a $(\mu+1)\times(\mu+1)$ matrix, and the $\gamma$'s are defined by the relations

$$\gamma_{10} = -\alpha_i^* + \bar{h}_1^*,$$
$$\gamma_{1,j+1} = -\alpha_i + \bar{h}_1 + \bar{h}_k \gamma_{ij}, \quad i = 0(1)\mu - 1.$$

Now using a result on pp. 33, 34 of (25), we have

$$\Phi(\xi, \eta) = \det(G - \xi I) \propto \det(\Gamma_0 + \Gamma_1 \xi + \ldots + \Gamma_{k-1} \xi^{k-1} - \xi^k) \quad (*)$$

Now by direct computation

$$\gamma_{1\mu} = [1 + \bar{h}_1 + (\bar{h}_1^*)^2 + \ldots + (\bar{h}_1^*)^{\mu-1}][(-\alpha_i + \bar{h}_1^*) + (\bar{h}_1^*)^\mu(-\alpha_i + \bar{h}_1^*)]$$

$$\gamma_{1\mu} = \frac{1 - (\bar{h}_1^*)^\mu}{1 - \bar{h}_1} = \frac{1 - (\bar{h}_1^*)^\mu}{\alpha_i + \bar{h}_1 + \mu(\bar{h})(-\alpha_i + \bar{h}_1^*)}.$$

(*) By "det$F(\xi) \propto$ det$G(\xi)$" we mean that det$F(\xi) = k\xi^q$det$G(\xi)$ for some constant $k$ and integer $q$. This implies that the solutions of the equation det$F(\xi) = 0$ are $\xi = 0$ $q$ times and the zeroes of det$G(\xi)$.
Hence \( \Pi \text{EC}^{(\mu)}_{\mu}(\zeta, h) \approx \xi^k \left( \frac{1 - (\tilde{h}\beta_k \xi)}{1 - \tilde{h}\beta_k} \right)^{k-1} \sum_j \left[ (-\alpha^*_j + \tilde{h}\beta_j) + M_{\mu}(\tilde{h})(\alpha^*_j + h\beta_j) \right] \xi^j \)

\[
= \xi^k \left( \frac{1 - (\tilde{h}\beta_k \xi)}{1 - \tilde{h}\beta_k} \right)^{k-1} \left[ (1 - \tilde{h}\beta_k \xi)^{-1} \xi^k \xi + \tilde{h}\sigma(\xi) \right. \\
\left. + M_{\mu}(\tilde{h})(\xi^k - \tilde{h}\sigma(\xi) + \tilde{h}\sigma^*(\xi)) \right] \\
= \frac{1 - (\tilde{h}\beta_k \xi)}{1 - \tilde{h}\beta_k} \left[ \xi(\xi^k - \tilde{h}\sigma(\xi) + M_{\mu}(\tilde{h})(\xi^k - \tilde{h}\sigma^*(\xi)) \right],
\]
and so the result follows.

**Proposition 4.2.7.**

For the predictor-corrector scheme (2.3.1) in \( \text{EC}^{(\mu)} \) mode the stability polynomial is \( \Pi \text{EC}^{(\mu)}_{\mu}(\zeta, h) = \beta_k \xi^k \left[ \xi(\xi^k - \tilde{h}\sigma(\xi)) \right. \\
\left. + M_{\mu}(\tilde{h})(\xi^k - \tilde{h}\sigma^*(\xi)) \right] \)

**Proof.**

Again taking \( \alpha^*_k = \alpha_k = 1 \) we now have in \( A_h - \tilde{h}B_h \) the repeated array

\[
\begin{array}{cccccc}
0 & \cdots & 0 & -\tilde{h}\beta_0 & \alpha^*_0 \\
\vdots & & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & -\tilde{h}\beta_1 & \alpha^*_1 \\
0 & \cdots & 0 & -\tilde{h}\beta_k & \alpha^*_k \\
\end{array}
\]

and this gives
\( G_1 = \begin{bmatrix}
0 & 1 & 1 & & & & & \\
\vdots & & & & & & & \\
0 & \cdots & 0 & -\alpha^* & 0 & \cdots & 0 & -\alpha^* & \cdots & 0 & \cdots & 0 & -\alpha^* & 1 \\
\end{bmatrix}
\)

\( G_\nu = \begin{bmatrix}
0 & 1 & 1 & & & & & \\
\vdots & & & & & & & \\
0 & \cdots & 0 & \bar{h}_0 & -\alpha & 0 & \cdots & 0 & \bar{h}_1 & -\alpha & \cdots & 0 & \cdots & 0 & \bar{h}_k & -\alpha_{k-1} \\
\end{bmatrix}
\)

\( \nu = 2(1)\mu+1 \).

Hence by direct computation \( G \) is the \([1 + k(\mu+1)] \times [1 + k(\mu+1)]\) matrix whose zero \( \nu \)-th column is composed entirely of zeroes, whose zero \( \mu \)-th row is composed of zeroes, except the \((\mu+1)\)-th element, which is 1, and whose remaining rows and columns are

\[
\begin{pmatrix}
0 & \mu+1 & I & \mu+1 \\
0 & \mu+1 & 0 & \mu+1 & \mu+1 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \mu+1 & \cdots & 0 & \mu+1 & I & \mu+1 \\
\end{pmatrix}
\]

where \( \Gamma^*_i \) is the \((\mu+1) \times (\mu+1)\) matrix

\[
\begin{pmatrix}
0 & \cdots & 0 & \delta_{i0} & \epsilon_{i0} \\
0 & \cdots & 0 & \delta_{i1} & \epsilon_{i1} \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & \delta_{i\mu} & \epsilon_{i\mu} \\
\end{pmatrix}
\]
and $\delta_{i,j} = h_{i}^{*} \cdot \delta_{i,j+1} = h_{i}^{*} + h_{i,k} \cdot \delta_{i,j}$.

Therefore again using the result on pp. 33, 34 of (25), we have

$$\Pi_{(BC)}^{\mu}(\xi, \eta) \propto \det \left( \Gamma_{1}^{*} + \ldots + \Gamma_{k-1}^{*} \xi_{1}^{k-1} - \Gamma_{k}^{*} \right)$$

$$= \begin{bmatrix} -\xi_{1}^{k} & x_{0} & z_{0} & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ x_{\mu-1} - \xi_{1}^{k} & \cdots & \cdots & x_{\mu} & z_{\mu} & -\xi_{1}^{k} \\ z_{\mu} & \cdots & \cdots & \cdots & \cdots & \cdots \\ \end{bmatrix},$$

where for $j = 0(1) \mu$, $x_{j} = \delta_{0,j} + \xi_{1,j} + \ldots + \xi_{k-1,j} \delta_{k-1,j}$.

and $z_{j} = \varepsilon_{0,j} + \varepsilon_{1,j} + \ldots + \varepsilon_{k-1,j} \varepsilon_{k-1,j}$.

Therefore $\Pi_{(BC)}^{\mu}(\xi, \eta) \propto (-\xi_{1}^{k})^{\mu-1} \left[ (\xi_{1}^{k} - x_{\mu-1})(\xi_{1}^{k} - z_{\mu}) - z_{\mu-1} x_{\mu} \right]^{(4.2.1)}$.

Now by direct computation,

$$\delta_{i,\mu-1} = h_{i}^{*} \frac{1 - (h_{i}^{*})^{\mu-1}}{1 - h_{i}^{*}} + (h_{i}^{*})^{\mu-1} \cdot (h_{i}^{*})^{\mu-1}$$

$$\delta_{i,\mu} = h_{i}^{*} \frac{1 - (h_{i}^{*})^{\mu}}{1 - h_{i}^{*}} + (h_{i}^{*})^{\mu} \cdot (h_{i}^{*})^{\mu}$$

$$\varepsilon_{i,\mu-1} = -\varepsilon_{i} \frac{1 - (h_{i}^{*})^{\mu-1}}{1 - h_{i}^{*}} - \varepsilon_{i}^{*} (h_{i}^{*})^{\mu-1}$$

$$\varepsilon_{i,\mu} = -\varepsilon_{i} \frac{1 - (h_{i}^{*})^{\mu}}{1 - h_{i}^{*}} - \varepsilon_{i}^{*} (h_{i}^{*})^{\mu}.$$

Hence $\xi_{1}^{k} - x_{\mu-1} = \xi_{1}^{k} \frac{1 - (h_{i}^{*})^{\mu}}{1 - h_{i}^{*}} - h_{i}^{*} (\xi_{1}^{k} - x_{\mu-1}) = \xi_{1}^{k} \cdot (h_{i}^{*})^{\mu-1} h_{i}^{*} (\xi_{1}^{k})$,

$$\xi_{1}^{k} - z_{\mu} = \frac{1 - (h_{i}^{*})^{\mu}}{1 - h_{i}^{*}} [\xi_{1}^{k} - h_{i}^{*} (\xi_{1}^{k}) + (h_{i}^{*})^{\mu} (\xi_{1}^{k})]$$. 
By substituting these last four equations into (4.2.1) we obtain the result.

We now go on to examine the absolute stability of some extrapolation methods considered by Dahlquist (7), beginning with example 2.6.1, for which $A_h - hB_h$ has the recurring array

$$
\begin{array}{cccc}
0 & 0 & -1 - \frac{h}{4} & 1 - \frac{h}{4} \\
0 & 0 & -1 - \frac{h}{4} & 1 - \frac{h}{4} \\
-1 - \frac{h}{2} & 0 & 0 & 1 - \frac{h}{2} \\
0 & -\sqrt{3} & 1/3 & 1 \\
\end{array}
$$

giving $G_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \frac{4+h}{4-h} \\ \frac{4/3}{4-h} \end{pmatrix}$, $G_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \frac{4+h}{4-h} \\ \frac{2+h}{2-h} \end{pmatrix}$, $G_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \frac{4+h}{4-h} \end{pmatrix}$, and $G = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & \frac{4+h}{4-h} \end{pmatrix}$.

Therefore the stability polynomial is $\zeta^3 - \zeta^2 \left[ \frac{4(4+h)}{4-h} - \frac{1}{3} \right]$, and so we have absolute stability for those $h$ for which $\left| \frac{4(4+h)}{4-h} - \frac{1}{3} \right| < 1$.

Unfortunately $\lim_{x \to -\infty} \left[ \frac{4(4+h)}{4-x} - \frac{1}{3} \right] = \frac{5}{3}$, so we do not have $A$-stability.

This then is the outcome when we use the extrapolated $y_{i-1}$ to compute $y_i$.

If instead we use one of the non-extrapolated $y_{i-1}^{[0]}$ or $y_{i-1}^{[1]}$, we do get $A$-stability, as we shall now see.
Example 4.2.7.

Using \( y_{i-1}^{[0]} \) the scheme is

\[
\begin{align*}
   y_{i-1}^{[0]} - y_{i-1}^{[0]} &= \frac{h}{4} f_{i-1}^{[1]} + f_{i-1}^{[1]} + f_{i-1}^{[0]}, \\
   y_{i}^{[0]} - y_{i-1}^{[0]} &= \frac{h}{4} f_{i}^{[1]} + f_{i-1}^{[0]}, \\
   y_{i}^{[1]} - y_{i}^{[0]} &= \frac{h}{2} f_{i}^{[1]} + f_{i-1}^{[0]}, \\
   y_{i} + 3 y_{i}^{[1]} - \frac{4}{3} y_{i}^{[0]} &= 0.
\end{align*}
\]

This has the array

\[
\begin{pmatrix}
   0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & 1/4 & 0 & 0 & 0 & \frac{1}{2} \\
   0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 1 \\
   0 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 1/2 & 0 \\
   0 & 0 & 0 & 0 & -1/3 & 1/3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix},
\]

and so in \( A_h - \tilde{h} B_h \) we have the recurring array

\[
\begin{pmatrix}
   0 & 0 & -1 & -\frac{h}{4} & 0 & 0 & 1 & -\frac{h}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 1 \\
   0 & 0 & 0 & 0 & -1 & \frac{h}{4} & 1 & 1 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 1 \\
   -1 & -\frac{h}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 1 \\
   0 & 0 & 0 & -1/3 & 1/3 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix},
\]

from which we obtain \( G =
\[
\begin{pmatrix}
   0 & 0 & 0 & 0 & 0 & 0 & 1 \\
   0 & 0 & \frac{4+\tilde{h}}{4-\tilde{h}} & 0 & 0 & 0 \\
   0 & 0 & 0 & \frac{4+\tilde{h}}{4-\tilde{h}} & 0 & 0 \\
   0 & 0 & \frac{2+\tilde{h}}{2-\tilde{h}} & 0 & 0 \\
   0 & 0 & \frac{4+\tilde{h}}{3(4-\tilde{h})} - \frac{1}{2} & 0 & 0 \\
\end{pmatrix},
\]

and the stability polynomial is \( \zeta^5 - \frac{4}{5} \left( \frac{4+\tilde{h}}{4-\tilde{h}} \right)^2 \). Since \( \left| \frac{4+\tilde{h}}{4-\tilde{h}} \right| < 1 \) for all complex \( \tilde{h} \) with negative real part, we have A-stability.

Example 4.2.8.

Using \( y_{i-1}^{[1]} \) and \( f_{i-1}^{[1]} \) instead of \( y_{i-1}^{[0]} \) and \( f_{i-1}^{[0]} \) we have the array
\[
\begin{pmatrix}
0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
0 & 1/4 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\
0 & 0 & 0 & 1/4 & 1/4 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 1
\end{pmatrix}
\]

and so in \( A_h - \frac{h}{h} \), we have the recurring array

\[
\begin{pmatrix}
0 & 0 & -1 & -\frac{h}{4} & 0 & 1 & -\frac{h}{4} \\
0 & 0 & 0 & -1 & -\frac{h}{4} & 1 & -\frac{h}{4} \\
-1 & -\frac{h}{2} & 0 & 0 & 0 & 0 & 1 & -\frac{h}{2} \\
0 & 0 & -\frac{h}{3} & 1/3 & 1 &
\end{pmatrix}
\]

from which we obtain \( G = \)

\[
\begin{pmatrix}
0 & 0 & \frac{h}{4} + \frac{h}{2} & 0 \\
0 & 0 & \frac{h}{4} & \frac{h}{2} \\
0 & 0 & 2h & \frac{h}{2} \\
0 & 0 & \frac{h}{3} + \frac{h}{2} & 0 & \frac{h}{3} & \frac{2h}{3} & \frac{h}{3} & \frac{2h}{3} & 0
\end{pmatrix}
\]

so that the stability polynomial is \( z^4 - \frac{3}{2} \left( \frac{2+3h}{2-h} \right) \). Since \( \left| \frac{2+h}{2-h} \right| < 1 \) for all complex \( h \) with negative real part, we have \( A \)-stability.

4.3 Zero Stability.

We recall that for block methods, zero stability is defined as for cyclic methods.

Let us first consider Runge-Kutta methods. For both the explicit case (2.4,2) and the implicit case (3.3,2)

\[
A_o = \begin{pmatrix} 0 & \cdots & 0 & -1 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & -1 \end{pmatrix} \quad \text{and} \quad A_1 = I_m.
\]

The companion matrices \( (1,3,3) \) are \( G = \)

\[
\begin{pmatrix}
0 & 1 & & & \\
& \ddots & \ddots & & \\
& & \ddots & \ddots & \\
0 & \cdots & 0 & 1 & \cdots & 0
\end{pmatrix}
\]
\[ V = \begin{pmatrix} 1(1) & \ldots & 0 \end{pmatrix} m, \text{ and hence } \mathbf{G} = \begin{pmatrix} 0 & \ldots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & 1 \end{pmatrix}, \] which has for eigenvalues \((m-1)\) zeroes and 1, and so any Runge-Kutta method is zero stable.

We now consider the zero stability of the predictor-corrector schemes whose numerical stability we looked at in section 4.2.

The matrix \( \mathbf{G} \) for zero stability may be obtained from the corresponding matrix for absolute stability by putting \( h = 0 \), since for absolute stability we demand that \( \| (A_h - \bar{h}B_h)^{-1} \|_{\infty} \leq M^* \), while for zero stability we stipulate \( \max_{i,j} |(A_h^{-1})_{i,j}| = O(1) \). Accordingly if we take the stability polynomial of a given method and put \( h = 0 \), we may test for zero stability by requiring that the resulting polynomial have none of its roots with modulus greater than one, and with those roots of modulus one having linear elementary divisors.

Doing this for \( P(\text{EC})^E \) mode we obtain from proposition 4.2.6 the polynomial \( C(k) \), and so the predictor-corrector scheme is zero stable if and only if the corrector as a linear multistep method is. From proposition 4.2.7 we obtain for \( P(\text{EC})^E \) mode the polynomial \( \beta_k C(k) \), so this scheme is also zero stable if and only if the corrector is.

4.4 Convergence Theorems.

In this section we provide convergence proofs for block methods based on definition 4.1.3 of consistency.

Definition 4.4.1.

For any vectors \( u, v \in \mathbb{R}^{r+nm} \) we define the vector \( f(u, v) \in \mathbb{R}^{r+nm} \) by \( [f(u, v)]_i = f(u_i, v_i) \) for all \( i \).
Thus $\Delta_h y'(t) = \Delta_h f(t, y(t)) = f(\Delta_h t, \Delta_h y(t))$ and $f^* = f(\Delta_h t, y^*)$.

**Theorem 4.4.2.**

If an $m$-block semi explicit method is zero stable, consistent of order $p$ and starting values are accurate of order $p$, then the method is convergent of order $p$.

**Proof.**

The method may be written $A_h y^* = hB f(\Delta_h t, y^*) + g$, and by definition 4.1.1 we have $A_h \Delta_h y(t) = hB f(\Delta_h t, \Delta_h y(t)) + g - \Theta^*$.

By subtraction we obtain $A_h e = \Theta^* + hB [f(\Delta_h t, y^*) - f(\Delta_h t, \Delta_h y(t))]$,

so $e = A_h^{-1} \Theta^* + hA_h^{-1} B_h [f(\Delta_h t, y^*) - f(\Delta_h t, \Delta_h y(t))]$.

Therefore for all $i$, by the triangle inequality and applying the Lipschitz condition (1.1.2) to $f$, $|e_i| \leq |[A_h^{-1} \Theta^*]_i| + hL \sum_{j=0}^{i} |(A_h^{-1} B_h)_{ij}| |e_j|$.

The summation is up to $i$ since $A_h$ and $B_h$ are both lower triangular and therefore so is $A_h^{-1} B_h$.

Now zero stability implies that there exists $M_o$ such that

$$\max_{0 \leq i, j \leq r+nm-1} \left| (A_h^{-1})_{ij} \right| = M_o.$$ Since $A_h$ is lower triangular, so is $A_h^{-1}$, and therefore $|[A_h^{-1} \Theta^*]_i| \leq M_o \sum_{j=0}^{i} |\Theta^*_j| = M_o \left[ \sum_{j=0}^{r-1} c_j h^p + \sum_{j=r}^{i} |\Theta^*_j| \right] = O(h^p)$ for all $i$, where $\sum_{j=r}^{i} = 0$ for $i < r$.

$$|(A_h^{-1} B_h)_{ij}| \leq \sum_{q=0}^{r+nm-1} \left| (A_h^{-1})_{iq} \right| \left| (B_h)_{qj} \right|$$ by the triangle inequality,

$$\leq M_o \sum_{q=0}^{r+nm-1} \left| (B_h)_{qj} \right| \leq M_o \|B_h\|_1 \quad (4.4.1),$$

which is bounded since all elements of $B_h$ are finite and $B_h$ has bandwidth independent of $h$. Therefore there exist $C'$ and $M'$ such that
\begin{align*}
|e_1| &\leq C'h^P + hM' \sum_{j=0}^{i-1} |e_j| + hM' |e_1| , \text{ therefore} \\
(1 - hM')|e_1| &\leq C'h^P + hM' \sum_{j=0}^{i-1} |e_j|. \text{ So for sufficiently small h there exist} \\\nC \text{ and } M \text{ such that} \\
|e_1| &\leq Ch^P + hM \sum_{j=0}^{i-1} |e_j| .
\end{align*}

We obtain the result \( \|e\|_\infty \leq C'h^P \) for some \( C \) by making use of the following lemma .

Lemma 4.4.3 .

Let \( x = (x_0, x_1, \ldots, x_n)^T \) be a real vector with
\[ |x_i| \leq \delta + hM \sum_{j=0}^{i-1} |x_j| , \text{ i = 1(1)W} , \text{ where M and \( \delta \) are positive constants .} \]

Then \( \|x\|_\infty \leq \left(ah |x_0| + \delta \right) \exp\left(M \omega h \right) \).

The proof of this lemma is elementary .

Theorem 4.4.4 .

If an \( m \)-block implicit method is zero stable, consistent of order \( p \) and starting values are accurate of order \( p \), then it is convergent of order \( p \).

Proof .

We may take \( A_h \) to be lower triangular without loss of generality .

As in theorem 4.4.2 , \( e = A_h^{-1} \Theta^* + hA_h^{-1}B_h [ f(A_h t, y^*) - f(A_h t, \Delta h y(t)) ] \).

Now \( B_h \) has the form (2.1.2) , where \( B \) is an \( m \times m \) matrix on the diagonal and is not lower triangular . Therefore due to the particular form of \( A^{-1} B_h \),
\[ |e_i| \leq |[A^{-1} \Theta^*]_i| + hL \sum_{j=0}^{i+m-v} |(A_h^{-1}B_h)_{ij}| |e_j| , \text{ where } v \equiv i-r \equiv 1 (\mod m) , \]
\[ 1 \leq v \leq m \text{ and } i \geq r . \text{ As in theorem 4.4.2 , } [A_h^{-1} \Theta^*]_i = O(h^P) \text{ for all } i , \text{ so we} \]
obtain, putting \( i = r + sm + \nu - 1 \),

\[
|e_{r+sm+\nu-1}| \leq C'h^p + hL \sum_{j=0}^{r+sm+\nu-1} \left| \left( A_{h}^{-1} B_{h} \right)_{ij} \right| e_j, \quad \nu = 1(1) m,
\]

\( s = O(1)n-1 \), for some \( C' \). Again \( \left| \left( A_{h}^{-1} B_{h} \right)_{ij} \right| \leq M_0 \|B_h\|_1 \), so there exists \( M' \) such that

\[
e_{r+sm+\nu-1} \leq C'h^p + hM' \sum_{j=0}^{r+sm+\nu-1} |e_j|, \quad \nu = 1(1) m, \quad s = O(1)n-1.
\]

Using lemma 4.4.5 below we can reduce this inequality to

\[
e_{r+sm+\nu-1} \leq Ch^p + hM \sum_{j=0}^{r+sm+\nu-1} |e_j|, \quad \nu = 1(1) m, \quad s = O(1)n-1 \text{ for sufficiently small } h,
\]

for some \( C \) and \( M \). That is,

\[
e_i \leq C'h^p + hM \sum_{j=0}^{i+m} |e_j|, \quad \text{which gives } \|e\|_\infty \leq C'h^p \text{ for some } C_0 \text{ as in theorem 4.4.2}.
\]

**Lemma 4.4.5.**

Take any positive integers \( i \) and \( m \) and suppose that the real vector \( (x_0, x_1, \ldots, x_{i+m})^T \) is such that for some real and positive \( M \) and \( \delta \)

\[
x_{i+m} \leq \delta + hM \sum_{j=0}^{i+m} |x_j| \quad \text{for } \nu = 1(1)m.
\]

Then for sufficiently small \( h \)

\[
x_{i+m-q} \leq \frac{hM}{1 - qhM} \sum_{j=0}^{i+m-q} |x_j| + \frac{\delta}{1 - qhM} \quad \text{for } q = O(1)m-1.
\]

**Proof.**

We shall proceed by an induction argument on \( m \). For \( m = 1 \) the result is immediate. Assume then that the lemma holds for \( m \). Then for \( m+1 \) we have

\[
x_{i+m} \leq \delta + hM \sum_{j=0}^{i+m+1} |x_j|, \quad \nu = 1(1)m+1 \quad (4.4.2).
\]

Therefore when \( \nu = m+1 \), \( |x_{i+m+1}| \leq \delta + hM \sum_{j=0}^{i+m+1} |x_j| \), and so for sufficiently small \( h \),
\[ |x_{i+m+1}| \leq \delta' + hM' \sum_{j=0}^{i+m} |x_j| , \text{ where } M' = M/(1 - hM) \geq 0 \text{ and } \]
\[ \delta' = \delta/(1 - hM) \geq 0 . \]
Therefore for \( \nu = 1(1)m \), eliminating \( |x_{i+m+1}| \) we obtain
\[ |x_{i+\nu}| \leq \delta + hM \sum_{j=0}^{i+m} |x_j| + hM(\delta' + hM' \sum_{j=0}^{i+m} |x_j|) \]
\[ \leq \delta + hM\delta' + hM(1 - hM') \sum_{j=0}^{i+m} |x_j| = \frac{\delta}{1 - hM} + \frac{hM}{1 - hM} \sum_{j=0}^{i+m} |x_j| \]
\[ \leq \delta' + hM' \sum_{j=0}^{i+m} |x_j| . \]
Therefore by the assumed hypothesis, for \( q = O(1)m-1 \) and for \( h \) sufficiently small,
\[ |x_{i+m-q}| \leq \frac{hM'}{1 - qhM} \sum_{j=0}^{i+m-q} |x_j| + \frac{\delta'}{1 - qhM} \]
\[ = \frac{\delta}{1 - (q+1)hM} + \frac{hM}{1 - (q+1)hM} \sum_{j=0}^{i+m-q} |x_j| . \]
That is,
\[ |x_{i+(m+1)-q}| \leq \frac{\delta}{1 - qhM} + \frac{hM}{1 - qhM} \sum_{j=0}^{i+(m+1)-q} |x_j| , \] \( q = 1(1)m \).

The remaining case \( q = 0 \) is simply \((4.4.2)\) .
CHAPTER FIVE.

TWO-SIDED ERROR BOUNDS.

5.1 Optimal Consistency.

Existing convergence proofs such as those of Stetter (21, pp. 220, 224, 226) and the previous section usually give an upper bound on the infinity norm of the global error vector of the form

\[ \| e \|_\infty \leq C_U h^p \quad (5.1.1) \]

This implies convergence of order at least p and holds provided the particular method is zero stable and consistent of order p and the r starting values are accurate of order p. Definition 4.1.3 of consistency however is not able to provide a two-sided error bound of the form

\[ C_L h^p \leq \| e \|_\infty \leq C_U h^p \quad (5.1.2) \]

as the following two examples will show.

Example 1.2.1.

\[ \theta_{2s+2}^* = -\frac{1}{90} h^5 y(t_{2s+2}) + O(h^6), \]

\[ \theta_{2s+3}^* = -\frac{1}{24} h^4 y(t_{2s+2}) + O(h^5), \quad s = O(1)n^{-1}. \]

Example 1.2.1.

Taking the array (3.2.1) we have

\[ \theta_{2s+1}^* = \frac{1}{24} h^4 y(t_{2s+1}) + O(h^5), \]

\[ \theta_{2s+2}^* = -\frac{1}{90} h^5 y(t_{2s+2}) + O(h^6), \quad s = O(1)n^{-1}. \]

Both examples then are consistent of order three, and yet computations show order four convergence, which we shall also show analytically.

A one-sided error bound of the form (5.1.1) then is not
sufficient to establish the exact order of convergence, whereas a two-sided error bound of the form (5.1.2) gives exact order of convergence $p$.

Albrecht (2) obtains a two-sided error bound by retaining a definition of consistency similar to definition 4.1.3 and stating conditions under which convergence is one order greater than consistency. Our strategy here will be different: we shall create a new concept, namely optimal consistency, which will permit a two-sided error bound, giving the correct order of convergence, $q$ say, if and only if the method is optimally consistent of order $q$.

**Definition 5.1.1.**

An $m$-block method is OPTIMALLY CONSISTENT of ORDERS $(p_1, p_2, \ldots, p_m)$ where $p_\nu \in \mathbb{N}$, $\nu = 1(1)m$, if for each $i \geq r$ there exists $C(i) \neq 0$ such that

$$
|([A_n^{-1}e^*_i]_i) = |[(\Delta_h y(t) - hA_n^{-1}B_h \Delta_h y'(t) - A_n^{-1}g)_i)| = C(i) h^{p_\nu} + o(h^{p_\nu+1}),
$$

$$
i = r + \sum^{m+\nu-1}_1, \ s = \alpha(1)n-1, \ \nu = 1(1)m.
$$

If $p = \min_{\nu = 1(1)m} p_\nu$ then we say that the method is optimally consistent of order $p$.

**Theorem 5.1.2.**

If an $m$-block method is optimally consistent of order $p$, zero stable and has $r$ starting values accurate of order $p$ then there exist $C_L$ and $C_U$ such that

$$
C_L h^p \leq \|e\|_\infty \leq C_U h^p.
$$

**Proof.**

(i) The upper bound.

As in theorem 4.4.2, $e = A_h^{-1}e^* + hA_h^{-1}B_h [f(\Delta_h t, y^*) - f(\Delta_h t, \Delta_h y(t))]$. Therefore for each $i$, $|e_i| \leq |([A_h^{-1}e^*]_i) + hL \sum_{j=0}^{r+nm-1} (A_h^{-1}B_h)_{ij} |e_j|$

$$
= |([A_h^{-1}e^*]_i) + hL \sum_{j=0}^{i+m-\nu} (A_h^{-1}B_h)_{ij} |e_j|,
$$
where \( \nu \equiv i - r + 1 \pmod{m} \) and \( 1 \leq \nu \leq m \).

This applies to both block implicit and block semi explicit methods: for a block semi explicit method \( (A_n^{-1}B_n)_{ij} = 0 \) whenever \( j > i \).

Therefore \( |e_i| \leq C^{(1)}h^\nu + hL \sum_{j=0}^{i+m-\nu} |(A_n^{-1}B_n)_{ij}| |e_j| \) and using (4.4.1) and lemmas 4.4.3 and 4.4.5 again we obtain \( ||e||_\infty \leq C y h^p \).

(ii) The lower bound.

\[ A_n^{-1}e^* = e - hA_n^{-1}B_n[f(\Delta_n t, y^*) - f(\Delta_n t, A_n y(t))] \] (5.1.3), therefore

\[ C^{(1)}h^\nu \leq ||e||_\infty + hL \sum_{j=0}^{r+n-1} |(A_n^{-1}B_n)_{ij}| |e_j| \]

\[ \leq ||e||_\infty + hM' \sum_{j=0}^{r+n-1} |e_j| \text{ for some } M' \text{ by (4.4.1)}, \]

\[ \leq ||e||_\infty + hM'(r + nm) ||e||_\infty, \text{ for all } i, \text{ where } \nu \equiv i - r + 1 \pmod{m} \]

and \( 1 \leq \nu \leq m \). Now \( (r + n \gamma - 1)h = T(2.1.3) \), so there exist \( C'' \) and \( M'' \) such that \( C''h^p \leq (1 + M'') ||e||_\infty \), and hence we obtain the result, taking \( C_L = C''/(1 + M'') \).

Let us now see how this result applies to various examples.

**Example 1.3.1.**

\[ [\Delta_n y(t) - hA_n^{-1}B_n A_n y'(t) - A_n^{-1}g]_i = y(t_i) - \tilde{y}_0 - h[\Delta_n^{-1}B_n A_n y'(t)]_i \]

\[ = \int_0^{t_i} y'(\eta) d\eta + O(h^4) - h[\Delta_n^{-1}B_n A_n y'(t)]_i, \]

assuming a starting value \( \tilde{y}_0 \) accurate of order four.

\( h[\Delta_n^{-1}B_n A_n y'(t)]_i \) is then a quadrature formula approximating

\[ \int_0^{t_i} y'(\eta) d\eta. \]
Let us now determine the form of $A_h^{-1} B_h$.

$$
A_h = \begin{pmatrix}
1 & 0 & 1 \\
-1 & 0 & 1 \\
0 & 0 & -1 & 1 \\
0 & 0 & -1 & 0 & 1 \\
0 & 0 & 0 & 0 & -1 & 1 \\
\cdots
\end{pmatrix}
$$

and by direct computation

$$
A_h^{-1} = \begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 & 1 & 1 \\
\cdots
\end{pmatrix}
$$

so $\max_{i,j} |(A_h^{-1})_{ij}| = 1$ and we have zero stability.

$$
B_h = \begin{pmatrix}
0 & 0 \\
0 & 0 \\
1/3 & 4/3 & 1/3 \\
0 & -1/12 & 8/12 & 5/12 \\
0 & 0 & 1/3 & 4/3 & 1/3 \\
0 & 0 & 0 & -1/12 & 8/12 & 5/12 \\
\cdots
\end{pmatrix}
$$

so

$$
A_h^{-1} B_h = \begin{pmatrix}
0 \\
0 \\
0 \\
1/3 & 4/3 & 4/3 \\
1/3 & 4/3 & 4/3 & 1 \\
1/3 & 4/3 & 4/3 & 1 \\
1/3 & 4/3 & 4/3 & 1 \\
1/3 & 4/3 & 4/3 & 1 \\
\cdots
\end{pmatrix}
$$
and we see that when \( i \) is even we have Simpson's rule applied globally, which has order four accuracy, and when \( i \) is odd we have again a global Simpson's rule with a local application of the quadrature rule

\[
\int_{x-h}^{x} y'(\eta) \, d\eta = h\left[ \frac{5}{12}y'(x) + \frac{3}{12}y''(x-h) - \frac{1}{12}y''(x-2h) \right],
\]

which has local accuracy of order four. Hence we have optimal consistency of order four, which implies convergence of order four, whereas consistency is only of order three.

**Example 2.2**

\( r = 1 \) and \( \theta_{2s+1}^* = \frac{1}{2}h^2 y''(t_{2s+1}) + O(h^3) \),

\( \theta_{2s+2} = -\frac{1}{2}h^2 y''(t_{2s+1}) + O(h^3) \), so we have order one consistency.

\[
A_h = \begin{pmatrix}
1 \\
-1 & 1 \\
0 & -1 & 1 \\
0 & 0 & -1 & 1 \\
\end{pmatrix}, \text{ so by direct computation}
\]

\[
A_h^{-1} = \begin{pmatrix}
1 \\
1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
\end{pmatrix}, \text{ and } \max_{i,j} |(A_h^{-1})_{ij}| = 1, \text{ giving zero stability.}
\]

Therefore \( [A_h y(t) - hA_h^{-1}A_h y'(t) - A_h^{-1} g]_i = y(t_i) - \tilde{y}_0 - h[A_h^{-1}B_h A_h y'(t)]_i \)

\[
= \int_{0}^{t_i} y'(\eta) \, d\eta + O(h^2)
\]

assuming a starting value accurate of order two.
and \( h[A^{-1}_h B_h A_h y'(t)]_4 \) is the trapezoidal quadrature approximation to 
\[
\int_0^t y'(\eta) d\eta,
\]
which has global accuracy of order two. Hence we have optimal consistency of order two, which implies convergence of order two.

**Example 3.2.1.**

Taking the array (3.2.1), \( A_h = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \), so

\[
A^{-1}_h = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}
\]

and again \( \max_{i,j} |(A^{-1}_h)_{ij}| = 1 \), giving zero stability.

\[
[B_h y(t) - hA^{-1}_h B_h A_h y'(t) - A^{-1}_h g]_4 = y(t_4) - \tilde{y}_o - h[A^{-1}_h B_h A_h y'(t)]_4,
\]

and again \( h[A^{-1}_h B_h A_h y'(t)]_4 \) is a quadrature formula approximating 
\[
\int_0^t y'(\eta) d\eta.
\]
For $i$ even we have a global application of Simpson's rule and for $i$ odd we have the same quadrature formula together with a local application of
\[
\int_{x-h}^{x} y(x) \, dx \approx h \left[ \frac{5}{12} y'(x-h) + \frac{8}{12} y'(x) - \frac{1}{12} y'(x+h) \right],
\]
which has fourth order accuracy locally. Thus we have optimal consistency and hence convergence of order four, whereas we saw previously that consistency is only of order three.

5.2 The form of $A_h^{-1}$.

We may obtain further convergence results by examining in greater detail the form of $A_h^{-1}$. With $A_h$ having the form (2.1.1) we define the $nm \times nm$ matrix $D_h = \begin{pmatrix} A_{\gamma} & A_{\gamma-1} & \cdots & A_1 & A_0 \end{pmatrix}$ for some matrix $E_h$. By equating $A_h A_h^{-1}$ to $I_{r+nm}$ we see that

\[
A_h^{-1} = \begin{pmatrix} I_r & 0 \\ 0 & D_h^{-1} E_h \end{pmatrix}
\]

for some matrix $E_h$. If $F_h$ is the $nm \times nm$ matrix

\[
\begin{pmatrix} A_{\gamma} & A_\tau & 0 \\ 0 & \cdots \\ A_\tau & 0 & A_\gamma \end{pmatrix},
\]

then $D_h = F_h G_h^\ast$. 

\[
I_{r+nm}
\]
where $G^*_n = \begin{bmatrix}
I_m \\
g_{r-1}^* & I_m \\
\vdots & \ddots & \ddots \\
g_1^* & \ldots & g_{r-1}^* & I_m \\
g_0^* & \ldots & g_{r-1}^* & I_m \\
\vdots & \ddots & \ddots & \ddots \\
g_0^* & \ldots & g_{r-1}^* & I_m \\
\end{bmatrix}
$

and $g_0^* = A^{-1}_T A_\theta$ for $\theta = O(1)r^{-1}$.

$D_h^{-1} = g_{h-1}^* h^{-1}$ and $h^{-1} = \begin{bmatrix}
A^{-1}_T \\
A^{-1}_r \\
A^{-1}_r \\
\vdots \\
A^{-1}_r \\
\end{bmatrix}$,

so it remains for us to consider the inversion of $g^*_n$.

**Lemma 5.2.1**

Let $W$ be an $(l+2)m \times (l+2)m$ matrix made up of $m \times m$ blocks

$X_0, X_1, \ldots, X_l : W = \begin{bmatrix}
I_m \\
X_0 & I_m \\
X_1 & X_0 & I_m \\
X_2 & X_1 & X_0 & I_m \\
\vdots & \vdots & \vdots & \vdots \\
X_l & \ldots & X_0 & I_m \\
\end{bmatrix}$

Let $W^{-1} = \begin{bmatrix}
I_m \\
H^{(0)}_o & I_m \\
H^{(1)}_o & H^{(1)}_1 & I_m \\
H^{(2)}_o & H^{(2)}_1 & H^{(2)}_2 & I_m \\
\vdots & \vdots & \vdots & \vdots \\
H^{(l)}_o & \ldots & H^{(l)}_1 & I_m \\
\end{bmatrix}$
Then $H^{(s)}_{\sigma} = H^{(s+1)}_{\sigma+1}$ for $\sigma = O(1)s$, $s = O(1)l-1$. (*)

Proof.

We shall use induction on $s$. For $s = 0$ we must prove that

$H^{(0)}_0 = H^{(1)}_1$. In performing the multiplication $WW^{-1} = I_{(l+2)m}$ we have

$X_0 + H^{(0)}_0 = 0_m$ and $X_0 + H^{(1)}_1 = 0_m$, giving $H^{(0)}_0 = H^{(1)}_1 = -X_0$.

Let us assume now that $H^{(k)}_{\sigma} = H^{(k+1)}_{\sigma+1}$ for $\sigma = O(1)k$, $k = O(1)s-1$.

Then in performing the multiplication $WW^{-1} = I_{(l+2)m}$ and considering rows $(s+1)m$ to $(s+2)m - 1$, we obtain

$$X_{s-k} + \sum_{q=0}^{s-k} X H^{(s-q)}_k + H^{(s)}_k = 0_m, \quad k = O(1)s \quad (5.2.1).$$

Considering rows $(s+2)m$ to $(s+3)m - 1$ we obtain

$$X_{s+1-k} + \sum_{q=0}^{s-k} X H^{(s-q)}_k + H^{(s+1)}_k = 0_m, \quad k = O(1)s+1 \quad (5.2.2).$$

Taking all the equations in (5.2.2) except the case $k = 0$ and putting $l = k-1$,

$$X_{s-l-1} + \sum_{q=0}^{s-l-1} X H^{(s-q)}_{l+1} + H^{(s+1)}_{l+1} = 0_m \quad \text{for } l = O(1)s \quad (5.2.3).$$

Then combining (5.2.1) and (5.2.3) and recalling that by assumption $H^{(k)}_{\sigma} = H^{(k+1)}_{\sigma+1}$ for $\sigma = O(1)k$, $k = O(1)s-1$, we obtain the result

$H^{(s)}_{\sigma} = H^{(s+1)}_{\sigma+1}$ for $\sigma = O(1)s$.

Corollary 5.2.2.

$G_n^{-1}$ has the form

$$\begin{pmatrix}
I_m \\
P_0 & I_m \\
P_1 & P_0 & I_m \\
P_2 & P_1 & P_0 & I_m \\
\vdots & \ddots & \ddots & \ddots \\
P_{n-2} & \ldots & P_0 & I_m
\end{pmatrix}$$

(*) This lemma is a generalisation of the result for isoclinical matrices given in theorem 2.2 of Leroy J. Derr, Pac. J. of Maths., 37, no. 1, 41-43 (1971).
This is an immediate consequence of lemma 5.2.1.

Performing the multiplication $G^*G^{-1}_h = I_n$ we see that

\[
\begin{align*}
G^*_{-1} + P_{o} &= 0_m \\
G^*_{-2} + G^*_{-1}P_{o} + P_{1} &= 0_m \\
G^*_{-3} + G^*_{-2}P_{o} + G^*_{-1}P_{1} + P_{2} &= 0_m \\
& \quad \vdots \\
G^*_{o} + G^*P_{1} + G^*P_{1} + \cdots + G^*_{-1}P_{-2} + P_{-1} &= 0_m \\
G^*P_{1} + G^*P_{1} + \cdots + G^*_{-2}P_{-2} + G^*_{-1}P_{-1} + P_{-2} &= 0_m \\
G^*P_{1} + \cdots + G^*_{-1}P_{-1} + P_{-1} &= 0_m \\
\quad \text{and so on up to}
\end{align*}
\]

\[
G^*P_{o} n_{-2} + \cdots + G^*_{-1}P_{n-3} + P_{n-2} = 0_m.
\]

Therefore for each $q = O(1)n-2$, 

\[
P_q = \sum_{1=1}^{u} P_{q} = \sum_{1=1}^{u} \left( \sum_{\theta=0}^{T-1} \sum_{1=1}^{u} \right)
\]

(5.2.4)

for some positive integers $u_q$, $\theta = O(1)T-1$, and real scalars $\beta_1, \beta_2, \ldots, \beta_{u_q}$.

\[
D^{-1}_h = G^*h^{-1}G^*_h = \\
\begin{pmatrix}
A_{-1}^{-1} & A_{-1}^{-1} & A_{-1}^{-1} \\
P_{o}A_{-1}^{-1} & A_{-1}^{-1} & A_{-1}^{-1} \\
P_{1}A_{-1}^{-1} & P_{o}A_{-1}^{-1} & A_{-1}^{-1} \\
\vdots & \vdots & \vdots \\
P_{n-2}A_{-1}^{-1} & \cdots & P_{o}A_{-1}^{-1} & A_{-1}^{-1}
\end{pmatrix}
\]

5.3 Improved Convergence.

Having determined the form of $A^{-1}_h$, we are now in a position to formulate a criterion analogous to that of Albrecht (2, theorem 4.3) for the order of convergence to exceed that of consistency by one.

This situation we refer to as improved convergence.
Theorem 5.3.1.
Suppose we have an m-block method which is consistent of order p and zero stable, and starting values are accurate of order p+1. Then we have exact order of convergence p+1 if optimal consistency is of order p+1.

Proof.
If the method is optimally consistent of order p+1 we have exact order of convergence p+1, due to theorem 5.1.2.

If the method converges with exact order p+1, assume that optimal consistency is of order p*. Therefore by theorem 5.1.2 the exact order of convergence is p*, so p* = p+1.

Definition 5.3.2.
The truncation error vector constants c(1), c(2), ... are defined for y(t) ∈ C([0,T]) by c(q) = (c_1(q), c_2(q), ..., c_m(q))^T, q ≥ 1 and

\[\theta^{p+q}_{r+nm+\nu-1} = \sum_{q=1}^{Q} c(q)[\Delta_h y^{(p+q)}(t)]_{r+nm} h^{p+q} + O(h^{p+q+1}), \nu = 1(1)m, s = O(1)n^{-1}\]

Theorem 5.3.3.
Suppose we have an m-block method which is consistent of order p and zero stable and starting values are accurate of order p+1. Then we have exact order of convergence p+1 if \(A_h^{-1}\Theta = 0\), the zero vector, for \(\Theta = O(1)^{r-1}\).

Proof.
By virtue of theorem 5.3.1 we must prove optimal consistency of order p+1. \(A_h^{-1}\Theta = \begin{pmatrix} I_r & 0 \\ -D_h^{-1}E_h & D_h^{-1} \end{pmatrix} \begin{pmatrix} \Theta_{r+nm-1} \\ \vdots \\ \Theta_{r} \\ \vdots \\ \Theta_{1} \\ 0 \end{pmatrix} \).
so \([A_h^{-1} e^*]_{r+sm+v-1}\) is the \((sm+v-1)\)th component of

\[
-D_h^{-1}E_h \begin{pmatrix} e^* \\ \vdots \\ e^* \\ \Theta^* \\ e^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \end{pmatrix} + D_h^{-1} \begin{pmatrix} e^* \\ \vdots \\ e^* \\ \Theta^* \\ e^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \end{pmatrix}.
\]

Because of zero stability the elements of \(D_h^{-1}\) are bounded, therefore

\[
-D_h^{-1}E_h \begin{pmatrix} e^* \\ \vdots \\ e^* \\ \Theta^* \\ e^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \end{pmatrix} = -D_h^{-1}E_h \begin{pmatrix} C^* \\ \vdots \\ C^* \\ \Theta^* \\ C^* \\ \vdots \\ C^* \\ \Theta^* \\ \vdots \\ C^* \\ \Theta^* \\ \vdots \\ C^* \\ \Theta^* \\ \vdots \\ C^* \end{pmatrix} h^{p+1} + O(h^{p+2})
\]

for some constants \(C^*, \ldots, C^*_{r+sm-1}\). Accordingly it remains to consider

\[
-D_h^{-1}E_h \begin{pmatrix} e^* \\ \vdots \\ e^* \\ \Theta^* \\ e^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \\ \Theta^* \\ \vdots \\ e^* \end{pmatrix},
\]

whose elements \(s_m\) to \(s_{m+m-1}\) are of the form

\[
A^{-1}_\tau c^{(1)}(1) \Delta_h \gamma^{(p+1)}(t) \big|_{r+sm} h^{p+1} + \sum_{\sigma=0}^{s-1} P_{\sigma} A^{-1}_\tau c^{(1)}(1) \Delta_h \gamma^{(p+1)}(t) \big|_{r+(s-\sigma-1)m} h^{p+1} + O(h^{p+2})
\]

Since the method is zero stable, the elements of \(A^{-1}_\tau\) are bounded, so

\[
A^{-1}_\tau c^{(1)}(1) \Delta_h \gamma^{(p+1)}(t) \big|_{r+sm} h^{p+1} = C' h^{p+1}
\]

for some vector \(C'\). Therefore we have order \(p+1\) convergence if and only if

\[
\left\| \sum_{\sigma=0}^{n-2} P_{\sigma} A^{-1}_\tau c^{(1)}(1) \Delta_h \gamma^{(p+1)}(t) \big|_{r+(n-\sigma-2)m} \right\|_\infty \leq C
\]

for some \(C\), for all \(n \geq 2\) (5.3.1).

Now for all \(\sigma\),

\[
P_{\sigma} A^{-1}_\tau c^{(1)}(1) = \sum_{l=1}^{u_{\sigma}} \sum_{\Theta=0}^{\tau-1} (A^{-1}_\tau A_{\Theta}) \Theta A^{-1}_\tau c^{(1)}(1)
\]

by (5.2.4)

\[
= \sum_{l=1}^{u_{\sigma}} \sum_{\Theta=0}^{\tau-1} (A^{-1}_\tau A_{\Theta}) \Theta A^{-1}_\tau c^{(1)}(1)
\]

= 0 if \(A_{\Theta} A^{-1}_\tau c^{(1)} = 0\) for \(\Theta = O(1) \tau^{-1}\).

Hence (5.3.1) holds if \(A_{\Theta} A^{-1}_\tau c^{(1)} = 0\) for \(\Theta = O(1) \tau^{-1}\).

We now examine some applications of this result.
Example 1.3.1.

\[ A_0 A_1^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \] and \( c(1) = \begin{bmatrix} 0 \\ -1/24 \end{bmatrix} \), so \( A_0 A_1^{-1} c(1) = 0 \), implying fourth order convergence.

Example 2.2.1.

\[ A_0 A_1^{-1} = \begin{bmatrix} -1 & -1 \\ 0 & 0 \end{bmatrix} \] and \( c(1) = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix} \), so \( A_0 A_1^{-1} c(1) = 0 \), implying second order convergence.

Example 2.2.2.

The companion matrices (1.3.3) are

\[ G_1 = \begin{bmatrix} 0 & 1 \\ -3/2 & 5/2 \end{bmatrix}, \quad G_2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad G_3 = \begin{bmatrix} 0 & 1 \\ 6/5 & -1/5 \end{bmatrix}, \]

giving \( G = \begin{bmatrix} -3/2 & 5/2 \\ -3/2 & 5/2 \end{bmatrix} \), which has eigenvalues 0 and 1, and so we have zero stability. Consistency is of order three and

\[ c(1) = \begin{bmatrix} -5/48 \\ -1/24 \\ 1/120 \end{bmatrix}, \quad A_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1/5 & 1 \end{bmatrix}, \quad \text{so } A_0 A_1^{-1} = \begin{bmatrix} -1 & 2 & -5/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \]

and hence \( A_0 A_1^{-1} c(1) = 0 \), which implies fourth order convergence.

Example 2.2.3.

The companion matrices (1.3.3) are

\[ G_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad G_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 2 \end{bmatrix}, \quad G_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 3/2 & -1/2 \end{bmatrix}, \]

\[ G = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 2 & 0 & 0 & -1 \\ 1/2 & 0 & 0 & 1/2 \end{bmatrix}, \] which has eigenvalues 0,0,\(-\frac{1}{2}\),1 and so we have zero stability. Consistency is of order three.
and \[ c^{(1)} = (2,-2,-2)^T \]. $A_0A_2^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $A_1A_2^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 1 \\ 0 & 0 & 0 \end{pmatrix}$,

so $A_0A_2^{-1}c^{(1)} = A_1A_2^{-1}c^{(1)} = 0$ and we have fourth order convergence.

**Example 5.3.4.**

Consider the Runge-Kutta method quoted in (15, p.118).

\[ y_i - y_{i-1} = h k_2, \]

\[ k_1 = f(t_{i-1}, y_{i-1}), \]

\[ k_2 = f(t_{i-1} + \frac{1}{2}h, y_{i-1} + \frac{1}{2}hk_1). \]

This is a two-block explicit method with array

\[
\begin{pmatrix}
0 & -1 & 1 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\
0 & -1 & 0 & 1 & 0 & 0 & 1 & 0 & 1
\end{pmatrix}.
\]

Any Runge-Kutta method is zero stable, as we saw in section 4.3. This example is consistent of order one and

\[ c^{(1)} = \begin{pmatrix} 1/8 \\ 0 \end{pmatrix}, \]

so $A_0A_1^{-1}c^{(1)} = 0$, giving second order convergence.

**Example 5.3.5.**

Another Runge-Kutta method (15, p.119) is

\[ y_i - y_{i-1} = h(\frac{1}{2}k_1 + \frac{1}{2}k_2), \]

\[ k_1 = f(t_{i-1}, y_{i-1}), \]

\[ k_2 = f(t_{i-1} + h, y_{i-1} + hk_1). \]

This is a 2-block explicit method with array

\[
\begin{pmatrix}
0 & -1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & -1 & 0 & 1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0
\end{pmatrix}.
\]

Consistency is of order one and

\[ c^{(1)} = \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}. \]

Again $A_0A_1^{-1} = \begin{pmatrix} 0 & -1 \\ 0 & -1 \end{pmatrix}$,

so $A_0A_1^{-1}c^{(1)} = 0$, giving convergence of order two.
Example 5.3.6.

For a predictor-corrector method in PBCE mode,

\[
A_\theta = \begin{bmatrix}
0 & \alpha^*_\theta \\
0 & \alpha^*_\theta 
\end{bmatrix}, \quad \theta = \mathcal{O}(1)k^{-1}
\]

and

\[
A_k = \begin{bmatrix}
\alpha^*_k & 0 \\
0 & \alpha^*_k 
\end{bmatrix}.
\]

Suppose we take a predictor consistent of order \( p-1 \) with error constant \( C^* \) and a corrector of order \( p \). Then the predictor-corrector scheme as a 2-block method is consistent of order \( p-1 \) and \( c(1) = \begin{bmatrix} C^* \\ 0 \end{bmatrix} \).

For each \( \theta = \mathcal{O}(1)k^{-1} \),

\[
A_\theta A_k^{-1} = \begin{bmatrix}
0 & \alpha^*_\theta \\
0 & \alpha^*_\theta 
\end{bmatrix} \begin{bmatrix}
1/\alpha^*_k & 0 \\
0 & 1/\alpha^*_k 
\end{bmatrix} = \begin{bmatrix}
0 & \alpha^*_\theta/\alpha^*_k \\
0 & \alpha^*_\theta/\alpha^*_k 
\end{bmatrix},
\]

so \( A_\theta A_k^{-1}c(1) = 0 \), and since any predictor-corrector method in P(EC) mode is zero stable if and only if the corrector is when considered as a linear multistep method, we have convergence of order \( p \) in both the predictor and the corrector, provided we take a zero stable corrector.

In PEC mode \( A_h \) is exactly as in PECE mode, and so we have the same result.

Example 3.2.1.

Consistency is of order three and \( c(1) = \begin{bmatrix} 1/24 \\ 0 \end{bmatrix} \).

\[
A_0 A_1^{-1} = \begin{bmatrix}
0 & -1 \\
0 & -1 
\end{bmatrix}, \quad \text{so } A_0 A_1^{-1}c(1) = 0, \text{ giving order four convergence.}
\]

Example 5.3.7.

Consider the following generalised Nordsieck procedure (2, p.248).

\[
y_{i-1} = \frac{1}{2} y_{i-1} + \frac{1}{2} y_{i-2} + \frac{h}{24} f_1 + 8f_{i-1} + 5f_{i-2},
\]

\[
y_i = \frac{1}{2} y_{i-1} + \frac{1}{2} y_{i-2} + \frac{h}{8} (3f_i + 8f_{i-1} + f_{i-2}).
\]
This is a 2-block implicit method with array
\[
\begin{pmatrix}
0 & 0 & -1/2 & -1/2 & 1 & 0 & 0 & 5/24 & 0 & 1/3 & 0 & -1/24 & 0 \\
0 & 0 & -1/2 & -1/2 & 0 & 1 & 0 & 1/8 & 0 & 1 & 0 & 3/8 & 1
\end{pmatrix}.
\]

Consistency is of order three and
\[c(1) = \begin{pmatrix} 1/48 \\ -1/48 \end{pmatrix}.
\]

\[A_0A_2^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \text{ so } A_0A_2^{-1}c(1) = 0.
\]

\[A_1A_2^{-1} = \begin{pmatrix} -1/2 & -1/2 \\ -1/2 & -1/2 \end{pmatrix}, \text{ so } A_1A_2^{-1}c(1) = 0.
\]

\[G_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1/2 & 1/2 \end{pmatrix} \text{ and } G_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}, \text{ so } G = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix},
\]

which has eigenvalues 0, 0, 1. Hence we have zero stability and fourth order convergence.

Example 5.3.8.

Consider a P_{2/3}EP^k hybrid method (5, p.658):
\[
y_{1-1/3} - \frac{5}{9}y_{i-1} - \frac{1}{9}y_{i-2} = \frac{10h}{9} f_{i-1},
\]
\[
y_i - y_{i-1} = \frac{h}{9} (3f_{i-1/3} + f_{i-1}),
\]

which is a two block explicit method with array
\[
\begin{pmatrix}
0 & -4/9 & 0 & -5/9 & 1 & 0 & 0 & 0 & 0 & 10/9 & 0 & 0 & 2/3 \\
0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 1/4 & 3/4 & 0 & 1
\end{pmatrix}.
\]

For zero stability the companion matrices are
\[G_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4/9 & 0 & 5/9 \end{pmatrix} \text{ and } G_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \text{ so } G = \begin{pmatrix} 0 & 0 & 1 \\ 4/9 & 0 & 5/9 \\ 0 & 0 & 1 \end{pmatrix},
\]

which has eigenvalues 0, 0, 1, so we have zero stability.

Consistency is of order two and
\[c(1) = \begin{pmatrix} 10/81 \\ 0 \end{pmatrix}.
\]
\[ A_0^{-1} A_2^{-1} = \begin{pmatrix} 0 & -4/9 \\ 0 & 0 \end{pmatrix}, \text{ so } A_0^{-1} A_2^{-1} c(1) = 0 \]

\[ A_1^{-1} A_2^{-1} = \begin{pmatrix} 0 & -5/9 \\ 0 & -1 \end{pmatrix}, \text{ so } A_1^{-1} A_2^{-1} c(1) = 0, \text{ and we have third order convergence.} \]

**Example 5.3.9.**

Cooper also quotes the following example (5).

\[ y_{i-1/3} + 3y_{i-1} - 4y_{i-4/3} = 2f_{i-1}, \]
\[ y_i + 3y_{i-1} - 4y_{i-4/3} = \frac{h}{12}(5f_{i-1/3} + 23f_{i-1}), \]

which is a 2-block explicit method with array

\[
\begin{pmatrix}
-4 & 3 & 1 & 0 \\
-4 & 3 & 0 & 1 \\
0 & 2 & 0 & 0 \\
0 & 0 & 23/12 & 5/12 \\
0 & 0 & 0 & 1
\end{pmatrix}.
\]

For zero stability the companion matrices are

\[ G_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 4 & -3 \end{pmatrix} \quad \text{and} \quad G_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -3 & 0 \end{pmatrix}, \text{ so } G = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 4 & -3 \\ 0 & 4 & -3 \end{pmatrix}, \]

which has eigenvalues 0, 0, 1, so we have zero stability.

Consistency is of order two and

\[ c(1) = \frac{2}{41} \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \quad A_0^{-1} A_1^{-1} = \begin{pmatrix} -4 & 3 \\ -4 & 3 \end{pmatrix}, \]

so \[ A_0^{-1} A_1^{-1} c(1) = 0, \text{ giving order three convergence}. \]

**Example 5.3.10.**

Another method quoted in (5) is

\[ y_{1-2/3} - y_{1-1} + \frac{2}{3} y_{1-1} - y_{1-5/3} + \frac{1}{3} y_{1-2} = \frac{2h}{3} f_{i-1}, \]
\[ y_{1-1/2} - y_{1-1} = \frac{h}{8}(\frac{1}{2} f_{i-1} + 3f_{i-2/3}), \]
\[ y_i - y_{1-1} = \frac{h}{6}(f_{i-1} + 4f_{i-1/2} + f_{1}), \]
\[ y_i - y_{1-1} = \frac{h}{6}(f_{i-1} + 4f_{i-1/2} + f_{1}), \]
which is a $4$-block explicit method with array

\[
\begin{bmatrix}
0 & 0 & 0 & 1/3 & -1 & 0 & 2/3 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2/3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/8 & 3/8 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & -3/2 & 2 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 1/2 & -1/6 & 0 & 1
\end{bmatrix}
\]

\[
G_1 = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
-1/3 & 1 & 0 & -2/3 & 1
\end{bmatrix}, \quad G_2 = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0
\end{bmatrix}, \quad G_3 = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0
\end{bmatrix}, \quad G_4 = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

\[
G = \begin{bmatrix}
0 & 0 & 0 & 0 & 1 \\
-1/3 & 1 & 0 & -2/3 & 1 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

\[
A_o^t A_2^{-1} = \begin{bmatrix}
0 & 0 & 0 & 1/3 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

\[
\delta(1) = \frac{1}{27 \cdot 3^4} \begin{bmatrix}
64 \\
3 \\
96 \\
0
\end{bmatrix}
\]

\[
A_o A_2^{-1} \delta(1) = 0.
\]
\[ A_1 A_2^{-1} = \begin{pmatrix} -1 & 0 & 2/3 & -1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \end{pmatrix} \], so \( A_1 A_2^{-1} e^{(1)} = 0 \), giving order four convergence.

Example 5.3.11.

Albrecht (1) quotes the method
\[ \tilde{y}_{i-1} = \frac{11}{30} y_{i-1} + \frac{19}{30} y_{i-2} + \frac{12 h}{720} f_1 + 13 f_{i-1} + 13 f_{i-2} - f_{i-3}, \]
\[ y_i = \frac{11}{30} y_{i-1} + \frac{19}{30} y_{i-2} + \frac{h}{720} (251 f_1 + 817 f_{i-1} + 97 f_{i-2} + 11 f_{i-3}). \]

This is a 2-block implicit method with array
\[
\begin{pmatrix}
0 & 0 & 0 & -19/30 & -11/30 & 1 & 0 & 0 & -19/720 & 247/720 & 0 & a_7/720 \\
0 & 0 & 0 & -19/30 & -11/30 & 0 & 1 & 0 & 11/720 & 97/720 & 0 & 817/720 \\
& & & & & & & & & & & 0 & -19/720 & 0 & 251/720 & 1
\end{pmatrix}.
\]

Consistency is of order four and \( c^{(1)} = \frac{19}{180.120} \begin{pmatrix} 11 \\ -19 \end{pmatrix} \).

For zero stability the companion matrices are
\[ G_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 19/30 & 11/30 \end{pmatrix} \quad \text{and} \quad G_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 19/30 & 11/30 & 0 \end{pmatrix}, \]
so \( G = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 19/30 & 11/30 \\ 0 & 19/30 & 11/30 \end{pmatrix} \), which has eigenvalues 0, 0, 1, so we have zero stability.

\[ A_0 A_3^{-1} = A_1 A_3^{-1} = A_2 A_3^{-1} = 0 \], so \( A_0 A_3^{-1} c^{(1)} = A_1 A_3^{-1} c^{(1)} = A_2 A_3^{-1} c^{(1)} = 0 \).

\[ A_0 A_3^{-1} = \begin{pmatrix} -19/30 & -11/30 \\ -19/30 & -11/30 \end{pmatrix} \], so \( A_0 A_3^{-1} c^{(1)} = 0 \) and we have fifth order convergence.
6.1 Two-Sided Error Bounds.

In all the examples considered so far we have had optimal consistency of orders $p(1,1,\ldots,1)$ for some $p \in \mathbb{N}$. A more detailed analysis of convergence is needed however when we encounter methods whose optimal consistency is of orders $(p_1,p_2,\ldots,p_m)$ with not all the $p_i$ equal.

Definition 6.1.1.

Let $a_1,a_2,\ldots,a_w$ be distinct positive integers:

$1 \leq a_1 < a_2 < \ldots < a_w \leq m$. Then we define the vector 

$\epsilon^{(a_1,a_2,\ldots,a_w)} \in \mathbb{R}^{r+nw}$ by 

$[\epsilon^{(a_1,a_2,\ldots,a_w)}]_i = \epsilon_i^1, i = O(1)r^1,$

$[\epsilon^{(a_1,a_2,\ldots,a_w)}]_i = \epsilon_i^{r+sw+q-1}, q = 1(1)w, s = O(1)n-1,$

and $\epsilon$ is the global error vector as in definition 4.1.4. Note that $\epsilon^{(1,2,\ldots,m)} = \epsilon$.

It is tempting to imagine that optimal consistency of orders $(p_1,p_2,\ldots,p_m)$ will imply $\|\epsilon^{(\nu)}\|_\infty \leq C_U^{(\nu)} h^{p^*}$ for some $C_U^{(\nu)}$, $\nu = 1(1)m$, but this is not generally the case. The only thing we can say at this stage is that optimal consistency of orders $(p_1,p_2,\ldots,p_m)$ implies

$C_L h^{p} \leq \|\epsilon\|_\infty \leq C_U h^{p^*}$, where $p = \min_{\nu=1(1)m} p_{\nu}$ (6.1.1).

Our aim in this section is to obtain two-sided bounds of the form

$C_L^{(a_1,a_2,\ldots,a_w)} h^{p^*} \leq \|\epsilon^{(a_1,a_2,\ldots,a_w)}\|_\infty \leq C_U^{(a_1,a_2,\ldots,a_w)} h^{p^*}$ (6.1.2)

for some $p^*$, for some proper subsets $\{a_1,a_2,\ldots,a_w\}$ of $\{1,2,\ldots,m\}$, and thus obtain more information than that given by (6.1.1). The analysis
we employ will retain the essential simplicity of the approach adopted in the preceding chapters, resorting only to optimal consistency and lemmas 4.4.3 and 4.4.5.

Example 6.1.2.

Let us consider a predictor-corrector scheme in P(EC)

with predictor \( y_i = y_{i-1} + hf_{i-1} \)

and corrector \( y_i = y_{i-1} + \frac{h}{12}(5f_i + 8f_{i-1} - f_{i-2}) \).

This is a 3-block explicit method with array

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1/12 & 0 & 0 & 8/12 & 5/12 & 0 & 0 \\
0 & -1/12 & 0 & 0 & 8/12 & 0 & 5/12 & 0 \\
\end{bmatrix}
\]

By direct computation

\[
A_h^{-1} = \begin{bmatrix}
1 \\
0 & 1 \\
0 & 1 & 1 \\
0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\
\end{bmatrix}
\]

Let us assume two starting values accurate of order three. Then \( \theta_0^* = C_0 h^3 \) and \( \theta_1^* = C_1 h^3 \).

\[
\theta_{s+2}^* = \frac{1}{2} h^2 y''(t_{s+2}) - \frac{3}{2} h^3 y'''(t_{s+2}) + \frac{3}{8} h^4 y''(t_{s+2}) + O(h^5),
\]

\[
\theta_{s+1+s+\nu}^* = -\frac{1}{24} h^4 y''(t_{s+2}) + O(h^5),
\]

\( \nu = 2, 3, s = O(1)n-1 \).

Hence by computing \( A_h^{-1} \theta^* \) directly we obtain optimal consistency of orders \((2, 3, 3)\).

Example 6.1.3.

Consider the 3-stage explicit Runge-Kutta method with array
Let us assume a starting value accurate of order three. Then $\Theta_0 = O(h^3)$.

\[
\Theta_{sm+1}^* = \frac{1}{18} h^2 y''(t_s) + \frac{1}{162} h^3 y'''(t_s) + \frac{1}{1944} h^4 y^4(v(t_s)) + O(h^5),
\]

\[
\Theta_{sm+2}^* = \frac{1}{81} h^3 y'''(t_s) + \frac{1}{243} h^4 y^4(v(t_s)) + O(h^5),
\]

\[
\Theta_{sm+3}^* = \frac{1}{216} h^4 y^4(v(t_s)) + O(h^5),
\]

Hence by computing $A_{h}^{-1} \Theta^*$ directly we obtain optimal consistency of orders $(2, 3, 3)$.  

6.2 A Special Class of Methods.

We shall be considering a class of methods characterised by $A_h$ and $B_h$ having certain forms, which will permit us to obtain bounds on the error of the form $(6.1.2)$.

Definition 6.2.1.

A square matrix $A$ is in the class $Q_{m, \alpha}$ for two positive integers $m$ and $\alpha$: $m \geq \alpha$ if

(i) $A$ is lower triangular;

(ii) if we denote the $(i,j)$th element of $A$ by $A_{ij}$, then there exist positive integers $r, n, \tau$ such that $A \in \mathbb{R}^{r \times n \times m} x \mathbb{R}^{r \times n \times m}$ and for each
\[ v = 1(1)m^\infty, s = O(1)n^{-1}, A_{i, r+sm+v = 1} = 0 \text{ for } i = r+sm+v(1)r+sm+m^{-1}, \]
\[ i = r+sm+q^{-1}, q = 2(1)m \text{ and } \sigma = s+1(1)s+\tau; \]

(iii) \( A_{ii} \neq 0 \) for all \( i \),

(iv) \( A \) is a band matrix: for each \( v = 1(1)m, s = O(1)n^{-1}, \)
\[ A_{i, r+sm+v = 1} = 0 \text{ for } i \geq r+(s+\tau+1)m \text{ and for } j = 0(1)r^{-1}, A_{ij} = 0 \text{ for } i \geq r + \gamma m. \]

By way of illustration, when \( m = 5, \alpha = 3, r = 2, n = 4 \) and \( \gamma = 2 \), \( A \in Q_{5,3} \) has the form

\[
\begin{bmatrix}
\text{X} \\
\text{X X} \\
\text{X X X} \\
\text{X X 0 X} \\
\text{X X 0 0 X} \\
\text{X X 0 0 X X} \\
\text{X X 0 0 X X X} \\
\text{X X 0 0 X X X X} \\
\text{X X 0 0 X X X X X} \\
\text{X X 0 0 X X X X X X}
\end{bmatrix}
\]

The zeroes represent elements which are necessarily zero, and all the elements on the diagonal are strictly non-zero.

We see that for an \( m \)-block method \( A_{h} \in Q_{m,\alpha} \) if \( A_{\tau} \) has the form

\[
\begin{bmatrix}
\text{X} \\
\text{... X} \\
\text{X X} \\
\text{X X X} \\
\text{X X X} \\
\text{X X X} \\
\text{X X X} \\
\text{X X X} \\
\text{X X X}
\end{bmatrix}
\]

, all elements on the diagonal being strictly non-zero,
and for $\theta = O(1)r^{-1}$, $A_\theta$ has the form
\[
\begin{pmatrix}
0 \\
\vdots \\
x \ldots x
\end{pmatrix}
\]

**Definition 6.2.2.**

A square matrix $B$ is in the class $Q_{m, \alpha}$ for three positive integers $m, \alpha, \beta : \alpha, \beta \leq m$ if

(i) it is lower triangular;

(ii) there exist positive integers $r, n$ such that $B \in \mathbb{R}^{r+nm \times r+nm}$ and for each $v = 1(1)m-\alpha$, $s = O(1)n-1$, $B_{i, r+\alpha s+n-1} = 0$ for $i = r+\alpha s+\beta(1)v+\alpha s+n-1$, $i = r+\alpha s+n-q-1$, $q = 2(1)m$ and $\sigma = s+1(1)n-1$.

When $m = 5$, $\alpha = 3$, $\beta = 2$, $r = 2$ and $n = 4$, $B \in Q_{5, 3}$ has the form

\[
\begin{pmatrix}
1 \\
\vdots \\
x \ldots x
\end{pmatrix}
\]

**Lemma 6.2.3.**

Let $U$ be any square $(1+1) \times (1+1)$ matrix and suppose that there exist positive integers $\rho, \sigma : 0 \leq \rho, \sigma \leq 1$ such that for all integers $i : 0 \leq i \leq 1$, $u_{i\sigma} = u_{i\sigma} \delta_{i\sigma}^\epsilon$, where $\delta_{i\sigma}$ is
the Kronecker delta. Then $|\text{det } U| = |u_{i\sigma}||\text{det } \tilde{U}|$, where $\tilde{U}$ is the matrix formed by deleting row $\rho$ and column $\sigma$ from $U$.

The proof follows trivially from considering how the determinant of a square matrix is defined.

**Lemma 6.2.4.**

If $A \in \mathbb{Q}^{m \times \alpha}$ then $A$ is nonsingular and $A^{-1} \in \mathbb{Q}^{(1)}_{m \times \alpha}$.

**Proof.**

$$\det A = \prod_{i=0}^{r+s+m-1} A_{ii} \text{ since } A \text{ is lower triangular. Therefore } \det A \neq 0$$

because of definition 6.2.1 (iii), and so $A$ is nonsingular.

Part (i) of definition 6.2.2 is fulfilled since the inverse of any nonsingular lower matrix is itself lower triangular.

Now take any $\psi, s: 1 \leq \psi \leq m-\alpha$, $0 \leq s \leq n-1$. Let us denote the $(i, j)^{th}$ element of $A^{-1}$ by $p_{ij}$, and consider $p_{i_1, i_1-r+s+m+1}$ for some $i_1 \geq r+s+m+1$ such that $i_1 - r \neq 0 \pmod{m}$.

Now $A^{-1} = \frac{1}{\det A} \text{adj } A$, so $p_{i_1, i_1-r+s+m+1} = \frac{1}{\det A} A_{i_1, i_1-r+s+m+1} = A_{i_1, i_1-r+s+m+1}$, where $A_{i_1, i_1}$ is the $(j, i)^{th}$ cofactor of $A$. Now $A_{i_1, i_1-r+s+m+1}$ is the determinant of the matrix formed by deleting row $r+s+m+1$ and column $i_1$ from $A$. Consider now the $(r+s+m+1)^{th}$ column of this resulting matrix $A$. $A_{i_1, i_1-r+s+m+1} = 0$ whenever $i - r + 1 \neq 0 \pmod{m}$ or $i < r + s + m + 1$.

That is, the only possible non-zero elements in the $(r+s+m+1)^{th}$ column of $A$ are $A_{i, i-r+s+m+1}$, where $\sigma = s+1(1)n-1$.

Consider now columns $r+s\sigma$, $\sigma = s+1(1)n-1$ of $A$. By definition, for each $r$ the only possible non-zero elements in column $r+s\sigma$ are

$$A_{r, r+s\sigma}, A_{r+(s+1)m, r+s\sigma}, \ldots, A_{r+(n-1)m, r+s\sigma}$$

Let us now exchange columns zero and $r+s+m+1$ of $A$, and then columns $\sigma-s$ and $r+s-1$ of $A$, $\sigma = s+1(1)n-1$. In the resulting matrix $A^{(1)}$, whose determinant has the same magnitude as that of $A$, the first $n-s$ columns are of the form...
By lemma 6.2.3 \( \det \hat{A}^{(1)} \propto \det \hat{A}^{(2)} \), where \( \hat{A}^{(2)} \) is the matrix formed by deleting column \( n-s-1 \) and row \( r+(n-1)m-1 \) of \( \hat{A}^{(1)} \). Since row \( r+(n-1)m-1 \) of \( \hat{A}^{(1)} \) is now deleted, column \( n-s-2 \) of \( \hat{A}^{(2)} \) has only one possibly non-zero element, in row \( r+(n-2)m-1 \). Again lemma 6.2.3 gives \( \det \hat{A}^{(2)} \propto \det \hat{A}^{(3)} \), where \( \hat{A}^{(3)} \) is formed by deleting column \( n-s-2 \) and row \( r+(n-2)m-1 \) of \( \hat{A}^{(2)} \). We continue in this manner until we obtain \( \hat{A}^{(n-s-1)} \), whose zero t h. column is composed entirely of zeroes. Therefore \( \det \hat{A} = 0 \), and so \( \hat{A}_{r+(s+\nu)m-1,i_1} = 0 \) and we have the result.

**Lemma 6.2.5.**

If \( A \in Q_{m, \alpha_1} \) and \( B \in Q_{m, \alpha_2}^{(\beta)} \), where \( \alpha_1, \alpha_2 \) and \( \beta \) are positive integers less than \( m \), then \( A^{-1}B \in Q_{m, \alpha}^{(\beta)} \) where \( \alpha = \min \{ m, \max (\alpha_2, \beta+\alpha_1-1) \} \).
Proof.

By lemma 6.2.4, $A^{-1} \in Q_{m \neq 1}^{(1)}$. Let us consider only rows and columns $r$ to $r+nm-1$ of $A^{-1}$ and $B$. We shall divide these elements of $A^{-1}$ into $m \times m$ matrices:

$$
\begin{bmatrix}
\bar{A}_{00} & \bar{A}_{10} & \cdots & \bar{A}_{n-1,0} \\
\bar{A}_{10} & \bar{A}_{11} & & \\
\vdots & \vdots & \ddots & \\
\bar{A}_{n-1,0} & \cdots & \bar{A}_{n-1,n-1}
\end{bmatrix}
$$

For $s = O(1)n^{-1}$, $\bar{A}_{ss}$ has the form

$$
\begin{bmatrix}
\cdots & \cdots & \cdots & x \\
\cdot & \cdot & \cdots & x \\
\cdot & \cdot & \cdots & x \\
0 & \cdots & \cdots & x
\end{bmatrix}
$$

and $\bar{A}_{ss'}$ for $0 \leq s' \leq s-1$ has the form

$$
\begin{bmatrix}
\cdots & \cdots & \cdots & x \\
0 & \cdots & \cdots & x \\
\vdots & \vdots & \ddots & x \\
0 & \cdots & \cdots & x
\end{bmatrix}
$$

In like manner we divide the corresponding elements of $B$ into $m \times m$ matrices:

$$
\begin{bmatrix}
B_{00} & B_{10} & \cdots & B_{n-1,0} \\
B_{10} & B_{11} & & \\
\vdots & \vdots & \ddots & \\
B_{n-1,0} & \cdots & B_{n-1,n-1}
\end{bmatrix}
$$

For $s = O(1)n^{-1}$, $B_{ss}$ has the form

$$
\begin{bmatrix}
x & \cdots & \cdots & x \\
x & \cdots & \cdots & x \\
0 & \cdots & \cdots & x \\
0 & \cdots & \cdots & x
\end{bmatrix}
$$

with $\alpha' = \alpha_2$.
and $B_{ss}$ for $0 \leq \sigma \leq s-1$ has the form

$$
\begin{pmatrix}
\cdots & \cdots & \cdots & \cdots & x \\
0 & \cdots & 0 & \cdots & x \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & \cdots & x \\
\end{pmatrix}
$$

Now by theorem 1.6.1 of (16) we obtain that rows $r+um$ to $r+(u+1)m-1$, columns $r+wm$ to $r+(v+1)m-1$, where $u > v$, of $A^{-1}B$ are composed of the $m \times m$ matrix

$$
\sum_{w=v}^{u} \tilde{A}_{uu} B_{uv}
$$

Therefore the matrix on the diagonal of $A^{-1}B$ is $\tilde{A}_{uu} B_{uu}$, which by direct computation we see to be of the same form as $B_{ss}$ above, only with $\alpha' = \alpha$. For $v < u$, which corresponds to one of the $m \times m$ matrices below the diagonal of $A^{-1}B$, we observe by direct computation that $\tilde{A}_{uu} B_{uv}$ has the form

$$
\begin{pmatrix}
\cdots & \cdots & \cdots & \cdots & x \\
0 & \cdots & 0 & \cdots & x \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & \cdots & x \\
\end{pmatrix}
$$

and $\tilde{A}_{uu} B_{uv}$ has the form

$$
\begin{pmatrix}
\cdots & \cdots & \cdots & \cdots & x \\
0 & \cdots & 0 & \cdots & x \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & \cdots & x \\
\end{pmatrix}
$$

$$
\min \left[ m, \max (\alpha'_2 \beta + \alpha_1 - 1) \right]
$$

Hence we obtain the result.

**Lemma 6.2.6.**

If $\{ \xi_i \}$ is an increasing sequence of real numbers and the sequence $\{ \eta_i \}$ is defined by $\eta_1 = \xi_1$, $\eta_i = \min (\xi_i, \eta_{i-1} + 1)$, $i \geq 2$, then $\{ \eta_i \}$ is also increasing.

**Proof.**

We shall show that $\eta_i \leq \eta_{i+1}$ for each $i$. 
\((i) \, i = 1.\)

\[ \eta_2 = \min (\xi_2, \eta_1 + 1) = \min (\xi_2, \xi_1 + 1). \]  
Now \(\xi_2 \geq \xi_1\) since \(\{\xi_i\}\) is increasing, so \(\eta_2 \geq \xi_1 = \eta_1.\)

\((ii) \, i \geq 2.\)

\[ \eta_i = \min (\xi_i, \eta_1 + 1) \text{ and } \eta_{i+1} = \min (\xi_{i+1}, \eta_1 + 1). \]

There are four possibilities:

(a) \(\eta_i = \xi_i\) and \(\eta_{i+1} = \xi_{i+1}\). Then \(\eta_{i+1} \geq \eta_i\) since \(\{\xi_i\}\) is increasing.

(b) \(\eta_i = \xi_i\) and \(\eta_{i+1} = \eta_1 + 1\). Then \(\xi_i = \min (\xi_i, \eta_1 + 1)\), and so \(\xi_i \leq \eta_1 + 1 = \eta_{i+1}\), that is, \(\eta_i \leq \eta_{i+1}\).

(c) \(\eta_i = \eta_1 + 1\) and \(\eta_{i+1} = \xi_{i+1}\). Then \(\eta_1 + 1 = \min (\xi_i, \eta_1 + 1)\), and so \(\eta_1 + 1 \leq \xi_i \leq \xi_{i+1}\) since \(\{\xi_i\}\) is increasing. That is, \(\eta_i \leq \eta_{i+1}\).

(d) \(\eta_i = \eta_1 + 1\) and \(\eta_{i+1} = \eta_1 + 1\). Then \(\eta_i = \eta_{i+1}\).

In all four cases then, \(\eta_i \leq \eta_{i+1}\).

**Lemma 6.2.7.**

If \(\{\xi_i\}\) is an increasing sequence of real numbers, \(c\) is a positive integer greater than 2 and the sequence \(\{\eta_i\}\) is defined by

\[ \eta_1 = \xi_1, \]
\[ \eta_i = \min (\xi_i, \eta_1 + 1), \, i = 2(1)c, \]
\[ \eta_i = \min (\xi_i, \eta_{i-c+1} + 1), \, i \geq c+1, \]

then \(\{\eta_i\}\) is also increasing.

**Proof.**

Lemma 6.2.6 gives \(\eta_1 \leq \eta_2 \leq \ldots \leq \eta_c\) immediately.

Assume now that \(\eta_1 \leq \eta_2 \leq \ldots \leq \eta_q (6.2.1)\) for some positive integer \(q \geq c\).

Then \(\eta_q = \min (\xi_q, \eta_{q-c+1} + 1)\) and \(\eta_{q+1} = \min (\xi_{q+1}, \eta_{q-c+2} + 1)\).
Again there are four possibilities:

(a) \( \eta_q = \xi_q \) and \( \eta_{q+1} = \xi_{q+1} \). Then \( \eta_{q+1} \geq \eta_q \) since \( \{ \xi_i \} \) is increasing.

(b) \( \eta_q = \xi_q \) and \( \eta_{q+1} = \eta_{q+1} > r \). Then \( \eta_q = \xi_q \leq \eta_{q+1} < r \).

\[ \eta_q = \xi_q \leq \eta_{q+1} + 1 \] by (6.2.1), so \( \eta_q \leq \eta_{q+1} \).

(c) \( \eta_q = \eta_{q+1} = \xi_{q+1} \). Then \( \eta_q = \eta_{q+1} + 1 \leq \xi_q \leq \xi_{q+1} \) since \( \{ \xi_i \} \) is increasing, so \( \eta_q \leq \eta_{q+1} \).

(d) \( \eta_q = \eta_{q+1} = \eta_{q+2} + 1 \). Then \( \eta_q = \eta_{q+1} \leq \eta_{q+2} \) by (6.2.1), so \( \eta_q \leq \eta_{q+1} \) and we have the result by induction.

**Definition 6.2.8.**

Take any set of positive integers \( \{ p_1, p_2, \ldots, p_m \} \) and \( 2 \leq \beta \leq \alpha \leq m \).

We define the set \( \{ \pi_1, \pi_2, \ldots, \pi_m \} \) by

\[
\pi_1 = p_1, \\
\pi_\nu = \min (p_\nu, \pi_1 + 1), \quad \nu = 2(1) \min (\beta, m-\alpha+1), \\
\pi_\nu = \min (p_\nu, \pi_{\nu-1} + 1), \quad \nu = \beta + 1(1) m-\alpha+1 \text{ if } \alpha \leq m-\beta, \\
\pi_\nu = \pi_{m-\alpha+1}, \quad \nu = m-\alpha+2(1)m.
\]

**Lemma 6.2.9.**

In definition 6.2.8 if \( p_1 \leq p_2 \leq \ldots \leq p_m \), then \( \pi_1 \leq \pi_2 \leq \ldots \leq \pi_m \).

**Proof.**

This result follows immediately from putting \( \xi = \min (\beta, m-\alpha+1) \) in lemma 6.2.7.

**Theorem 6.2.10.**

Suppose we have an \( m \)-block semi explicit method which is zero stable and optimally consistent of orders \( (p_1, p_2, \ldots, p_m) \) with \( p_1 \leq p_2 \leq \ldots \leq p_m \) with starting values accurate of order \( p_m \). Suppose also that \( A_n \in Q_{m,1} \) and \( B_n \in Q_{m,\beta} \) with \( 2 \leq \beta \leq \alpha < m \). Then for each \( \nu = 1(1)m \)
there exists $c_{\mu}^{(v)}$ such that $\|e^{(v)}\|_\infty \leq c_{\mu}^{(v)} \mu$, where $\pi_1, \pi_2, \ldots, \pi_m$ are defined as in definition 6.2.8.

Proof.

By lemma 6.2.5, $A^{-1}_h B_h \in Q_{m, \alpha, \beta}$, where $\alpha = \min [m, \max (\alpha, \beta)] = \max (\alpha, \beta) = \alpha$ since $\beta \leq \alpha < m$. That is, $A^{-1}_h B_h$ has the form

There are two cases to consider. Case I is when $\alpha > m - \beta$ and case II is when $\alpha \leq m - \beta$.

In case I the $m \times m$ matrices on the diagonal of $A^{-1}_h B_h$ have the form
As in theorem 4.4.2, for each \( \nu = 1(1)m \), \( s = O(1)n-1 \), putting \( i = r+sm+\nu-1 \)

\[
|e_{r+sm+\nu-1}| \leq |(A_h^{-1}e^*)_{r+sm+\nu-1}| + hL \sum_{j=0}^{r+sm+\nu-1} |(A_h^{-1}B_h)_{r+sm+\nu-1,j}| |e_j|
\]

\[
= c(r+sm+\nu-1)_h P_{\nu} + hL \sum_{j=0}^{r+sm+\nu-1} |(A_h^{-1}B_h)_{r+sm+\nu-1,j}| |e_j|.
\]

Now because of zero stability \( |(A_h^{-1}B_h)_{i,j}| \leq M_0 \|B_h\|_1 \) for all \( i,j \) (4.4.1),
so \( \sum_{j=0}^{r-1} |(A_h^{-1}B_h)_{r+sm+\nu-1,j}| |e_j| \leq M_0 \|B_h\|_1 \sum_{j=0}^{r-1} c_j p^m \), and so for sufficiently small \( h \), for each \( s = O(1)n-1 \) there exist constants \( c_1(r+sm+\nu-1)_h, M_1 \) such that

\[
|e_{r+sm}| \leq c_1(r+sm)_h p_1 + hM_1 \sum_{j=r}^{r+sm+1} |e_j|,
\]

\[
|e_{r+sm+\nu-1}| \leq c_1(r+sm+\nu-1)_h P_{\nu} + hM_1 \sum_{\sigma=0}^{s-1} \sum_{q=m-\sigma+1}^{m} |e_{r+sm+q-1}| + hM_1 \sum_{\nu}^\prime,
\]
where
\[ \sum_{\nu} = \begin{cases} 
\sum_{q=1}^{\nu-1} |e_{r+sm+q-1}|, & \nu = 2(1)\beta, \\
\sum_{q=\nu-\beta+1}^{\nu-1} |e_{r+sm+q-1}|, & \nu = \beta + 1(1)m-\alpha+\beta-1, \text{ if } \alpha < m-2, \\
\sum_{q=m-\alpha+1}^{\nu-1} |e_{r+sm+q-1}|, & \nu = m-\alpha+\beta(1)m.
\end{cases} \]

(6.2.4)

taking into account the form of \( A_{h}^{-1}B_{h} \).

The result will follow once we show that for each \( \nu = 1(1)m \) there exists \( C_{U}^{(\nu)} \) such that
\[ |e_{r+sm+\nu-1}| \leq C_{U}^{(\nu)} h, \quad s = O(1)n^{-1} \quad (6.2.5). \]

To prove this we shall use induction on \( s \). First however let us define the set \( \{\pi_{1}^{*}, \pi_{2}^{*}, \ldots, \pi_{m}^{*}\} \) by \( \pi_{1}^{*} = p_1 \),
\[ \pi_{\nu}^{*} = \min (p_{\nu}, \pi_{\nu}^{*} + 1), \quad \nu = 2(1)\beta, \]
\[ \pi_{\nu}^{*} = \min (p_{\nu}, \pi_{\nu}^{*} - \nu + 1), \quad \nu = \beta + 1(1)m. \]

By lemma 6.2.7, \( \pi_{1}^{*} \leq \pi_{2}^{*} \leq \ldots \leq \pi_{m}^{*} \quad (6.2.6). \)

There are two possibilities: case I — \( \alpha + \beta > m \);

\[ \begin{cases} \text{case I — } \alpha + \beta > m; \\
\text{case II — } \alpha + \beta \leq m.
\end{cases} \]

Case I.

In definition 6.2.8 we have \( \pi_{1} = p_{1} \),
\[ \pi_{\nu} = \min (p_{\nu}, \pi_{1}^{*} + 1), \quad \nu = 2(1)m-\alpha+1, \]
\[ \pi_{\nu} = \pi_{m-\alpha+1}, \quad \nu = m-\alpha+2(1)m. \]

Therefore \( \pi_{\nu} = \pi_{\nu}^{*} \) for \( \nu = 1(1)m-\alpha+1 \) (6.2.7) and \( \pi_{\nu}^{*} \geq \pi_{\nu} \) for \( \nu = m-\alpha+1(1)m \)

(6.2.8).

(1) \( s = 0 \).

We must show that \( |e_{r+\nu-1}| \leq C_{U}^{(0, \nu)} h \) (6.2.9), \( \nu = 1(1)m \), for some constants \( C_{U}^{(0, \nu)} \). To do this, we use induction on \( \nu \).
(a) $\forall \epsilon_1 = 1$.

(6.2.4) gives $|e_{r^+q}| \leq c_1(r)h^{p_{q+1}} = c_1(r)\pi_1$.

(b) Assume that (6.2.9) holds for $\forall = 1(1)q : 1 \leq q < m - \alpha + 1$.

$\alpha + \beta > m$, so $\beta \geq m - \alpha + 1$ and (6.2.4) gives

\[
|e_{r^+q}| \leq c_1(r+q)h^{p_{q+1}} + h\sum_{\lambda=1}^{q} |e_{r^+\lambda-1}| \leq c_1(r+q)h^{p_{q+1}} + h\sum_{\lambda=1}^{q} c_U(0,\lambda)\pi_\lambda
\]

\[
\leq c_1(r+q)h^{p_{q+1}} + M_{12}^{(0,q)}h^{\pi_1} \text{ for some } M_{12}^{(0,q)}, \text{ since by lemma 6.2.9}
\]

$\pi_1 \leq \pi_2 \leq \ldots \leq \pi_q$. So $|e_{r^+q}| \leq c_U(0,q+1)h^{\pi_1} \text{ for some } c_U(0,q+1)$. This gives (6.2,5) for $s = 0$, $\forall = 1(1)m - \alpha + 1$.

(c) Assume that (6.2.2) holds for $\forall = 1(1)q : m - \alpha + 1 \leq q < \beta$.

(6.2.4) gives

\[
|e_{r^+q}| \leq c_1(r+q)h^{p_{q+1}} + h\sum_{\lambda=1}^{q} |e_{r^+\lambda-1}|
\]

\[
\leq c_1(r+q)h^{p_{q+1}} + h\sum_{\lambda=1}^{q} c_U(0,\lambda)\pi_\lambda \leq c_U(0,q+1)h^{\pi_1}
\]

for some $c_U(0,q+1) \cdot \pi_1 \geq \pi_{m - \alpha + 1}$ by (6.2.6). $\pi_{m - \alpha + 1} = \pi_{m - \alpha + 1}$ (6.2.7).

$\pi_{m - \alpha + 1} = \pi_{q+1}$ by definition, so we have $|e_{r^+q}| \leq c_U(0,q+1)h^{\pi_{q+1}}$.

(d) Assume that (6.2.9) holds for $\forall = 1(1)q : \beta \leq q < m - \alpha + 1$.

(6.2.4) gives

\[
|e_{r^+q}| \leq c_1(r+q)h^{p_{q+1}} + h\sum_{\lambda=1}^{\pi_{q+1}} |e_{r^+\lambda-1}|
\]

\[
\leq c_1(r+q)h^{p_{q+1}} + h\sum_{\lambda=1}^{\pi_{q+1}} c_U(0,\lambda)\pi_\lambda
\]

\[
\leq c_1(r+q)h^{p_{q+1}} + M_{12}^{(0,q)}h^{\pi_{q+1}} \text{ for some } M_{12}^{(0,q)},
\]

since $\pi_{q+1} \leq \pi_{q+2} \leq \ldots \leq \pi_q$ by lemma 6.2.9.

Now for $\beta \leq q < m - \alpha + 1$, $2 \leq q_2 + 1 \leq m - \alpha + 1$, so $\pi_{q_2 + 1} = \pi \text{ for some } c_1(r+q)$. (6.2.8) implies
that \( \pi_{q+1} \geq \pi_{q+1} \), so \( \|e_{r+q}\| \leq c_{U}^{(0, q+1)} h^{q+1} \).

(e) Assume that (6.2.9) holds for \( \nu = 1(1) q : \pi_{q+1} \leq q \leq \pi \). (6.2.4) gives
\[
|e_{r+q}| \leq c_{1}^{(r+q)} h^{q+1} + h M_{1} \sum_{\lambda = m-\alpha+1}^{m} |e_{r+\lambda-1}|
\]
\[
\leq c_{1}^{(r+q)} h^{q+1} + h M_{1} \sum_{\lambda = m-\alpha+1}^{m} c_{U}^{(0, \lambda)} h^{\pi_{\lambda}}
\]
\[
\leq c_{1}^{(r+q)} h^{q+1} + M_{12}^{(0, q)} h^{m-\pi_{\alpha+1} + 1} \text{ for some } M_{12}^{(0, q)}
\]
since \( \pi_{m-\pi_{\alpha+1}} = \pi_{m-\pi_{2}} = \ldots = \pi_{q} \) by definition.

So by (6.2.7),
\[
|e_{r+q}| \leq c_{1}^{(r+q)} h^{q+1} + M_{12}^{(0, q)} h^{m-\pi_{\alpha+1} + 1}.
\]

By lemma 6.2.7, \( \pi_{m-\pi_{\alpha+1}} \leq \pi_{\alpha+2} \), so
\[
|e_{r+q}| \leq c_{1}^{(r+q)} h^{q+1} + M_{12}^{(0, q)} h^{q+1} + 1 \leq c_{U}^{(0, q+1)} \pi_{q+1} \text{ for some } c_{U}^{(0, q+1)}.
\]

Thus by (6.2.8),
\[
|e_{r+q}| \leq c_{U}^{(0, q+1)} h^{q+1}.
\]

(ii) Assume that \( |e_{r+\sigma+\nu-1}| \leq c_{U}^{(\sigma, \nu)} h^{\nu}, \ \nu = 1(1) m \).

\[
\sigma = O(1) s \text{ for some constants } c_{U}^{(\sigma, \nu)}.
\]

We then need to show that \( |e_{r+(s+1) m-\nu-1}| \leq c_{U}^{(s+1, \nu)} h^{\nu}(6.2.10), \ \nu = 1(1) m \),

for some constants \( c_{U}^{(s+1, \nu)} \). Let us use induction on \( \nu \).

(a) \( \nu = 1 \).

(6.2.4) gives
\[
|e_{r+(s+1) m}| \leq c_{1}^{(r+(s+1) m)} h^{P_{1}} + h M_{1} \sum_{j=r}^{r+(s+1) m-1} |e_{j}|
\]
\[
\leq c_{1}^{(r+(s+1) m)} h^{P_{1}} + M_{12}^{(s+1, q)} h^{r+1} \text{ for some } M_{12}^{(s+1, q)}.
\]
\[
|e_{r+(s+1) m}| \leq c_{U}^{(s+1, 0)} h^{\nu} \text{ for some } c_{U}^{(s+1, 0)}.
\]
(b) Assume that (6.2.10) holds for $v = 1(1)g : 1 \leq g \leq m-\alpha+1$.

(6.2.4) gives

$$|e_{r^+(s+1)m+q}| \leq c_1(r^+(s+1)m+q)h_{p+1} + hM_1 \sum_{\sigma=0}^{s} \sum_{\lambda=m-\alpha+1}^{m} |e_{r^+\sigma m+\lambda-1}|$$

$$+ hM_1 \sum_{\lambda=1}^{q} |e_{r^+(s+1)m+\lambda-1}|$$

$$\leq c_1(r^+(s+1)m+q)h_{p+1} + hM_1 \sum_{\sigma=0}^{s} \sum_{\lambda=m-\alpha+1}^{m} c_U^{(\sigma,\lambda)}h_{\pi^\lambda}$$

$$+ hM_1 \sum_{\lambda=1}^{q} c_U^{(s+1,\lambda)}h_{\pi^\lambda}$$

So $|e_{r^+(s+1)m+q}| \leq c_1(r^+(s+1)m+q)h_{p+1} + M_{12}^{(s+1,q)}h_{m-\alpha+1} + M_{12}^{(s+1,q)}h_{\pi^\lambda+1}$

for some $M_{12}^{(s+1,q)}$, and hence

$$|e_{r^+(s+1)m+q}| \leq M_{13}^{(s+1,q)}h_{\pi^\lambda+1} + M_{12}^{(s+1,q)}h_{m-\alpha+1}$$

for some $M_{13}^{(s+1,q)}$, since $\pi^\lambda+1 \leq m-\alpha+1$.

(c) Assume that (6.2.10) holds for $v = 1(1)g : m-\alpha+1 \leq g \leq \beta$.

(6.2.4) gives

$$|e_{r^+(s+1)m+q}| \leq c_1(r^+(s+1)m+q)h_{p+1} + hM_1 \sum_{\sigma=0}^{s} \sum_{\lambda=m-\alpha+1}^{m} |e_{r^+\sigma m+\lambda-1}|$$

$$+ hM_1 \sum_{\lambda=1}^{q} |e_{r^+(s+1)m+\lambda-1}|$$

$$\leq c_1(r^+(s+1)m+q)h_{p+1} + M_{12}^{(s+1,q)}h_{m-\alpha+1} + M_{12}^{(s+1,q)}h_{\pi^\lambda+1}$$

as before,

$$\leq M_{13}^{(s+1,q)}h_{\pi^\lambda+1} + M_{12}^{(s+1,q)}h_{m-\alpha+1}.$$
\[ \left| e_{r+1}(s+1)_{m+q} \right| \leq c_1^{(s+1,q)} h^{q+1} \text{ for some } c_1^{(s+1,q)}. \]

(d) Assume that (6.2,10) holds for \( \alpha = 1(1) q : 0 \leq q \leq m-\alpha+1 \).

(6.2,4) gives

\[ \left| e_{r+1}(s+1)_{m+q} \right| \leq c_1^{(r+1)(s+1)_{m+q}} h^{q+1} + hM_{12} \sum_{\sigma=0}^{\pi} \sum_{\lambda=m-\alpha+1}^{\infty} \left| e_{r+1} \sigma m \lambda-1 \right| \]
\[ + hM_{12} \sum_{\lambda=q-\alpha+1}^{q+1} \left| e_{r+1} \sigma m \lambda-1 \right| \]
\[ \leq c_1^{(r+1)(s+1)_{m+q}} h^{q+1} + M_{12}^{(s+1,q)} h^{m-\alpha+1} + M_{12}^{(s+1,q)} h^{q-\alpha+2} \text{ for some } M_{12}^{(s+1,q)}, \]
\[ \leq M_{13}^{(s+1,q)} h^{q+1} + M_{12}^{(s+1,q)} h^{m-\alpha+1} \text{ for some } M_{13}^{(s+1,q)}, \]

applying the same reasoning as in (i) (d).

Then since \( \pi \geq \pi_{m-\alpha+1} \) there exists \( c_1^{(s+1,q+1)} \) such that

\[ \left| e_{r+1}(s+1)_{m+q} \right| \leq c_1^{(s+1,q+1)} h^{q+1} \text{ for some } c_1^{(s+1,q+1)}. \]

(e) Assume that (6.2,10) holds for \( \alpha = 1(1) q : m-\alpha+1 \leq q \leq m-1 \).

(6.2,4) gives

\[ \left| e_{r+1}(s+1)_{m+q} \right| \leq c_1^{(r+1)(s+1)_{m+q}} h^{q+1} + hM_{12} \sum_{\sigma=0}^{\pi} \sum_{\lambda=m-\alpha+1}^{\infty} \left| e_{r+1} \sigma m \lambda-1 \right| \]
\[ + hM_{12} \sum_{\lambda=m-\alpha+1}^{q+1} \left| e_{r+1} \sigma m \lambda-1 \right| \]
\[ \leq c_1^{(r+1)(s+1)_{m+q}} h^{q+1} + M_{12}^{(s+1,q)} h^{m-\alpha+1} + M_{12}^{(s+1,q)} h^{q-\alpha+2} \text{ for some } M_{12}^{(s+1,q)}, \]
\[ \leq M_{13}^{(s+1,q)} h^{q+1} + M_{12}^{(s+1,q)} h^{m-\alpha+1} \text{ for some } M_{13}^{(s+1,q)}, \]
applying the same reasoning as in (i) (e). Then since \( \pi_{q+1} > \pi_{m-\alpha+1} \),
there exists \( c_{U}^{(s+1,q+1)} \) such that \( |e_{r+s+1,m+q}| \leq c_{U}^{(s+1,q+1)} h^{q+1} \).

This completes the proof for case I.

Case II.

In definition 6.2.8 we have

\[
\pi_{1} = p_{1},
\]
\[
\pi_{\nu} = \min(p_{\nu}, \pi_{1} + 1), \quad \nu = 2(1)\beta,
\]
\[
\pi_{\nu} = \min(p_{\nu}, \pi_{\nu-\beta+1} + 1), \quad \nu = \beta+1(1)m-\alpha+1,
\]
\[
\pi_{\nu} = \pi_{m-\alpha+1}, \quad \nu = m-\alpha+2(1)m.
\]

Again \( \pi_{\nu} = \pi_{0} \) for \( \nu = 1(1)m-\alpha+1 \) (6.2.11) and \( \pi_{\nu} > \pi_{\nu} \) for \( \nu = m-\alpha+1(1)m \)
(6.2.12).

(i) \( s = 0 \).

(a) Assume that (6.2.9) holds for \( \nu = 1(1)q : 0 \leq q < \beta \).

Then we obtain
\[
|e_{r}| \leq c_{U}^{(0,1)} h^{q+1} \text{ as in case I}.
\]

(b) Assume that (6.2.9) holds for \( \nu = 1(1)q : q \leq q < \beta+1 \).

We obtain
\[
|e_{r+q}| \leq c_{U}^{(0,q+1)} h^{q+1} \text{ as in case I, (i) (d)}.
\]

(c) Assume that (6.2.9) holds for \( \nu = 1(1)q : m-\alpha+\beta-1 \leq q < m \).

We obtain
\[
|e_{r+q}| \leq c_{U}^{(0,q+1)} h^{q+1} \text{ as in case I, (i) (e)}.
\]

(d) Assume that (6.2.9) holds for \( \nu = 1(1)q : m-\alpha+\beta-1 \leq q < m \).

We obtain
\[
|e_{r+q}| \leq c_{U}^{(0,q+1)} h^{q+1} \text{ as in case I, (i) (d)}.
\]

(ii) Assume that \( |e_{r+q+m+q-1}| \leq c_{U}^{(\sigma,\nu)} h^{\nu}, \quad \nu = 1(1)m, \sigma = O(1)s, \)
for some constants \( c_{U}^{(\sigma,\nu)} \).

Going through the steps (a), (b), (c), (d) as above we obtain
the required result as in case I, and so we have the result for both cases.

This theorem gives a sequence of upper bounds on the various stages of the global error vector, and we can now apply this to various examples, and shall go on in the next section to give two-sided error bounds for some cases.

Example 6.1.2.

$A_h \in Q_{3,1}$ and $B_h \in Q_{3,1}^{(2)}$ and we saw previously that optimal consistency is of orders $(2,3,3)$. We have also noted in section 4.3 that a predictor corrector scheme is zero stable if and only if the corrector is when considered as a single linear multistep method. This is the case in the method under consideration. Thus if we take starting values accurate of order three, the conditions of theorem 6.2.10 are fulfilled. $\beta = 2$ and hence we obtain $\pi_1 = 2$, $\pi_2 = \pi_3 = 3$, and

\[
\| e^{(1)} \|_{\infty} \leq c^{(1)}_U h^2 ,
\]

\[
\| e^{(2)} \|_{\infty} \leq c^{(2)}_U h^3 ,
\]

\[
\| e^{(3)} \|_{\infty} \leq c^{(3)}_U h^3 \quad \text{for some } c^{(1)}_U , c^{(2)}_U , c^{(3)}_U .
\]

Example 6.1.3.

$A_h \in Q_{3,1}$ and $B_h \in Q_{3,1}^{(2)}$ and optimal consistency is of orders $(2,3,3)$. We have also seen that any Runge-Kutta method is zero stable. Thus if we take a starting value accurate of order three, the conditions of theorem 6.2.10 are fulfilled. Again $\beta = 2$, $\pi_1 = 2$, $\pi_2 = \pi_3 = 3$, and so
Example 6.2.11.

Suppose we have a predictor-corrector method in $P(EC)^{\mu}(E)$ mode using a predictor of order $p-\mu$ and a corrector of order $p$, with $\mu \geq 2$.

\[ A_h \in Q_{\mu+1,1} \quad \text{and in } P(EC)^{\mu} \text{ mode } B_h \in Q_{\mu+1,2} \quad \text{and in } P(EC)^{\mu} \text{ mode } B_h \in Q_{\mu+1,2} \]

so in both cases $B_h \in Q_{\mu+1,2}$.

By lemma 6.2.4 $A_h^{-1} \in Q_{\mu+1,1}^{(1)}$ and has the form

\[
\begin{bmatrix}
\text{x} & \text{x} & \ldots & \text{x} \\
\text{x} & \text{0}^* & \ldots & \text{x} \\
\text{x} & \text{x} & \ldots & \text{x} \\
\text{x} & \text{x} & \ldots & \text{x} \\
\text{x} & \text{x} & \ldots & \text{x} \\
\end{bmatrix}
\]

Taking starting values accurate of order $p$,

\[ \Theta^* = (C_0^P, \ldots, C_{\mu-1}^P, C_{\mu+1}^P, \ldots, C_{\mu+1}^P, \ldots)^T \]

where $C^*$ and $C$ are the error constants of the predictor and the corrector respectively.

Therefore by computing $A_h^{-1}\Theta^*$ directly we obtain optimal consistency of orders $(p-\mu+1, p, \ldots, p)$. Then provided we take a zero stable corrector, the conditions of theorem 6.2.10 are fulfilled with $\beta = \alpha = 2$, and so we obtain

\[ \begin{align*}
\|e^{(1)}\|_\infty & \leq c_u^{(1)} h^2, \\
\|e^{(2)}\|_\infty & \leq c_u^{(2)} h^3, \\
\|e^{(3)}\|_\infty & \leq c_u^{(3)} h^3.
\end{align*} \]
\|e^{(1)}\|_{\infty} \leq c^{(1)}_{U} h^{p-\mu+1},
\|e^{(2)}\|_{\infty} \leq c^{(2)}_{U} h^{p-\mu+2},
\|e^{(\mu-1)}\|_{\infty} \leq c^{(\mu-1)}_{U} h^{p-1},
\|e^{(\mu)}\|_{\infty} \leq c^{(\mu)}_{U} h^{p},
\|e^{(\mu+1)}\|_{\infty} \leq c^{(\mu+1)}_{U} h^{p}.

We note that both \( y_{r+s(\mu+1)+\mu} \) and \( y_{r+s(\mu+1)+\mu-1} \) converge with order at least \( p \). This implies that on the final application of the predictor-corrector scheme the \( \mu \)th iteration of the corrector may be omitted and the same order of convergence obtained.

Example 6.2.12.

If we take as predictor \( y_{i} = y_{i-1} + hf_{i-1} \), and as corrector
\( y_{i} = y_{i-1} + \frac{h}{720} (251 f_{i} + 646 f_{i-1} - 264 f_{i-2} + 106 f_{i-3} - 19 f_{i-4}) \)
in \( P(EC) \) mode we have a special case of example 6.2.11 with \( \mu = 4 \) and \( p = 5 \). Hence
\[ \|e^{(1)}\|_{\infty} \leq c^{(1)}_{U} h^{2}, \]
\[ \|e^{(2)}\|_{\infty} \leq c^{(2)}_{U} h^{3}, \]
\[ \|e^{(3)}\|_{\infty} \leq c^{(3)}_{U} h^{4}, \]
\[ \|e^{(4)}\|_{\infty} \leq c^{(4)}_{U} h^{5}, \]
\[ \|e^{(5)}\|_{\infty} \leq c^{(5)}_{U} h^{5}. \]

Example 6.2.13.

Consider a method composed of an explicit predictor of order \( p-\mu \),
followed by $\mu$ different implicit correctors $C_1, C_2, \ldots, C_\mu$, of orders $p-\mu+1, p-\mu+2, \ldots, p$.

We then have a predictor-corrector scheme in $P \bigoplus_{\omega=1}^{\mu} (EC(2))(E)$ mode. Again $A_h \in Q_{\mu+1,1}^1$ and $B_h \in Q_{\mu+2,2}$ and we have zero stability. Taking starting values accurate of order $p$ we obtain

$$\theta^* = (C_0 h^p, \ldots, C_{r-1} h^p, C_* h^{p-\mu+1}, C_1 h^{p-\mu+2}, \ldots, C_{\mu-1} h^p, C_\mu h^{p+1})^T,$$

where $C^*$ is the error constant of the predictor and $C_1, \ldots, C_\mu$ are the error constants of the correctors. Again we obtain by direct computation $A_h^{-1} \theta^*$ optimal consistency of orders $(p-\mu+1, p-\mu+2, \ldots, p-1, p, p)$. The conditions of theorem 6.2.10 are again fulfilled with $\beta = \alpha = 2$, giving

$$\|e^{(1)}\|_\infty \leq C_U^{(1)} h^{p-\mu+1},$$
$$\|e^{(2)}\|_\infty \leq C_U^{(2)} h^{p-\mu+2},$$
$$\|e^{(\mu-1)}\|_\infty \leq C_U^{(\mu-1)} h^{p-1},$$
$$\|e^{(\mu)}\|_\infty \leq C_U^{(\mu)} h^p,$$
$$\|e^{(\mu+1)}\|_\infty \leq C_U^{(\mu+1)} h^p.$$

6.3 Two-Sided Error Bounds.

In some cases we may continue the analysis to obtain sequences of two-sided error bounds of the form (6.1.2).

Theorem 6.3.1. Suppose we have an $m$-block semi explicit method for which the conditions of theorem 6.2.10 are fulfilled and $m \geq 3$.

Then $C_L h^{p_1} \leq \|e\|_\infty \leq C_U h^{p_1}$ for some $C_L$ and $C_U$ (6.3.1).
If in addition there exists \( q \geq \beta + 1 \) such that \( p_q = \pi_{q-\beta+1} \), then we have the two-sided error bound

\[
C_L(q^*, q^*+1, \ldots, m-1, m) h^{q^*} \leq \| e(q^*, q^*+1, \ldots, m-1, m) \|_\infty
\]

\[
\leq C_U(q^*, q^*+1, \ldots, m-1, m) h^{q^*}
\]

(6.3.2)

for some \( C_L(q^*, q^*+1, \ldots, m-1, m) \) and \( C_U(q^*, q^*+1, \ldots, m-1, m) \), where

\[ q^* = \min(q-\beta+1, m-\alpha+1) \]

Proof.

(6.3.1) follows immediately from (6.1.1). For each \( q \geq \beta + 1 \),

\[
\| e(q^*, q^*+1, \ldots, m-1, m) \|_\infty = \max_{\nu=q^*(1)} \| e(\nu) \|_\infty \leq C_U(q^* h^{q^*}
\]

by theorem 6.2.10, since \( \pi_q \leq \pi_{q^*+1} \leq \ldots \leq \pi_m \) by lemma 6.2.9. So we have the upper bound in (6.3.2), taking \( C_U(q^*, q^*+1, \ldots, m-1, m) = C_U(q^*) \).

Now \( A_h^{-1} e^* = e - \frac{1}{h} A_h^{-1} B^h [f(A_h t, y^*) - f(A_h t, \Delta_h x(t))] (5.1.3) \), so for each \( s = \Omega(1)n-1 \),

\[
C_r(s + m + q-1) h^p \leq \| e_{r+s+m+q-1} \| + h \| [A_h^{-1} B^h [f(A_h t, y^*) - f(A_h t, \Delta_h x(t))] ]_{r+s+m+q-1} \|
\]

\[
\leq \| e_{r+s+m+q-1} \| + h M \sum_{j=0}^{r-1} \| e_j \| + h M' \sum_{s=0}^{s-1} \sum_{\lambda=m-\alpha+1}^{s} \| e_{r+s+m+\lambda-1} \|
\]

\[
+ h M' \sum_{\lambda=q^*}^{l} \| e_{r+s+m+\lambda-1} \| ,
\]

taking into account the particular form of \( A_h^{-1} B_h \). Now

\[
| e_{r+s+m+q-1} | \leq \| e(q^*, q^*+1, \ldots, m-1, m) \|_\infty ,
\]

\[
| e_i | \leq \| e(q^*, q^*+1, \ldots, m-1, m) \|_\infty \text{ for } i = \Omega(1)r-1 ,
\]
\[ |e_{r+\sigma m+\lambda-1}| \leq \|e(q^*, q^*+1, \ldots, m-1, m)\|_\infty \text{ for } \sigma = O(1)n-1, \lambda = m-\lambda+1, m, \]
\[ |e_{r+\lambda m+\lambda-1}| \leq \|e(q^*, q^*+1, \ldots, m-1, m)\|_\infty \text{ for } \lambda = O(1)n-1, m = q^*(1)m, \]
so we obtain \[ C^* \|e(q^*, q^*+1, \ldots, m-1, m)\|_\infty \text{ for some } C^* . \]

Now \( q^* = \min(q^*-1, m-\lambda+1) \) and \( \pi_{\lambda} = \pi_{m-\lambda+1} \) for \( \lambda = m-\lambda+2(1)m, \)
so \( \pi_{q^*} = \pi_{q^*-1} = \pi_q \), and the lower bound in (6.3.2) follows, if we take \( C^* \left( q^*, q^*+1, \ldots, m-1, m \right) = \frac{1}{C^*} \min_{m=O(1)n-1} \left\{ C^* \left( r+\sigma m+q-1 \right) \right\} . \)

The first part of this theorem gives a two-sided bound on the maximum of all the global errors, and the second part will give two-sided bounds on the error in some of the more accurate stages in certain methods.

Example 6.1.2.

(6.3.1) gives \( C_L h^2 \leq \|e\|_\infty \leq C_U h^2 \). \( \beta = 2 \), so \( q^*-1 = q-1 \), and for \( q = 3 \) we have \( p_3 = q^* = \min(2, 3) = 2 \), \( \pi_{q^*} = \pi_2 = 3 \).

Hence (6.3.2) gives \( C_L (2, 3)^3 \leq \|e(2, 3)\|_\infty \leq C_U (2, 3)^3 \) (6.3.3).

That is, we have exactly order three convergence in the second and third stages.

Example 6.1.3.

Again we have \( C_L h^2 \leq \|e\|_\infty \leq C_U h^2 \) and \( C_L (2, 3)^3 \leq \|e(2, 3)\|_\infty \leq C_U (2, 3)^3 \) (6.3.4) by applying theorem 6.3.1 exactly as in the previous example.

Example 6.2.13.

(6.3.1) gives \( C_L h^{p-\mu+1} \leq \|e\|_\infty \leq C_U h^{p-\mu+1} \). \( \beta = 2 \), so for \( q = \mu + 1 \), \( p_q = \pi_{q-1} = p \). \( q^* = \min(\mu, \mu) = \mu \), so \( \pi_{q^*} = \pi_{\mu} = p \). Hence (6.3.2) gives \( C_L (\mu, \mu+1)^P \leq \|e(\mu, \mu+1)\|_\infty \leq C_U (\mu, \mu+1)^P \) (6.3.5), and so the \((\mu-1)^{\text{th}}\) and
\( \mu \) th. correctors converge with order exactly \( p \).

Example 6.1.3 is Heun's third order formula (15, p.119), a Runge-Kutta method for which we have obtained the correct orders of convergence at each of the stages through optimal consistency, based on the calculation only of the local truncation errors of the following three linear multistep methods based on steplength \( h/3 \):

\[
\begin{align*}
y_{i-2/3} - y_{i-1} &= \frac{h}{2} f_{i-1}, \\
y_{i-1/3} - y_{i-1} &= \frac{h}{2} f_{i-2/3}, \\
y_{i} - y_{i-1} &= \frac{h}{4} (f_{i-1} + 3f_{i-1/3}).
\end{align*}
\]

Our approach not only gives two-sided error bounds, but is also far simpler than the standard means of calculating the local truncation error \( y(t_i) - y(t_{i-1}) = h\delta(t_{i-1}, y(t_{i-1})) \), where \( \delta(t_{i-1}, y_{i-1}) = \frac{1}{4}(k_1 + 3k_3) \) and

\[
\begin{align*}
k_1 &= f(t_{i-1}, y_{i-1}), \\
k_2 &= f(t_{i-1} + h/3, y_{i-1} + hk_1/3), \\
k_3 &= f(t_{i-1} + 2h/3, y_{i-1} + 2hk_2/3).
\end{align*}
\]

Example 6.3.2.

\[
\begin{align*}
y_{i-1} - y_{i-2} &= \frac{h}{8} (9f_{i-1} + 3f_{i-2}), \\
y_{i} - \frac{1}{5}(28y_{i-1} - 23y_{i-2}) &= \frac{h}{75} (-255f_{i-1} - 160f_{i-2} + 136f_{i-3} + 9f_{i-4}), \\
y_{i} - y_{i-1} &= \frac{h}{720} (251f_{i} + 646f_{i-1} - 264f_{i-2} + 106f_{i-3} - 19f_{i-4}).
\end{align*}
\]

This is a 3-block explicit method with array

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 23/5 & 0 & 0 & -28/5 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & -19/720 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 3/8 & 0 & 0 & 9/8 & 0 & 0 & 0 & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -32/15 & 0 & 0 & -17/5 & 136/75 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 53/360 & 0 & 0 & -11/30 & 0 & 343/360 & 0 & 251/720 & 0 & 1
\end{pmatrix}
\]
For zero stability we obtain
\[
G = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 \\
0 & -23/5 & 0 & 0 & 28/5 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix},
\]
which has eigenvalues 0, 0, 0, 0, 1, giving zero stability. Assuming starting values accurate of order 5,
\[
\varphi^* = \left( C_0 h^5, C_1 h^5, C_2 h^5, C_3 h^5, \frac{3}{128} h^4 y_{i+1}(t_{7/2}) + o(h^5), \frac{-31}{1200} h^5 y_{i+1}(t_{7/2}) + o(h^7), \ldots \right)^T.
\]

\[A_h \in Q^{(1)}_{3,1}, \text{ so by lemma 6.2.4 } A_h^{-1} \in Q^{(1)}_{3,1} \text{ and by direct computation of } A_h^{-1} \varphi^* \text{ we obtain optimal consistency of orders } (4, 5, 5).
\]

\[B_h \in Q^{(2)}_{3,1}, \text{ so theorem 6.2.10 gives }
\]
\[
c_L h^p \leq \|e\|_\infty \leq c_U h^p
\]
and since \( \beta = 2 \), \( p_3 = \pi_2 = 5 \) and \( q^* = 2 \),
\[
c_L^{(2,3)} h^5 \leq \|e(2,3)\|_\infty \leq c_U^{(2,3)} h^5 \quad (6.3.6).
\]

**Example 6.2.11.**

(6.3.1) gives \( c_L h^{p-\mu+1} \leq \|e\|_\infty \leq c_U h^{p-\mu+1} \). \( \beta = 2 \) and \( p_{\mu+1} = \pi_\mu = p \), \( q^* = \min(\mu, \mu) = \mu \). So \( p_{\mu+1} = \pi_\mu = p \), and (6.3.2) gives
\[
c_L^{(\mu, \mu+1)} h^p \leq \|e(\mu, \mu+1)\|_\infty \leq c_U^{(\mu, \mu+1)} h^p \quad (6.3.7).
\]

**Example 6.3.3.**

Consider the following \( P_{2/3}^{EP_{1/2}} H \) hybrid method:
\[
y_{i-1/3} - y_{i-1} = \frac{b_r}{9}(8f_{i-1} - 2f_{i-2}),
\]
\[
y_{i-1/2} - y_{i-2} = \frac{b_r}{8}(9f_{i-1} + 3f_{i-2}),
\]
\[
y_{i-1} - y_{i-1} = \frac{b_r}{4}(3f_{i-1/3} + f_{i-1}),
\]
\[
y_i - y_{i-1} = \frac{b_r}{6}(f_{i} + f_{i-1} + 4f_{i-1/2}).
\]
This is a 4-block explicit method with array

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 0 & -2/9 & 0 & 0 & 0 & 8/9 \\
0 & 0 & 0 & 3/8 & 0 & 0 & 0 & 9/8 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/4 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 2/3 \\
0 & 0 & 0 & 0 & 0 & 1/2 \\
3/4 & 0 & 0 & 0 & 1 \\
0 & 2/3 & 1/6 & 0 & 1
\end{bmatrix}
\]

\[
G = \begin{bmatrix}
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

, which has eigenvalues 0, 0, 0, 0, 0, 1, giving zero stability.

Assuming starting values accurate of order 4, we have

\[
g^* = (C^* h^4, C^* h^4, 13 \frac{1}{8} h^3 y''(t_{5/3}) + 0(h^4), \frac{3}{128} h^3 y'(t_{5/3}) + 0(h^5),
\]

\[
\frac{1}{216} h^4 y'(t_{5/3}) + 0(h^5), \frac{1}{2880} h^5 y(t_{5/3}) + 0(h^6), \cdots )^T.
\]

\[A_h \in Q_{4,1}, \text{ so } A_h^{-1} \in Q_{4,1}^{(1)} \text{ and direct computation of } A_h^{-1} g^* \text{ we obtain optimal consistency of orders (3, 4, 4, 4)}.
\]

\[B_h \in Q_{4,1}^{(3)} \subset Q_{4,3}^{(3)}, \text{ and theorem 6.2.10 gives}
\]

\[\|e^{(1)}\|_{\infty} \leq C_U^U h^3 \text{ and } \|e^{(\upsilon)}\|_{\infty} \leq C_U^{(\upsilon)} h^4, \upsilon = 2, 3, 4 \ (6.3.8).
\]

Also, theorem 6.3.1 gives \( C_{L} h^3 \leq \|e\|_{\infty} \leq C_{U} h^3 \) and

\[C_{L}^{(2, 3, 4)} h^4 \leq \|e^{(2, 3, 4)}\|_{\infty} \leq C_{U}^{(2, 3, 4)} h^4 \ (6.3.9).
\]
Example 6.3.4.

The seventh example given in the concluding section of (5) is a 4-block explicit method with array

\[
\begin{bmatrix}
0 & -4/9 & 0 & -5/9 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 5/9 & 0 & 0 & 0 & 0 & 1/3 \\
0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 1/8 & 3/8 & 0 & 0 & 0 & 1/2 \\
0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 1/2 & -3/2 & 2 & 0 & 0 & 1 \\
0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1/6 & 0 & 2/3 & 1/6 & 0 & 1
\end{bmatrix}
\]

\[G = \begin{bmatrix}
0 & 4/9 & 0 & 5/9 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix}\]

which has eigenvalues 0, 0, 0, 1, giving zero-stability.

Assuming starting values accurate of order four, we have
\[\theta^* = (C_0^4 h^4, C_1^4 h^4, \frac{-1}{324} h^3 y''(t_{5/6}) + O(h^4), \frac{1}{3456} h^5 y''(t_{5/6}) + O(h^5), \frac{1}{108} h^4 y'''(t_{5/6}) + O(h^5), \frac{-1}{2880} h^5 y''(t_{5/6}) + O(h^6), \ldots)^T.
\]

\[A_n^{-1} \in Q_{4,1}^1, \text{ so } A_n^{-1} \in Q_{4,1}^{1}, \text{ and by direct computation of } A_n^{-1} \theta^* \text{ we obtain optimal consistency of orders } (3,4,4,4).
\]

Definition 6.4.1.

An m-block method is CONSISTENT of ORDER ZERO if there exist \(C_1(1), C_2(1), \ldots, C_m(1)\), some, but not all of which may be zero, such that for each \(y \in C(1)[0,T]\).
We note that a method which is consistent of order zero is not consistent at all in the sense of definition 4.1.3.

**Definition 6.4.2.**

For any positive integer \( m \) we define the vectors \( \delta^e_\nu, x_\nu = 1 \) \( (1) m \), by \( \delta^e_\nu = \delta_\nu \), \( q = 1 \) \( (1) m \), where \( \delta_\nu \) is the Kronecker delta, and we are here beginning the vector suffices at one rather than zero.

**Theorem 6.4.3.**

Suppose we have an \( m \)-block semi explicit method which is zero stable and consistent of order \( p \geq 0 \). Suppose also that

(i) \( A_n^{-1} (j) = 0 \), \( \theta = O(1)^{T-1} \), \( j = 1 \) \( (1) q \), where \( q \geq 1 \),

(ii) \( \delta^{T A_n^{-1} c}_m (j) = 0 \), \( j = 1 \) \( (1) q-1 \),

(iii) \( \delta^{T A_n^{-1} c}_m (1) \neq 0 \), \( \nu = 1 \) \( (1) m-1 \),

(iv) \( \delta^T A_\nu A_\nu \gamma^T B_\beta = \lambda(\theta, \lambda) \delta^T_m \) for some constants \( \lambda(\theta, \lambda) \), independent of \( h \),

\( \nu = 1 \) \( (1) m \), \( \beta = O(1)^{T-1} \), \( \theta = O(1)^{T-1} \),

(v) \( \delta^T_m \) is a left eigenvector of \( A_\gamma B_\beta \), \( \beta = O(1)^{T-1} \).

Then there exist \( C_L \), \( C_U \) such that \( C_L h^{p+1} \leq \|e\|_\infty \leq C_U h^{p+1} \) and \( C_L h^{p+q'} \leq \|e^{(m)}\|_\infty \leq C_U h^{p+q'} \), where \( q' \) is an integer not less than \( q \), assuming starting values accurate of order \( p+q' \).

**Proof.**

(a) Let us first determine the orders of optimal consistency.

As in the proof of theorem 5.3.3 we need only consider

\[
\begin{pmatrix}
\theta_r \\
\vdots \\
\theta_{r+n-1} \\
\theta_{r+nm-1}
\end{pmatrix}
\]

For each \( s = O(1)n^{-1} \) rows \( sm \) to \( sm+m-1 \) of this vector are

\[
\sum_{\sigma=0}^{s-1} P_{\sigma} A_r^{-1} \sum_{j=1}^{\infty} o(j) [\Delta_h y^{(p+j)}(t)]^r_{s-\sigma} h^{p+j} + A_r^{-1} \sum_{j=1}^{\infty} o(j) [\Delta_h y^{(p+j)}(t)]^r_{s+nm} h^{p+j},
\]
assuming \( y(t) \) has sufficient continuity.

Now by (i) \( \sum_{\sigma=0}^{s-1} \sum_{j=1}^{\infty} c(j)[A_h y(p+j)(t)]_{r+s-1} h^{p+j} = 0(1) \), so

\[
\sum_{\sigma=0}^{s-1} \sum_{j=1}^{\infty} c(j)[A_h y(p+j)(t)]_{r+(s-\sigma-1)m} h^{p+j}
\]

if we define \( a = O(h^b) as h \to 0 \) for vector \( a \) and scalar \( b \) to mean that

\[ ||a||_\infty = O(h^b) as h \to 0. \]

(iii) implies that the first \( m-1 \) components of \( A_h^{-1} c(1) \) are non-zero. Elements \( sm \) to \( sm+m-2 \) of

\[
D_h^{-1} \begin{pmatrix}
\Theta^*_{r+1} \\
\vdots \\
\Theta^*_{r+sm-1}
\end{pmatrix}
\]

are elements \( sm \) to \( sm+m-2 \) of

\[ O(h^{p+q}) + A_h^{-1} c(1)[A_h y(p+1)(t)]_{r+sm} h^{p+1} + O(h^{p+2}), \]

and for \( s = O(1)n-1 \) and \( \nu = 1(1)m-1 \) there exist \( C^{(r+sm+\nu-1)} \) such that

\[ ||[A_h^{-1} \Theta^*]_{r+sm+\nu-1}|| = C^{(r+sm+\nu-1)} h^{p+1}. \]

(ii) implies that the \( m \)th component of \( A_h^{-1} c(j) \) is zero for

\[ j = 1(1)q-1, \]

hence the \( m \)th component of

\[
A_h^{-1} \sum_{j=1}^{\infty} c(j)[A_h y(p+j)(t)]_{r+sm} h^{p+j} is O(h^{p+q}).
\]

Therefore for each \( s = O(1)n-1 \), \( [A_h^{-1} \Theta^*]_{r+sm+m-1} = O(h^{p+q}) \), and there exist \( C^{(r+sm+m-1)} \) and some \( q' \geq q \) such that

\[ ||[A_h^{-1} \Theta^*]_{r+sm+m-1}|| = C^{(r+sm+m-1)} h^{p+q'}. \]

So optimal consistency is of orders \((p+1,p+1, \ldots, p+1,p+q')\).

The result \( C_L h^{p+1} \leq ||e||_\infty \leq C_U h^{p+1} \) for some \( C_L, C_U \) now follows by (6.1.1).
(b) $B_h$ has the form (2.1.2) and so we may write

$$B_h = \begin{pmatrix} 0 & \vdots \\ V_h & U_h \end{pmatrix}, \quad U_h = \begin{pmatrix} B_t \\ \vdots \\ B_o \end{pmatrix}.$$

Hence $A_h^{-1} B_h = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -D^{-1} E_h & D^{-1} \\ V_h & U_h \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -D^{-1} V_h & D^{-1} U_h \end{pmatrix}$.

Let us now investigate $D_h^{-1} U_h$.

$$D_h^{-1} U_h = \begin{pmatrix} A^{-1} \\ P_o A^{-1} \\ \vdots \\ P_{n-2} A^{-1} \\ P_{n-1} A^{-1} \end{pmatrix} A_h^{-1} = \begin{pmatrix} B_t \\ \vdots \\ B_o \end{pmatrix}.$$

Consider rows $\sigma m$ to $\sigma m + m - 1$, columns $\sigma m$ to $\sigma m + m - 1$ of $D_h^{-1} U_h$ for $\sigma \leq s$.

I. When $\sigma = s$, we have $A^{-1}_s B_t$.

II. When $s - \tau \leq \sigma \leq s - 1$, we have $A_s^{-1} B_{\tau - s + \sigma} + \sum_{\xi = 0}^{s-\sigma-1} P_{\xi} A_s^{-1} B_{\xi - s + \sigma + \tau + 1}$.

III. When $\sigma \leq s - \tau - 1$, we have $\sum_{\xi = s - \sigma - \tau - 1}^{s-\sigma-1} P_{\xi} A_s^{-1} B_{\xi - s + \sigma + \tau + 1}$.

Now (iv) implies that for $\Theta = O(1) \tau - 1$, $\phi = O(1) \tau$,

$$A_0 A_s^{-1} B_{\phi} = \begin{pmatrix} 0 & \cdots & 0 & \lambda_{1}^{(\Theta, \phi)} \\ \vdots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda_{m}^{(\Theta, \phi)} \end{pmatrix}.$$
(v) implies that for each $\theta = O(1)\tau$ there exists $\lambda(\theta)$ such that
$$\delta_m^T A^{-1}_{\theta} B = \lambda(\theta) \delta_m^T,$$
which in turn implies that the $m$th row of $A^{-1}_{\theta} B$ is $(0, \ldots, 0, \lambda(\theta)).$

Therefore for each $s,$ row $sm+m-1$ of $D^{-1}_h$ is of the form
$$(0, \ldots, 0, x, 0, \ldots, 0, x, \cdots, 0, \cdots, 0)$$
where the crosses denote non-zero terms. Now
$$|e_i| \leq |(A^{-1}_h e^*)_i| + h^s \sum_{j=0}^{\alpha} |(A^{-1}_h B)_i|^2 |e_j|$$
for all $i.$

Therefore for each $s$
$$|e_{r+sm+m-1}| \leq C^{(r+sm+m-1)} P^{p+q'} + h_{r+s}^{r-1} \sum_{j=0}^{q'} |e_{r+\sigma s+m-1}| + h_{r+s}^s \sum_{j=0}^{q'} |e_{r+\sigma s+m-1}|.$$

Hence there exists $C$ such that
$$|e_{r+s}| \leq C^{h \tau P^{p+q'}} + h_{r+s}^{r-1} \sum_{j=0}^{q'} |e_{r+\sigma s+m-1}|,$$
and lemma 4.4.3 gives $\|e_{r+s}\|_\infty \leq C_U h^{P+q'}$ for some $C_U.$

Now for all $i,$
$$|(A^{-1}_h e^*)_i| \leq |e_i| + h^s \sum_{j=0}^{\alpha} |(A^{-1}_h B)_i|^2 |e_j|$$
by (5.1.3).

Therefore for each $s,$
$$C^{(r+sm+m-1)} P^{p+q'} \leq |e_{r+sm+m-1}| + h_{r+s}^{r-1} \sum_{j=0}^{q'} |e_{r+\sigma s+m-1}| + h_{r+s}^s \sum_{j=0}^{q'} |e_{r+\sigma s+m-1}|.$$

So there exists $C^*$ such that
$$C^* P^{p+q'} \leq |e_{r+s}| + h_{r+s}^{r-1} \sum_{j=0}^{q'} |e_{r+\sigma s+m-1}| + h_{r+s}^s \sum_{j=0}^{q'} |e_{r+\sigma s+m-1}|.$$
\[\leq \left[ 1 + \frac{\text{hM}}{\text{r+n}} \right] \| e^{(m)} \|_{\infty} \leq (1 + M') \| e^{(m)} \|_{\infty} \text{ for some } M' .\]

Hence \[\| e^{(m)} \|_{\infty} \geq C_p \| e^{(m)} \|_{\infty} \] where \( C_p = C'/(1 + M') \).

So the result follows, taking \( C_L = \min (C'_L, C''_L) \) and \( C_U = \max (C'_U, C''_U) \).

Note that the order of convergence in the \( m \)th stage is exactly the same as the order of optimal consistency, that is \( p+q' \).

We may apply this result to various examples.

**Example 2.2.4.**

\( G = 1 \), so we have zero stability. Consistency is of order one

and \( c(1) = \begin{pmatrix} 1/6 \\ -1/6 \end{pmatrix}, c(2) = \begin{pmatrix} -2/3 \\ 2/3 \end{pmatrix}, c(3) = \begin{pmatrix} 29/72 \\ -77/72 \end{pmatrix} \).

\( A^n_h = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \end{pmatrix} \), so \( A^n_h^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ \vdots & \vdots & \vdots \end{pmatrix} \)

and by direct computation of \( A^n_h^{-1} \theta^n \) we obtain optimal consistency of orders \((2,3)\). Let us take starting values accurate of order three.

\( A^n_0 A^n_2^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \) and \( A^n_1 A^n_2^{-1} = \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix} \).

(i) \( A^n_0 A^n_2^{-1} (1) = A^n_0 A^n_2^{-1} (2) = A^n_1 A^n_2^{-1} (1) = A^n_1 A^n_2^{-1} (2) = 0 \).

(ii) \( A^n_2^{-1} (1) = (1/6,0)^T \), so \( A^n_2^{-1} (1) = 0 \).

(iii) \( A^n_1 A^n_2^{-1} (1) = 1/6 \).

(iv) \( A^n_0 A^n_2^{-1} B^n_\rho = 0_2 \) for \( \rho = 0, 1, 2 \), so \( \lambda^{(0,\rho)} = 0 \) for each \( \rho, \nu \).

\( A^n_1 A^n_2^{-1} B^n_0 = \begin{pmatrix} 0 & 1/6 \\ 0 & 0 \end{pmatrix}, A^n_1 A^n_2^{-1} B^n_1 = \begin{pmatrix} 0 & -1/3 \\ 0 & 0 \end{pmatrix}, A^n_1 A^n_2^{-1} B^n_2 = \begin{pmatrix} 0 & -5/6 \\ 0 & 0 \end{pmatrix} \).
Hence \( \lambda_1^{(1,0)} = 1/6 \), \( \lambda_1^{(1,1)} = -1/3 \), \( \lambda_1^{(1,2)} = -5/6 \),
\( \lambda_2^{(1,0)} = 0 \), \( \lambda_2^{(1,1)} = 0 \), \( \lambda_2^{(1,2)} = 0 \).

(v)
\[
A_2^{-1}B_0 = \begin{pmatrix} 0 & -1/6 \\ 0 & -1/6 \end{pmatrix}, \quad \text{so } \delta_2^{T}A_2^{-1}B_0 = \begin{pmatrix} -1/6 \\ 0 \end{pmatrix}.
\]
\[
A_2^{-1}B_1 = \begin{pmatrix} 1 & -5/6 \\ 0 & 1/3 \end{pmatrix}, \quad \text{so } \delta_2^{T}A_2^{-1}B_1 = \begin{pmatrix} 4/3 \\ 0 \end{pmatrix}.
\]
\[
A_2^{-1}B_2 = \begin{pmatrix} 1 & 0 \\ 0 & 5/6 \end{pmatrix}, \quad \text{so } \delta_2^{T}A_2^{-1}B_2 = \begin{pmatrix} 5/6 \\ 0 \end{pmatrix}.
\]

Hence by theorem 6.4.3 \( C_L h^2 \leq ||e||_\infty \leq C_U h^2 \) and \( C_L h^3 \leq ||e^{(2)}||_\infty \leq C_U h^3 \).

\[ (6.4.1) \]

Example 2.2.5.

\( G = 1 \), so we have zero stability. Consistency is of order zero and
\[
c^{(1)} = \begin{pmatrix} -19/6 \\ 19/6 \end{pmatrix}, \quad c^{(2)} = \begin{pmatrix} 7 \\ -7 \end{pmatrix}, \quad c^{(3)} = \begin{pmatrix} -133/12 \\ 133/12 \end{pmatrix}, \\
c^{(4)} = \begin{pmatrix} 103/8 \\ -325/24 \end{pmatrix}.
\]

\[
A_n^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ \cdots \end{pmatrix}
\]
and by computing \( A_n^{-1} \) we obtain optimal consistency of orders \((1,3)\). Let us take starting values accurate of order three.

\[
A_0A_2^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and } A_1A_2^{-1} = \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix}.
\]

(i) \( A_0A_2^{-1}c^{(1)} = A_0A_2^{-1}c^{(2)} = A_0A_2^{-1}c^{(3)} = A_1A_2^{-1}c^{(1)} = A_1A_2^{-1}c^{(2)} = A_1A_2^{-1}c^{(3)} = 0 \).

(ii) \( A_2^{-1}c^{(1)} = (-19/6,0)^T \), so \( \delta_2^{T}A_2^{-1}c^{(1)} = 0 \).

\( A_2^{-1}c^{(2)} = (7,0)^T \), so \( \delta_2^{T}A_2^{-1}c^{(2)} = 0 \).

(iii) \( \delta_2^{T}A_2^{-1}c^{(1)} = -19/6 \).
Example 2.2.4 is a cyclic method which is consistent of order one, and yet convergent of order two in the first stage and order three in the second stage.

Example 2.2.5 is not consistent at all in the sense of definition 4.1.3, and yet is convergent of order one in the first stage and order three in the second.

In both cases once the orders of optimal consistency are known we may use theorem 6.4.3 to give the correct orders of convergence.

Definition 4.1.3 of consistency is fitting for linear multistep methods, since it is able by theorem 1.2.5 to provide the correct order of convergence. When we consider general classes of $m$-block semi
explicit or implicit methods however, we see that it is an inadequate means of determining orders of convergence in general.

We saw in examples 1.3.1, 2.2.1, 3.2.1, 5.3.4,5,6,7,8,9,10,11 that there are methods which have better convergence than consistency and theorems 4.4.2 and 4.4.4 would lead us to expect. In each of these cases however optimal consistency with theorems 5.1.2, 5.3.1 and 5.3.3 give us the correct orders of convergence.

Moreover in example 2.2.4 we have convergence two orders better than consistency in one of the stages, and example 2.2.5 has order three convergence in one stage and is not even consistent according to definition 4.1.3. Thus for m-block methods consistency is not a necessary condition for convergence, and indeed we shall proceed to show that it is possible to construct inconsistent methods with arbitrarily high orders of convergence in some of their stages.

We also see that optimal consistency is a superior concept to consistency for m-block methods in general, as it gives a two-sided error bound and a necessary and sufficient condition for convergence.

The importance of having a two-sided error bound of the form $C_L h^p \leq \|e\|_\infty \leq C_U h^p$ (6.4.3) rather than just the bound $\|e\|_\infty \leq C_U h^p$ (6.4.4) is that (6.4.3) gives us the exact order of convergence $p$. (6.4.4) on the other hand only tells us that we have order of convergence at least $p$, since there is always the possibility that also $\|e\|_\infty \leq C_U h^{p_1}$ for some $p_1 > p$, since $C_U h^p \leq C_U h^{p_1}$.

Proposition 6.4.4.

For any positive integer $z$ it is possible to construct a zero
stable m-block method which is not consistent, and yet for which
\( C_L h \leq ||e||_{\infty} \leq C_U h \) and \( C_L h^{z'} \leq ||y(m)||_{\infty} \leq C_U h^{z'} \) for some \( C_L, C_U \) and
integer \( z' \geq z \), assuming sufficient continuity in \( y(t) \).

Proof.

Let us take \( m = 2 \) and consider a cyclic method composed of linear multistep methods with stepnumber \( k = 2z - 3 \).

We take
\[
y_i - y_{i-1} = h(\beta_{k}^{(1)} f_i + \beta_{k-1}^{(1)} f_{i-1} + \ldots + \beta_0^{(1)} f_{i-k}) ,
\]
\[
y_{i+1} - y_i = h(\beta_{k}^{(2)} f_{i+1} + \beta_{k-1}^{(2)} f_{i} + \beta_{k-2}^{(2)} f_{i-1} + \ldots + \beta_0^{(2)} f_{i-k+1}) .
\]

\( G = 1 \), and hence we have zero stability.

\( \tau = 1 + \left[ \frac{k-1}{2} \right] = 1 + [z-2] = z-1 \).

\( A_0 = 0 \), \( \Theta = O(1) z-3 \), \( A_{z-2} = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \), \( A_{z-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \).

\( B_0 = \begin{bmatrix} 0 & (1) \\ 0 & 0 \end{bmatrix} \), \( b = \begin{bmatrix} (1) \\ (1) \end{bmatrix} \), \( b = 1(1) z-2 \).

\( B_{z-1} = \begin{bmatrix} (1) & 0 \\ (2) & (2) \end{bmatrix} \).

By considering the local truncation error of the first linear multistep method we see that consistency is of order zero if

\[
k \sum_{j=0}^{j=k-j} \beta_{k}^{(1)} \neq 1 .
\]

We then have \( c(1) = \begin{bmatrix} 1 - (\beta_{o}^{(1)} + \beta_1^{(1)} + \ldots + \beta_k^{(1)}) \\ 1 - (\beta_{o}^{(2)} + \beta_2^{(2)} + \beta_4^{(2)} + \ldots + \beta_{k-1}^{(2)} + \beta_k^{(2)}) \end{bmatrix} \).
and
\[ c(l) = \frac{(-1)^l}{l!} \left( 1(\beta_{k-1}^{(1)} + 2^{l-1}\beta_{k-2}^{(1)} + 3^{l-1}\beta_{k-3}^{(1)} + \ldots + k^{l-1}\beta_0^{(1)}) - 1 \right) \]
\[ \cdot \left( 1((-1)^{l+1}\beta_k^{(2)} + 2^{l+1}\beta_{k-2}^{(2)} + 3^{l+1}\beta_{k-3}^{(2)} + \ldots + (k-1)^{l-1}\beta_o^{(2)}) + (-1)^l \right) \]
for all \( l \geq 2 \).

We shall make use of theorem 6.4.3 to obtain the result.

\( (v) \) \[ A_{z-1}^{-1} B_0 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & \beta_o^{(1)} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \beta_o^{(1)} \\ 0 & \beta_0^{(1)} \end{pmatrix}, \]
so
\[ \delta_2^T A_{z-1}^{-1} B_0 = \beta_o^{(1)} \delta_2^T. \]

For \( \phi = 1(1)z-2 \), \( A_{z-1}^{-1} B_\phi = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_2^{(1)} & \beta_2^{(1)} \\ \beta_2^{(2)} & 0 \end{pmatrix} \]
\[ = \begin{pmatrix} \beta_2^{(1)} + \beta_2^{(2)} & \beta_2^{(1)} \\ \beta_2^{(1)} + \beta_2^{(2)} & \beta_2^{(1)} \end{pmatrix}, \]
so
\[ \delta_2^T A_{z-1}^{-1} B_\phi = \beta_2^{(1)} \delta_2^T \text{ if } \beta_2^{(1)} + \beta_2^{(2)} = 0. \]

Then \( (v) \) holds if
\[ \beta_k^{(1)} + \beta_{k-1}^{(2)} = 0 \]
\[ \beta_k^{(1)} + \beta_{k-2}^{(2)} = 0 \]
\[ \beta_k^{(1)} + \beta_{k-3}^{(2)} = 0 \]
\[ \vdots \]
\[ \beta_1^{(1)} + \beta_0^{(2)} = 0 \]
(6.4.5).
(iv) $A_\Theta = 0_2$ for $\Theta = \Theta(1)z = 3$; so $\lambda^{(\Theta,0)} = 0$ for all $\beta, \nu, \Theta = \Theta(1)z = 3$.

$$A_{z-2}^{-1}A_{z-1}^{-1} = \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix},$$

so $A_{z-2}^{-1}A_{z-1}^{-1}B_0 = \begin{pmatrix} 0 & -\beta_0(1) \\ 0 & 0 \end{pmatrix}$ and $\delta^T_{\gamma} A_{z-2}^{-1}A_{z-1}^{-1}B_0 = -\beta_0(1)\delta_2$.

For $\beta = 1(1)z-2$, $A_{z-2}^{-1}B_\beta = \begin{pmatrix} -\beta_0(1) & -\beta_0(2) & -\beta_0(1) \\ 0 & 0 & 0 \end{pmatrix}$, so $\delta^T_{\gamma} A_{z-2}^{-1}B_\beta = -\beta_2\delta_2$ by (6.4.5) and $\delta^T_{\gamma} A_{z-2}^{-1}B_\beta = 0$.

$$A_{z-2}^{-1}B_{z-1}^\beta = \begin{pmatrix} -\beta_k(1) & -\beta_k(2) & -\beta_k(1) \\ 0 & 0 & 0 \end{pmatrix},$$

so $\delta^T_{\gamma} A_{z-2}^{-1}B_{z-1}^\beta = -\beta_k(2)\delta_2$ by (6.4.5) and $\delta^T_{\gamma} A_{z-2}^{-1}B_{z-1}^\beta = 0$.

Thus (iv) holds.

(i) Since $A_\Theta = 0_2$ for $\Theta = \Theta(1)z = 3$, $A_{z-2}^{-1}c^{(j)} = 0$ for all $j$, $\Theta = \Theta(1)z = 3$.

Now for any vector $v = (v_1, v_2)^T \in \mathbb{R}^2$, $A_{z-2}^{-1}v = 0$ if and only if

\[ v_1 + v_2 = 0, \text{ so } A_{z-2}^{-1}c^{(1)} = 0 \text{ if } \]

\[ (\beta_0(1) + \beta_1(1) + \ldots + \beta_k(1)) + (\beta_0(2) + \beta_2(2) + \beta_4(2) + \ldots + \beta_k(2) + \beta_k(2)) = 2. \]

That is, by (6.4.5), $\beta_0(1) + \beta_1(1) + \beta_2(1) + \ldots + \beta_k(1) + \beta_k(2) = 2$ (6.4.6).

For $l \geq 2$, $A_{z-2}^{-1}c^{(1)} = 0$ if

\[ \beta_{k-1}^{(1)} + 2^{l-1}\beta_{k-2}^{(1)} + 3^{l-1}\beta_{k-3}^{(1)} + \ldots + k^{l-1}\delta_0^{(1)} + (-1)^{l+1}\beta_k^{(2)} + 2^{l-1}\beta_k^{(2)} + \]

\[ 4^{l-1}\beta_{k-5}^{(2)} + 6^{l-1}\beta_{k-7}^{(2)} + \ldots + (k-1)^{l-1}\beta_k^{(2)} = 1 + (-1)^{l+1} \]

\[ \frac{1}{l} . \]
That is, due to (6.4.5),
\[
\beta_k^{(1)} + 3^{1-1} \beta_{k-1}^{(1)} + 5^{1-1} \beta_{k-3}^{(1)} + \ldots + k^{1-1} \beta_{1}^{(1)} + (-1)^{1+1} \beta_k^{(2)} = \frac{1 + (-1)^{1+1}}{1}
\]
(6.4.7).

So (i) holds provided
\[
\begin{bmatrix}
1 & 1 & 1 & 1 & \ldots & 1 \\
-1 & 1 & 3 & 5 & \ldots & k \\
1 & 1 & 3^2 & 5^2 & \ldots & k^2 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
(-1)^{z-1} & 1 & 3^{z-1} & 5^{z-1} & \ldots & k^{z-1}
\end{bmatrix}
\begin{bmatrix}
\beta_k^{(2)} \\
\beta_{k-1}^{(1)} \\
\beta_{k-3}^{(1)} \\
\vdots \\
\beta_1^{(1)} \\
1 + (-1)^{z+1}
\end{bmatrix}
= 
\begin{bmatrix}
2 \\
0 \\
2/3 \\
0 \\
2/5 \\
\frac{1}{z}
\end{bmatrix}
\]
(6.4.8)

The matrix is square, since \( k = 2z - 3 \).

For any \( v \in \mathbb{R}^2 \), \( \delta^{T}_{z-z-1} v = (0,1) \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = v_1 + v_2 = 0 \)

if and only if \( v_1 + v_2 = 0 \). Thus (ii) holds by virtue of (6.4.6) and (6.4.7).

(iii) \( \delta^{T}_{1-z-1} c^{(1)} = (1,0) \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 - (\beta_0^{(1)} + \beta_1^{(1)} + \ldots + \beta_k^{(1)}) \\ 1 - (\beta_0^{(2)} + \beta_2^{(2)} + \beta_4^{(2)} + \ldots + \beta_{k-1}^{(2)} + \beta_k^{(2)}) \end{bmatrix} \)

\[
= 1 - (\beta_0^{(1)} + \beta_1^{(1)} + \ldots + \beta_k^{(1)}) \neq 0 ,
\]
as was stipulated by consistency of order zero.

Therefore by theorem 6.4.3 we have the result provided we can
find a solution to (6.4.8). Now the matrix in that equation is the Vandermonde matrix \( V(-1,1,3,5,\ldots,2z-3) \), which is nonsingular. Hence

\[
\begin{bmatrix}
\beta_2 \\
\beta_1 \\
\beta_{k-1} \\
\beta_{k-3} \\
\vdots \\
\beta_0
\end{bmatrix} = [V(-1,1,3,5,\ldots,k)]^{-1}
\begin{bmatrix}
2 \\
0 \\
2/3 \\
\vdots \\
1 + (-1)^{z+1}
\end{bmatrix}
\]

6.5 Collocation.

We saw in section 3.5 that any collocation method may be represented as an implicit Runge-Kutta method, and that Galerkin and interpolatory quadrature methods may be considered as collocation. Thus collocation, Galerkin and interpolatory quadrature methods all fall within the scope of our m-block methods, and we may analyse their convergence in a manner similar to the approach adopted so far.

For collocation methods the superconvergence result has been quite widely quoted (10,11,12,19,20) and we now give a proof of it using the notation of m-block implicit methods, considering the collocation in Runge-Kutta form.

**Theorem 6.5.1.**

Suppose that for the collocation method (3.5.1) we have the quadrature formulas

\[
\int_t^{t+h} \xi(\eta) d\eta = h \sum_{q=1}^{m} a_q \xi(t + a_q h) + C_q h^{p+1} \quad (6.5.1),
\]
\[
\int_t^{t+a_n h} \xi(\eta) d\eta = h \sum_{q=1}^{m} b_{q} \xi(t + a_q h) + c_{q}^{(1)} h^{p_q} \quad (6.5.2),
\]

for some \( c_{q}^{(1)}, \ldots, c_{q}^{(\mu)} \), where

\[
b_{q} = \int_0^{1} a_{q} \xi(\eta) d\eta \quad \text{and} \quad c_{q} = \int_0^{1} \xi(\eta) d\eta.
\]

Let \( p_{q}^{*} = \min(p_{q}, p) \), \( \nu = 1(1) \mu \), and assume a starting value accurate of order \( p \). Then there exist \( C_{L}, C_{U}, C_{L}^{(\nu)}, C_{U}^{(\nu)} \), \( \nu = 1(1) \mu \), such that

\[
C_{L}^{(\nu)} h^{p_{q}^{*}} \leq ||e(\nu)||_{\infty} \leq C_{U}^{(\nu)} h^{p_{q}^{*}}, \quad \nu = 1(1) \mu,
\]

\[
C_{L} h^{p} \leq ||e(\mu+1)||_{\infty} \leq C_{U} h^{p}.
\]

**Proof.**

Let us first obtain the orders of optimal consistency. We recall that the collocation method \((3.5.1)\) is equivalent to the implicit Runge-Kutta method \((3.3.2)\), for which

\[
A_{h} = \begin{bmatrix}
1 \\
-1 & 1 \\
-1 & 0 & 1 \\
& & & \ddots \\
-1 & 0 & 0 & \cdots & 0 & 1 \\
0 & 0 & 0 & \cdots & 0 & -1 & 1 \\
& & & & & \ddots & \ddots \\
\end{bmatrix}
\]

and

\[
B_{h} = \begin{bmatrix}
0 \\
o_{11} & o_{12} & \cdots & o_{1\mu} \\
0 & b_{11} & b_{12} & \cdots & b_{1\mu} \\
& & & & \ddots \\
0 & b_{\mu 1} & b_{\mu 2} & \cdots & b_{\mu \mu} \\
o_{1} & o_{2} & \cdots & c_{\mu} \\
0 & 0 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1\mu} \\
& & & & & \ddots
\end{bmatrix}.
\]
\[ B_1 = \begin{bmatrix} b_{11} & \cdots & b_{1\mu} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ b_{\mu1} & \cdots & b_{\mu\mu} & 0 \\ c_1 & \cdots & c_{\mu} & 0 \end{bmatrix} \quad \text{and } B_0 = 0_{\mu+1}. \]

Let \( \tilde{B} = \begin{bmatrix} c_1 & c_2 & \cdots & c_{\mu} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_1 & c_2 & \cdots & c_{\mu} & 0 \end{bmatrix} \).

Then by direct computation we obtain

\[ A_n^{-1} B_n = \begin{bmatrix} 0 \\ \vdots \\ \tilde{B} \quad B_1 \\ \vdots \\ \tilde{B} \quad \tilde{B} \quad \tilde{B} \\ \vdots \\ \tilde{B} \\ 0 \end{bmatrix}. \]

\[ \Delta_h y(t) = (y(0), y(a_1 h), y(a_2 h), \ldots, y(a_\mu h), y(h_0), y(h_0 + a_1 h), \ldots)^T, \]

and since any Runge-Kutta method is zero stable, we have zero stability.

\[ m = \mu + 1, \]

and hence for \( s = O(1) n-1 \) and \( \nu = 1(1) \mu+1 \),

\[ [\Delta_h y(t) - h A_n^{-1} B_n \Delta_h y'(t) - A_n^{-1} \xi_{\nu}]_{\nu m+1} \]

\[ y((s+a_\nu)h) - \tilde{y}_h - h \sum_{q=0}^{\mu} \sum_{q=1}^{s-1} c_q y'(\sigma+q)h - h \sum_{q=1}^{\mu} b_q y'(s+a_q)h, \quad \nu = 1(1) \mu, \]

\[ y((s+1)h) - \tilde{y}_h - h \sum_{\sigma=0}^{s} \sum_{q=1}^{\mu} c_q y'(\sigma+q)h, \quad \nu = \mu+1, \]

\[ \int_0^{s+a_\nu} y'(\eta) d\eta + C_0^h P - \int_0^{s-1} \sum_{\sigma=0}^{\mu} \sum_{q=1}^{s-1} c_q y'(\sigma+q)h - h \sum_{q=1}^{\mu} b_q y'(s+a_q)h, \quad \nu = 1(1) \mu, \]

\[ \int_0^{s+1} y'(\eta) d\eta + C_0^h P - h \sum_{\sigma=0}^{s} \sum_{q=1}^{\mu} c_q y'(\sigma+q)h, \quad \nu = \mu+1. \]
At the $m^{th}$ stage then we have a global application of the quadrature (6.5.1), and at the previous stages we have a global application of (6.5.1) followed by a local application of (6.5.2). Hence we have optimal consistency of orders $(p_1^*, p_2^*, \ldots, p_n^*, p)$. Note that by the Lagrange interpolation theorem $p_{ij} \geq \mu + 1$, $i = 1, 2, \ldots, n$, and $p \geq \mu + 1$. Hence $p_{ij} \geq \mu + 1$, $i = 1, 2, \ldots, n$.

Now $e = A_{h1}^{-1} \Theta^* + h A_{h1}^{-1} \Delta_h [f(\Delta_h, t, y^*) - f(\Delta_h, t, \chi(t))]$.

The Runge-Kutta method gives a discrete set of solutions represented here in the vector $y^*$ which coincide exactly with the values taken by the continuous collocation polynomial $u(t)$ at the points $0, a_1 h, a_2 h, \ldots, a_{\mu} h, (a_1 + a_1) h, \ldots$, and so $y^* = \Delta_h u(t)$.

Therefore for each $j$,

$$[f(\Delta_h, t, y^*) - f(\Delta_h, t, \chi(t))]_j = f([\Delta_h, t]_j, [\Delta_h, u(t)]_j) - f([\Delta_h, t]_j, [\Delta_h, \chi(t)]_j)
= f([\Delta_h, t]_j, u([\Delta_h, t]_j)) - f([\Delta_h, t]_j, \chi([\Delta_h, t]_j))
= u'([\Delta_h, t]_j) - \chi'([\Delta_h, t]_j)
= \Delta_h (u'(t) - \chi'(t))_j.
$$

Thus $e = A_{h1}^{-1} \Theta^* + h A_{h1}^{-1} B_{h1} \Delta_h (u'(t) - \chi'(t))$. Due to the form of $A_{h1}^{-1} B_{h1}$ we have, for $s = O(1)n - 1$ and $\nu = 1(1)\mu$,

$$[A_{h1}^{-1} B_{h1} \Delta_h (u'(t) - \chi'(t))]_{\nu + \mu} = \sum_{\sigma=0}^{s-1} \sum_{q=1}^{\mu} c_q (u'([\Delta_h, t]_{\nu + \mu + q}) - \chi'([\Delta_h, t]_{\nu + \mu + q}))
+ \sum_{q=1}^{\mu} b_q (u'([\Delta_h, t]_{\nu + \mu + q}) - \chi'([\Delta_h, t]_{\nu + \mu + q}))
= \sum_{\sigma=0}^{s-1} \int_{t_{\sigma}}^{t_{\sigma} + h} (u'(\eta) - \chi'(\eta)) d\eta + C_q h^{\nu + 1}
+ \int_{t_{\nu}}^{t_{\nu} + h} (u'(\eta) - \chi'(\eta)) d\eta + C_q h^{\nu + 1}
+ C_q h^{\nu + 1}$$
Accordingly for $s = O(1) n^{-1}$ and $\nu = 1(1) \mu$,

$$e_{s \nu + \nu} = [\lambda_{h}^{-1} \Theta^{*}]_{s \nu + \nu} h_{s \nu + \nu} - h C_{h} P_{h} + s C_{q} h^{P+2} + C_{q}^{(\nu)} P_{h} P_{h+1},$$

and so

$$(1 - h)e_{s \nu + \nu} = [\lambda_{h}^{-1} \Theta^{*}]_{s \nu + \nu} - C_{h}^{(\nu)} h^{P+1} + s C_{q} h^{P+2} + C_{q}^{(\nu)} P_{h+1},$$

which gives

$$|(1 - h)e_{s \nu + \nu}| \leq C_{s \nu + \nu}^{(\nu)} P_{h}^{P_{h}+1} + |C_{h}^{(\nu)} h^{P+1} + (n-1)| C_{q} h^{P+2} + |C_{q}^{(\nu)}| P_{h+1}.$$  

Therefore there exists $C_{s \nu + \nu}^{(\nu)}$ such that for $h$ sufficiently small,

$$|e_{s \nu + \nu}| \leq C_{s \nu + \nu}^{(\nu)} P_{h}^{P_{h}+1},$$

and hence we obtain $\|e^{(\nu)}\|_{\infty} \leq C_{s \nu + \nu}^{(\nu)} P_{h}^{P_{h}+1}$.  

Also,

$$|(1 - h)e_{s \nu + \nu}| \geq |C_{s \nu + \nu}^{(\nu)} h^{P_{h}+1} - C_{h}^{(\nu)} h^{P+1} - s C_{q} h^{P+2} - C_{q}^{(\nu)} h^{P+1}|$$

and therefore there exists $C_{s \nu + \nu}^{(\nu)}$ such that for $h$ sufficiently small,

$$|e_{s \nu + \nu}| \geq C_{s \nu + \nu}^{(\nu)} P_{h}^{P_{h}+1},$$

and so $\|e^{(\nu)}\|_{\infty} \geq C_{s \nu + \nu}^{(\nu)} P_{h}^{P_{h}+1}$.  

Thus we have for each $\nu = 1(1) \mu$,

$$C_{s \nu + \nu}^{(\nu)} P_{h}^{P_{h}+1} \leq \|e^{(\nu)}\|_{\infty} \leq C_{s \nu + \nu}^{(\nu)} P_{h}^{P_{h}+1}.$$

As for the case $\nu = \mu + 1$,
Therefore \( e = e + (s+1)C_0h^{p+1} \)

Thus there exists \( C_U \) such that for \( h \) sufficiently small,

\[
|e| \leq C_U h^p,
\]

and hence \( \|e^{(\mu+1)}\|_\infty \leq C_U h^P \).

Also,

\[
|(1 - h)e| \geq C_0 h^{p+1} - |C_0 h^{p+1} - (s+1)C_q h^{p+2}|,
\]

and so there exists \( C_L \) such that for \( h \) sufficiently small,

\[
|e| \geq C_L h^P,
\]

giving \( \|e^{(\mu+1)}\|_\infty \geq C_L h^P \).

So we have \( C_L h^P \leq \|e^{(\mu+1)}\|_\infty \leq C_U h^P \).
CHAPTER SEVEN

CONCLUSIONS

7.1 The Work of Spijker

We shall refer to (22) and (23), two papers of M.N. Spijker.

We consider the problem (1.1.1) with the Lipschitz condition (1.1.2),

to which we apply the m-block method

\[ A_h y^* = hB_h f(A_h t, y^*) + g \]  (7.1.1).

Definition 7.1.1.

For each \( h > 0 \) an m-block method may be represented by the

\[ \Phi_h : \mathbb{R}^{r+nm-1} \rightarrow \mathbb{R}^{r+nm-1}, \]

where

\[ \Phi_h y^* = A_h y^* - hB_h f(A_h t, y^*) - g = 0. \]

Suppose now that \( \tilde{y} \in \mathbb{R}^{r+nm-1} \) is a perturbed solution of the

equation (7.1.1) with a perturbation \( \delta \in \mathbb{R}^{r+nm-1} \) such that \( \Phi_h \tilde{y} = \delta \)

Note that in the special case in which \( \bar{y} = \Delta_h y(t) \), we have \( \delta = \bar{\delta} \), the
local truncation error vector, as in definition 4.1.1.

Definition 7.1.2.

The discretisation \( \Phi_h \) is STABLE with respect to the functional \( \psi \) if there exist positive constants \( K_U \) and \( h_1 \) such that for all vectors \( y^*, \tilde{y}, \delta \in \mathbb{R}^{r+nm-1} \) satisfying \( \Phi_h y^* = 0 \) and \( \Phi_h \tilde{y} = \delta \) with \( h \leq h_1 \), we

have \[ \| \tilde{y} - y^* \|_{\infty} \leq K_U \psi(\delta). \]

Definition 7.1.3.

Let \( \phi \) and \( \psi \) be real functionals on \( \mathbb{R}^{r+nm-1} \) for each \( h > 0 \). Then \( \phi < \psi \) if there exist \( K_U \) and \( h_1 > 0 \) such that \( \phi(\delta) \leq K_U \psi(\delta) \) for all \( \delta \in \mathbb{R}^{r+nm-1} \) and \( h \leq h_1 \).
Definition 7.1.4.

A functional $\psi$ on $\mathbb{R}^{n+m-1}$ is a STABILITY FUNCTIONAL for $\Phi_h$ if $\Phi_h$ is stable with respect to $\psi$. A stability functional $\phi$ for $\Phi_h$ is a MINIMAL STABILITY FUNCTIONAL for $\Phi_h$ if all stability functionals $\psi$ for $\Phi_h$ satisfy $\phi < \psi$.

We may show that assuming zero stability the stability functional $\psi(\bar{e}) = \|A_h^{-1}\bar{e}\|_\infty$ allows us to obtain a two-sided error bound of the form

$$K_u \psi(\bar{e}) \leq \|\bar{y} - y^*\|_\infty \leq K_u \psi(\bar{e}) \quad (7.1.2)$$

by making use of the same argument as in the proof of theorem 5.1.2, only replacing $A_h y(t)$ and $\theta^*$ by $\bar{y}$ and $\bar{e}$ respectively. The following theorem also shows that $\|A_h^{-1}\bar{e}\|_\infty$ is a minimal stability functional for $\Phi_h$, provided $\Phi_h$ is zero stable.

Theorem 7.1.5.

$$\psi(\bar{e}) = \|A_h^{-1}\bar{e}\|_\infty$$

is a minimal stability functional for $\Phi_h$, assuming $\Phi_h$ to be zero stable.

Proof.

(i) We first show that $\psi$ is a stability functional for $\Phi_h$.

We have $\Phi_h y^* = 0 (7.1.3)$ and $\Phi_h \bar{y} = \bar{e} (7.1.4)$. Using the same analysis as in the proof of theorem 5.1.2 we have for each $i$

$$|\bar{y}_i - y^*_i| \leq \left| (A_h^{-1}\bar{e})_i \right| + h [A_h^{-1}B_h (f(A_h t, \bar{y}) - f(A_h t, y^*))]_i$$

$$\leq \|A_h^{-1}\bar{e}\|_\infty + h M_\infty \|B_h\|_1 \sum_{j=0}^{1+m-1} |\bar{y}_j - y^*_j|$$

where $i-r+1 \equiv j \pmod{m}$,
1 \leq \nu \leq m and for an m-block semi explicit method the summation goes only as far as i. Let \( M = M_1 B_h \). Then by lemma 4.4.5,

\[
|\bar{y}_i - y^*_i| < \frac{1}{1 - (m-1)hM} \frac{hM}{1 - (m-1)hM} \sum_{j=0}^{i-1} |\bar{y}_j - y^*_j| ,
\]

since \( \frac{1}{1 - qhM} \leq \frac{1}{1 - (m-1)hM} \) for \( q = O(1) m^{-1} \), if \( h \) is sufficiently small. Hence

\[
|\bar{y}_i - y^*_i| < \frac{1}{1 - mhM} \frac{hM}{1 - mhM} \sum_{j=0}^{i-1} |\bar{y}_j - y^*_j| ,
\]

and \( \frac{1}{1 - mhM} > 0 \) for \( h < 1/mM \). Therefore by lemma 4.4.3

\[
||\bar{y} - y^*||_\infty \leq \frac{1}{1 - mhM} ||\bar{y}_0 - y_0^*|| + ||A^{-1}_h(\bar{\bar{\sigma}})\exp[M(r+nm-1)h]/1 - mhM|| .
\]

Now by (2.1.3) there exists a positive \( Q \) independent of \( h \) such that

\[
(r+nm-1)h = QT , \quad \text{so since } ||\bar{y}_0 - y_0^*|| < ||A^{-1}_h(\bar{\bar{\sigma}})||, \quad ||\bar{y} - y^*||_\infty \leq K_0 \psi(\bar{\bar{\sigma}}) \text{ where } K_0 = \frac{1 + hM}{1 - mhM} \exp\left(\frac{MT}{1 - mhM}\right) .
\]

(ii) We now show that \( \psi(\bar{\bar{\sigma}}) \) is a minimal stability functional for \( \bar{\Phi}_h \).

By (7.1.3) and (7.1.4), \( \bar{\Phi}_h \bar{y} - \bar{\Phi}_h y^* = \bar{\bar{\sigma}} \), and so

\[
A_h(\bar{y} - y^*) - hA_h[f(\Delta_h t, \bar{y}) - f(\Delta_h t, y^*)] = \bar{\bar{\sigma}} ,
\]

and premultiplying through by \( A^{-1}_h \) we have

\[
||A^{-1}_h(\bar{\bar{\sigma}})|| \leq ||\bar{y} - y^*||_\infty + h||A^{-1}_h[f(\Delta_h t, \bar{y}) - f(\Delta_h t, y^*)]||_\infty .
\]

Now since we have zero stability \(, \quad \max_{0 \leq i, j \leq r+nm-1} |(A^{-1}_h)_{ij}| = M_0 \), and so

\[
\text{for any } w \in \mathbb{R}^{r+nm-1} , \quad h||A^{-1}_h w||_\infty \leq hM_0 (r+nm-1)||w||_\infty \leq M_0 QT ||w||_\infty .
\]
Therefore 
\[ \| A_h^{-1} \bar{\theta} \|_\infty \leq \| \bar{y} - y^* \|_\infty M_0 Q T \| B_h \| \| f(\Delta_h t, \bar{y}) - f(\Delta_h t, y^*) \|_\infty \]
\[ \leq \| \bar{y} - y^* \|_\infty M_0 Q T \| B_h \| \| \bar{y} - y^* \|_\infty \]
\[ = (1 + M_0 Q T \| B_h \|_\infty L) \| \bar{y} - y^* \|_\infty \]
\[(7.1.7).\]

Suppose now that \( \beta_1(\bar{\theta}) \) is any stability functional for \( \Phi_h \). Then
\[ \| \bar{y} - y^* \|_\infty \leq K_1 \beta_1(\bar{\theta}) \]
for some \( K_1 \) by definition \( 7.1.4 \), and so
\[ \psi(\bar{\theta}) = \| A_h^{-1} \bar{\theta} \|_\infty \leq (1 + M_0 Q T \| B_h \|_\infty L) K_1 \beta_1(\bar{\theta}) \].

Thus \( \psi < \beta_1 \) and so \( \psi \) is a minimal stability functional for \( \Phi_h \).

Spijker's approach in (23) is to introduce the functional
\[ \psi_s(\bar{\theta}) = \sum_{i=0}^{r-1} |\bar{\theta}_i| + \max_{r \leq i \leq r + n - 1} \| h \sum_{j=r}^i \bar{\theta}_j \| \]
and then put \( A_h = P_h Q_h \) where \( P_h \) has a special form. He shows that \( \psi_s \) is a minimal stability functional for any \( \Phi_h \) having the form \( \Phi_h y^* \equiv P_h y^* - h B_h f(\Delta_h t, y^*) - g \), and that for an "optimally stable" \( \Phi_h \), \( \psi_s \) is a minimal stability functional subject to a certain condition on \( Q_h \). It turns out however that the condition of "optimal stability" is too restrictive for a general discretisation \( \Phi_h \).

This is the motivation behind the work of Albrecht, which we turn to next.

7.2 The Work of Albrecht.

We refer in this section to (2). Albrecht alludes to the work of Spijker and develops his ideas to find a minimal stability functional which is optimal, in a sense to be defined.
Definition 7.2.1.

Let $\psi$ be a stability functional for $\Phi_h$ with $\psi$ independent of the right-hand side of (7.1.1). Then $\psi$ is OPTIMAL if the following conditions hold for all $\bar{\theta} \in \mathbb{R}^{r+nm-1}$ and all $h > 0$ sufficiently small:

(i) $\|\bar{y} - y^*\|_\infty = \psi(\bar{\theta})$ for all $\Phi_h$ in which the right-hand side of (7.1.1) is independent of $y^*$.

(ii) For any $\epsilon > 0$ there exists $\chi > 0$ such that

$$(1 - \epsilon)\psi(\bar{\theta}) < \|\bar{y} - y^*\|_\infty < (1 + \epsilon)\psi(\bar{\theta}) \text{ for all } f \text{ with } L < \chi.$$ 

Theorem 7.2.2.

$\psi(\bar{\theta}) = \|A_h^{-1}\bar{\theta}\|_\infty$ is optimal.

Proof.

(i) For all $\Phi_h$ in which the right-hand side of (7.1.1) is independent of $y^*$, (7.1.6) gives $A_h(\bar{y} - y^*) = \bar{\theta}$ and so $\|\bar{y} - y^*\|_\infty = \|A_h^{-1}\bar{\theta}\|_\infty = \psi(\bar{\theta})$.

(ii) (7.1.5) gives

$$\|\bar{y} - y^*\|_\infty \leq \frac{1 + hM}{1 - hmM} \exp[\frac{M\eta T}{1 - hmM}] \psi(\bar{\theta})$$

$$\leq (1 + hM)\exp(M\eta T)\psi(\bar{\theta}) \text{ if } 0 < h < 1/mM,$$

and so the right-hand side of (ii) follows since $M = \omega\|\mathcal{B}_h\|_{1,L}$.

(7.1.7) gives $\|\bar{y} - y^*\|_\infty \geq \frac{1}{1 + M\omega L\|\mathcal{B}_h\|_\infty} \psi(\bar{\theta})$

and the left-hand side of (ii) follows immediately.

The fact that $\|A_h^{-1}\bar{\theta}\|_\infty$ is an optimal stability functional is the motivation behind the term "optimal consistency" in section 5.1. By using optimal consistency in our formulation of $m$-block methods we are able to analyse correctly the convergence of more methods than if we use Albrecht's $\hat{\lambda}$-norm in his $\hat{\lambda}$ formulation of discretisation algorithms.
Albrecht's theorem 4.3 in (2) is covered by our theorems 5.1.2 and 5.3.3, which deal with optimal consistency of orders \( p(1,1,\ldots,1) \).

In section 6 we have gone further than this, obtaining error bounds, some of which are two-sided, for methods having optimal consistency of orders \( (p_1, p_2, \ldots, p_m) \) with not all the \( p_j \) being equal. Albrecht's theorem 4.3 could not obtain the bounds (6.3.3) to (6.3.11) inclusive, nor (6.4.1), (6.4.2), for instance.

Moreover, our theorem 6.2.10 gives upper bounds on the various stages of the global error for an easily recognisable class of methods, which goes beyond the work of Albrecht.

To take an illustrative case, let us consider example 2.2.5.

This is an \( \hat{A} \)-method with
\[
\begin{pmatrix}
y^*_{2s+4} \\
y^*_{2s+5}
\end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^*_{2s+2} \\
y^*_{2s+3}
\end{pmatrix} + h \begin{pmatrix} f^*_{2s+4} + f^*_{2s+3} + f^*_{2s+2} + \frac{1}{6} f^*_{2s+1} + f^*_{2s} \\
\frac{5}{6} f^*_{2s+5} + \frac{8}{6} f^*_{2s+3} - \frac{1}{6} f^*_{2s+1}
\end{pmatrix}
\]
and so the truncation error \( \delta \) is given by
\[
\delta = \begin{pmatrix} -19/6 \\ 0 \end{pmatrix} y'(t_{2s+5}) + o(h).
\]

The left eigenvector of \( \hat{A} \) corresponding to the eigenvalue 1 is \((0,1)^T\), and so since \((0,1) \begin{pmatrix} -19/6 \\ 0 \end{pmatrix} = 0\) the condition of Albrecht's theorem 4.3 is satisfied, giving \( C_L h \leq \|\varepsilon\|_\alpha \leq C_U h \).

Using our theorem 6.4.3 however we are able to obtain both this bound and \( C_L h^3 \leq \|\varepsilon^{(2)}\|_\alpha \leq C_U h^3 \).
7.3 Comparison with Some Results of Stetter.

In this section we refer to Stetter's book (24). It appears that Stetter alludes just once to the possibility of two-sided error bounds, mentioning on p. 84 that these can be obtained using Spijker's norm. As Albrecht has pointed out however, there are ways of examining convergence which are superior to this.

Generally, Stetter is concerned only with one-sided bounds on the error, defining convergence of order \( p \) at the \( i^{th} \) stage to hold when \( e_i = O(h^p) \) as \( h \to 0 \).

Stetter's theorems 4.3.6 and 4.3.7 on pp. 224 and 226 deal with cyclic methods, giving a criterion which if fulfilled gives convergence one order better than consistency. He adopts the approach of Donelson and Hansen (9). For an \( m \)-cyclic linear \( k \)-step method the \( m \)-cyclic linear \( mk \)-step auxiliary method is constructed and it is proved by Stetter that the order of convergence of the original method is the same as that of its auxiliary. This approach appears to be less concise and natural than ours, which constructs no auxiliary method and which for cyclic methods gives the exact order of convergence to be that of optimal consistency.

Consider for instance example 2.2.3. This is a 3-cyclic 4-step method which we saw in section 5.3 to be consistent of order three and yet convergent of order four. The auxiliary method is a 3-cyclic 12-step method and to derive it and then determine its local truncation error to find its order of consistency is a cumbersome process.

In his section 5.2 Stetter studies predictor-corrector methods
and in theorem 5.2.3 on pp.288,289 obtains a result which gives the orders of convergence at the various stages in such schemes. The same is achieved also by our theorem 6.2.10, as is seen in examples 6.2.11 and 6.2.13. Our theorem 6.2.10 however is more general, as it applies to all methods which have $A_n \in Q_{m,1}$ and $B_n \in Q_{m,\alpha}^{(\beta)}$, for instance examples 6.1.2, 6.1.3, 6.3.2, 6.3.3 and 6.3.4. Moreover our theorem 6.3.1 gives in certain cases two-sided error bounds such as (6.3.3) to (6.3.7) inclusive, (6.3.9) and (6.3.11).

7.4 The Work of Skeel.

We refer in this section to Skeel's paper (21). He formulates his methods in the form

$$u_n = Su_{n-1} + h(t_{n-1}, u_{n-1}; h) (7.4.1), 1 \leq n \leq N, u_n \text{ and } u_{n-1} \in \mathbb{R}^k, S \in \mathbb{R}^{k \times k} \text{ for some integer } k.$$  

Each of the methods which he quotes can be represented as an $m$-block method. Now an $m$-block method has a unique representation, whereas this is not the case for Skeel's formulation above. For instance, Skeel's example 2 on pp.667,668 is a predictor-corrector scheme in PEC mode, and he gives two alternative formulations of it.

Or we could look at his example 3 on p.668, which in our notation has the array

$$
\begin{pmatrix}
0 & 0 & -1 & 1 & 0 & 0 & -1/4 & 0 & 3/4 & 0 & 0 & 0 & 1/2 \\
0 & 0 & -1 & 0 & 1 & 0 & 1 & 0 & -2 & 2 & 0 & 0 & 1 \\
0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1/6 & 2/3 & 1/6 & 0 & 1
\end{pmatrix}
(7.4.2).

Skeel's representation however lacks this compactness:
\[ \begin{pmatrix} u_n \\ hu'_n \end{pmatrix} \frac{1}{2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_{n-1} \\ hu'_{n-3/2} \end{pmatrix} + \begin{pmatrix} f_{n-1} \\ f_{n-1/2} \end{pmatrix} \begin{pmatrix} 1/6 & 2/3 & 1/6 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} f_n \end{pmatrix} \]

where
\[ f_{n-1} = f(t_{n-1}, u_{n-1}), \]
\[ f_{n-1/2} = f(t_{n-1/2}, u_{n-1} + \frac{3}{4}h f_{n-1} - \frac{1}{4}hu'_{n-3/2}), \]
\[ f_n = f(t_n, u_{n-1} + 2hf_{n-1/2} - 2hf_{n-1} + hu'_{n-3/2}) \quad (7.4.3). \]

Moreover this formulation is not unique, as we could equally well write
\[ \begin{pmatrix} u_{n-3/2} \\ u_{n-1} \\ u_{n-1/2} \\ u_n \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_{n-5/2} \\ u_{n-2} \\ u_{n-3/2} \\ u_{n-1} \\ u_n \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1/4 & 0 & 3/4 & 0 & 0 \end{pmatrix} \begin{pmatrix} f_{n-3/2} \\ f_n \\ f_n \end{pmatrix} \quad (7.4.4). \]

Our representation \[ A_n y^* = hB_n f^* + g \] is essentially linear and gives a natural and unique formulation of any method which can be represented in this way. \( (7.4.3) \) is a more complicated representation and the
question may be asked, by which criteria in Skeel's formulation do we decide which of the representations is the best.

**Definition 7.4.1.**

If for each \( n \) the vector \( d_n \) denotes the local discretisation error then a method is **QUASI-CONSISTENT of order** \( p \) if

\[
E = \lim_{n \to \infty} \mathbb{R}^n \quad \text{and} \quad \max_n |d_n| = O(h^p) \quad \text{and} \quad \max_n |E(d_0 + d_1 + \ldots + d_n)| = O(h^p).
\]

\( E \) is a \( k \times k \) matrix and we define the \((N+1)k \times (N+1)k\) matrix \([E]\) by

\[
[E] = \begin{pmatrix}
I_k & \cdots & 0 \\
-E & I_k & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & -E & I_k
\end{pmatrix}
\]

which gives

\[
[E]^{-1} = \begin{pmatrix}
I_k & \cdots & 0 \\
E & I_k & \cdots & 0 \\
E & \cdots & I_k \\
E \cdots E & \cdots & E & I_k
\end{pmatrix}
\]

Skeel takes various stability concepts, including a "moderately strict root condition" (MSRC), and obtains his theorem 3.6, which says that given a method satisfying the MSRC there exist positive constants \( c \) and \( C \) such that for sufficiently small \( h \) and any \( R \in \mathbb{R}^{(N+1)k} \),

\[
c \left\| [E]^{-1} R \right\|_{\infty} \leq \| \hat{U} - U \|_{\infty} \leq C \left\| [E]^{-1} R \right\|_{\infty}, \quad \text{where} \quad \| \hat{U} - U \|_{\infty} = \max_{n \geq N} \left| u_n - \hat{u}_n \right|,
\]

\( u_n \) being the solution of (7.4.1) and \( \hat{u}_n \) being a perturbed solution with perturbation \( r_n \) and \( R = (r_o, r_1, \ldots, r_n)^T \).
Thus \( ||E^{-1}R||_\infty \) is a minimal stability functional, but as Albrecht (2) points out, it is not optimal.

Skeel uses the above result to show that a method satisfying the MSRC is convergent of order \( p \) if and only if it is quasi-consistent of order \( p \), and he gives four examples of applications of this. First he considers his example 2 on pp.667,668 and shows that although the method is consistent of order three, it is quasi-consistent and hence convergent of order four. Considered as a 2-block method it has the array

\[
\begin{pmatrix}
0 & 0 & 0 & 1 & 0 & -1/2 & 0 & 3/2 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & -1 & 0 & 1 & -1/12 & 0 & 2/3 & 0 & 5/12 & 0 & 1
\end{pmatrix}
\]

and although we also get third order consistency, we see that this is just a special case of example 5.3.6 and so we get the "improved" convergence of order four.

Skeel's example 4 on p.668 is just our example 5.3.7, and we have third order consistency but obtain fourth order convergence as he does. On p.680 he also quotes our example 1.3.1, for which he obtains fourth order convergence as we do.

When it comes to Skeel's example 3 on p.668, he obtains truncation error

\[
d_n = h^4 \left[ \begin{array}{c}
0 \\
-\frac{5}{96} \frac{\partial^2 F}{\partial y} (t_n, y_n, y'_n, y''_n)
\end{array} \right] + O(h^5)
\]

if the method is given the formulation (7.4.3), and he shows that quasi-consistency is of order four.

When we write this example as a 3-block method (7.4.2) we obtain our local truncation error by considering that of each of the hybrid methods.
We are unable by this means to show fourth order convergence even using optimal consistency. This is because the \( m \)-block representation is entirely linear, whereas the method under consideration is non-linear, being essentially of Runge-Kutta form. By looking at the method in Skeel's way we are forced to perform complicated computations in order to derive the local truncation error, as is the case for example with an explicit Runge-Kutta method written in the form

\[
y_i - y_{i-1} = h \phi(t_{i-1}, y_{i-1})
\]

where \( \phi \) is a non-linear function. Thus the simplicity of looking at the local truncation error through just the individual hybrid methods is lost.

Moreover if we choose the formulation (7.4.4) of the method we obtain quasi-consistency only of order one, and so the effectiveness of Skeel's convergence theorem in giving the correct order of convergence for a particular method depends on the way chosen to represent it.

One of the advantages of our approach is the comparatively small amount of work needed to calculate the local truncation errors, as we saw with Heun's third order Runge-Kutta formula, example 6.1.3. For this method we are able to obtain the bounds

\[
\|e^{(1)}\|_\infty \leq c_U^{(1)} h^2, \quad \|e^{(2)}\|_\infty \leq c_U^{(2)} h^3, \quad \|e^{(3)}\|_\infty \leq c_U^{(3)} h^3,
\]

\[
c_L h^2 \leq \|e\|_\infty \leq c_U h^2, \quad c_L^{(2,3)} h^3 \leq \|e^{(2,3)}\|_\infty \leq c_U^{(2,3)} h^3
\]

by using theorems 6.2.10 and 6.3.1, and the considering the local truncation errors of just the individual hybrid methods.
7.5 The Work of Cooper.

In this section we consider Cooper's approach as presented in (5). He looks at linear methods, using the $(A,B)$ formulation of Butcher in (4). The class of methods of this kind is identical to our class of $m$-block methods.

To illustrate Cooper's concept of consistency we consider an $(A,B)$ method in our $m$-block formulation with $A_h^* = hB_h f^* + g$. We define the vector $z \in \mathbb{R}^{r+nm-1}$ by

$$z = (I_{r+nm-1} - A_h) \Delta_h y(t) + hB_h f(\Delta_h t, z) + g,$$

where we may stipulate without loss of generality that $A_h$ is lower triangular and has every element on the diagonal equal to one.

For each $s = 0(1)n-1$ the components $r + sm + \nu - 1, \nu = 1(1)m$ of $\Delta_h y(t) - z$ give the errors at each stage when all $y^*_{r+sm+\nu-1}$ are assumed to have exact solution values for $\sigma = 0(1)s-1, \nu = 1(1)m$. It is this vector $\Delta_h y(t) - z$ from which Cooper obtains the order of consistency rather than our approach of examining the linear expression $\Theta^*$ in definition 4.1.1. Although $\Delta_h y(t) - z$ provides us with more information than $\Theta^*$, the calculation of the former vector presents more problems than that of the latter, and the algebraic conditions for consistency presented in Cooper's lemma 4 and theorem 5 are complicated.

In his final section Cooper looks at nine examples to which his theory is applied.

His first, second and sixth examples are our examples 5.3.8, 5.3.9 and 5.3.10 respectively and he obtains the orders of convergence using his non-linear consistency concept as we do using optimal consistency.
Cooper's seventh example is our example 6.3.4 for which we obtained the bounds
\[ \|e^{(1)}\|_{\infty} \leq C_U^{(1)} h^3, \|e^{(2)}\|_{\infty} \leq C_U^{(2)} h^4, \quad \nu = 2, 3, 4, \]
\[ c_L h^3 \leq \|e\|_{\infty} \leq C_U c_L h^4, \quad c_U^{(2, 3, 4)} h^4 \leq \|e^{(2, 3, 4)}\|_{\infty} \leq C_U^{(2, 3, 4)} h^4, \]
giving the same orders of convergence as he does.

Note that in these examples we get two-sided error bounds, whereas all of Cooper's bounds are only one-sided.

Cooper's third example is a 3-block explicit method with array
\[
\begin{pmatrix}
0 & 0 & -1/4 & 0 & 0 & -3/4 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & -2 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0
\end{pmatrix}
\]
It is optimally consistent of orders \((3, 3, 4)\) and so by \((6.1.1)\)
\[ c_L h^3 \leq \|e\|_{\infty} \leq C_U h_{3/4}^3. \]
We are unable however to obtain fourth order convergence which Cooper's analysis shows in the third stage.

His fourth example is a 4-block explicit method with array
\[
\begin{pmatrix}
0 & 0 & 0 & -11/27 & 0 & 0 & 0 & -16/27 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 20/27 & 0 & 0 & 0 & -47/27 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & -23/10 & 0 & 0 & 0 & 13/10 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1/49 & 0 & 0 & 0 & -48/49 & 0 & 0 & 0 & 1
\end{pmatrix}
\]
It is optimally consistent of orders \((4, 4, 4, 5)\) and so by \((6.1.1)\)
\[ c_L h^4 \leq \|e\|_{\infty} \leq C_U h_4^4, \]
but we are unable to show the sixth order convergence which Cooper
demonstrates in the fourth stage.

His eighth example is a 3-block explicit method with array
\[ \begin{pmatrix} 0 & 0 & -1 & 1 & 0 & 0 & -1/4 & 0 & 3/4 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & -1 & 0 & 1 & 0 & 1 & 0 & -2 & 2 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1/6 & 2/3 & 1/6 & 0 & 1 \end{pmatrix} \]

It is optimally consistent of orders \((3,3,4)\) and so as before
\[ C_L h^3 \leq \|e\|_\infty \leq C_U h^3, \]
but we are unable to show the fourth order convergence which Cooper demonstrates in the third stage.

His ninth example is a 4-block implicit method with array
\[ \begin{pmatrix} -5/81 & 0 & 0 & -76/81 & 1 & 0 & 0 & 0 & 0 & 0 & -14/27 & 0 & 1/9 & 0 & 0 & 1/2 \\ -1/9 & 0 & 0 & -8/9 & 0 & 1 & 0 & 0 & 0 & 0 & 2/3 & 0 & 0 & 0 & 1/2 \\ 4/9 & 0 & 0 & -13/9 & 0 & 0 & 1 & 0 & 0 & 0 & 5/3 & 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1/6 & 0 & 2/3 & 1/6 & 0 & 1 \end{pmatrix} \]

It is optimally consistent of orders \((3,3,3,4)\) and so again we have
\[ C_L h^3 \leq \|e\|_\infty \leq C_U h^3, \]
but we are unable to show fourth order convergence at the fourth stage as Cooper does. Cooper however points out that computations indicate fourth order convergence in the first stage, and like us he is unable to demonstrate this.

As we might have expected then Cooper's analysis gives correct orders of convergence for some methods which are beyond the scope of our approach as it now stands, because unlike \(\Theta^*\) the vector \(\Delta_h y(t) - z\) contains non-linear information. The advantage of our approach however is the comparative simplicity of the determination of the orders of optimal consistency, and in any case there are methods, such as the last one considered, for which neither Cooper nor we are able to obtain theoretically all the correct orders of convergence.
BIBLIOGRAPHY.


2. P. Albrecht: "Explicit, Optimal Stability Functionals and Their Application to Cyclic Discretisation Methods."
   Comp. 19, 233-249 (1978).

   JACM 12, 124-135 (1965).


7. G. Dahlquist: "A Special Stability Problem for Linear Multistep Methods."


11. B. L. Hulme: "One-Step Piecewise Polynomial Galerkin Methods for Initial Value Methods."

12. B. L. Hulme: "Discrete Galerkin and Related One-Step Methods for Ordinary Differential Equations."

13 J. J. Kohfeld, G. T. Thompson: "Multistep Methods with Modified Predictors and Correctors."
    JACM 14, 155-166 (1967).


15. J. D. Lambert: "Computational Methods in Ordinary Differential Equations."


Submitted for publication.


Here we verify computationally the orders of convergence claimed in the various methods we have quoted.

We consider the differential equation (1.1.1) with \( f(t,y(t)) = -y(t) \) and \( y_0 = 1 \), which has the solution \( y(t) = \exp(-t) \).

Using each particular method we integrate this using two steplengths \( h_1 \) and \( h_2 \) in turn and compare the computed solutions at corresponding points to determine the order of convergence. In the tables below the errors in the computed solutions are given at certain points for the pre-assigned \( h_1 \) and \( h_2 \), and the orders of convergence at each of the stages are deduced.

**Example 1.3.1.**

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t = 4 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h_1 = 1/10 )</td>
<td>-2.6x10^-7</td>
<td>-1.6x10^-7</td>
</tr>
<tr>
<td>( h_2 = 1/30 )</td>
<td>-3.1x10^-9</td>
<td>-2.0x10^-9</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Example 2.2.1.**

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t = 1.9 )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( h_1 = 1/10 )</td>
<td>-1.7x10^-3</td>
<td>-9.0x10^-4</td>
</tr>
<tr>
<td>( h_2 = 1/30 )</td>
<td>-1.9x10^-4</td>
<td>-1.0x10^-4</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
### Example 2.2.2.

<table>
<thead>
<tr>
<th></th>
<th>1&lt;sup&gt;st&lt;/sup&gt; stage</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; stage</th>
<th>3&lt;sup&gt;rd&lt;/sup&gt; stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = 3.8$</td>
<td>$t = 3.9$</td>
<td>$t = 4$</td>
</tr>
<tr>
<td>$h_1 = 0.1$</td>
<td>$-6.4 \times 10^{-7}$</td>
<td>$-4.9 \times 10^{-7}$</td>
<td>$-5.7 \times 10^{-7}$</td>
</tr>
<tr>
<td>$h_2 = 0.01$</td>
<td>$5.4 \times 10^{-11}$</td>
<td>$5.9 \times 10^{-11}$</td>
<td>$4.4 \times 10^{-11}$</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

### Example 2.2.3.

<table>
<thead>
<tr>
<th></th>
<th>1&lt;sup&gt;st&lt;/sup&gt; stage</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; stage</th>
<th>3&lt;sup&gt;rd&lt;/sup&gt; stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = 2.9$</td>
<td>$t = 2.95$</td>
<td>$t = 3$</td>
</tr>
<tr>
<td>$h_1 = 0.05$</td>
<td>$-2.6 \times 10^{-6}$</td>
<td>$-3.0 \times 10^{-6}$</td>
<td>$-1.5 \times 10^{-6}$</td>
</tr>
<tr>
<td>$h_2 = 0.005$</td>
<td>$-2.0 \times 10^{-10}$</td>
<td>$-1.8 \times 10^{-10}$</td>
<td>$-1.2 \times 10^{-10}$</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

### Example 2.2.4.

<table>
<thead>
<tr>
<th></th>
<th>1&lt;sup&gt;st&lt;/sup&gt; stage</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = 0.95$</td>
<td>$t = 1$</td>
</tr>
<tr>
<td>$h_1 = 1/20$</td>
<td>$1.6 \times 10^{-4}$</td>
<td>$-1.4 \times 10^{-5}$</td>
</tr>
<tr>
<td>$h_2 = 1/60$</td>
<td>$1.8 \times 10^{-5}$</td>
<td>$-5.6 \times 10^{-7}$</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

### Example 2.2.5.

<table>
<thead>
<tr>
<th></th>
<th>1&lt;sup&gt;st&lt;/sup&gt; stage</th>
<th>2&lt;sup&gt;nd&lt;/sup&gt; stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = 0.9$</td>
<td>$t = 0.95$</td>
</tr>
<tr>
<td>$h_1 = 1/20$</td>
<td>$6.2 \times 10^{-2}$</td>
<td>$-1.3 \times 10^{-5}$</td>
</tr>
<tr>
<td>$h_2 = 1/60$</td>
<td>$2.1 \times 10^{-2}$</td>
<td>$-5.4 \times 10^{-7}$</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>
Example 3.2.1.

<table>
<thead>
<tr>
<th></th>
<th>1st stage</th>
<th>2nd stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 = \frac{1}{10} )</td>
<td>( 9.1 \times 10^{-8} )</td>
<td>( 1.6 \times 10^{-7} )</td>
</tr>
<tr>
<td>( h_2 = \frac{1}{30} )</td>
<td>( 1.1 \times 10^{-9} )</td>
<td>( 2.0 \times 10^{-9} )</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Example 3.3.4.

<table>
<thead>
<tr>
<th></th>
<th>1st stage</th>
<th>2nd stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 = \frac{1}{10} )</td>
<td>( 5.0 \times 10^{-4} )</td>
<td>( 4.9 \times 10^{-4} )</td>
</tr>
<tr>
<td>( h_2 = \frac{1}{30} )</td>
<td>( 3.2 \times 10^{-5} )</td>
<td>( 5.1 \times 10^{-5} )</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Example 5.3.5.

<table>
<thead>
<tr>
<th></th>
<th>1st stage</th>
<th>2nd stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 = 0.1 )</td>
<td>( 2.3 \times 10^{-5} )</td>
<td>( 6.1 \times 10^{-5} )</td>
</tr>
<tr>
<td>( h_2 = 0.01 )</td>
<td>( 2.3 \times 10^{-7} )</td>
<td>( 5.7 \times 10^{-7} )</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Example 5.3.6.

Let us take the third order predictor

\[ y_1 - y_{i-1} = \frac{h}{12} (23f_{i-1} - 16f_{i-2} + 5f_{i-3}) \]

with the fourth order corrector

\[ y_1 - y_{i-1} = \frac{h}{12} (9f_i + 19f_{i-1} - 5f_{i-2} + f_{i-3}) \]

The numerical results which we obtain are as follows.
<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t = 3</td>
<td>t = 3</td>
</tr>
<tr>
<td>h₁ = 0.1</td>
<td>-4.0x10^{-7}</td>
<td>1.9x10^{-6}</td>
</tr>
<tr>
<td>h₂ = 0.01</td>
<td>-1.7x10^{-11}</td>
<td>1.7x10^{-10}</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Example 5.3.7.**

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t = 4.9</td>
<td>t = 5</td>
</tr>
<tr>
<td>h₁ = 1/10</td>
<td>-1.2x10^{-7}</td>
<td>7.8x10^{-8}</td>
</tr>
<tr>
<td>h₂ = 1/30</td>
<td>-1.4x10^{-9}</td>
<td>9.4x10^{-10}</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Example 5.3.8.**

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t = 3.96</td>
<td>t = 4</td>
</tr>
<tr>
<td>h₁ = 0.1</td>
<td>-5.5x10^{-6}</td>
<td>-8.0x10^{-6}</td>
</tr>
<tr>
<td>h₂ = 0.01</td>
<td>-5.0x10^{-9}</td>
<td>-7.2x10^{-9}</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

**Example 5.3.9.**

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t = 4</td>
<td>t = 4.03</td>
</tr>
<tr>
<td>h₁ = 0.1</td>
<td>-6.5x10^{-6}</td>
<td>-5.9x10^{-6}</td>
</tr>
<tr>
<td>h₂ = 0.01</td>
<td>-5.6x10^{-9}</td>
<td>-5.0x10^{-9}</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
### Example 3.3.10.

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
<th>3rd. stage</th>
<th>4th. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = 3.06$</td>
<td>$t = 3.05$</td>
<td>$t = 3.1$</td>
<td>$t = 3.1$</td>
</tr>
<tr>
<td>$h_1 = 0.2$</td>
<td>$3.2 \times 10^{-7}$</td>
<td>$1.9 \times 10^{-8}$</td>
<td>$-2.5 \times 10^{-8}$</td>
<td>$2.0 \times 10^{-8}$</td>
</tr>
<tr>
<td>$h_2 = 0.02$</td>
<td>$2.4 \times 10^{-11}$</td>
<td>$2.1 \times 10^{-10}$</td>
<td>$-3.1 \times 10^{-10}$</td>
<td>$2.2 \times 10^{-10}$</td>
</tr>
<tr>
<td>Orders of</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>convergence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Example 5.3.11.

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = 3.9$</td>
<td>$t = 4$</td>
</tr>
<tr>
<td>$h_1 = 1/10$</td>
<td>$1.8 \times 10^{-8}$</td>
<td>$1.1 \times 10^{-8}$</td>
</tr>
<tr>
<td>$h_2 = 1/30$</td>
<td>$7.3 \times 10^{-11}$</td>
<td>$4.7 \times 10^{-11}$</td>
</tr>
<tr>
<td>Orders of</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>convergence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Example 6.1.2.

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
<th>3rd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = 3$</td>
<td>$t = 3$</td>
<td>$t = 3$</td>
</tr>
<tr>
<td>$h_1 = 0.1$</td>
<td>$-2.7 \times 10^{-4}$</td>
<td>$4.5 \times 10^{-6}$</td>
<td>$-7.0 \times 10^{-6}$</td>
</tr>
<tr>
<td>$h_2 = 0.01$</td>
<td>$-2.5 \times 10^{-6}$</td>
<td>$3.7 \times 10^{-9}$</td>
<td>$-6.8 \times 10^{-9}$</td>
</tr>
<tr>
<td>Orders of</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>convergence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Example 6.1.3.

<table>
<thead>
<tr>
<th></th>
<th>1st. stage</th>
<th>2nd. stage</th>
<th>3rd. stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = 2.93$</td>
<td>$t = 2.96$</td>
<td>$t = 3$</td>
</tr>
<tr>
<td>$h_1 = 0.1$</td>
<td>$-3.8 \times 10^{-5}$</td>
<td>$-4.1 \times 10^{-5}$</td>
<td>$-6.7 \times 10^{-5}$</td>
</tr>
<tr>
<td>$h_2 = 0.01$</td>
<td>$-3.0 \times 10^{-7}$</td>
<td>$-3.8 \times 10^{-9}$</td>
<td>$-6.3 \times 10^{-9}$</td>
</tr>
<tr>
<td>Orders of</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>convergence</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example 6.2.12

<table>
<thead>
<tr>
<th>Orders of convergence</th>
<th>$1^{st.}$ stage</th>
<th>$2^{nd.}$ stage</th>
<th>$3^{rd.}$ stage</th>
<th>$4^{th.}$ stage</th>
<th>$5^{th.}$ stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1 = 0.5$</td>
<td>$-1.2 \times 10^{-3}$</td>
<td>$2.3 \times 10^{-4}$</td>
<td>$-1.6 \times 10^{-5}$</td>
<td>$2.7 \times 10^{-5}$</td>
<td>$2.0 \times 10^{-5}$</td>
</tr>
<tr>
<td>$h_2 = 0.05$</td>
<td>$-0.87 \times 10^{-5}$</td>
<td>$1.5 \times 10^{-7}$</td>
<td>$-2.5 \times 10^{-9}$</td>
<td>$1.7 \times 10^{-10}$</td>
<td>$1.3 \times 10^{-10}$</td>
</tr>
</tbody>
</table>

Example 6.2.13

Let us take the first order predictor

$y_i - y_{i-1} = h f_{i-1}$ with the second order corrector

$y_i - y_{i-1} = \frac{h}{2} (f_i + f_{i-1})$ followed by the third order

$y_i - y_{i-1} = \frac{h}{12} (5f_i + 8f_{i-1} - f_{i-2})$, the fourth order

$y_i - y_{i-1} = \frac{h}{24} (9f_i + 19f_{i-1} - 5f_{i-2} + f_{i-3})$ and the fifth order

$y_i - y_{i-1} = \frac{h}{720} (251f_i + 646f_{i-1} - 264f_{i-2} + 106f_{i-3} - 19f_{i-4})$.

The numerical results we obtain are as follows.

<table>
<thead>
<tr>
<th>Orders of convergence</th>
<th>$1^{st.}$ stage</th>
<th>$2^{nd.}$ stage</th>
<th>$3^{rd.}$ stage</th>
<th>$4^{th.}$ stage</th>
<th>$5^{th.}$ stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1 = 0.5$</td>
<td>$-1.2 \times 10^{-3}$</td>
<td>$2.4 \times 10^{-4}$</td>
<td>$1.8 \times 10^{-5}$</td>
<td>$2.6 \times 10^{-5}$</td>
<td>$4.0 \times 10^{-5}$</td>
</tr>
<tr>
<td>$h_2 = 0.05$</td>
<td>$-0.87 \times 10^{-5}$</td>
<td>$1.5 \times 10^{-7}$</td>
<td>$0.9 \times 10^{-9}$</td>
<td>$2.3 \times 10^{-10}$</td>
<td>$2.7 \times 10^{-10}$</td>
</tr>
</tbody>
</table>

Example 6.3.2

<table>
<thead>
<tr>
<th>Orders of convergence</th>
<th>$1^{st.}$ stage</th>
<th>$2^{nd.}$ stage</th>
<th>$3^{rd.}$ stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1 = 1/10$</td>
<td>$-4.0 \times 10^{-8}$</td>
<td>$1.3 \times 10^{-8}$</td>
<td>$9.1 \times 10^{-9}$</td>
</tr>
<tr>
<td>$h_2 = 1/30$</td>
<td>$-5.3 \times 10^{-10}$</td>
<td>$5.6 \times 10^{-11}$</td>
<td>$4.2 \times 10^{-11}$</td>
</tr>
</tbody>
</table>

Orders of convergence

<table>
<thead>
<tr>
<th>$1^{st.}$ stage</th>
<th>$2^{nd.}$ stage</th>
<th>$3^{rd.}$ stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 3.95$</td>
<td>$t = 4$</td>
<td>$t = 4$</td>
</tr>
<tr>
<td>$h_1 = 1/10$</td>
<td>$-4.0 \times 10^{-8}$</td>
<td>$1.3 \times 10^{-8}$</td>
</tr>
<tr>
<td>$h_2 = 1/30$</td>
<td>$-5.3 \times 10^{-10}$</td>
<td>$5.6 \times 10^{-11}$</td>
</tr>
</tbody>
</table>
### Example 6.3.3.

<table>
<thead>
<tr>
<th>1st. stage ( t = 2.46 )</th>
<th>2nd. stage ( t = 2.45 )</th>
<th>3rd. stage ( t = 2.5 )</th>
<th>4th. stage ( t = 2.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 = 0.1 )</td>
<td>1.5x10^{-5}</td>
<td>( h_1 = 1/10 )</td>
<td>6.1x10^{-7}</td>
</tr>
<tr>
<td>( h_2 = 0.01 )</td>
<td>1.4x10^{-8}</td>
<td>( h_2 = 1/30 )</td>
<td>7.1x10^{-9}</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Example 6.3.4.

<table>
<thead>
<tr>
<th>1st. stage ( t = 2.483 )</th>
<th>2nd. stage ( t = 2.5 )</th>
<th>3rd. stage ( t = 2.55 )</th>
<th>4th. stage ( t = 2.55 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 = 1/10 )</td>
<td>1.3x10^{-6}</td>
<td>( h_1 = 1/10 )</td>
<td>-2.7x10^{-8}</td>
</tr>
<tr>
<td>( h_2 = 1/70 )</td>
<td>3.7x10^{-9}</td>
<td>( h_2 = 1/30 )</td>
<td>-3.0x10^{-10}</td>
</tr>
<tr>
<td>Orders of convergence</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>