J. Sand: On Convective Linear Multistep and One-Leg Methods

August 1993
DAIMI PB 164

Jørgen Sand

Computer Science Department
Aarhus University
Aarhus University

ISSN 0105-8817
On Contractive Linear Multistep and One-Leg Methods

by

Jørgen Sand

Abstract

Contractive variable-formula methods for the integration of $y' = \lambda(t)y$ ($\lambda(t)$ restricted to some region of interest) are formed by using polyhedral norms for showing the contractivity of several one-leg or linear multistep formulas. When possible the results are given for variable step-size.

1. Introduction

This paper contains the main results of the author's thesis ([1], [2] and [3]) concerning stability and boundedness of one-leg and linear multistep methods applied to initial-value problems of the form

$$y' = f(t,y), \quad y(t_0) = y_0 \in \mathbb{R}^s, \quad t \geq t_0 \in \mathbb{R}. \quad (1.1)$$

The numerical solution of (1.1) is a finite sequence $(y_0, y_1, \ldots, y_N)$ containing approximative values of the exact solution $\{y(t)|t \geq t_0\}$ (assumed to exist) on a grid $\{t_i|t_0 \leq t_i < t_{i+1} \leq \ldots \leq t_N\}$. We will consider two classes of methods for the production of such a sequence, viz. the linear multistep methods (LMM) and the one-leg methods (OLM). Let $k$ be a positive integer. Then a $k$-step LMM consists of finding $y_{n+k}$, $n = 1-k, 2-k, \ldots$ so that these satisfy the difference equations given by $k_n$-step linear multistep formulas (LMF):

$$\sum_{j=0}^{k} \alpha_{j,n} \cdot y_{n+j} = \bar{r}_{n+k} \sum_{j=0}^{k} \beta_{j,n} \cdot f(t_{n+j}, y_{n+j}), \quad (1.2a)$$
where \( k \geq k_n \geq 1 \),

\[
\alpha_{k,n} \neq 0, \quad |\alpha_{k-k_n,n}| + |\beta_{k-k_n,n}| = 0 \quad \text{and} \quad (1.2b)
\]

\[
\alpha_{j,n} = \beta_{j,n} = 0 \quad \text{for} \quad 0 \leq j < k-k_n.
\]

Likewise, a \( k \)-step OLM consists of solving for \( n=1-k,2-k,... \) the equations given by \( k \)-step one-leg formulas (OLF):

\[
\sum_{j=0}^{k} \alpha_{j,n} \cdot y_{n+j} = \bar{h}_{n+k} \cdot f(\sum_{j=0}^{k} \beta_{j,n} \cdot t_{n+j}, \sum_{j=0}^{k} \beta_{j,n} \cdot y_{n+j}), \quad (1.3a)
\]

where

\[
\alpha_{k,n} \neq 0, \quad |\alpha_{k-k_n,n}| + |\beta_{k-k_n,n}| = 0 \quad \text{and} \quad (1.3b)
\]

\[
\alpha_{j,n} = \beta_{j,n} = 0 \quad \text{for} \quad 0 \leq j < k-k_n.
\]

The LMF's and OLF's are chosen so that

\[
k_n \leq n+k \quad (1.4a)
\]

and the consistency (and normalization) requirements

\[
\bar{h}_{n+k} = \sum_{j=0}^{k} \alpha_{j,n} \cdot t_{n+j}, \quad \sum_{j=0}^{k} \alpha_{j,n} = 0, \quad \sum_{j=0}^{k} \beta_{j,n} = 1 \quad (1.4b)
\]

are met for \( n=1-k,2-k,... \). Most often \( \bar{h}_{n+k} \neq 0 \) and may be replaced by the step-size \( h_{n+k} = t_{n+k} - t_{n+k-1} \) used in the \((n+k)\)'th step.

On certain mild assumptions the question of stability and boundedness of these methods can be dealt with by considering their application to pseudo-linear differential systems, i.e. problems (1.1), where \( f(t,y) = \Lambda(t,y)y \) for some matrix-valued function \( \Lambda \). If \( \Lambda \) for all relevant values of \( t \) and \( y \) can be chosen sufficiently diagonally dominant, we may restrict ourselves to considering scalar equations only, i.e. problems of the form
\[ y' = \lambda(t, y)y, \quad y(t_0) = y_0 \in \mathbb{C}, \quad t \geq t_0 \in \mathbb{R}. \quad (1.5) \]

Let \( C_n \) denote the companion matrix of a \( k \)-step LMM or OLM applied to (1.5)

\[
C_n = \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
\gamma_{0,n} & \gamma_{1,n} & \ddots & \gamma_{k-1,n}
\end{bmatrix}, \quad n = 1-k, 2-k, \ldots,
\]

where

\[
\gamma_{j,n} = -\left(\frac{\alpha_j,n - q_j,n \cdot \beta_j,n}{\alpha_k,n - q_k,n \cdot \beta_k,n}\right), \quad j = 0(1)k-1,
\]

and

\[
q_{j,n} = \begin{cases}
\gamma_{n+k} \cdot \lambda(t_{n+j}, y_{n+j}) & \text{if an LMM is considered}
\end{cases}
\]

\[
\begin{cases}
\gamma_{n+k} \cdot \lambda(\sum_{j=0}^{k} \beta_j,n t_{n+j}, \sum_{j=0}^{k} \beta_j,n y_{n+j}) & \text{if an OLM is considered}
\end{cases}
\]

We are then interested in showing the finiteness of

\[
M = \sup_{1-k \leq N_1 \leq N_2 < \infty} \| \Pi \sum_{i=1}^{N_2} C_i \|\quad (1.7)
\]

where \( \| \cdot \| \) denotes some operator norm. In the case of OLM's the contractivity condition

\[
\| C_n \| \leq 1 \quad \text{for all } n \geq 1-k \quad (1.8)
\]

becomes necessary (and sufficient) for the finiteness of \( M \) in (1.7), unless we place some restrictions on the order in which the different formulas and step-sizes are used.

In this paper we consider contractivity of OLM's and LMM's with respect to operator norms corresponding to vector norms with "corners" on their unit sphere, viz.

(i) The max-norm: \( \| X \| = \max_{1 \leq i \leq k} | X^{(i)} | \).
(ii) Polyhedral norms: $||X|| = \alpha$, where $X = \alpha X^*$, 

$X^* \in \partial W$. 

(W a convex balanced polyhedral
neighbourhood of the origin)

(iii) Scaled max-norms: $||X|| = \max_{1 \leq i \leq k} |(T^{-1}X)^{(i)}|$ 

The reason for not choosing e.g. an inner-product norm is
the following fact that if all the formulas satisfy the
algebraic condition for strong zero-stability

$$\rho_n(r) = 0 \Rightarrow (r = 1 \text{ or } |r| < 1), \quad \rho_n(r) \leq \sum_{j=0}^{k} \alpha_j r^j,$$

then (cf. theorem 3.1 in [4]) condition (1.8) cannot be
satisfied in a matrix norm corresponding to a differentiable
vector norm if $\lambda \equiv 0$, unless $\rho_n$ is constant.

In section 2, we find that there exist LMM's with formulas
of arbitrarily high order for which (1.8) is satisfied in
the max-norm, as long as $q_{j, n} \in \{z | z + a | \leq 0 \}$ for some con-
stant $a > 0$ and the step-ratios $h_{n+k+1}/h_{n+k}$ do not exceed
some upper limit greater than one. For fixed step-size a
comparison of some of these formulas with the popular Adams-
Moulton formulas shows that the "strongly 0-contractive"
multistep formulas are reasonably stable and accurate, too.

For the integration of stiff differential equations, the
backward differentiation formulas (the BDF's) are widely
used and in section 3, we construct a polyhedral norm with
respect to which (1.8) is satisfied for the 3-step LMM con-
sisting of the BDF's of order one to three, as long as
$q_{j, n} \leq 0$ and certain combinations of order and step-ratio
are excluded.

Scaled norms are discussed in section 4, and by imbedding
a one-parameter coordinate-transformation of $c^k$ in the max-
norm, we develop a matrix-norm (called the $(b, k)$-norm),
which is a generalization of a norm used by Brayton and
Conley in [5]. By appropriate adjustment of $k$, which for a
given set of formulas is only limited from below, and the
parameter b, we are (in section 5) e.g. able to find for fixed step-size a 5-step OLM with formulas of order one to five for which (1.8) is satisfied in a (b, 5)-norm, as long as \( |q_{j,n}^+ + a| \leq a \) for some \( a > 0 \). Furthermore, we find a 4-step OLM with fixed-step formulas of order one to four for which (1.8) is satisfied in a (b, 4)-norm as long as

\[
\begin{align*}
|\arg (-q_{j,n}^+)| & \leq 90^\circ \text{ (if the first-order formula is used)} \\
|\arg (-q_{j,n}^+)| & \leq 89^\circ \text{ (- second-order - - -)} \\
|\arg (-q_{j,n}^+)| & \leq 45^\circ \text{ (- third-order - - -)} \\
|q_{j,n} + a| & \leq a, a \approx 0.2 \text{ (- fourth-order - - -)}
\end{align*}
\]

2. Strongly 0-contractive LMM's

In [1], we used the phrase "the LMM is \( l_\infty \)-decreasing in \( Q \subset \mathbb{C} \)" if \( \|C_n\|_\infty \leq 1 \) was valid for all \( q_{j,n} \in Q \). The new concept \( l_\infty \)-decrease was introduced to emphasize the variability of \( \lambda \) in our test-equation \( y' = \lambda(t)y \). Since we will not recommend the concept for general use, we shall only describe the LMM's by the adjective "contractive", although the ability of \( \lambda \) to vary relatively freely prevents most LMM's from satisfying the contractivity condition (1.8).

Consider, for example, the principal root \( r \) of a zero-stable linear multistep formula. If \( q_{j,n} = -a \cdot \delta_{1,j} \), then

\[
r = 1 - a \beta_{1,n} / \rho_n' \eta_n(1) + O(a^2), \quad \rho_n(\xi) = \sum_{j=0}^k \alpha_{1,n} \xi^j
\]

and thus

\[
\alpha_{k,n} \cdot \beta_{1,n} \geq 0 \quad \text{for all } l,n
\]

is a necessary requirement for the formula to be "strongly 0-contractive" in some norm (allowing \( q_{j,n} \) to vary freely in some disk, tangential to the origin, lying in the left half-plane). We note that (2.1) rules out most of the well-known classes of linear multistep formulas. Among these the Adams formulas of order greater than two.
Condition (2.1) is clearly not sufficient for this kind of strong 0-contractivity, but if we choose to operate with the max-norm, necessary and sufficient conditions are easily derived.

Theorem 2.1

Let $C_n$ in (1.8) denote the companion matrix of a $k$-step LMM with $\alpha_{k,n} > 0$. Then (1.8) is valid in the max-norm for all $(q_0, n, q_1, n, \ldots, q_k, n) \in [-a_n, 0]^{k+1}$ ($a_n$ is a some positive real number), if and only if $\beta_{k,n} \geq 0$ and

$$
\begin{align*}
\alpha_j, n \leq 0, \quad \beta_j, n \geq 0 & \quad \text{for } j = 0(1)k-1. \\
\alpha_j, n = 0 \Rightarrow \beta_j, n = 0 &
\end{align*}
$$

If $\beta_j, n = 0$ for all $j \in [0, k-1]$ then $\| C_n \|_\infty \leq 1$ for all $(q_0, n, q_1, n, \ldots, q_k, n) \in \{ z | z-(\alpha_{k,n}/\beta_{k,n}) \leq \alpha_{k,n}/\beta_{k,n} \}^{k+1}$. Otherwise $\| C_n \|_\infty \leq 1$ is valid for all $(q_0, n, q_1, n, \ldots, q_k, n) \in \{ z | z+r_n \leq r_n \}^{k+1}$, where $r_n = \min_{j=0(1)k-1} \alpha_j, n / \beta_j, n$.

Proof:

Dropping the subscript $n$ we now find necessary conditions for

$$
\begin{align*}
\sum_{j=0}^{k-1} |a_j - q_j \beta_j| & \leq |a_k - q_k \beta_k| \\
\end{align*}
$$

to hold for all $q_j \in [-a, 0]$, a being some positive constant. First let $q_j = 0, j = 0(1)k$. We then find that

$$
\begin{align*}
\sum_{j=0}^{k-1} |a_j| & = |a_k| \quad \text{i.e. } \alpha_j \leq 0.
\end{align*}
$$

Letting all the $q_j$'s but one be zero we find that

$$
\begin{align*}
|a_j - q_j \beta_j| & \leq |a_j| \quad \text{for } j = 0(1)k-1.
\end{align*}
$$

(2.3)
It follows that either $\beta_j$ is zero or

$$|q_j - (a_j/\beta_j)| \leq |a_j/\beta_j|,$$

i.e. $\alpha_j \beta_j < 0$.

Letting all the $q_j$'s apart from $q_k$ be zero we find that

$$|a_k - q_k \beta_k | \geq |\alpha_k|$$

and $\beta_k \geq 0$ follows. We have now found all the conditions in the theorem to be necessary. That they are sufficient for $||C_n||_{\infty} \leq 1$ to hold, when $(q_0, q_1, \ldots, q_k)$ belongs to the regions described, is easily seen since (2.3) and (2.4) will be valid and thus

$$\Sigma_{j=0}^{k-1} |a_j - q_j \beta_j| \leq \Sigma_{j=0}^{k-1} |a_j| = |\alpha_k| \leq |a_k - q_k \beta_k|.$$ 

\[\Box.\]

In order to check (2.2) for some linear multistep formulas we would like to have explicit expressions for the coefficients $\alpha_{j,n}$ and $\beta_{j,n}$, even in the case of variable step-size. By modifying the derivation of fixed-step formulas of order $2k$ ($k$=the step-number) made in [6], we obtain a lemma that can be applied for that purpose.

**Lemma 2.2**

For given sets $J_{\alpha,n} \subseteq J_{\beta,n}$ satisfying $J_{\beta,n} \subseteq J_{\alpha,n} \subseteq \{0,1,\ldots,k\}$ and $\max J_{\alpha,n} \leq k$, there exists a linear multistep formula

$$\sum_{j=0}^{k} \alpha_{j,n} Y_{n+j} = h_{n+k} \sum_{j=0}^{k} \beta_{j,n} Y_{n+j}$$

of order $r_{n+s_{n+2}}$,
\[ r_n \triangleq \text{number of elements in } J_{\alpha,n} \text{ (assumed } \geq 2), \]
\[ s_n \triangleq \text{number of elements in } J_{\beta,n} \text{ (assumed } \geq 1), \]
\[ \{ j \mid a_{j,n} = 0 \} = J_{\alpha,n} \text{ and } \{ j \mid b_{j,n} = 0 \} = J_{\beta,n}. \]

Apart from a normalization factor, this formula is uniquely determined by \( J_{\alpha,n} \) and \( J_{\beta,n} \) and given by

\[
\beta_{j,n} = (\prod_{i \in J_{\alpha,n} \setminus \{j\}} \tau_{j,i})(\prod_{i \in J_{\beta,n} \setminus \{j\}} \tau_{j,i}), \quad \forall j \in J_{\beta,n},
\]

\[
\alpha_{j,n} = \begin{cases} 
\sum_{i \in J_{\alpha,n} \setminus \{j\}} \tau_{j,i} + \sum_{i \in J_{\beta,n} \setminus \{j\}} \tau_{j,i} & \text{if } j \in J_{\beta,n}, \\
- \left( \prod_{i \in J_{\alpha,n} \setminus \{j\}} \tau_{j,i} \right) \cdot \left( \prod_{i \in J_{\beta,n} \setminus \{j\}} \tau_{j,i} \right) & \text{if } j \in J_{\alpha,n} \setminus J_{\beta,n},
\end{cases}
\]

where \( \tau_{j,i} = h_{n+k}/(t_{n+j} - t_{n+i}) \) for \( i \neq j \).

**Proof:**

Given \( r_n + s_n \) numbers \( y_{n+i}', f_{n+j}', i \in J_{\alpha,n}', j \in J_{\beta,n}' \), it is well-known from interpolation theory that there exists a unique polynomial \( p \) of degree at most \( r_n + s_n - 1 \) satisfying

\[
p(t_{n+i}) = y_{n+i}', \quad p'(t_{n+j}) = f_{n+j}', \quad i \in J_{\alpha,n}', j \in J_{\beta,n}'. \quad (2.5)
\]

If \( p \) is of degree \( r_n + s_n - 1 \) exactly, (2.5) cannot be satisfied by a polynomial of degree less than \( r_n + s_n - 1 \). In other words: all polynomials \( q \) of degree \( r_n + s_n - 2 \) or less satisfying (2.5) will also satisfy

\[
\sum_{j \in J_{\alpha,n} \setminus J_{\beta,n}} \\{ \phi_j(t_{n+j}) \psi(t_{n+j}) \}_1^{-1} \cdot q(t_{n+j}) + \sum_{j \in J_{\beta,n}} \{ \phi_j(t_{n+j}) \psi'(t_{n+j}) \}_1^{-1} \cdot q(t_{n+j}) = 0
\]

\[
[q'(t_{n+j}) - \phi_j(t_{n+j})/\psi_j(t_{n+j}) + \psi_j(t_{n+j})/\psi(t_{n+j}) \psi(t_{n+j})] = 0
\]

\[
(2.6)
\]
(2.6) is obtained by setting the leading coefficient of \( p \) equal to zero and using \( q(t_{n+1}) = y_{n+i}, q'(t_{n+j}) = f_{n+j} \).

Here

\[
\phi_j(x) = \prod_{i \in J_{\alpha,n} \setminus \{j\}} (x - t_{n+i}), \quad j \in J_{\alpha,n},
\]

\[
\psi(x) = \prod_{i \in J_{\beta,n}} (x - t_{n+i}), \quad \psi_j(x) = \psi(x)/(x - t_{n+j}), \quad j \in J_{\beta,n}.
\]

The formula is now obtained by setting \( q(t_{n+1}) = y_{n+i}, q'(t_{n+j}) = f_{n+j} \) in (2.6) and then multiplying on both sides by \( h_{n+k} \). \( \square \)

Combining theorem 2.1 and lemma 2.2 we arrive at the following theoretical result.

**Theorem 2.3**

For any \( p \geq 1 \) there exists a linear multistep formula (1.2) of order \( p \) with the following properties:

1) The number of non-zero coefficients is \( p+2 \).

2) For all step-ratios \( h_{n+k+1}/h_{n+k} \) belonging to a certain interval \( (0, 1+\varepsilon) \), \( \varepsilon \) a positive constant, (1.8) is satisfied in the max-norm as long as

\[
(q_0, n, q_1, n, \ldots, q_k, n) \in \{ z \mid z + r \leq r \}^{k+1}
\]

for some \( r > 0 \).

**Proof:**

For \( p \) even we may consider the formulas derived in lemma 2.2 with \( J_{\alpha,n} = J_{\beta,n} \). It follows that

\[
\beta_j,n = \left( \prod_{i \in J_{\beta,n} \setminus \{j\}}^\tau j, i \right)^2 > 0, \quad \forall j \in J_{\beta,n}
\]

and
\[ a_{j,n} = 2 \cdot \beta_{j,n} \cdot \left( \sum_{i \in J_{\beta,n} \setminus \{j\}} \tau_{j,i} \right), \quad \forall j \in J_{\alpha,n}. \quad (2.7) \]

Since \( \text{sgn} (\tau_{j,i}) = \text{sgn} (j-i) \), \( a_{k,n} > 0 \) and we have only to ensure that \( a_{j,n} < 0 \) for all \( j \in J_{\alpha,n} \cap \{0, 1, \ldots, k-1\} \). For fixed step-size this can be achieved by e.g. choosing \( J_{\alpha,n} = J_{\beta,n} = \{k, k-1, k-1-2^1, \ldots, 0\} \) with \( s \geq 0 \). For \( s=0 \) we obtain the trapezoidal rule which satisfies the given conditions. For \( s>0 \) we get formulas with \( k = 2^{s+1} \) and \( p = 2(s+1) \), i.e. \( p=k+1 \) only if \( p=6 \). Let us now show these formulas to satisfy condition (2.2) for fixed step-size. From (2.7) we have

\[
a_{k-1,n} = 2 \cdot \beta_{k-1,n} \cdot \sum_{i \in J_{\beta,n} \setminus \{k-1\}} (k-1-i)^{-1} = 2 \cdot \beta_{k-1,n} \cdot (-1 + \sum_{i=1}^{s} 2^{-i}) < 0
\]

and for \( l \in \{1, 2, \ldots, s\} \),

\[
a_{k-1-2^l,n} / (2 \beta_{k-1-2^l,n}) = \sum_{i \in J_{\beta,n} \setminus \{k-1-2^l\}} (k-1-2^l-i)^{-1} = -(1+2^{-l})^{-1} - 2^{-l} - \sum_{i=1}^{l-1} (2^{l-2^i})^{-1} + \sum_{i=l+1}^{s} (2^{l-2^i})^{-1}. \quad (2.8)
\]

Since

\[
\sum_{i=1}^{s} (2^{i-2^l})^{-1} \leq \sum_{i=1}^{s} (2^{i-1})^{-1} = 2^{-1} + 2^{-2} + \cdots + 2^{-s} < 2^{-1} + (2^{-1} - 1)^{-1} = \frac{1}{2} - 2^{-1} + 1 - 2
\]

this term is "outweighed" by the \((l-1)\)'th term in

\[
\sum_{i=1}^{s} (2^{i-2^l})^{-1} \text{ if } l > 1. \text{ For } l=s=1 \text{ the term is zero whereas } i=1.
\]

for \( s > 1 \) and \( l = 1 \)

\[
\sum_{i=1}^{s} (2^{i-2^l})^{-1} = \frac{1}{2} + \sum_{i=3}^{s} (2^{i-2})^{-1} \leq \frac{1}{2} + \frac{1}{3} \cdot \sum_{i=3}^{s} (2^{i-2})^{-1} < \frac{1}{2} + \frac{1}{3}
\]

and the term is "outweighed" by the first two terms in (2.8).
We have now shown that our formulas with even order
and fixed step-size satisfy the conditions in theorem 2.1.
In order to show 2) note first that, since the coefficients
of the formula depend continuously on the step-ratios, the
conditions \( a_{j,n} < 0 \) for \( j \in J_{\alpha,n} \) will remain satisfied for all
step-ratios belonging to a certain interval around 1. If
any of these step-ratios decreases it is easily seen that
the quantities

\[
\frac{h_{n+j}}{|t_{n+j} - t_{n+i}|}, \quad j = 1(1)k-1,
\]

increase (or remain fixed) if \( i > j \) and decrease (or remain
fixed) if \( i < j \). It follows from this observation that

\[
\frac{h_{n+j} \cdot a_{j,n}}{h_{n+k}} \quad \text{(and thus } a_{j,n} \text{), } 1 \leq j \leq k-1, \text{ remain negative}
\]

and the proof is immediate since \( a_{k,n} > 0, a_{0,n} < 0 \) and \( \beta_{j,n} > 0 \)
(for \( j \in J_{\beta,n} \)) hold for all step-sequences.

To obtain formulas of odd order we may simply remove 0
from the set \( J_{\beta,n} \) above. Let the new non-vanishing coeffi-
cients be denoted by \( \tilde{a}_{j,n} \) and \( \tilde{\beta}_{j,n} \). Then

\[
\tilde{\beta}_{j,n} = \beta_{j,n} / \tau_{j,0} > 0, \quad \tilde{\alpha}_{0,n} = -\tilde{\beta}_{0,n} < 0,
\]

\[
\tilde{\alpha}_{k,n} = \tilde{\beta}_{k,n} \cdot (2 \cdot \sum_{i \in J_{\beta,n} \setminus \{k\}} \tau_{k, i + \tau_{k,0}}) > 0,
\]

\[
\tilde{\alpha}_{j,n} = \tilde{\beta}_{j,n} \cdot (2 \cdot \sum_{i \in J_{\beta,n} \setminus \{j\}} \tau_{j, i - \tau_{j,0}}) < 0 \quad \text{for } j \in J_{\alpha,n} \setminus \{0,k\}
\]

and all the arguments above can be repeated.

\[\square\]

Although the theorem above is mainly of theoretical in-
terest the proof gives us an idea of how formulas fulfilling
condition 1) and 2) in theorem 2.3 may be derived. Using the
computer we have examined all formulas with step-number not
exceeding \( 2 \cdot [(p-1)/2] + 1 \) (\( p = \text{the order } \geq 12 \)) and only \( p+2 \)
non-zero coefficients. For even order we chose \( J_{\beta} = J_{\alpha} \) and
for odd order we selected \( J_{\beta} = J_{\alpha} \setminus \{0\} \).
The formulas are listed for fixed step-size only, but corresponding variable step-size formulas are easily obtained from lemma 2.2 knowing $J_\alpha$ and $J_\beta$. For $p \leq 8$ the formulas with the largest radius $r$ (among those examined!) are listed. ($C_{p+1}$ = the error-constant corresponding to $\sum_{j \in J_\beta}^n y_{j+1} = 1$)

$p = 1$:

$$y_{n+1} = y_n + h f_{n+1}$$

$r = \infty$, $C = -\frac{1}{2}$.

$p = 2$:

$$y_{n+1} = y_n + \frac{h}{2} (f_{n+1} + f_n)$$

$r = 2$, $C = -\frac{1}{12}$.

$p = 3$:

$$y_{n+2} = \frac{1}{5} (4 y_{n+1} + y_n) + \frac{2h}{5} (f_{n+2} + 2f_{n+1})$$

$$y_{n+3} = \frac{1}{28} (27 y_{n+2} + y_n) + \frac{3h}{14} (4f_{n+3} + 3f_{n+2} + f_n)$$

$r = 1$, $C = -\frac{1}{36}$.

$p = 4$:

$$y_{n+3} = \frac{1}{32} (27 y_{n+2} + 5y_n) + \frac{3h}{32} (4f_{n+3} + 9f_{n+2} + 3f_n)$$

$r = 1$, $C = -\frac{3}{140}$.

$p = 5$:

$$y_{n+4} = \frac{1}{35} (24 y_{n+3} + 8y_{n+1} + 3y_n) + \frac{12h}{35} (f_{n+4} + 3f_{n+3} + 3f_{n+1})$$

$r = \frac{2}{3}$, $C = -\frac{1}{100}$.

$$y_{n+5} = \frac{1}{2752} (2025 y_{n+4} + 700 y_{n+2} + 27 y_n)$$

$$+ \frac{15h}{688} (16 f_{n+5} + 45 f_{n+4} + 10 f_{n+2})$$

$r = \frac{3}{4}$, $C = -\frac{1}{71}$.

$p = 6$:

$$y_{n+5} = \frac{1}{5888} (3375 y_{n+4} + 2000 y_{n+2} + 513 y_n)$$

$$+ \frac{15h}{2944} (64 f_{n+5} + 225 f_{n+4} + 100 f_{n+2} + 9 f_n)$$

$r = \frac{1}{2}$, $C = -\frac{10}{1393}$.

$p = 7$:

$$y_{n+6} = \frac{1}{3104} (1215 y_{n+5} + 1000 y_{n+3} + 729 y_{n+1} + 160 y_n)$$

$$+ \frac{15h}{1552} (32 f_{n+6} + 135 f_{n+4} + 100 f_{n+3} + 27 f_{n+1})$$

$r = \frac{3}{10}$, $C = -\frac{5}{1372}$.

$$y_{n+8} = \frac{1}{197632} (120393 y_{n+8} + 74088 y_{n+6} + 2808 y_{n+2} + 343 y_n)$$

$$+ \frac{63h}{24704} (128 f_{n+9} + 441 f_{n+8} + 147 f_{n+6} + 9 f_{n+2})$$

$r = \frac{13}{24}$, $C = -\frac{63}{3625}$.

$p = 8$:

$$y_{n+7} = \frac{1}{180224} (42875 y_{n+6} + 42875 y_{n+4} + 83349 y_{n+2} + 11125 y_n)$$

$$+ \frac{105h}{90112} (256 f_{n+7} + 1225 f_{n+6} + 1225 f_{n+4} + 441 f_{n+2} + 25 f_n)$$

$r = \frac{1}{6}$, $C = -\frac{35}{14274}$.
\[ y_{n+9} = \frac{1}{928000}(413343y_{n+8} + 432000y_{n+6} + 76032y_{n+3} + 6625y_n) \]
\[ + \frac{9}{46400}(1600f_{n+9} + 6561f_{n+8} + 3600f_{n+6} + 576f_{n+3} + 25f_n), \]
\[ r = \frac{7}{20}, \quad c_9 = \frac{405}{43267}. \]

\[ p = 9: \]
\[ y_{n+9} = \frac{1}{3386096}(455625y_{n+8} + 592704y_{n+6} + 2185596y_{n+4} + 109296y_{n+1} + 42875y_n) \]
\[ + \frac{45h}{60466}(392f_{n+9} + 2025f_{n+8} + 2352f_{n+6} + 882f_{n+4} + 72f_{n+1}), \]
\[ r = \frac{5}{56}, \quad c_9 = \frac{14}{5723}. \]

\[ p = 10: \]
\[ y_{n+11} = \frac{1}{1078400000}(235782657y_{n+10} + 561515625y_{n+8} + 183997440y_{n+5} + 70709375y_{n+2} + 26394903y_n) \]
\[ + \frac{99h}{53920000}(160000f_{n+11} + 793881f_{n+10} + 680625f_{n+8} + 278784f_{n+5} + 75625f_{n+2} + 6516f_n), \]
\[ r = \frac{3}{20}, \quad c_{10} = \frac{7506675}{1145403224}. \]

\[ p = 11: \]
\[ y_{n+12} = \frac{1}{276200000}(25332021y_{n+11} + 137259375y_{n+9} + 45999360y_{n+6} + 20796875y_{n+3} + 40212369y_{n+1} + 66000000y_n) \]
\[ + \frac{99h}{13010000}(40000f_{n+12} + 216513f_{n+11} + 226875f_{n+9} + 139392f_{n+6} + 75625f_{n+3} + 19683f_{n+1}), \]
\[ r = \frac{13}{220}, \quad c_{11} = \frac{22275}{5026616}. \]

\[ p = 12: \]
\[ y_{n+14} = \frac{1}{1301272115000}(29894619132y_{n+13} + 626971072000y_{n+11} + 243729729243y_{n+9} + 292876876292y_{n+8} + 85153523000y_{n+6} + 22646295333y_n) \]
\[ + \frac{63h}{454990250}(2044900f_{n+14} + 11573604f_{n+13} + 13249600f_{n+11} + 9018000f_{n+8} + 4008004f_{n+5} + 828100f_{n+2} + 59049f_n), \]
\[ r = \frac{41}{2860}, \quad c_{13} = \frac{243243}{40781266}. \]
The main feature of the formulas above is that they all are "strongly 0-contractive" in the same norm and may thus be combined to form "contractive" LMM's (1.2). The Adams-Moulton formulas of order \( p \geq 3 \) do not satisfy condition (2.1), but they have reasonable absolute stability and accuracy properties for fixed step-size, and we therefore (in [1]) compared the "size" of the region of absolute stability and the error-constant of these formulas (denoted AM) to those of the formulas listed above with minimal step-number \( k_n \) (denoted SC). In the table below

\[ p_n, \ k_n \text{ and } C_n \text{ denote the order, step-number and error-constant, respectively, and} \]

\( S_{I_n} \text{ denotes the stability interval.} \)

<table>
<thead>
<tr>
<th>( p_n )</th>
<th>( k_n )</th>
<th>( C_n \cdot 10^3 )</th>
<th>( S_{I_n} )</th>
<th>( k_n )</th>
<th>( C_n \cdot 10^3 )</th>
<th>( S_{I_n} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>2</td>
<td>-41.666</td>
<td>[−6 ,0]</td>
<td>2</td>
<td>-27.777</td>
<td>[−4 ,0]</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-26.388</td>
<td>[−3 ,0]</td>
<td>3</td>
<td>-21.429</td>
<td>[−3.56 ,0]</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>-18.750</td>
<td>[−1.84 ,0]</td>
<td>4</td>
<td>-10.000</td>
<td>[−1.78 ,0]</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>-14.269</td>
<td>[−1.18 ,0]</td>
<td>5</td>
<td>7.179</td>
<td>[−1.45 ,0]</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>-11.367</td>
<td>[−0.769 ,0]</td>
<td>6</td>
<td>3.644</td>
<td>[−0.853 ,0]</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>-9.357</td>
<td>[−0.493 ,0]</td>
<td>7</td>
<td>2.452</td>
<td>[−0.645 ,0]</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>-7.893</td>
<td>[−0.310 ,0]</td>
<td>9</td>
<td>2.446</td>
<td>[−0.542 ,0]</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>-6.786</td>
<td>[−0.191 ,0]</td>
<td>11</td>
<td>6.554</td>
<td>[−0.808 ,0]</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>-5.924</td>
<td>[−0.115 ,0]</td>
<td>12</td>
<td>4.431</td>
<td>[−0.629 ,0]</td>
</tr>
<tr>
<td>12</td>
<td>11</td>
<td>-5.237</td>
<td>[−0.068 ,0]</td>
<td>14</td>
<td>5.965</td>
<td>[−0.624 ,0]</td>
</tr>
</tbody>
</table>
3. A₀-Contractivity of the BDF's of order ≤ 3.

In 1972 Brayton and Conley ([5]) showed that the usual variable-step version of the second-order BDF was contractive in a scaled max-norm when applied to any differential equation of the form

\[ y' = \lambda(t) \cdot y, \quad |\arg(-\lambda(t))| \leq \arctan(2\sqrt{2}) \approx 70^\circ 32', \]

with step-ratios \( \gamma \leq \frac{1}{2}(1+\sqrt{3}) \approx 1.366 \). If we allow larger values of \( \gamma \) the formula is not A₀-contractive in any norm. We shall sketch a proof of their result and show a similar result for another variable-step version of the second-order BDF since these proofs illustrate

a. A technique which in certain cases can be used for extending contractivity results in the fixed step-ratio case to results concerning variable step-ratios. This technique (a certain splitting of the companion matrix) will, for example, be used when proving the A₀-contractivity of the BDF's of order \( p = 1(1)3 \) (theorem 3.4).

b. The strength of the scaled max-norm derived in section 4. Note that even for constant step-size the second-order BDF is not 0-contractive in the max-norm itself and when the step-size varies, the usual variable-step version does not remain 0-contractive in any (constant) inner-product norm (cf. [4, theorem 3.1]).

c. It may be worthwhile considering variable-step versions of the BDF's other than the usual one (cf. [1]).

To demonstrate the technique mentioned in a. consider the companion matrix of the second-order BDF:

\[
\frac{1+2\gamma}{1+\gamma} y_{n+2} - (1+\gamma) y_{n+1} + \frac{\gamma^2}{1+\gamma} y_n = h_{n+2} f_{n+2}, \quad \gamma = \frac{h_{n+2}}{h_{n+1}},
\]

i.e.
\[ C(q, \gamma) = \begin{bmatrix} 0 & 1 \\ \frac{\gamma^2}{(1+2\gamma)-q(1+\gamma)} & \frac{(1+\gamma)^2}{(1+2\gamma)-q(1+\gamma)} \end{bmatrix} \]

A simple calculation shows that

\[ C(q, \gamma) = z(q, \gamma) \cdot C(0, \gamma) + [1 - z(q, \gamma)] \cdot C(\infty, \gamma_{\text{max}}), \]

\[ C(0, \gamma) = \frac{\gamma^2}{1+2\gamma} \cdot \varepsilon^{-1} \cdot C(0, \gamma_{\text{max}}) + [1 - \frac{\gamma^2}{1+2\gamma} \cdot \varepsilon^{-1}] \cdot C(0, 0), \]

where

\[ z(q, \gamma) = \frac{1+2\gamma}{(1+2\gamma)-q \cdot (1+\gamma)} \quad \text{and} \quad \varepsilon = \frac{\gamma_{\text{max}}^2}{1+2\gamma_{\text{max}}}. \]

If \( q \in [-\infty, 0] \), we see that \( C(q, \gamma) \) for \( \gamma \leq \gamma_{\text{max}} \) is a convex combination of \( C(0, \gamma_{\text{max}}) \), \( C(\infty, \gamma_{\text{max}}) \) and \( C(0, 0) \) (the latter may be regarded as the companion matrix of the first-order BDF). For most families of one-leg formulas with the same (constant) leading coefficient or the same (constant) ratio between the two leading coefficients in the second characteristic polynomial a similar splitting is possible, but to ensure that the combination is convex certain pairs (formula, step-ratios) may have to be excluded (as we shall see in the proof of theorem 3.4).

**Lemma 3.1**

For \( i = 1(1)k \) let \( \tilde{C}_i(q, \gamma_{k+1-i}, \gamma_{k+2-i}, \ldots, \gamma_{k-1}) \) denote the companion matrix of an \( i \)-step consistent one-leg formula with step-ratios \( \gamma_s = h_{n+s+1}/h_{n+s} \) and coefficients \((h_{n+k}/\bar{h}_n)^{a_{ij}}/\bar{\beta}_{ij}^{a_{ij}}\), \( j = k-i(1)k \). Assume that either \( \beta_{k}^1 = b \) or \( \beta_{k-1}^1/\beta_{k}^1 = B \) (\( b \) and \( B \) non-zero constants) for all \( i \), and that \( \bar{h}_n \cdot (\alpha_{k-i}^1/\alpha_k^i) \cdot (\beta_{k-1}^i/\beta_k^i) \neq 0 \) when \( \gamma_s \) are all equal to some number \( \eta > 0 \). Then, for any \( m \in \{1, 2, \ldots, k\} \) and any step-ratios \( \gamma_s \geq 0 \),
\[ C_m(q, \gamma_{k+1-m}, \ldots, \gamma_{k-1}) = \sum_{i=1}^{k} a_i \cdot C_i(0, \eta) + (1-z) \sum_{i=1}^{k} b_i \cdot C_i(\infty, \eta) \]

for some scalars \( a_i, b_i \), where \( \sum a_i = \sum b_i = 1 \), and \( z \in (0, 1] \) for \( (\beta_k^m/\alpha_k^m) \cdot \delta_n \cdot q \leq 0 \).

Here

\[
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
\vdots & \ddots & \ddots \\
0 & \cdots & 0
\end{bmatrix}
\]

and \( C_i(q, n) = C_i(q, n, \ldots, n) \) denote \( k \times k \) companion matrices.

**Proof:**

Let \( z \) be \( \frac{a_k^m}{(a_k^m \cdot \bar{h}_n \cdot q \cdot \beta_k^m/\alpha_k^m)} \). Then we only have to show the existence of scalars \( \{a_i\}_{i=1}^{k} \), \( \{b_i\}_{i=1}^{k} \) so that \( \sum a_i = \sum b_i = 1 \),

\[
C_m(0, \gamma_{k+1-m}, \ldots, \gamma_{k-1}) = \sum_{i=1}^{k} a_i \cdot C_i(0, \eta)
\]

and

\[
C_m(\infty, \gamma_{k+1-m}, \ldots, \gamma_{k-1}) = \sum_{i=1}^{k} b_i \cdot C_i(\infty, \eta).
\]

By looking at the last row of these matrices we find that \( a_k = a_{k-1} = \ldots = a_{m+1} = b_k = b_{k-1} = \ldots = b_{m+1} = 0 \) and the rest of the scalars can be found by solving two non-singular triangular systems of equations. Premultiplying the coefficient matrix and the vector on the right-hand side in these systems by either the vector \((1, 1, \ldots, 1)\) or the vector \((1, 0, \ldots, 0)\) it follows that \( \sum a_i = \sum b_i = 1 \).

\[ \square. \]
For the usual BDF's we need no splitting of $C_m(\infty, \gamma_{k+1-m}, \ldots, \gamma_{k-1})$ and the lemma will be used (with $q=0$) in this section for splitting $C_m(0, \gamma_{k+1-m}, \ldots, \gamma_{k-1})$ only. (Here $\beta_k^i \neq 0$ is not needed, but $\alpha_k^i \neq 0$ is).

Proving $A_0$-contractivity of the usual BDF's of order $p \leq 2$ for all step-ratios $\gamma \leq \gamma_{\text{max}}$ is thus equivalent to finding some norm in which $C(0, \gamma_{\text{max}})$, $C(\infty, \gamma_{\text{max}})$ and $C(0,0)$ all become contractive! Brayton and Conley succeeded in doing this by noting that for $\varepsilon = \gamma_{\text{max}}^2 / (1 + 2\gamma_{\text{max}}) \leq \frac{1}{2}$ all three matrices map a symmetric convex region of the form

![Diagram](image)

into itself. For $\varepsilon > \frac{1}{2}$ (i.e. $\gamma_{\text{max}} > \frac{1}{2}(1+\sqrt{3}) \approx 1.366$), an unfortunate combination of the 3 matrices will produce an unstable matrix, since e.g.

$$C(0, \gamma_{\text{max}})^n \cdot C(\infty, \gamma_{\text{max}}) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{\varepsilon - \varepsilon^n}{\varepsilon - 1} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \varepsilon^n \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

By noticing that the norm derived was of the form

$$|| C || = || T^{-1} CT ||_{\infty}, T = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \end{bmatrix},$$

they succeeded furthermore in establishing $A(\alpha)$-contractivity with $\alpha = \arctan(2\sqrt{2}) \approx 70.23'$. The $(b,k)$-norm developed in section 4 is a generalization of this norm to matrices of arbitrary dimension $k$.

Let us try the approach of Brayton and Conley on another variable-step version of the second-order BDF.
Proposition 3.2

The one-leg formula \((\gamma \hat{h}_{n+2}/h_{n+1})\)

\[
\frac{3}{2} y_{n+2} - 2 y_{n+1} + \frac{1}{2} y_n = \frac{3\gamma - 1}{2\gamma} \cdot h_{n+2} \mp (t_{n+1} + \beta h_{n+2}, \beta y_{n+2} + (1-\beta) y_{n+1}),
\]

where \(\beta = (3\gamma^2 + 1)/(2\gamma(3\gamma - 1))\), is contractive in a scaled max-norm (constructed below) for all step-ratios \(\frac{1}{3} < \gamma \leq \frac{1}{6}(3 + \sqrt{33}) \approx 1.457\) when applied to any scalar equation of the form

\[y' = \lambda(t) \cdot y, \quad |\arg(-\lambda(t))| \leq \alpha(\gamma), \alpha(\gamma) > 0,\]

\[\lceil \alpha(\gamma) \leq \arccos \left(\frac{1}{5}\right) \approx 78.027° \text{ for } \gamma \geq \frac{1}{12}(3 + \sqrt{33}) \rceil\]

By allowing smaller and larger values of \(\gamma\), the formula can be made unstable even for \(\alpha(\gamma) = 0\).

Proof:

The companion matrix is

\[C(q, \gamma) = z \cdot C(0, \gamma) + (1 - z) \cdot C(\infty, \gamma),\]

where

\[z = 6\gamma^2 / [6\gamma^2 - q(3\gamma^2 + 1)],\]

\[C(0, \gamma) = \begin{bmatrix} 0 & 1 \\ -1/3 & 4/3 \end{bmatrix} \quad \text{and} \quad C(\infty, \gamma) = \begin{bmatrix} 0 & 1 \\ 0 & -\varepsilon \end{bmatrix}, \quad \varepsilon = 3\gamma^2 - 2\gamma - 1 / 3\gamma^2 + 1.
\]

In this case, further splitting is not needed since it is easily seen that for \(-1 \leq \varepsilon \leq 1/3\) the matrices \(C(0, \gamma)\) and \(C(\infty, \gamma)\) map a symmetric region of the form
into itself.

For $\varepsilon < -1$, $C(\gamma, \gamma)$ is unstable and for $\varepsilon > \frac{1}{3}$ an unfortunate combination of the two matrices may produce an unstable matrix:

$$C(0, \gamma)^n \cdot C(\gamma, \gamma_{\text{max}})^{n \rightarrow \infty} \rightarrow -\frac{1}{2} (3\varepsilon+1) \cdot 1.$$ 

We note that the formula has now been proved $A_0$-contractive for $\gamma \in \left(\frac{1}{3}, \frac{1}{6} (3+\sqrt{33})\right]$ (excluded per definition of $\beta$) in the matrix norm

$$\|C\| = \|T^{-1}CT\| \infty, \quad T = \begin{bmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$ 

The fact that the companion matrix is contractive in this norm may be expressed as

$$|q(6\gamma^2 - 3\gamma - 1)| + |18\gamma^2 + 2q(6\gamma^2 - 3\gamma - 1)| \leq |18\gamma^2 - 3q(3\gamma^2 + 1)|.$$ 

For $\gamma \in \left[(3+\sqrt{33})/12, (3+\sqrt{33})/6\right]$ we find that $0 \leq 6\gamma^2 - 3\gamma - 1 \leq 3\gamma^2 + 1$, and a sufficient condition for contractivity in this case is thus $[z \in q(6\gamma^2 - 3\gamma - 1), a \leq 18\gamma^2]$:

$$|z| + |a + 2z| \leq |a - 3z|$$

$$|z| \cdot |a + 2z| \leq 2|z|^2 - 5a \cdot \text{Re } z.$$ 

If $|\text{arg}(-z)| \leq \arccos \left(\frac{1}{5}\right)$ then $-5 \cdot \text{Re } z \geq |z|$ and the condition is satisfied. In general, $A(0)$-contractivity follows from a
criterion in [7,p.468] (concerning the max-norm) since 
\[ |6\gamma^2-3\gamma-1| \leq |3\gamma^2+1| \]
for the step-ratios in question.

Remark

Although the BDF is not \( A_0 \)-contractive in the max-norm, we see from the polygons shown in this section that

\[
\| \Pi_n \sum_{n=N}^{N+2} C(q_n, \gamma_n) \|_\infty \leq 3(2, \text{ respectively}) \quad \forall N_1, N_2
\]

when \( q_n \leq 0 \) and

\[
\gamma_n \leq \frac{1}{2}(1+\sqrt{3}) \quad \left( \frac{1}{3} \leq \gamma_n \leq \frac{1}{6}(3+\sqrt{33}), \text{ respectively} \right)
\]

for the usual and the "unusual" variable-step version, respectively, i.e. only a moderate expansion (measured in the max-norm) can occur in these cases.

In the notation of section 4, the norm used by Brayton and Conley is the \((0.5,2)\)-norm, whereas the one constructed in the proof of proposition 3.2 is the \((1/3,2)\)-norm. The figures on the next page show for the fixed-step second-order BDF its region of absolute stability (the dashed line) and the regions of contractivity w.r.t. the two norms. The formula is not \( A \)-contractive in these norms.

We shall now apply the techniques demonstrated in this section to derive an \( A_0 \)-contractivity result for a method consisting of the BDF's of order \( p = 1(1)3 \). We start by using lemma 3.1 for splitting \( C_3(0, \gamma_1, \gamma_2) \) (although \( \beta_k \neq 0 \) is not satisfied).

The linear system for determination of \( a_1, a_2, a_3 \) (cf. lemma 3.1) is the following:
Fig. 3.1. The region of absolute stability and the region of contractivity in the \((1/3,2)\)-norm of the fixed-step second-order BDF.

Fig. 3.2. The region of contractivity in the \((0.5,2)\)-norm of the fixed-step second-order BDF.
\[
\begin{bmatrix}
1 & \frac{(1+n)^2}{1+2n} & \frac{(1+n)(1+n+n^2)^2}{3n^3+4n^2+3n+1} \\
0 & \frac{n^2}{1+2n} & \frac{n(1+n+n^2)^2}{3n^3+4n^2+3n+1} \\
0 & 0 & \frac{n^5(1+n)}{3n^3+4n^2+3n+1}
\end{bmatrix}
\begin{bmatrix}
a_1 \\
a_2 \\
a_3
\end{bmatrix}
= \begin{cases}
\text{last row in} \\
\{C_3(0, \gamma_1, \gamma_2)\} \\
\text{transposed}
\end{cases}
\]

Setting \( \gamma_1 = n \cdot \frac{a}{1+a} \) and \( \gamma_2 = n \cdot \frac{b}{1+b} \) where \( a, b \geq 0 \), the solution of this system will be

\[
a_3 = \frac{p_3}{q_3}, \text{ where}
\]

\[
p_3 = a^3 b^2 (3n^3+4n^2+3n+1) (b^2 n^2+2b(1+b)n+(1+b)^2),
\]

\[
q_3 = (1+a)(1+b)^2 (1+n) \left[ 3a^2 b^2 n^4 + ab (4a+7ab+3b) n^3 + a(1+b)(a+7ab+b)n^2 \\
+ 2(1+a)(1+b)(a+2ab+b)n+(1+a)^2 (1+b)^2 \right].
\]

\[
a_2 = \frac{p_2}{q_3}, \text{ where}
\]

\[
p_2 = b^2 (1+2n) \left[ a^2 b^2 n^5 + a^2 b (2a+5b+4)n^4 + a((1+b)(3a+2a^2+11ab+2b) \\
+ 2a^2 b)n^3 + a((1+b)(3+9a+3a^2+5b+13ab)+a^2 b)n^2 \\
+ (1+b)((1+b)(1+6a+9a^2)+2a^3)n+(1+b)^2(1+3a+3a^2) \right].
\]

\[
a_1 = \frac{p_1}{q_3}, \text{ where}
\]

\[
p_1 = (1+n) \cdot \Sigma_{i=0}^{3} \Sigma_{j=0}^{4} C_{ij} a^i b^j,
\]

\[
C_{00} = 1, \quad C_{10} = 2n+3, \quad C_{01} = 2n+4, \quad C_{20} = n^2+4n+3,
\]

\[
C_{11} = 6n^2+14n+12, \quad C_{02} = 4n+5, \quad C_{30} = (n+1)^2,
\]

\[
C_{21} = 4n^3+16n^2+22n+12, \quad C_{12} = 3n^3+12n^2+21n+15,
\]
\[ C_{03} = 2(n+1), \quad C_{31} = 2(n+1)(2n^2+3n+2), \]
\[ C_{22} = 3n^4+12n^3+27n^2+30n+15, \quad C_{13} = 2(n+1)(n^2+n+3), \quad C_{04} = 0, \]
\[ C_{32} = (n+1)(3n^3+8n^2+9n+5), \quad C_{23} = 2(n+1)(-n^3+n^2+2n+3), \]
\[ C_{14} = -n(n+1)^2, \quad C_{33} = 2(n^2+n+1)(n+1)^2, \]
\[ C_{24} = -2n(n^2+n+1)(n+1)^2, \quad C_{34} = 0. \]

Thus we find that \( a_2, a_3 \) are non-negative for all \( (a,b) \), whereas some requirements on the relation between \( a \) and \( b \) (i.e. \( \gamma_1 \) and \( \gamma_2 \)) are needed in order to ensure \( a_1 \geq 0 \). Let us first find \( \eta \), so as to make \( C_3(0,\eta), C_2(0,\eta), C_1(0,\eta), C_1(\infty,\eta) \) a stable family of matrices.

**Lemma 3.3**

\( C_3(0,\eta), C_2(0,\eta), C_1(0,\eta) \) and \( C_1(\infty,\eta) \) is a stable family of matrices if \( \eta = 1.011588... \) and not if \( \eta \) is larger.

**Proof:**

Let \( \varepsilon \) denote \( \eta^2/(1+2\eta) \). Then for \( |\varepsilon| < 1 \)

\[ C_2(0,\eta)^n \xrightarrow{n \to \infty} \frac{1}{1-\varepsilon} (1,1,1)^T \cdot (0,-\varepsilon,1). \]

If we for notational convenience write \( C_3(0,\eta) \) as

\[
C_3(0,\eta) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \gamma_0 & \gamma_1 & \gamma_2 \end{bmatrix}
\]

we find that for \( |\varepsilon| < 1 \)

\[
C_2(0,\eta)^n \cdot C_3(0,\eta)^2 \cdot C_1(\infty,\eta) \cdot (1,1,1)^T \xrightarrow{n \to \infty} [\gamma_0+\gamma_2(\gamma_0+\gamma_1)-\varepsilon(\gamma_0+\gamma_1)]/(1-\varepsilon) \cdot (1,1,1)^T.
\]
If \( \eta > 1.011588 \ldots \) we will have \( \gamma_0 + \gamma_2 (\gamma_0 + \gamma_1) - \varepsilon (\gamma_0 + \gamma_1) < \varepsilon - 1 \), and hence for \( n \) sufficiently large \( C_2 (0, \eta) \eta_n \cdot C_3 (0, \eta)^2 \cdot C_1 (\infty, \eta) \) will possess an eigenvalue of modulus larger than one. On the other hand, if \( \eta = 1.011588 \ldots \), it is possible to find a symmetric convex neighbourhood \( W \) of the origin so that all the matrices in question map \( W \) into itself. The following region will do:

\[
W = \text{the convex hull of the set}
\[
\left\{ \pm \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \pm \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \pm \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \pm \begin{bmatrix} 0 \\ c_0 + c_1 \end{bmatrix}, \pm \begin{bmatrix} 0 \\ c_0 \end{bmatrix}, \pm \begin{bmatrix} 0 \\ (c_0 + c_1 + 1) - 1 \\ u (c_0 + c_1 + 1) - 1 \end{bmatrix} \right\}
\]

Let \( \pm e_1, \pm e_2, \ldots, \pm e_6 \) denote the extreme points of \( W \). Then we can illustrate \( W \) in \( \mathbb{R}^3 = \{(x, y, z)^T | x, y, z \in \mathbb{R}\} \) by considering intersections with different planes. (Note, however, that \( W \) may be larger than indicated by these intersections).

\[
\begin{align*}
\text{x=1:} & \quad z & \text{x=0:} & \quad z \\
\text{e}_1 & \quad \text{e}_2 & \quad \text{e}_5 \\
\text{e}_3 & \quad \text{e}_4 & \quad \text{e}_6 \\
\text{x=c_0+c_1:} & \quad z & \text{x=-1:} & \quad z \\
\text{p}_1 & \quad \text{p}_2 & \quad \text{p}_3 & \quad \text{p}_4, \\
p_1 &= (1+x)e_5 + xe_3, & p_2 &= (1+x)e_6 + xe_1, \\
p_3 &= -(1+x)e_5 + xe_2, & p_4 &= -(1+x)e_6 + xe_4.
\end{align*}
\]
Using these intersections, a straight-forward (but tedious) calculation will show that $W$ is mapped into itself by any of the matrices in question if $\eta = 1.011588\ldots$.

By means of this lemma we can now show the main result of this section.

**Theorem 3.4**

Consider a variable step-size method based on the first, second and third order BDF. Let $k_n$ and $\gamma_n$ denote the order and the (last) step-ratio, respectively, used in the $n$'th integration step. If the variation of $(k_n, \gamma_n)$ for all $n$ is confined as shown below, the method will produce bounded solutions to any (scalar) equation of the form $y' = \lambda(t, y) \cdot y$, $\lambda(t, u) \leq 0 \forall t, u$.

\[
\begin{array}{c}
\gamma_{n-1} \\
\downarrow \\
k_n \leq 2 \\
\downarrow \\
k_n \leq 3 \\
\downarrow \\
k_n \leq 2 \\
\downarrow \\
k_n \leq 1 \\
\downarrow \\
\gamma_n
\end{array}
\]

($\eta = 1.011588\ldots$).

**Proof:**

The coefficients $C_{ij}$ on p. 23-24 are (rounded)

<table>
<thead>
<tr>
<th></th>
<th>1.000</th>
<th>6.023</th>
<th>9.046</th>
<th>4.023</th>
<th>0.000</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.023</td>
<td>32.302</td>
<td>51.629</td>
<td>20.256</td>
<td>-4.093</td>
<td></td>
</tr>
<tr>
<td>4.046</td>
<td>28.490</td>
<td>51.087</td>
<td>24.561</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>
We see that if \( b(\text{cf. p.23}) \) is "very large" \( p_1 \geq 0 \) is possible only if \( a \) is "very large" or "very small". Using a root-seeking algorithm we find that \( p_1 \geq 0 \) if

i) \( (\gamma_n/\eta) \leq 0.809 \)

ii) \( (\gamma_n/\eta) = 0.81 \) and \( |(\gamma_{n-1}/\eta) - 0.475| \geq 0.036 \)

iii) \( (\gamma_n/\eta) = 0.82 \) and \( |(\gamma_{n-1}/\eta) - 0.478| \geq 0.137 \)

iv) \( (\gamma_n/\eta) = 0.85 \) and \( |(\gamma_{n-1}/\eta) - 0.485| \geq 0.259 \)

v) \( (\gamma_n/\eta) = 0.90 \) and \( |(\gamma_{n-1}/\eta) - 0.493| \geq 0.371 \)

vi) \( (\gamma_n/\eta) = 0.99 \) and \( |(\gamma_{n-1}/\eta) - 0.4999| \geq 0.4899 \)

vii) \( (\gamma_n/\eta) = 0.999 \) and \( |(\gamma_{n-1}/\eta) - 0.5| \geq 0.499 \)

For \( k_n = 2 \), \( a = 0 \) and \( p_1 \geq 0 \) is satisfied for all \( b \geq 0 \) (all \( \gamma_n \in [0,\eta] \)). If \( k_n = 1 \) the companion matrix will map \( W \) (cf. the proof of lemma 3.3) into itself regardless of \( (\gamma_{n-1},\gamma_n) \). \( \square \).

Corollary 3.5

On the assumptions in theorem 3.4, the companion matrices satisfy

\[
\| \prod_{n=N_1}^{N_2} C(q_n) \|_{\infty} \leq 3.83, \quad \forall N_1, N_2.
\]

Proof:

We shall show that \( A = \{(x,y,z)^T \ | \ |(x,y,z)^T|_{\infty} \leq 0.261414 \} \) is a subset of \( W \) (cf. the proof of lemma 3.3) if \( \delta \leq 0.261414 \), and the proof will follow, since \( \delta \leq 3.83 \). \( A \) is convex and hence we only consider its extreme points. From the intersections shown p. 25 it is evident that \( z(0,\delta,\delta,0)^T, z(\delta-\delta,0)^T \in W \).

Furthermore, for \( \delta \approx 0.261414 \), we obtain with \( e_\delta = (0,-a,-b) \):
\[
\begin{bmatrix}
\delta \\
-\delta \\
\delta
\end{bmatrix} = \delta \cdot \begin{bmatrix} 1 \\
1 \\
1
\end{bmatrix} + \frac{2\delta-(b-a)\cdot c}{1-\gamma_0} \cdot \begin{bmatrix} 0 \\
-1 \\
-\gamma_0
\end{bmatrix} + c \cdot \begin{bmatrix} 0 \\
a \\
b
\end{bmatrix},
\]

\[
c = \frac{\delta(\gamma_0+3)+\gamma_0-1}{2b-a(\gamma_0+1)-(1-\gamma_0)}
\]

and hence \(\pm(\delta, -\delta, \delta)^T\) also lie in \(W\). The last extreme points to be considered are \(\pm(\delta, \delta, -\delta)^T\), but these points evidently lie in \(W\), since for \(\delta \approx 0.261414\):

\(\pm(0, \delta, -\delta)^T, \pm(1, \delta, -\delta)^T \in W\).

In the light of theorem 3.4, it is not surprising that unboundedness (or instability if \(\lambda\) is independent of \(\gamma\)) may occur even for fixed step-size if the 4'\textsuperscript{th} order BDF is included in the method (and no restrictions are imposed on the variation of \(\lambda\)). This is a fact which Brayton and Conley pointed out in [5]. In our notation they showed that for a \(k\)-step consistent one-leg formula \((\rho_n, \sigma_n)\) the pair of matrices

\[H = \{C(0, n, n, \ldots, n), C(q, n, n, \ldots, n)\}\]

is unstable if \(q\) satisfies

\[
\left| \frac{\alpha_{k,n} q}{\alpha_{k,n} - q \beta_{k,n}} + \rho_n'(1) \right| > \rho_n'(1),
\]

where \((\rho_n, \sigma_n)\) is assumed constant for fixed step-ratio \(n\). In the case of fixed step-size \((n=1)\), the \(k\)-step BDF has \(\beta_k = 1, \alpha_k = \sum_{j=1}^{k} \frac{1}{j}, \rho'(1) = 1\) and hence we have an unstable pair if \(k \geq 4\) and \(q < -2 \cdot \left[1 + 24 \cdot (1 + \sum_{j=5}^{k} \frac{1}{j})^{-1}\right]\).
4. Contractivity in Scaled Norms

In this section we shall define a generalization of the scaled max-norm used by Brayton and Conley when proving $A(a)$-contractivity of the second-order BDF [5]. This generalized norm has the form $\|C\|_b = \|T_b^{-1}CT_b\|_{\infty}$, where $T_b$ is a matrix dependent only on the parameter $b$. As an important by-product of our considerations, we will obtain a transformation which may be useful in showing contractivity even in the case of non-linear problems (cf. theorem 4.4). Before specifying $T_b$, let us indicate how powerful a tool similarity transformations is in connection with companion matrices.

Lemma 4.1

Let $C$ be a companion matrix and $A$ a matrix with the same characteristic polynomial. Then $C$ is similar to $A$ if and only if the minimal polynomial of $A$ has the same degree as the characteristic polynomial.

Proof: Follows from Householder [8,p.150 and p.18].

We may express the condition in lemma 4.1. in another way if $A$ is some "well behaved" transformation of a companion matrix. The transformations, $g$, which we will consider are analytic in some (open) region $\Omega$ containing the spectrum, $\Lambda$, of the companion matrix and univalent on $\Lambda$(i.e. $g(\lambda_i) \neq g(\lambda_j)$ when $\lambda_i \neq \lambda_j$ and $\lambda_i, \lambda_j \in \Lambda$). $(*)$

Lemma 4.2

Let $C$ be a companion matrix with eigenvalues $\{\lambda_i\}_{i \in I}$ and $C_g$ the companion matrix with the eigenvalues $\{g(\lambda_i)\}_{i \in I}$.

$(*)$ In the literature univalence is most often used in connection with regions.
where $g$ is analytic in a region $\Omega \supset \{\lambda_i\}_{i \in I}$ and univalent on $\{\lambda_i\}_{i \in I}$. Then

$$C_g = T^{-1} g(C) T \text{ for some matrix } T$$

if and only if

$$g'(\lambda) \neq 0 \text{ for all multiple eigenvalues of } C.$$ 

**Proof:**

We start by considering the Jordan Canonical Form (JCF) of the companion matrix $C$ (cf. e.g. [2,p.3]):

$$C = V \cdot J \cdot V^{-1}, \text{ i.e. } g(C) = V \cdot g(J) \cdot V^{-1}.$$ 

Since $g$ is univalent on the spectrum of $C$, $C_g$ can be represented by a JCF of the same block-structure as the JCF of $C$, say

$$C_g = V_g J_g V_g^{-1}.$$ 

It follows that $C_g$ is similar to $g(C)$ iff $J_g$ is similar to $g(J)$. We therefore consider the equation $J_g \cdot X = X \cdot g(J)$. Let $J = \text{blockdiag}(J_g, i)$, then $X = \{X_{i,j}\}_{i,j=1}^s$, where $J_g, i \cdot X_{i,j} = X_{i,j} \cdot g(J_{i,j})$. These equations have the form

$$\begin{bmatrix}
g(\lambda_i)
1
0
0
1
0
0
0
0
\end{bmatrix}
\begin{bmatrix}
x_1^T
\vdots
x_s^T
\end{bmatrix}
= \begin{bmatrix}
x_1^T
\vdots
\end{bmatrix}
\begin{bmatrix}
g(\lambda_j)
\cdots
\frac{1}{(m_j-1)} \cdot g^{(m_j-1)}(\lambda_j)
\end{bmatrix},$$

i.e.
\[
\begin{bmatrix}
\bar{x}_2^T \\
\vdots \\
\bar{x}_i^T \\
\cdot^T \\
\bar{x}_{m_i}^T \\
0^T
\end{bmatrix}
= 
\begin{bmatrix}
\bar{x}_1^T \\
\vdots \\
\bar{x}_i^T \\
\cdot^T \\
\bar{x}_{m_i}^T \\
0^T
\end{bmatrix}
\cdot 
\begin{bmatrix}
g(\lambda_j) - g(\lambda_i) \\
\vdots \\
\frac{1}{(m_j-1)!}g^{(m_j-1)}(\lambda_j)
\end{bmatrix}
\quad (4.1)
\]

Comparing elements in the matrices on both sides (starting with the last row) we observe that \( X_{ij} = 0 \) if \( i \neq j \). If \( i = j \) we find by comparing the elements (now starting with the first row) that \( X_{i_i} \) is upper triangular with the diagonal elements \( g'(\lambda_i)^{\ell-1} \cdot (1,0,\ldots,0) \cdot x_i \), \( \ell = 1(1)m_i \). Since the multiplicity of \( \lambda_r \) is \( m_r \), we see that \( J_g \cdot X = X \cdot g(J) \) has a non-singular solution iff \( g'(\lambda_r) \neq 0 \) for all multiple eigenvalues \( \lambda_r \). On this assumption, \( C_g = T^{-1}g(C)T \) follows for

\[
T = VX^{-1}V^{-1}g \quad (4.2)
\]

If the similarity transformation in the lemma above is to be incorporated in a norm, it will be most useful if \( T \) is independent of \( C \).

**Theorem 4.3**

The matrix \( T \) in lemma 4.2 can be chosen independent of \( C \) if and only if \( g(C) \) has the form

\[
g(C) = (-c \cdot C + a \cdot I)^{-1}(d \cdot C - b \cdot I), \quad ad \neq bc . \quad (4.3)
\]

In this case \( T = \{ t_{ij} \} \) can be chosen so that for \( k = \text{dim } (C) \):

\[
\sum_{j=1}^{k} t_{ij} \lambda_j^{j-1} = (c \lambda + d)^{k-1} \frac{(a \lambda + b)^{i-1}}{(c \lambda + d)} , \quad i = 1(1)k, \forall \lambda . \quad (4.4)
\]

**Proof:**

In order to show that \( g \) must be of the stated form, split \( V_g, X \) and \( V \) into blocks:
\[ V_g = [V_{g,1}, V_{g,2}, \ldots, V_{g,s}], \quad \lambda = \text{blockdiag} \{ x_{\lambda} \}, \quad V = [V_1, V_2, \ldots, V_s]. \]

According to (4.2) \( T = U = \{ u_{ij} \} \) must satisfy

\[ V_{g,\lambda} \cdot X_{\lambda} = U \cdot V_{\lambda} = \{ u_{i}^{(j-1)}(\lambda_{,})/(j-1)! \}_{i,j}, \quad \lambda = 1(1)s, \]

where \( u_{i} \) denotes the polynomial \( u_{i}(\lambda) = \sum_{j=1}^{k} u_{ij} \lambda^{j-1} \).

Since \( X_{\lambda} \) is triangular (cf. the proof of lemma 4.2), we find by comparing the elements in the first column of \( V_{g,\lambda} \cdot X_{\lambda} \) and \( U \cdot V_{\lambda} \) that, for \( i = 1(1)k \):

\[ g(\lambda) i^{-1} \cdot x_{1,1} = u_{i}(\lambda), \quad \text{i.e.} \quad u_{i}(\lambda) = u_{1}(\lambda) \cdot u_{1}(\lambda) i^{-1} - u_{1}(\lambda) \cdot u_{2}(\lambda) i^{-1} = 0 \]

Since \( u_{i}, \quad i = 1(1)k \), are independent of \( \lambda \), \( \lambda = 1(1)s \), it follows that \( u_{1} \cdot u_{1}^{i-1} - u_{1} \cdot u_{2}^{i-1} = 0 \), \( i = 1(1)k \), and, in particular, since \( u_{i} \neq 0 \) is necessary for \( U \) to be non-singular,

\[ u_{k} = \left( \frac{V_{2}}{V_{1}} \right)^{k-1} \cdot u_{1}, \quad \text{where} \quad V_{i} = u_{i}/\gcd(u_{1},u_{2}), \quad i = 1,2. \]

We see that \( V_{1}^{k-1} \) must divide \( u_{1} \) and thus deg \( (V_{1}) \leq 1 \). Likewise, \( V_{2} \) is at most of degree 1 since deg \( (u_{k}) \leq k-1 \). Hence \( u_{2}/u_{1} \) must be of the form

\[ \frac{u_{2}(\lambda)}{u_{1}(\lambda)} = \frac{d\lambda - b}{-c\lambda + a}, \]

where \( ad \neq bc \) because \( U \) must be non-singular.

If we now compare the first two rows of \( V_{g,\lambda} \cdot X_{\lambda} \) with those of \( U \cdot V_{\lambda} \), we notice that \( (X_{\lambda} = :\{ x_{i,j}^{(\lambda)} \}) \):

\[ x_{1,j}^{(\lambda)} = u_{1}^{(j-1)}(\lambda_{,})/(j-1)! \quad \text{and} \quad x_{2,j}^{(\lambda)} = u_{2}^{(j-1)}(\lambda_{,})/(j-1)! - x_{1,j}^{(\lambda)} g(\lambda). \]

On the other hand, (4.1) with \( \lambda_{i} = \lambda_{j} = \lambda \) tells us that
\( x_{2j}^{(l)} = \sum_{i=1}^{j-1} x_{1i}^{(l)} \cdot g(j-i)(\lambda_\ell)/(j-i)!, \ j = 1(1)m_\ell. \)

Combining these relations, one finds that (using the chain rule)

\[ u_2^{(j-1)}(\lambda_\ell) = (u_1 \cdot g)^{(j-1)}(\lambda_\ell), \ j = 1(1)m_\ell, \]

and

\[ g(C) = u_1(C)^{-1} \cdot u_2(C) = (-c \cdot C + aI)^{-1} \cdot (dC - bI) \]

follows easily.

In order to show that

\[ (-cC + aI) \cdot T \cdot C_g = (dC - bI) \cdot T \quad (4.5) \]

holds for the matrix \( T \) given in (4.4), apply the matrix on the right-hand-side of (4.5) to the vector \((1, r, \ldots, r^{k-1})^T\).

It is easily seen that this gives the vector \( v = (v_1, v_2, \ldots v_k)^T \), where

\[ v_i = (ad-bc) \cdot r \cdot (cr+d)^{k-2} \cdot \left( \frac{ar+b}{cr+d} \right)^{i-1}, \ i = 1(1)k-1, \]

\[ v_k = (cr+d)^{k-1} \left[ (ad-bc) \cdot \frac{r}{cr+d} \cdot \frac{ar+b}{cr+d}^{k-1} \cdot \frac{d}{p_k} \cdot p\left( \frac{ar+b}{cr+d} \right) \right]. \]

Here \( p(\lambda) = \sum_{j=0}^{k} p_j \lambda^j \) denotes the polynomial \( p_k \cdot (1)^k \cdot \det (C - \lambda I) \).

Likewise, we see that

\[ C_g \cdot (1, r, \ldots, r^{k-1})^T = (r, r^2, \ldots, r^{k-1}, r \cdot \frac{1}{\tilde{p}_k} \cdot (cr+d)^k \cdot p\left( \frac{ar+b}{cr+d} \right))^T, \]

where \( \tilde{p}_k = \sum_{j=0}^{k} p_j a^j c^{k-j} \). It follows that \( T \cdot C_g \cdot (1, r, \ldots, r^{k-1})^T = w = (w_1, w_2, \ldots w_k)^T \), where

\[ w_i = (cr+d)^{k-1} \cdot \left[ r \cdot \frac{(ar+b)^{i-1}}{cr+d} \cdot -c^{-i} \cdot a^{-i} \cdot (cr+d) \cdot p\left( \frac{ar+b}{cr+d} \right) / \tilde{p}_k \right]. \]
Some elementary calculations show that \((-cC+aI)w = v\), and (4.5) is proven. The theorem now follows from (4.5) since
\(-cC+aI\) and \(T\) are non-singular. \(T^{-1} = (u_{ij})\) is obtained by inserting \(\lambda = (du-b)/(cu+a)\) in (4.4):

\[
\sum_{j=1}^{k} u_{ij}^*u_j^j = \left(\frac{-cu+a}{ad-bc}\right)^{k-1} \cdot \left(\frac{du-b}{cu+a}\right)^{i-1}, i = 1(1)k, \forall \mu. \tag{4.6}
\]

It follows from (4.5) that

\[T^{-1}CT = (aC_g + bI)(cC_g + dI)^{-1}.\]

In a Hilbert space \(||C_g|| \leq 1\) will thus imply \(||T^{-1}CT|| \leq 1\) if

\[
\left|\frac{az+b}{cz+d}\right| \leq 1 \text{ for all } |z| \leq 1. \tag{4.7}
\]

If \(a, b, c\) and \(d\) are real, (4.7) is equivalent to

\[|a+b| \leq |c+d|, |a-b| \leq |c-d| \text{ and } |c| < |d|.\]

For a general operator norm \(||C_g|| \leq 1\) will imply

\[|d|-|c| \cdot |T^{-1}CT|| \leq |a|+|b|,\]

and thus \(||T^{-1}CT|| \leq 1\) if \(|a|+|b| \leq |d|-|c|\). It is interesting to find a similar result even in the non-linear case:

**Theorem 4.4**

For any integer \(s \geq 1\) let \(Z_n^s\) denote the super-vector
\((z_n^T, z_{n+1}^T, \ldots , z_{n+s-1}^T)^T\). Furthermore, let \(a, b, c\) and \(d\) denote scalars satisfying

\[ad-bc, |a|+|b| \leq |d|-|c| \text{ and } c^k \cdot \rho_n(a/c) \cdot c^k \cdot \sigma_n(a/c) \geq 0,\]

where \((\rho_n, \sigma_n)\) is a k-step one-leg formula.
Assume that, for some function $f$ and some sequences $\{\tilde{r}_n\}$, $\{t_n\}$, any solution of the $k$-step one-leg equation

$$\tilde{r}_n(E)x_n = \tilde{r}_n(E)f(t_n, \tilde{r}_n(E)x_n),$$

(4.8)

$$\tilde{r}_n(r) = (cr+d)^k \cdot \rho_n \left( \frac{ar+b}{cr+d} \right),$$

$$\tilde{r}_n(r) = (cr+d)^k \cdot \rho_n \left( \frac{ar+b}{cr+d} \right),$$

satisfies

$$|| x_{n+1}^k || \leq || x_n^k ||,$$

where $|| \cdot ||$ is some vector-norm. Then any solution of the one-leg equation

$$\rho_n(E)y_n = \tilde{r}_n(E)f(t_n, \rho_n(E)y_n),$$

(4.9)

will satisfy

$$|| (T^{-1} \otimes I) \cdot y_{n+1}^k || \leq || (T^{-1} \otimes I) \cdot y_n^k ||,$$

where $T=\{t_{ij}\}$ is given by (4.4) and $\otimes$ denotes the Kronecker product.

**Proof:**

For any integer $s \geq 1$ let $T_s=\{t_{ij}^s\}$ denote the matrix of dimension $s$ given by

$$\sum_{j=1}^{s} t_{ij}^s \cdot \lambda^{j-1} = (c\lambda+d)^{s-1} \cdot \left( \frac{a\lambda+b}{c\lambda+d} \right)^{i-1}, \ i=1(1)s, \ \forall \lambda.$$

From (4.6) we have that $U_s=\{u_{ij}^s\}=T_s^{-1}$ is given by

$$\sum_{j=1}^{s} u_{ij}^s \xi^{j-1} = \left( \frac{-c\xi+a}{ad-bc} \right)^{s-1} \left( \frac{d\xi-b}{-c\xi+a} \right)^{i-1}, \ i=1(1)s, \ \forall \xi.$$

Let $u_{ij}^s(\xi)$ denote the polynomials $\sum_{j=1}^{s} u_{ij}^s \xi^{j-1}$. Then
\[ u_i^{s+1}(\xi) = \frac{(-c\xi + a)}{(ad-bc)} u_i^s(\xi), \quad i = 1(1)s, \]
\[ u_i^{s+1}(\xi) = \frac{(d\xi - b)}{(ad-bc)} u_{i-1}^s(\xi), i = 2(1)s+1. \]

In other words \( T_{s+1}^{-1} \) can be decomposed in two ways

\[
\begin{align*}
(ad-bc) \cdot T_{s+1}^{-1} & = \frac{[-c \cdot [0 \mid T_s^{-1}] + a \cdot [T_s^{-1} \mid 0]}{a_1 a_2 \ldots \ldots \ldots a_{s+1}} \\
& = \begin{bmatrix}
\begin{array}{c}
b_1 b_2 \ldots \ldots \ldots b_{s+1} \\
\end{array}
\end{bmatrix},
\end{align*}
\]

(4.10)

Let now \( y_n^{k+1} \) satisfy (4.9). Then \( x_n^{k+1} := (T_k^{-1} \otimes I) \cdot y_n^{k+1} \) satisfies (4.8) and thus \( \| x_{n+1}^k \| \leq \| x_n^k \| \).

Making use of the decompositions in (4.10), we find that

\[
|d| \cdot \| (T_k^{-1} \otimes I) y_n^k \| - |b| \cdot \| (T_k^{-1} \otimes I) y_n^k \| \leq \]
\[
|d \cdot (T_k^{-1} \otimes I) y_n^k - b \cdot (T_k^{-1} \otimes I) y_n^k| = |ad-bc| \cdot \| x_{n+1}^k \| \leq \]
\[
|ad-bc| \cdot \| x_n^k \| = |c \cdot (T_k^{-1} \otimes I) y_n^k + a \cdot (T_k^{-1} \otimes I) y_n^k| \leq \]
\[
|c| \cdot \| (T_k^{-1} \otimes I) y_n^k + |a| \cdot \| (T_k^{-1} \otimes I) y_n^k|,
\]

and the proof follows, since \( 0 < |a| + |b| \leq |d| - |c| \).

\[ \Box \]

Remark

It seems rather essential that our coordinate transformation represented by the matrix \( T_k^{-1} \) is such that \( T_k^{-1} \) can be decomposed in two ways

\[
T_k^{-1} = \begin{bmatrix}
-\tilde{c} \cdot [0 \mid R] + \tilde{a} \cdot [R \mid 0] \\
\tilde{a}_1 \tilde{a}_2 \ldots \ldots \tilde{a}_{k+1}
\end{bmatrix} = \begin{bmatrix}
\tilde{b}_1 \tilde{b}_2 \ldots \ldots \tilde{b}_{k+1} \\
\tilde{a}_1 \tilde{a}_2 \ldots \ldots \tilde{a}_{k+1}
\end{bmatrix} = \begin{bmatrix}
\tilde{b}_1 \tilde{b}_2 \ldots \ldots \tilde{b}_{k+1} \\
\tilde{a}_1 \tilde{a}_2 \ldots \ldots \tilde{a}_{k+1}
\end{bmatrix}.
\]
where $R = \{r_{ij}\}$ is an arbitrary non-singular matrix. It is therefore interesting to note that - apart from a normalization factor - the particular matrix $T_k^{-1}$ used in theorem 4.4 is the only matrix with this property. This fact follows from the necessary identities $(r_i(\xi) := \sum_{j=1}^{k} r_{ij} \cdot \xi^{j-1})$:

$$(d\tilde{\xi} - \tilde{b}) \cdot r_{i-1}(\xi) = (-c\tilde{\xi} + \tilde{a}) \cdot r_i(\xi), \quad i = 2(1)k.$$  

From these we obtain that $R$ must be such that

$$r_i(\xi) = \left(\frac{d\tilde{\xi} - \tilde{b}}{-c\tilde{\xi} + \tilde{a}}\right)^{i-1} \cdot r_1(\xi) \cdot \text{const} \cdot (-c\tilde{\xi} + \tilde{a})^{k-1} \cdot \left(\frac{d\tilde{\xi} - \tilde{b}}{-c\tilde{\xi} + \tilde{a}}\right)^{i-1}.$$  

As in the proof of theorem 4.3, we may calculate

$$T_k^{-1} CT \cdot (1, \lambda, \ldots, \lambda^{k-1})^T = (\mu_1, \mu_2, \ldots, \mu_k)^T.$$  

We find that

$$\mu_i = \frac{1}{c\lambda + d} \left[ (a\lambda + b) \cdot \lambda^{i-1} - \frac{(c\lambda + d)^{k-i} \cdot (c\lambda + d)^{k-1}}{p_k \cdot (ad-bc)^k} \cdot \frac{(a\lambda + b)}{(c\lambda + d)} \right].$$  

We notice that if and only if $c \cdot d = 0$, the matrix $T_k^{-1} CT$ will contain only one row dependent on $p$. Since we here wish to consider the criterion $\|T_k^{-1} CT\|_\infty \leq 1$, we thus see that this criterion is greatly simplified by choosing $c \cdot d = 0$. Since $d = 0$ makes it impossible to satisfy the conditions of theorem 4.4 we shall choose $c = 0$ and (without loss of generality) $d = 1$. The conditions of theorem 4.4 then read $a \neq 0$ and $|a| + |b| \leq 1$. If the theorem, however, is to be applied to a consistent formula $(\rho, \sigma)$, the polynomial $\hat{p}_n$ will have a root in $(1-b)/a$ and we therefore require $|1-b| \leq |a|$ to be satisfied, too. These conditions only leave the possibility $|a| = 1-b$, $0 \leq b < 1$. Since $\|T_k^{-1} CT\|_\infty$ and $\|C_q\|_\infty$ are independent of the orientation of $a$ we choose $a = 1-b$ and obtain the $(b,k)$-norm:
Definition 4.5

A norm is said to be a \((b,k)\)-norm if it for some \(b \in [0,1)\) has the form

\[
\| A \|_b = \| T_b^{-1} A T_b \|_\infty,
\]

where \(T_b = (t_{ij})\) is the triangular matrix of dimension \(k\) given by

\[
t_{ij} = \begin{cases} 
(i-1) (i-1) b^{i-j} (1-b)^{j-1}, & 1 \leq j \leq i \leq k
\end{cases}
\]

\[4.12\]

It is known that for any two ordered triples \(\{\lambda_1, \lambda_2, \lambda_3\}\) and \(\{\mu_1, \mu_2, \mu_3\}\) of distinct elements in \(\mathbb{C}\) there is one (and only one) linear fractional transformation which maps \(\lambda_i\) to \(\mu_i\), \(i = 1(1)3\). Although it is intuitively clear from this that we lose some "adaptability" of our norm by withholding only one parameter in the fractional transformation, we shall see that this choice simplifies the analysis very much without eliminating all "adaptability" of the norm.

In certain cases we may have to shift the value of \(b\) in the \((b,k)\)-norm during the integration, but the effect of this is easily calculated.

Lemma 4.6

Assume that

\[
\| (T_n^{-1} \otimes I) \cdot Y_{n+1} \| \leq \| (T_n^{-1} \otimes I) \cdot Y_n \|, \quad n = 0, 1, \ldots,
\]

where \(T_n = (t_{ij}^{(n)})\) is the \(k \times k\) matrix determined by

\[
\sum_{j=1}^{k} t_{ij}^{(n)} \cdot \lambda_j^{-1} = (c_n \lambda + d_n) k^{-1} \cdot \left( \frac{a_i \lambda + b_i}{c_n \lambda + d_n} \right), \quad i = 1(1)k, \quad \forall \lambda,
\]

where \(a_n d_n \neq b_n c_n, \forall n\). Then

\[
\| Y_{n+1} \| \leq \sum_{s=0}^{n-1} \Pi \| (T_{s+1}^{-1} \otimes I) \| \cdot \| T_s \otimes I \| \cdot \| T_0^{-1} \otimes I \| \cdot \| Y_0 \|,
\]
and if \( \| \cdot \| = \| \cdot \|_\infty \), \( c_n \equiv 0, d_n \equiv 1, a_n \equiv 1 - b_n, b_n \in [0, 1) \)

\[
\| Y_{n+1} \|_\infty \leq \prod_{s=0}^{n-1} (1 + \varepsilon_s) \cdot \left( \frac{1 + b_s}{1 - b_s} \right)^{k-1} \cdot \| Y_0 \|_\infty,
\]

\[
\varepsilon_s = \begin{cases} 
0 & \text{if } b_{s+1} \leq b_s \\
2 \cdot \frac{b_{s+1} - b_s}{1 - b_{s+1}} & \text{otherwise}.
\end{cases}
\]

**Proof:**

For \( n = 1, 2, \ldots \)

\[
\| (T_n^{-1} \otimes I) \cdot Y_{n+1} \| \leq \| (T_n^{-1} \otimes I) (T_{n-1}^{-1} \otimes I) (T_{n-1}^{-1} \otimes I) \cdot Y_n \|
\]

\[
\leq \| (T_n^{-1} \otimes I) \cdot I \| \cdot \| (T_{n-1}^{-1} \otimes I) \cdot Y_n \|.
\]

Hence

\[
\| Y_{n+1} \| = \| (T_n \otimes I) (T_n^{-1} \otimes I) \cdot Y_{n+1} \| \leq \| (T_n \otimes I) \|
\]

\[
\cdot \prod_{s=0}^{n-1} \| (T_{s+1}^{-1} \cdot T_s) \otimes I \| \cdot \| T_0^{-1} \otimes I \| \cdot \| Y_0 \|.
\]

From (4.4) and (4.6) we find that \( T_{s+1}^{-1} \cdot T_s =: R_s = \{ r_{ij}^{(s)} \} \) is determined by the requirements

\[
k \sum_{j=1}^{k} r_{ij}^{(s)} j^{-1} = \left[ \frac{-c_{s+1} (a_s \lambda + b_s) + a_{s+1} (c_s \lambda + d_s)}{a_{s+1} d_{s+1} - b_{s+1} c_{s+1}} \right]^{k-1} \cdot \\
\left[ \frac{d_{s+1} (a_s \lambda + b_s) - b_{s+1} (c_s \lambda + d_s)}{-c_{s+1} (a_s \lambda + b_s) + a_{s+1} (c_s \lambda + d_s)} \right]^{i-1}, \quad i = 1(1) k, \quad \forall \lambda.
\]

If \( c_n \equiv 0, d_n \equiv 1, a_n \equiv 1 - b_n, b_n \in [0, 1) \) we thus have that

\[
\| (T_{s+1}^{-1} \cdot T_s) \otimes I \|_\infty = \| T_{s+1}^{-1} \cdot T_s \|_\infty = \left[ 1 + \frac{(b_{s+1} - b_s) + b_{s+1} b_s}{1 - b_{s+1}} \right]^{k-1},
\]
\[ ||T_n \otimes I||_\infty = ||T_n||_\infty = \left( |1-b_n| + |b_n| \right)^{k-1} = 1, \]

and

\[ ||T_0^{-1} \otimes I||_\infty = ||T_0^{-1}||_\infty = \left[ (1+b_0)/(1-b_0) \right]^{k-1}, \text{q.e.d.} \]

For the sake of reference we state some of the properties of the \((b,k)\)-norm:

**Theorem 4.7**

Let \( C \) be the companion matrix of a polynomial \( p \) of degree \( k \geq 2 \). If there exists a real number \( b \in [0,1) \) so that

\[
\sum_{j=0}^{k-2} \frac{(1-b)^j}{j!} |p(j)(b)| + \frac{(1-b)^{k-1}}{(k-1)!} \cdot |p(k-1)(b) - \frac{b}{k} \cdot p(k)(b)| \leq \frac{(1-b)^{k-1}}{k!} \cdot |p(k)(b)|,
\]

then \( ||C||_b = ||(1-b)C + b \cdot I||_\infty \leq 1 \), where \( C_b \) is the companion matrix of the polynomial \( q(r) = p((1-b)r + b) \), and

\[ ||C^n||_\infty \leq ||T_b||_\infty \cdot ||T_b^{-1}||_\infty = ||T_b^{-1}||_\infty = \left( \frac{1+b}{1-b} \right)^{k-1}, \forall n \geq 0. \]

where \( T_b \) is given by (4.12).

**Proof:**

(4.13) follows from (4.11) because

\[
\mu_k = (1-b) \cdot \lambda^k + b \lambda^{k-1} - p((1-b) \lambda + b) / (p_k \cdot (1-b)^{k-1}) \\
= - \left( p_k \cdot (1-b)^{k-1} \right)^{-1} \cdot \left\{ (1-b)^{k-1} \cdot \left[ \frac{p(k-1)(b) - b \lambda^k}{(k-1)!} \right] \cdot \lambda^{k-1} + \sum_{j=0}^{k-2} \frac{(1-b)^j}{j!} p(j)(b) \lambda^j \right\}.
\]

From the proof of theorem 4.4 we see that \( T_b^{-1} \) is the triangular matrix with the elements
\[ u_{ij} = \binom{i-1}{j-1} (-b)^{i-j} / (1-b)^{i-j}, \quad 1 \leq j \leq i \leq k, \]

and the theorem follows easily from the previous discussion. □.

Because of the similarity between \( \|C\|_b \leq 1 \), \( \|C_b\|_\infty \leq 1 \) and \( \|C\|_\infty \leq 1 \) we may utilize some of the results in [7] and [9] to derive some relevant criteria.

**Lemma 4.8**

Let \( C(q) \) denote the companion matrix of the polynomial

\[ p(r, q) = \rho(r) - q \cdot \sigma(r), \text{ where } \rho(r) = \sum_{j=0}^{k} \alpha_j r^j, \quad \sigma(r) = \sum_{j=0}^{k} \beta_j r^j \]

and \( p(1, q) = -q \). Furthermore, let \( C_b(q) \) denote the companion matrix of the polynomials \( p_b(r, q) = p((1-b) \cdot r + b, q), \; 0 \leq b < 1 \).

Then

\[ \|C(0)\|_b \leq 1 \text{ if and only if } \alpha_k \cdot \rho^{(j)}(b) \leq 0, \; j = 0(1)k-2, \text{ and } b \leq -\alpha_{k-1} / [(k-1) \alpha_k]; \]

\[ \|C_b(0)\|_\infty \leq 1 \text{ if and only if } \alpha_k \cdot \rho^{(j)}(b) \leq 0, \; j = 0(1)k-1 \quad (4.15) \]

If \( \alpha_k > 0 \), and (4.14) or (4.15), respectively, holds with strict inequalities, then the following is valid

\[ \forall \alpha \in [0, \pi/2] \exists a > 0 : \|C(q)\|_b \leq 1 \text{ or } \|C_b(q)\|_\infty \leq 1, \text{ respectively,} \]

holds for all \( q \in \{z | \arg(-z) \leq \alpha \text{ and Re}(-z) \leq a\} \).

**Proof:**

\[ \|C(q)\|_b \leq 1 \text{ and } \|C_b(q)\|_\infty \leq 1 \text{ are both equivalent to criteria of the form} \]

\[ \sum_{j=0}^{k-1} |a_j - q \cdot b_j| \leq |a_k - q \cdot b_k|, \quad (4.16) \]

where
\[ \tilde{p}(r,q) \overset{1}{\equiv} \sum_{j=0}^{k} (a_j - q \cdot b_j) r^j = p_b(r,q) + b \cdot (\alpha_k - q^2) \cdot (r-1)^{k-1} \]

or \( p_b(r,q) \), respectively (cf. theorem 4.7). We note that \( \tilde{p}(1,0) = 0 \), and thus \( \|C(0)\|_b \leq 1 \) and \( \|C_b(0)\|_m \leq 1 \) are equivalent to \( a_k \cdot a_j \leq 0 \) for \( j = 0,1,k-1 \), which gives (4.14) and (4.15).

Assume now that \( \alpha_k > 0 \) (i.e. \( \alpha_k > 0 \)) and that \( a_k \cdot a_j \leq 0 \) for \( j = 0,1,k-1 \). Let \( \alpha \in [0,\pi/2] \) be given. We must then find \( a > 0 \) so that (4.16) holds for all \( q = c \cdot (-1 + ib) \), where \( 0 \leq c \leq a \) and \( |b| \leq \tan(\alpha) \). But if \( q = c \cdot (-1 + ib) \) we have as \( a \) tends to zero,

\[ |a_j - q b_j| = |a_j| \cdot (1 + c \cdot b_j/a_j + O(c^2)), \]

and thus (4.16) reads

\[ 0 \leq (|a_k| - \sum_{j=0}^{k-1} |a_j|) + \text{sgn}(a_k) \cdot c \cdot (b_j + \sum_{j=0}^{k-1} b_j) + O(c^2) = c + O(c^2). \]

\[ \square. \]

**Remark**

The conditions for strong 0-contractivity in the \((b,k)\)-norm (i.e. \( \exists r > 0: \|C(q)\|_b \leq 1 \) for all \( q \in \{z: |z+r| \leq r\} \)) have later been found by Nevanlinna and the writer ([10]) to be:

\[ \alpha_k > 0, \]

\[ \rho(j)(b) \leq 0 \quad \text{and} \quad \rho(j)(b) = 0 \Rightarrow \alpha_j(b) = 0 \quad \text{for} \quad j = 0,1,k-2, \]

\[ b \leq -\alpha_{k-1}/[(k-1)\alpha_k] \quad \text{and} \quad b = -\alpha_{k-1}/[(k-1)\alpha_k] \Rightarrow \]

\[ b = -\beta_{k-1}/[(k-1)\beta_k] \]

\[ \square. \]

Necessary and sufficient criteria for \( \|C(q)\|_b \leq 1 \) or \( \|C_b(q)\|_m \leq 1 \) to hold for all \( q \in \tilde{C}^- \) are found by using techniques very similar to those applied in [9] concerning the case \( b = 0 \):
Lemma 4.9

With the notation of lemma 4.8, assume that $\alpha_k/\beta_k > 0$. Then $\|C(q)\|_b \leq 1$ holds for all $q \in \hat{C}$ if and only if

1) $\|C(0)\|_b \leq 1$,

2) $\rho^{(j)}(b) = 0 \Rightarrow \sigma^{(j)}(b) = 0$ for $j = 0(1)k-2$, and

$b = -\alpha_k/(k-1)\alpha_k \Rightarrow b = -\beta_k/(k-1)\beta_k$, and

3) $\alpha_k \cdot \left[ \sum_{j=0}^{k-2} \frac{(1-b)^j}{j!} \left( \frac{\sigma^{(j)}(b)}{\rho^{(j)}(b)} \right)^2 + (1-b)^{k-1} \frac{\beta_k (k-1)b + \beta_{k-1}^2}{\alpha_k (k-1)b + \alpha_{k-1}} + (1-b)^{k-1} \frac{\beta_k^2}{\alpha_k} \right] \geq 0$.

$\|C_b(q)\|_\infty \leq 1$ holds for all $q \in \hat{C}$ if and only if

1') $\|C_b(0)\|_\infty \leq 1$,

2') $\rho^{(j)}(b) = 0 \Rightarrow \sigma^{(j)}(b) = 0$ for $j = 0(1)k-1$, and

3') $\alpha_k \cdot \left[ \sum_{j=0}^{k-2} \frac{(1-b)^j}{j!} \left( \frac{\sigma^{(j)}(b)}{\rho^{(j)}(b)} \right)^2 \right] \geq 0$

In 3) and 3') terms with zero denominator are to be removed.

Proof:

As mentioned in the proof of lemma 4.8, $\|C(q)\|_b \leq 1$ and $\|C_b(q)\|_\infty$ both have the form

$$\sum_{j=0}^{k-1} |a_j - qb_j| \cdot |a_k - qb_k|,$$

where $\sum_{j=0}^{k} a_j = \sum_{j=0}^{k} b_j = 0$. (4.18)

Let $J$ be the index set $\{j|a_j \neq 0\}$. Then for $q = iy$ purely imaginary, we have that

$$|a_j - qb_j| = \begin{cases} \frac{b_j^2}{2a_j^2} \cdot y^2 + O(y^4), & \text{if } j \in J \\ |b_j| \cdot y, & \text{otherwise.} \end{cases}$$
From the proof of lemma 4.8 we know that 1) and 1') imply
\[ \sum_{j=0}^{k-1} b_j = a_k, \quad \text{and on this assumption we see that} \ (4.18) \]
for \( q = iy \) has the form
\[ y \cdot \sum_{j \neq J} b_j + \sum_{j=0}^{k-1} \frac{b_j^2}{a_j^2} \leq \frac{y^2}{2} \sum_{j \neq J} b_j^2 + O(y^4). \]

Hence, all conditions in the lemma are recognized as necessary. That these are sufficient follows from theorem 2.1 in [7], since they imply that (4.18) holds for all \( q = iy \) (including \( \infty \)):
\[ \sum_{j=0}^{k-1} |a_j - iy b_j| = \sum_{j \in J} |a_j|^\frac{1}{2} \cdot \left( |a_j| + y^2 \cdot \frac{b_j^2}{a_j^2} \right)^\frac{1}{2} \leq \left( \sum_{j \in J} |a_j|^\frac{1}{2} \right) \leq \left( \sum_{j \in J} \left( |a_j|^\frac{1}{2} + y^2 \cdot \frac{b_j^2}{a_j^2} \right)^\frac{1}{2} \right) \]
\[ = \left( \prod_{j=0}^{k-1} |a_j|^\frac{1}{2} \cdot \left( \sum_{j \in J} b_j^2 \right)^\frac{1}{2} \right) \leq |a_k - iy b_k|. \]

\[ \square. \]

5. Certain Contractive Variable-Formula Methods

It is well-known that the order of a one-leg formula
\[ \rho(E)y_n = \hbar^* f(\Theta t_n + (1-\Theta) t_{n+k-1}, \sigma(E)y_n), \]
is \( p = \min \{ l, m \} \), where \( l \) and \( m \) are the largest integers for which the following hold:
\[ \hbar^{-1} \rho(E)y(t_n) = y'(\Theta t_n + (1-\Theta) t_{n+k-1}) \text{ is exact for all polynomials } y \text{ of degree } \leq l \]
and
\[ \sigma(E)y(t_n) = y(\Theta t_n + (1-\Theta) t_{n+k-1}) \text{ is exact for all polynomials } y \text{ of degree } \leq m-1. \]
It follows that a \( k_n \)-step one-leg formula is of order \( k_n \) iff
\[
\bar{h}^{-1} \rho(E) y_n = \sum_{j=k-k_n}^{k} \frac{\phi_j'(\theta t_n + k(1-\theta) t_n + k-1)}{\phi_j(t_n + j)} \cdot y_{n+j}
\]
and
\[
\sigma(E) y_n = \sum_{j=k-k_n}^{k} \frac{\phi_j(\theta t_n + k(1-\theta) t_n + k-1) + c}{\phi_j(t_n + j)} \cdot y_{n+j},
\]
where \( c \) is independent of \( j \) and \( \phi_j(t) \) denotes the polynomial
\[
\prod_{i=k-k_n}^{k} \frac{(t-t_n+i)}{(t-t_n+j)}. \]
We replace the parameter \( c \) by \( B \) satisfying \( B \cdot \beta_k + \beta_{k-1} = 0 \) and obtain the following fixed-step formulas:

\( k_n = 1: \)
\[
\rho(r)/r^{k-1} = r-1,
\]
\[
(1-B) \sigma(r)/r^{k-1} = r-B.
\]

\( k_n = 2: \)
\[
2 \cdot \rho(r)/r^{k-2} = (2\theta+1)r^2 - 4\theta r + (2\theta - 1),
\]
\[
(2-B) \sigma(r)/r^{k-2} = [((\theta+1)r^2 - (\theta - 1)] - B \cdot [(\theta+1)r-\theta].
\]

\( k_n = 3: \)
\[
6 \cdot \rho(r)/r^{k-3} = (3\theta^2 + 6\theta + 2)r^3 - (9\theta^2 + 12\theta - 3)r^2 + (9\theta^2 + 6\theta - 6)r - (3\theta - 1),
\]
\[
2(3-B) \sigma(r)/r^{k-3} = [(\theta^2 + 3\theta + 2)r^3 - 3(\theta^2 + \theta - 2)r + 2(\theta - 1)]
\]
\[
-B \cdot [(\theta^2 + 3\theta + 2)r^2 - 2(\theta^2 + \theta)r + (\theta^2 + \theta)].
\]

\( k_n = 4: \)
\[
12 \cdot \rho(r)/r^{k-4} = (2\theta^3 + 9\theta^2 + 11\theta + 3)r^4 - (8\theta^3 + 30\theta^2 + 20\theta - 10)r^3 +
\]
\[
+ (12\theta^3 + 36\theta^2 + 6\theta - 18)r^2 - (8\theta^3 + 18\theta^2 - 4\theta - 9)r + (2\theta^3 + 3\theta^2 - \theta - 1),
\]
The reason for replacing the parameter $c$ by $B$ is two-fold. Firstly, setting $B$ to the same constant in several formulas, we may apply lemma 3.1 to the resulting OLM (1.3). Secondly, our analysis in [2] (pp.34-44) shows that in most cases (even for variable step-sizes) the $\omega$-contractive two- or three-step formulas with minimal $\omega$-value satisfy

$$(k-1)b \beta_k + \beta_{k-1} = 0.$$ 

Since the $\omega$-values for which a one-leg formula is $\omega$-contractive w.r.t some $(b,k)$-norm usually are smaller than the

$\theta$-values for which it becomes $\omega$-contractive, we shall here only consider OLM's with $B = (k-1)b$, $k \geq 3$ and refer the reader to [2] for a more thorough (but not that successful) treatment of one- and two-step formulas, $k = 2, 3$.

The following three figures illustrate the set of $(b, \theta)$ for which the fixed-step $k$'th order formulas with $B = (k-1)b$ are $0$-contractive in the $(b, k)$-norm. Moreover, the minimal $\theta$-value for which the formulas are $\omega$-contractive in the $(b, k)$-norm is displayed. Below each figure the non-differentiable behaviour of some of the curves is explained.

We observe that for $k = 3$, no 3-step 3'rd order formula is $A_\theta$-contractive in a $(b, k)$-norm with $b \leq 1/(k-1)$. Fortunately, this is not the case for $k = 4$ or 5 and we find a second- and a third-order formula being $A(a)$-contractive in the same $(b, k)$-norm and with a reasonably large $a$ (approx. $89^\circ$ and $45^\circ$, respectively), cf. fig. 5.1 and fig. 5.2.

To construct a 4-step OLM we also need a one-step formula and from lemma 4.9 it is easily seen that choosing $\beta_k = 1.1$ (say) will make the formula $A$-contractive in the $(0.2, 4)$-norm. Finally, we may add a four-step formula although they seem to have a relatively small region of contractivity in the $(0.2, 4)$-norm, cf. fig. 5.3.
\( k=3 \): \( \rho(b) = 2\alpha_3 b + \alpha_2 = 0 \) for \( \Theta \approx 0.3193995 \) and \( b \approx 0.2073370 \).
\[ k=3: \quad \rho'(b) = 3\alpha_3 b + \alpha_2 = 0 \quad \text{for} \quad \theta = 0.3797959 \quad \text{and} \quad b = 0.2020410. \]

\[ \rho''(b) = 0 \quad \text{for} \quad \theta = 1.1542916 \quad \text{and} \quad b = 0.2725283. \]

\[ k=4: \quad \rho'(b) = 3\alpha_4 b + \alpha_3 = 0 \quad \text{for} \quad \theta = 0.4648768 \quad \text{and} \quad b = 0.2139326. \]
$k=3$: $\rho(b) = 4\alpha_3 b + \alpha_2 = 0$ for $\theta \approx 0.4363371$ and $b \approx 0.1902780$.

$k=4$: $\rho^{(2)}(b) = 4\alpha_4 b + \alpha_3 = 0$ for $\theta \approx 0.4489212$ and $b \approx 0.1446741$. 
Fig. 5.1. The region of absolute stability (dashed line) and the region of contractivity in the $(0,2,4)$-norm of the second-order 2-step one-leg formula with $\theta=0.54$, $B=0.6$.

Fig. 5.2. The region of absolute stability (dashed line) and the region of contractivity in the $(0,2,4)$-norm of the third-order 3-step one-leg formula with $\theta=\sqrt{5.53-1.5}$, $B=0.6$. 
Fig. 5.3. The region of absolute stability (dashed line) and the region of contractivity in the \((0.2,4)\)-norm of the fourth-order 4-step one-leg formula with \(\rho=B=0.6\).

Remark

The \((b,2)\)-norm has turned out to be an appropriate measure of strong 0-contractivity of 2-step OLM's ([10]) and we hope to publish this result soon.

Acknowledgement

The thesis was written under the supervision of Professor G. Dahlquist at the Royal Institute of Technology, Stockholm, and I am very grateful to him as well as to the Danish Natural Science Research Council, which supported me financially during most of my stay at the institute.
References


