stichting
mathematisch
centrum

MC

DEPARTMENT OF NUMERICAL MATHEMATICS NW 30/76 JUNE

J.G. VERWER

MULTIPOINT MULTISTEP RUNGE-KUTTA METHODS I: ON A CLASS OF TWO-STEP METHODS FOR PARABOLIC EQUATIONS

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS(MOS) subject classification scheme (1970): 65L05, M10, M20

Multipoint multistep Runge-Kutta methods I: On a class of two-step methods for parabolic equations

by

J.G. Verwer

ABSTRACT

Multipoint multistep Runge-Kutta methods are discussed for the numerical solution of initial value problems for systems of ordinary differential equations $y' = f(y)$. These systems are supposed to originate from parabolic partial differential equations by applying the method of lines.
In this report the discussion is concentrated on a class of multipoint two-step methods. The main objective is to develop stabilized formulas of first and second order. Numerical examples are discussed.

KEY WORDS & PHRASES: *Numerical analysis, multipoint multistep Runge-Kutta methods, parabolic partial differential equations, method of lines.*

CONTENTS

1. INTRODUCTION

In this paper we discuss *multipoint multistep Runge-Kutta methods* for the numerical solution of initial value problems for systems of ordinary differential equations of the type

$$y' = f(y). \tag{1.1}$$

The systems we have in mind are supposed to originate from *parabolic partial differential equations* by applying the *method of lines*. In general, these systems are very large. As a consequence, an efficient scheme for (1.1) should possess limited storage requirements.

Multipoint multistep Runge-Kutta methods belong to the class of the so called *hybrid methods*. This class of methods is very wide. Generally speaking, an hybrid method shares certain linear multistep characteristics with the Runge-Kutta property of utilizing data at non-step points. When compared with linear multistep or Runge-Kutta methods, the hybrid methods did not yet have very much attention in literature. The class of methods we consider is similar tot that of GEAR [4]. This class is also discussed by WATT [13] and, more recently, in VAN DER HOUWEN [11], whose work inspired the author to the present investigation. Moreover, it belongs to a still wider class of methods originally dicussed by BUTCHER [2].

Until further notice the discussion is mainly concentrated on a class of two-step methods which need at most five arrays of storage in a computer implementation. We have only considered first and second order formulas. For partial differential equations higher order methods are usually not required.

A still more important aspect in integrating partial differential equations is the *stability* of the difference scheme. An efficient integration method for (1.1) should possess a stability region which contains a very long strip along the negative axis. Our main objective is to develop integration formulas which possess such a stability region.
A great deal of this paper is devoted to the stability analysis of the class of two-step methods mentioned above. This paper presents only partial results on this subject. In the near future we intend to investigate also

multipoint k-step methods for k > 2.

The class of two-step methods we consider contains a particular two-step scheme already discussed by VAN DER HOUWEN [10]. It also contains a class of stabilized Runge-Kutta methods which is also given by VAN DER HOUWEN [9].

In the last section of this report a comparison is made between a stabilized one-step and a stabilized two-step Runge-Kutta method by applying these methods to a non-linear diffusion problem. The numerical calculations have been carried out on a CYBER 73-28 computer using 14 significant digits.

2. THE MULTIPOINT MULTISTEP RUNGE-KUTTA METHOD

We define an m-point k-step Runge-Kutta formula to be a formula of the type

$$
\begin{aligned}
y_{n+1}^{(0)} &= y_n, \\
y_{n+1}^{(j)} &= \sum_{\ell=1}^{k} a_{j,\ell} y_{n+1-\ell} + \\
&\qquad h \sum_{\ell=2}^{k} b_{j,\ell} f(y_{n+1-\ell}) + h \sum_{\ell=0}^{j-1} \lambda_{j,\ell} f(y_{n+1}^{(\ell)}), \qquad j = 1,\ldots,m, \\
y_{n+1} &= y_{n+1}^{(m)}.
\end{aligned}
\tag{2.1}
$$

The points x_j, $j = n+1,\ldots, n+1-k$, denote the reference points of the k-step formula and h denotes the steplength, i.e. $x_{n+1} = x_n + h$, $n = 0,1,\ldots$. Unless otherwise stated, h is supposed to be constant. The vector y_n always represents a numerical approximation to the analytical solution y(x) at $x = x_n$.

The method of the above type belongs to the class of the so-called hybrid methods. This class of methods is very wide. It contains Runge-Kutta methods, as well as linear multistep methods.

The class of methods we consider is similar to that of GEAR [4]. This class is also discussed by WATT [13], and more recently, by VAN DER HOUWEN [11]. Moreover, (2.1) belongs to a still wider class of methods discussed

by BUTCHER [2] (see also STETTER [8]).
Butcher's class contains implicit methods, where as we consider only *explicit* ones. Following VAN DER HOUWEN [11] we shall represent formula (2.1) by the parameter matrix

$$(2.2)\qquad \left(\begin{array}{ccc|ccc|ccc} a_{1,1} & \cdots & a_{1,k} & b_{1,2} & \cdots & b_{1,k} & \lambda_{1,0} & & \\ a_{2,1} & \cdots & a_{2,k} & b_{2,2} & \cdots & b_{2,k} & \lambda_{2,0} & \lambda_{2,1} & \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots \\ a_{m,1} & \cdots & a_{m,k} & b_{m,2} & \cdots & b_{m,k} & \lambda_{m,0} & \lambda_{m,1} & \cdots \lambda_{m,m-1} \end{array}\right).$$

As already noticed, multistep Runge-Kutta methods are not so familiar as classical multistep or Runge-Kutta methods. Therefore, we shall briefly discuss the basic concepts of convergence, consistency and zero-stability for method (2.1). For a more thoroughly theoretical analysis of method (2.1) the interested reader is referred to the paper of WATT [13] (see also BUTCHER [2] and STETTER [8]).

It is convenient to associate multistep method (2.1) with a non-linear difference operator

$$(2.3)\qquad y(x_{n+1}) - E[y(x_n),\ldots, y(x_{n+1-k})],$$

where $y(x)$ now denotes a vector function of sufficient differentiability. It is also convenient to define the polynomials

$$(2.4)\qquad \rho_j(\zeta) = \zeta^k - \sum_{\ell=1}^{k} a_{j,\ell}\zeta^{k-\ell}, \quad j = 1,\ldots,m.$$

Now the usual definitions of convergence, consistency and zero-stability apply (cf. HENRICI [5,6]):

DEFINITION 2.1. The method is said to be *convergent* if, for every problem $y' = f(y)$, $y(a) = s$, $x \in [a,b]$, f an L-function, we have that

$$\lim_{\substack{h \to 0 \\ nh = x - a}} y_n = y(x_n)$$

holds for all $x \in [a,b]$, and for all solutions $\{y_n\}$ of the method satisfying starting conditions $y_\ell \to s$ as $h \to 0$, $\ell = 0,\ldots,k-1$.

DEFINITION 2.2. The method is said to be *consistent* of order p, at $x = x_n$, if p is the largest integer such that

$$y(x_{n+1}) - E[y(x_n),\ldots,y(x_{n+1-k})] = O(h^{p+1}), \quad h \to 0, \tag{2.5}$$

where y(x) is a sufficiently differentiable function. The method is said to be consistent if $p \geq 1$.

DEFINITION 2.3. The method is said to be *zero-stable* if no root of the polynomial ρ_m has modulus greater than one, and if every root with modulus one is simple.

Before we proceed with the consistency conditions we first make an assumption about the parameters $a_{j,\ell}$. From now on it is always assumed that

$$\rho_j(1) = 0, \quad j = 1,\ldots,m. \tag{2.6}$$

As a consequence of (2.6) we have that $y_{n+1}^{(j)}$, $j = 1,\ldots,m$, is always an approximation of order $p \geq 0$. Although relations (2.6) are not strictly necessary, they are usually made.

Expanding $y(x_{n+1-\ell})$ about x_n in (2.3) yields

$$\begin{aligned} y(x_{n+1}) - E[y(x_n),\ldots, y(x_{n+1-k})] &= (1 - \sum_{\ell=1}^{k} a_{m,\ell})y(x_n) + \\ h[(1 + \sum_{\ell=1}^{k} (\ell-1)a_{m,\ell})y'(x_n) &- (\sum_{\ell=2}^{k} b_{m,\ell} + \sum_{\ell=0}^{m-1} \lambda_{m,\ell})\, f(y(x_n))] + \\ O(h^2), \quad h \to 0. & \end{aligned} \tag{2.7}$$

LEMMA 2.1. *The method is consistent, if*

$$\rho_m(1) = 0,$$

$$1 + \sum_{\ell=1}^{k} (\ell-1)a_{m,\ell} = \sum_{\ell=2}^{k} b_{m,\ell} + \sum_{\ell=0}^{m-1} \lambda_{m,\ell}. \tag{2.8}$$

We also see from expansion (2.7) that a necessary condition for convergence is that

$$\sum_{\ell=2}^{k} b_{m,\ell} + \sum_{\ell=0}^{m-1} \lambda_{m,\ell} \neq 0. \tag{2.9}$$

If condition (2.9) is not satisfied the difference scheme may approximate a wrong differential equation. In fact, (2.9) is a necessary condition for a consistent scheme to be zero-stable. This follows easily from relation

$$1 + \sum_{\ell=1}^{k} (\ell-1)a_{m,\ell} = \rho_m'(1) + (1-k)\rho_m(1). \tag{2.10}$$

Just as in the case for linear multistep methods, the error constants in the truncation error (2.5) should be normalized with (2.9).

Next we give the convergence theorem which is similar to the convergence theorem for linear multistep methods.

THEOREM 2.1. *The method is convergent if and only if it is zero-stable and consistent.*

The proof closely parallels the proof of the convergence theorem for linear multistep methods (see HENRICI [5,6]). The interested reader is referred to the papers of GEAR [4] and WATT [13] (see also BUTCHER [2]).

3. A MULTIPOINT TWO-STEP RUNGE-KUTTA METHOD

We now confine the discussion to an m-point two-step scheme which can efficiently be used for the integration of large systems arising from semi-discretization of parabolic partial differential equations. As a consequence, we shall concentrate on the stability of the scheme when applied to the test-model

$$y' = \delta y, \tag{3.1}$$

where

$$\delta \in \mathbb{R}, \quad \delta < 0.$$

Moreover, the scheme should possess limited storage requirements in order to cope with the usually very large systems as a result of the semi-discretization.

3.1. The difference scheme

The formula reads

$$\begin{aligned}
y_{n+1}^{(0)} &= y_n, \\
y_{n+1}^{(1)} &= (1-b_1)y_n + b_1 y_{n-1} + c_1 hf(y_{n-1}) + \lambda_{1,0} hf(y_n), \\
y_{n+1}^{(j)} &= (1-b_j)y_n + b_j y_{n-1} + c_j hf(y_{n-1}) + \lambda_{j,0} hf(y_n) + \\
&\qquad \lambda_{j,j-1} hf(y_{n+1}^{(j-1)}), \quad j = 2,\ldots,m, \\
y_{n+1} &= y_{n+1}^{(m)}, \quad m \geq 2.
\end{aligned} \tag{3.2}$$

The parameter matrix of (3.2) is given by

$$
(3.3)\qquad
\left(\begin{array}{cc|c|cccc}
1-b_1 & b_1 & c_1 & \lambda_{1,0} & & & \\
1-b_2 & b_2 & c_2 & \lambda_{2,0} & \lambda_{2,1} & & \\
1-b_3 & b_3 & c_3 & \lambda_{3,0} & & \lambda_{3,2} & \\
\vdots & \vdots & \vdots & \vdots & & & \ddots \\
1-b_m & b_m & c_m & \lambda_{m,0} & 0 & & \lambda_{m,m-1}
\end{array}\right)
$$

By the choice

$$
(3.4)\qquad \lambda_{j,\ell} = 0, \quad j > \ell + 1, \quad \ell \neq 0,
$$

the scheme uses at most five arrays of storage in an actual computer implementation. If $b_i = 0$, $i = 1,\ldots,m-1$; $c_i = 0$, $i = 1,\ldots,m$; and $\lambda_{j,0} = 0$, $j = 2,\ldots,m$; we have the method discussed in VAN DER HOUWEN [10].

For scheme (3.1) the polynomial ρ_m (cf. (2.4)) is given by

$$
(3.5)\qquad \rho_m(\zeta) = \zeta^2 - (1-b_m)\zeta - b_m,
$$

which has the roots $\zeta_1 = 1$ and $\zeta_2 = -b_m$. The second equation of (2.8) yields the consistency condition

$$
(3.6)\qquad 1 + b_m = c_m + \lambda_{m,0} + \lambda_{m,m-1}.
$$

Thus, according to theorem 2.1, we have:

THEOREM 3.1. *Method* (3.2) *is convergent, if and only if* $-1 < b_m \leq 1$ *and* $1 + b_m = c_m + \lambda_{m,0} + \lambda_{m,m-1}$.

Observe, however that when b_m is very close to -1, $c_m + \lambda_{m,0} + \lambda_{m,m-1}$ is very close to zero. That means that the convergence condition (2.7) is almost violated. In practice such a situation must be avoided in order to get accurate results. For an extensive discussion about the behaviour of

the global discretization error of multistep Runge-Kutta methods, the interested reader is referred to WATT [13].

3.2. Consistency conditions

Before analyzing the real stability properties of the method, we shall give consistency conditions for orders p = 1,2 and 3. We intend to develop first and second order methods. The conditions for p = 3 may be used to obtain some information about the error constants occurring in the principal local truncation error. The conditions are listed in table 3.1, and are obtained in the usual way by means of Taylor expansions about x_n.

TABLE 3.1. Consistency conditions for scheme 3.2.

p = 1	$-b_m + c_m + \lambda_{m,0} + \lambda_{m,m-1} = 1$;
p = 2	$\frac{1}{2}b_m - c_m + \lambda_{m,m-1}(-b_{m-1} + c_{m-1} + \lambda_{m-1,m-2} + \lambda_{m-1,0}) = \frac{1}{2}$;
p = 3	$-\frac{1}{6}b_m + \frac{1}{2}c_m + \frac{1}{2}\lambda_{m,m-1}(-b_{m-1} + c_{m-1} + \lambda_{m-1,m-2} + \lambda_{m-1,0})^2 = \frac{1}{6}$,
	$-\frac{1}{6}b_m + \frac{1}{2}c_m + \lambda_{m,m-1}(\frac{1}{2}b_{m-1} - c_{m-1} +$
	$\lambda_{m-1,m-2}(-b_{m-2} + c_{m-2} + \lambda_{m-2,m-3} + \lambda_{m-2,0})) = \frac{1}{6}$;

3.3. Stability properties

Let us apply method (3.2) to the linear test-model (3.1). This yields the recursion

$$y_{n+1} = S(z)y_n + P(z)y_{n-1}, \quad n = 1,2,\ldots, \tag{3.7}$$

where $z = h\delta$, and where S and P are polynomials of degree m. Denote

$$S(z) = \sum_{i=0}^{m} s_i z^i \quad \text{and} \quad P(z) = \sum_{i=0}^{m} p_i z^i. \tag{3.8}$$

Then the coefficients s_i and p_i are defined by

$$
\begin{aligned}
s_0 &= 1 - b_m, \\
s_1 &= \lambda_{m,0} + \lambda_{m,m-1}(1-b_{m-1}), \\
s_i &= \prod_{j=m-i+2}^{m} \lambda_{j,j-1}(\lambda_{m-i+1,0} + \lambda_{m-i+1,m-i}(1-b_{m-i})), \quad i = 2,\ldots,m-1, \\
s_m &= \prod_{j=1}^{m} \lambda_{j,j-1},
\end{aligned} \tag{3.9}
$$

and

$$
\begin{aligned}
p_0 &= b_m, \\
p_1 &= \lambda_{m,m-1} b_{m-1} + c_m, \\
p_i &= \Big(\prod_{j=m-i+1}^{m} \lambda_{j,j-1}\Big) b_{m-i} + \Big(\prod_{j=m-i+2}^{m} \lambda_{j,j-1}\Big) c_{m-i+1}, \quad i = 2,\ldots,m-1, \\
p_m &= \Big(\prod_{j=2}^{m} \lambda_{j,j-1}\Big) c_1.
\end{aligned} \tag{3.10}
$$

Among other things we are interested in the *absolute stability* properties of our method. Let α_i, $i = 1,2$, be a root of the characteristic equation

$$\alpha^2 - S(z)\alpha - P(z) = 0 \tag{3.11}$$

of recursion (3.7). We shall use the following definition:

DEFINITION 3.1. Method (3.2) is said to be absolutely stable for a given $z = h\delta$, if $|\alpha_i| \le 1$ and if in case of $|\alpha_i| = 1$ α_i is simple.

By applying the Routh-Hurwitz criterion to (3.11), it is easily seen that method (3.12) is absolutely stable, for a given $z < 0$, if and only if

$$
\begin{aligned}
&|S(z)| \le 1 - P(z), \\
&P(z) \ge -1, \\
&|\alpha_i| = 1 \Rightarrow \alpha_i \text{ is simple.}
\end{aligned} \tag{3.12}
$$

As already noticed, we want to construct stabilized schemes. Thus the problem we are faced with is:

PROBLEM 3.1. Determine the coefficients s_i and p_i, which are supposed to be compatible with an imposed order of consistency, in such a way that (3.12) is valid for

$$-\beta \le z < 0, \quad \beta \text{ maximal.}$$

According to definition 3.1, β is called *the real boundary of absolute stability*, while the interval $[-\beta,0)$ is called *the real interval of absolute stability*.

An alternative to the requirement of absolute stability is the requirement that for a given $z < 0$ the roots of (3.11) are within or on the circle with radius $\rho(z)$, $0 < \rho(z) < 1$. This is the case, if and only if

$$\begin{aligned} &|S(z)| \le \rho(z) - \rho^{-1}(z)\, P(z), \\ &P(z) \ge -\rho^2(z). \end{aligned} \tag{3.13}$$

Thus we are led to

PROBLEM 3.2. Let $\rho : (-\infty,0] \to (0,1]$, $\rho(0) = 1$ be given, Determine the coefficients s_i and p_i, which are supposed to be compatible with an imposed order of consistency, in such a way that (3.13) is valid for

$$-\beta_\rho \le z \le 0, \quad \beta_\rho \text{ maximal.}$$

The function $\rho(z)$ shall be called a *damping function*. It can be used in order to obtain a stronger decay for the higher harmonics, which are almost always negligible in the true solution.

The function $\rho(z)$ may also be considered as an aid to enlarge

$$\text{minimum } \mathrm{Im}(z), \quad -\beta_\rho \le \mathrm{Re}(z) \le 0,$$

where z belongs to the region of absolute stability. This is of importance for problems where the eigenvalues of the Jacobian matrix of (1.1) are not purely real.

In the present report we confine the discussion to problem 3.1, i.e. we only discuss absolute stability. Problem 3.2 is subject of further investigations (results will be published in the near future).

The consistency conditions for orders p = 1 and 2 can be expressed in terms of the coefficients s_i and p_i. From relations (3.9)-(3.10) and table 3.1 we have:

TABLE 3.2. Consistency conditions in terms of s_i and p_i

p = 1	$s_0 = 1 - p_0$,
	$s_1 = 1 + p_0 - p_1$;
p = 2	$s_2 = \frac{1}{2} - \frac{1}{2}p_0 + p_1 - p_2$;

REMARK 3.1. These conditions can also be obtained by substituting the second order Padé-approximation

$$1 + z + \tfrac{1}{2}z^2$$

to the exponential into equation (3.11).

Before proceeding with problem 3.1, we first remark that no optimal solutions to this problem are obtained. The author intends to discuss optimal solutions to problem 3.1 in a following report, where also problem 3.2 shall be discussed. Here we shall give approximate solutions to problem 3.1. However, these solutions are very satisfactory. We still have to observe that with respect to stability the parameter p_0 may vary between -1 and +1. However, as $p_0 = b_m$, the convergence condition (2.7) requires that p_0 is not allowed to be close to -1 (see Theorem 3.1). Therefore, in the solutions discussed the parameter p_0 is fixed beforehand.

3.4. Two approximate solutions

Our starting point is the following theorem:

THEOREM 3.2. *Let* $Q(z)$ *be a given boundary curve for the inequality* $|V(z)| \leq Q(z)$, *where* V *denotes a polynomial of maximum degree* $m \geq 2$ and

$z \le 0$. *Of all polynomials* V, *where* $V(0) = Q(0)$, $V'(0)$ *prescribed,* V *non-constant, the polynomial* V *which has* m - 1 *alternating points of tangency to the curves* $\pm Q(z)$, $z \le 0$, *maximizes (if it exists) the negative interval on which the inequality is satisfied.*

The proof of this theorem follows the lines along which the minimax property of the Chebyshev polynomials is proved. Observe however, that this theorem does not guarantee the existence of a polynomial V with m - 1 alternating points of tangency. Nevertheless, it can be useful for the construction of an optimal polynomial. The property characterized in theorem 3.2 is known as the "*equal ripple*" property.

Thus the idea is to prescribe the polynomial P, and after that to apply the "equal ripple" property in order to find the accompanying optimal S, provided such an S exists. We begin with the first order case:

The case p = 1. For p = 1 we give a solution which is similar to the optimal solution given by VAN DER HOUWEN [10] for his special scheme.
Let p_0 be given. According to table 3.2 we have

$$
\begin{aligned}
P(z) &= \sum_{i=0}^{m} p_i z^i, \\
S(z) &= 1 - p_0 + (1+p_0-p_1)z + \sum_{i=2}^{m} s_i z^i,
\end{aligned}
\tag{3.14}
$$

where $p_1,\ldots,p_m$ and $s_2,\ldots,s_m$ are free parameters. Now set $p_i = 0$, $i = 1,\ldots,m$, i.e. $P(z) \equiv p_0$, and write

$$
\begin{aligned}
S(z) &= (1-p_0)\bar{S}(w), \\
\bar{S}(w) &= 1 + w + \sum_{i=2}^{m} \frac{(1-p_0)^{i-1}}{(1+p_0)^i} s_i w^i, \qquad w = \frac{1+p_0}{1-p_0} z .
\end{aligned}
\tag{3.15}
$$

According to theorem 3.2 the accompanying optimal $\bar{S}$ is the polynomial which has m - 1 alternating points of tangency to the curves ± 1, $w \le 0$. This polynomial is well-known;

one has

$$\bar{S}(w) = T_m\left(1+\frac{w}{m^2}\right), \tag{3.16}$$

where $T_m(w)$ denotes the Chebyshev polynomial of degree m in w, i.e.

$$T_m(w) = \cos\,(m \arccos w). \tag{3.17}$$

For $\bar{S}$, given by (3.16), we have

$$|\bar{S}(w)| \le 1, \quad -2m^2 \le w \le 0. \tag{3.18}$$

Thus the real boundary of absolute stability β is given by

$$\beta = \frac{2(1-p_0)m^2}{1+p_0}, \tag{3.19}$$

provided that $p_0 > -1$. Observe that $\beta \to \infty$ as $p_0 \to -1$. However, the normalized error constants also tend to infinity as $p_0 \to -1$. We have a similar situation as with the well-known scheme of Du Fort and Frankel (see RICHTMYER and MORTON [7]). For $p_0 = 0$, $\beta = 2m^2$, i.e. the stability boundary of the stabilized Euler method (see VAN DER HOUWEN [9]).

According to ABRAMOWITZ and STEGUN [1] (formulas 15.1.1 and 15.4.3), the polynomial T_m can be written as

$$T_m(w) = \sum_{i=0}^{m} \frac{(-m)_i (m)_i}{(\frac{1}{2})_i\, i!} \left(\frac{1-w}{2}\right)^i, \tag{3.20}$$

where, for $a \in \mathbb{R}$, $(a)_i$ is defined as

$$(a)_0 = 1, \quad (a)_i = a(a+1)\ldots(a+i-1), \quad i \ge 1. \tag{3.21}$$

By means of (3.20) we then find

$$T_m\left(1+\frac{w}{m^2}\right) = \sum_{\ell=0}^{m} c_{\ell,m}\, w^{\ell}, \tag{3.22}$$

where $c_{\ell,m}$ is defined by the recursion

$$
\begin{aligned}
& c_{o,m} = 1, \\
& c_{\ell,m} = \frac{1 - (\ell-1)^2/m^2}{\ell(2\ell-1)}\, c_{\ell-1,m}, \qquad \ell = 1,\ldots,m.
\end{aligned}
\tag{3.23}
$$

Thus, using (3.22), the coefficients s_i and p_i are given by:

$$
\begin{aligned}
& -1 < p_0 < 1, \\
& p_i = 0, \quad i = 1,\ldots,m; \\
& s_0 = 1 - p_0, \\
& s_j = \frac{(1+p_0)^j c_{j,m}}{(1-p_0)^{j-1}}, \qquad j = 1,\ldots,m.
\end{aligned}
\tag{3.24}
$$

The case p = 2 Let p_0 be given. According to table 3.2 we have

$$
\begin{aligned}
& P(z) = \sum_{i=0}^{m} p_i z^i, \\
& S(z) = (1-p_0) + (1+p_0-p_1)z + (\tfrac{1}{2}-\tfrac{1}{2}p_0+p_1-p_2)z^2 + \sum_{i=3}^{m} s_i z^i.
\end{aligned}
\tag{3.25}
$$

First we prove the following result:

THEOREM 3.3. *Let* $p_0 = -1$. *Let* P *and* S *be defined by* (3.25) *and consider the inequalities* $|S(z)| \le 1 - P(z)$ *and* $|P(z)| \le 1$. *The interval* $[-\beta^*,0]$, *on which these inequalities are satisfied, is maximized by*

$$
P(z) = -T_m(1+\frac{z}{m^2}), \qquad S(z) = 1 - P(z), \qquad \beta^* = 2m^2,
$$

where $T_m(z) = \cos(m \arccos z)$, *i.e. the Chebyshev polynomial of degree* m *in* z.

PROOF: Because $p_0 = -1$ the polynomials S and P are given by

$$S(z) = 2 - p_1 z + (1+p_1-p_2)z^2 + \sum_{i=3}^{m} s_i z^i,$$

$$P(z) = -1 + \sum_{i=1}^{m} p_i z^i.$$

Thus a necessary condition in order to have $S(z) \leq 1 - P(z)$ for $-\varepsilon \leq z \leq 0$, $\varepsilon > 0$ and arbitrary, is $1 + p_1 - p_2 \leq -p_2$ or $p_1 \leq -1$.
The optimal P which maximizes the negative interval, say $[-\beta^{**},0]$, on which $|P(z)| \leq 1$ is well-known; one has

$$P(z) = -T_m(1-p_1 z/m^2), \qquad \beta^{**} = -2m^2/p_1,$$

where $T_m(z) = \cos(m \arccos z)$. As a consequence the optimal choice for p_1 is $p_1 = -1$. Further if $p_0 = p_1 = -1$ we can set $S(z) = 1 - P(z)$ by choosing $s_i = p_i$, $i = 3,\ldots,m$. This establishes the proof of the theorem. □

Conjecture: Suppose $p_0 > -1$ and fixed. Then the conjecture exists that for the optimal solution of problem 3.1 holds

$$\beta < 2m^2,$$

in the second order case. Until now we did not succeed to prove this conjecture.

For $p_0 > -1$ and p_0 fixed we shall construct an approximate solution in such a way that this solution tends to the solution of theorem 3.3 as $p_0 \to -1$. Define the polynomial

$$\bar{P}(w) = -p_0 T_m(1+\frac{w}{m^2}), \tag{3.26}$$

and set

$$w = \frac{p_1 z}{p_0},$$

$$p_i = \frac{c_{i,m}\, p_1^i}{p_0^{i-1}}, \quad i = 2,\ldots,m. \tag{3.27}$$

Thus there holds $P(z) = \bar{P}(w)$, while p_1 is still a free parameter. We also write S as a function of w, i.e. $S(z) = \bar{S}(w)$, where

$$
\begin{aligned}
&\bar{S}(w) = \sum_{i=0}^{m} \bar{s}_i w^i, \\
&\bar{s}_0 = 1 - p_0, \\
&\bar{s}_1 = \frac{(1+p_0-p_1)p_0}{p_1}, \\
&\bar{s}_2 = \frac{(\frac{1}{2}-\frac{1}{2}p_0+p_1-c_{2,m}p_1^2/p_0)p_0^2}{p_1^2}, \\
&\bar{s}_i = \frac{s_i p_0^i}{p_1^i}, \quad i = 3,\ldots,m.
\end{aligned}
\tag{3.28}
$$

The parameters $\bar{s}_i$, $i = 3,\ldots,m$, are free and $\bar{s}_1$ and $\bar{s}_2$ both depend upon the free parameter p_1.

Because the polynomial $\bar{P}$ is fixed, we can now proceed with the "equal ripple" property. That means we try to construct a polynomial $\bar{S}$, which has $m - 1$ alternating points of tangency to the curves $\pm(1-\bar{P}(w))$, $w < 0$. If such a polynomial indeed exists we have, for $i = 1,\ldots,m - 1$,

$$
\begin{aligned}
&\bar{S}(w_i) = (-1)^i(1-\bar{P}(w_i)), \\
&\bar{S}'(w_i) = (-1)^{i-1}\bar{P}'(w_i),
\end{aligned}
\tag{3.29}
$$

where for each i, w_i represents the point where $\bar{S}(w)$ touches the boundary curves. Relations (3.29) constitute a non-linear system of $2m - 2$ equations for the $2m - 2$ unknowns: $p_1, \bar{s}_3,\ldots,\bar{s}_m, w_1,\ldots,w_{m-1}$.

As already observed theorem 3.2 does not guarantee the existence of an "equal ripple" polynomial. In fact, we have the situation that system (3.29) does not always has a solution. Let us illustrate this for $m = 2$. The two unknowns, to be solved from (3.29), are p_1 and w_1. Solving (3.29) yields a quadratic equation for p_1, i.e.

$$(3.30) \qquad (2+2p_0)p_1^2 + (4p_0^2-12p_0)p_1 + (-3p_0^3+10p_0^2-3p_0) = 0.$$

The discriminant of (3.30) can be written as

$$40p_0(p_0-1)(p_0-3)(p_0+0.2).$$

It now easily follows that no solution exists if $-0.2 < p_0 < 0$. The same situation arises for values of m greater than 2. This has been verified by numerical experimentation.

Because relations (3.29) are not sufficient for the "equal ripple" property, a polynomial $\bar{S}$ belonging to a solution of (3.29) does not necessarily satisfy this property. So each solution must be verified. If a solution found satisfies the property we have that

$$(3.31) \qquad \beta = \max\{\min\{z \mid |P(z)| \le 1\}, \min\{z \mid |S(z)| \le 1 - P(z)\}\}.$$

By means of a Newton-Raphson method system (3.29) has been solved numerically for m = 2,...,10, while

$$p_0 = -\frac{3}{4}.$$

The polynomials P and S of theorem 3.3 did serve as an initial guess. As a measure of safety, the first equation of (3.29) was replaced by

$$(3.22) \qquad \bar{S}(w_i) = (-1)^i(0.99-\bar{P}(w_i)), \quad i = 1,...,m-1.$$

The solutions found all satisfy the "equal ripple" property. The coefficients $-p_1$ and $s_3,...s_m$, corresponding to these solutions, are listed in table 3.3. The corresponding p_i, i = 2,...,m, can be determined from (3.27), while the corresponding s_i, i = 0,1,2, can be calculated from the consistency relations (see table 3.2).

TABLE 3.3. Coefficients $-p_1$; s_i, i = 3,...,m.

m	$-10^{13}p_1$	$10^{14}s_3$	$10^{16}s_4$	$10^{18}s_5$
2	8433976470221			
3	8373943414819	714642946011		
4	8353287170311	1010977435660	1726749099618	
5	8343487258568	1156801510216	2890156512230	2529810379359
6	8338338202996	1237615568887	3619850449730	4882090890394
7	8335088244243	1287488484636	4099170910850	6704819396726
8	8333109733929	1319746351067	4421028523838	8046191949864
9	8331630767474	1342367929599	4652101448364	9062951609280
10	8293222925118	1395517005412	5018542084218	10362223955442

m	$10^{20}s_6$	$10^{22}s_7$	$10^{24}s_8$	$10^{27}s_9$
6	2469972407288			
7	5442314391295	1736916306222		
8	8115614961054	4263796094047	910317207146	
9	10378688540331	6931995019678	2498621414458	3755585480498
10	13021686763735	10125630113776	4757383942238	12373496908462

m	$10^{30}s_{10}$
10	13676409585179

In table 3.4 we list the corresponding values of β, β/m and β/m^2 for $m = 2,\ldots,10$. From this table we see that $\beta/m^2 \sim 1.80$. By way of comparison we mention that for the second order two-step formulas given by VAN DER HOUWEN [10] there holds $\beta/m^2 \sim 1.16$. The second order one-step formulas, given by VAN DER HOUWEN [9], yield $\beta/m^2 \sim 0.81$.

TABLE 3.4.

m	β	β/m	β/m^2
2	7.3	3.65	1.82
3	16.2	5.40	1.80
4	29.0	7.25	1.81
5	45.2	9.04	1.80
6	65.0	10.83	1.80
7	88.2	12.60	1.80
8	115.4	14.42	1.80
9	144.9	16.10	1.78
10	181.1	18.11	1.81

In order to illustrate the behaviour of the polynomials S and P, the curves $\pm(1-P(-z))$, $0 \leq -z \leq -\beta$ and $S(-z)$, $0 \leq -z \leq -\beta$ are given for $m = 6$ (see fig. 3.1). In fig. 3.2 we have plotted the curve $\max_{i=1,2} |\alpha_i(-z)|$, $0 \leq -z \leq -\beta$ for $m = 6$.

3.5. Internal stability

Internal stability deals with the *propagation of round-off errors in a single integration step*. For methods of the Runge-Kutta type, which use a relatively *large number of stages* and which have relatively *large stability boundaries*, the amplification of round-off errors in a single step may be of a considerable magnitude. Therefore, for these methods it is necessary to analyze the internal stability behaviour.

In VAN DER HOUWEN [11], section 2.6.10, the internal stability is discussed for a class of one-step Runge-Kutta methods which is contained in class (3.2). He defines a so-called *internal stability function*, i.e. a function which approximately controls the propagation of round-off errors

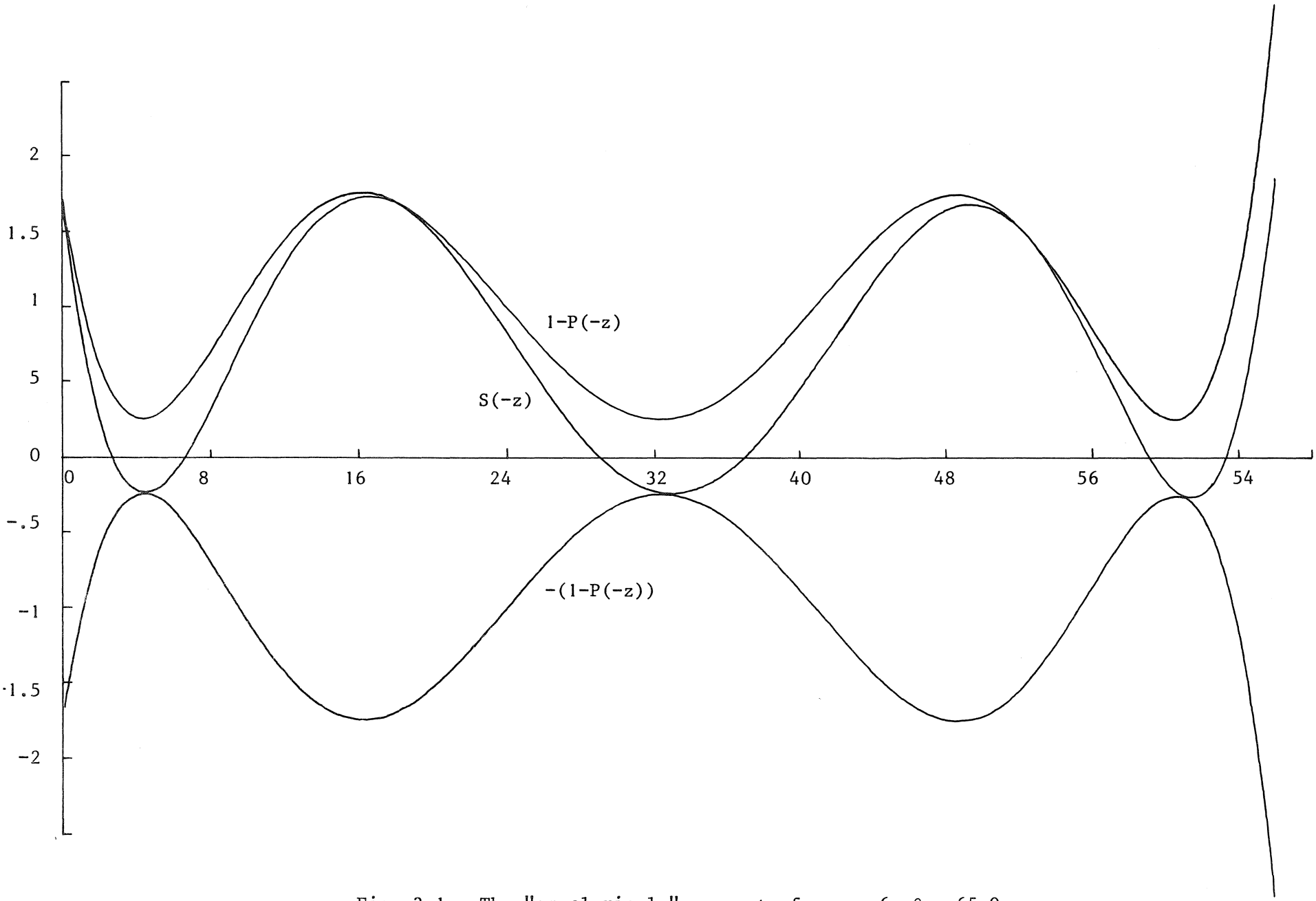

Fig. 3.1. The "equal ripple" property for m = 6, β = 65.0.

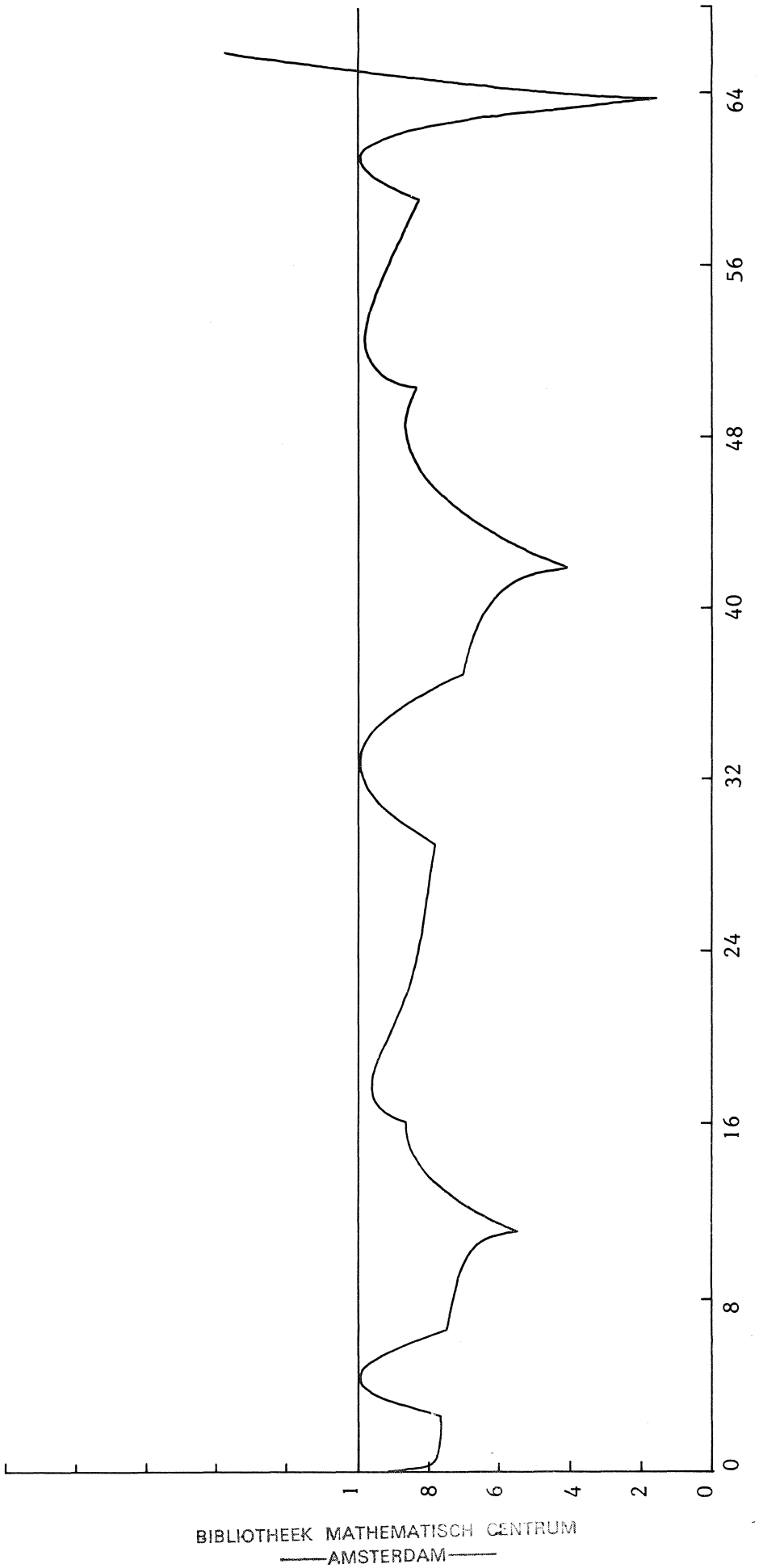

Fig. 3.2. $|\alpha_{\max}|$ for $m = 6$, $\beta = 65.0$.

in a single step. We shall also define such a function for method (3.2). It turns out that this function is equivalent to the function of Van der Houwen.

Let $\bar{y}_{n+1}^{(j)}$ denote the perturbed solution to $y_{n+1}^{(j)}$. Define

$$\varepsilon_{n+1}^{(j)} = \bar{y}_{n+1}^{(j)} - y_{n+1}^{(j)}, \quad j = 1,\ldots,m \; ; \; \varepsilon_{n+1} = \varepsilon_{n+1}^{(m)} . \tag{3.33}$$

Instead of (3.2) we now consider the recurrence

$$\begin{aligned}
\bar{y}_{n+1}^{(0)} &= y_n, \\
\bar{y}_{n+1}^{(1)} &= (1-b_1)y_n + b_1 y_{n-1} + c_1 hf(y_{n-1}) + \lambda_{1,0} hf(y_n) + \rho_{n+1}^{(1)}, \\
\bar{y}_{n+1}^{(j)} &= (1-b_j)y_n + b_j y_{n-1} + c_j hf(y_{n-1}) + \lambda_{j,0} hf(y_n) + \\
&\qquad \lambda_{j,j-1} hf(\bar{y}_{n+1}^{(j-1)}) + \rho_{n+1}^{(j)}, \quad j = 2,\ldots,m, \\
\bar{y}_{n+1} &= \bar{y}_{n+1}^{(m)}, \quad m \geq 2.
\end{aligned} \tag{3.34}$$

The errors $\varepsilon_{n+1}^{(j)}$ then satisfy the recurrence relation

$$\begin{aligned}
\varepsilon_{n+1}^{(1)} &= \rho_{n+1}^{(1)}, \\
\varepsilon_{n+1}^{(j)} &= \lambda_{j,j-1} h[f(y_{n+1}^{(j-1)} + \varepsilon_{n+1}^{(j-1)}) - f(y_{n+1}^{(j-1)})] + \rho_{n+1}^{(j)}, \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j = 2,\ldots,m, \\
\varepsilon_{n+1} &= \varepsilon_{n+1}^{(m)}.
\end{aligned} \tag{3.55}$$

By assuming that the Jacobian matrix, say $J(y)$, of $f(y)$ is slowly varying, we find the approximate relation

$$\varepsilon_{n+1}^{(j)} \simeq \lambda_{j,j-1} hJ(y_n) \, \varepsilon_{n+1}^{(j-1)} + \rho_{n+1}^{(j)}, \quad j = 2,\ldots,m.$$

We thus arrive at the estimate

$$\|\varepsilon_{n+1}\| \leq [1 + \sum_{k=1}^{m-1} \prod_{j=m+1-k}^{m} |\lambda_{j,j-1}| \, \|(hJ(y_n))^k\|] \max_{1\leq k\leq m} \|\rho_{n+1}^{(k)}\|. \tag{3.36}$$

Following Van der Houwen we now define the *internal stability function*

$$(3.37) \qquad Q(z) = 1 + \sum_{k=1}^{m-1} \prod_{j=m+1-k}^{m} |\lambda_{j,j-1}| \, |z|^k .$$

In case of *normal matrices* $J(y_n)$ we then have

$$(3.38) \qquad \|\varepsilon_{n+1}\| \leq Q(h\sigma(J(y_n))) \max_{1\leq k\leq m} \|\rho_{n+1}^{(k)}\| ,$$

where σ denotes the spectral radius. As a consequence, in actual computation the steplength h should at least satisfy the *internal stability condition*

$$(3.39) \qquad Q(h\sigma(J(y_n))) \leq \frac{\text{tolerance}}{\text{machine precision}} ,$$

where tolerance is understood to be the maximal allowable local truncation error. When the parameters of the scheme have positive signs, we know from practical experience that the internal stability behaviour can reasonably be controlled by condition (3.39). In case of opposite signs, however, this condition may be too optimistic because of a possible cancellation of digits. Therefore, we shall try to construct schemes with positive parameters, at least for $\lambda_{j,j-1}$, and in such a way that (3.39) is satisfied for relevant values of h.

4. A CLASS OF TWO-STEP METHODS OF SECOND ORDER

In this section we give a number of second order formulas belonging to class (3.2). We shall require that the principal local truncation error of these formulas can be represented as (see (2.5))

$$(4.1) \qquad Ch^3 \frac{d^3y(x)}{dx^3}\Bigg|_{x = x_n} , \qquad C \text{ constant} .$$

The reason for this representation is that in the near future we intend to develop methods which incorporate both automatic error and steplength control. For such methods it is very convenient when the local errors can be

approximated by expressions of type (4.1).

Using the tensor notation in the Taylor expansion of the local truncation error (2.5) we can write (compare HENRICI [5], p. 118)

$$y(x_{n+1}) - E[y(x_n), y(x_{n-1})] = C_1 h^3 f_j f_k^j f^k + C_2 h^3 f_{jk} f^j f^k + O(h^4),$$

as $h \to 0$, where the error constants C_1 and C_2 are given by (see table 3.1):

$$C_1 = \frac{1}{6} - [- \frac{1}{6} b_m + \frac{1}{2} c_m + \lambda_{m,m-1} (\frac{1}{2} b_{m-1} - c_{m-1} + \lambda_{m-1,m-2}$$

$$(-b_{m-2} + c_{m-2} + \lambda_{m-2,m-3} + \lambda_{m-2,0}))],$$

$$C_2 = \frac{1}{6} - [- \frac{1}{6} b_m + \frac{1}{2} c_m + \frac{1}{2} \lambda_{m,m-1} (-b_{m-1} + c_{m-1} + \lambda_{m-1,m-2} + \lambda_{m-1,0})^2].$$

The third derivative of y can be expressed as

$$y''' = f_j f_k^j f^k + f_{jk} f^j f^k.$$

Thus the principal local truncation error is of type (4.1) if we can satisfy relation

$$C_1 = C_2. \tag{4.2}$$

We observe that for linear equations the term $f_{jk} f^j f^k$ vanishes and that the scheme is completely determined by the coefficients s_i and p_i of the polynomials S and P. This means that $C = C_1$ and C depends completely on s_i and p_i.

It is convenient to express the parameters of the scheme into the coefficients s_i and p_i. From relations (3.9) - (3.10) it is clear that there exists more than one solution. Unfortunately, in case of negative p_i no solution exists for which all the parameters are positive. As, in our situation, the coefficients p_i are all negative (see section 3.4), we select a solution which reduces the computational effort. To that end we set

$$b_i = 0, \; i = 1,\ldots,m-2; \; \lambda_{i,0} = 0, \; i = 2,\ldots,m. \tag{4.3}$$

By using the relations of table 3.1 and relations (3.9) - (3.10) it now easily follows that (4.2) is satisfied, if and only if

$$c_m = \frac{(1+p_0)(p_1-2p_2+2p_3+2s_3) - \frac{1}{4}(1-p_0)^2}{2 + p_1 - 2p_2 + 2p_3 + 2s_3} \tag{4.4}$$

By performing some elementary calcuations, the remaining parameters can be solved from (3.9) - (3.10). Summarizing, we have:

$$\begin{aligned}
&b_i = 0, \quad i = 1,\ldots,m-2,\\
&b_{m-1} = \frac{p_1-c_m}{1+p_0-c_m},\\
&b_m = p_0 ;\\
&c_i = \frac{p_{m+1-i}}{s_{m-i}}, \quad i = 1,\ldots,m-2,\\
&c_{m-1} = \frac{p_2}{1+p_0-c_m},\\
&c_m = \frac{(1+p_0)(p_1-2p_2+2p_3+2s_3) - \frac{1}{4}(1-p_0)^2}{2 + p_1 - 2p_2 + 2p_3 + 2s_3} ;\\
&\lambda_{i,0} = 0, \quad i = 2,\ldots,m ;\\
&\lambda_{i,i-1} = \frac{s_{m+1-i}}{s_{m-i}}, \quad i = 1,\ldots,m-2,\\
&\lambda_{m-1,m-2} = \frac{s_2}{1+p_0-c_m},\\
&\lambda_{m,m-1} = 1 + p_0 - c_m.
\end{aligned} \tag{4.5}$$

In table 4.2 we give for m = 10 the parameter matrix (3.3) for this set of parameters (the coefficients s_i and p_i are taken from table 3.3). In the next section numerical results are given of the method generated by this matrix.

Finally, in table 4.1 we list the values (see (3.37))

(4.6) $\quad Q(\beta), \quad m = 2,\ldots,10,$

for parameters (4.5) (again the coefficients s_i and p_i are taken from table 3.3). Observe that the parameters $\lambda_{i,i-1}$ are all positive. By using relations (3.9) it easily follows that Q(z) can be written as

$$(4.7) \qquad Q(z) = 1 + (1+p_0-c_m)|z| + \sum_{k=2}^{m-1} s_k |z|^k.$$

m	Q(β)
2	$0.76_{10}1$
3	$0.61_{10}2$
4	$0.43_{10}3$
5	$0.27_{10}4$
6	$0.16_{10}5$
7	$0.97_{10}5$
8	$0.57_{10}6$
9	$0.32_{10}7$
10	$0.22_{10}8$

TABLE 4.1

1	0	$-0.8481243492344_{10}-3$	$0.11052986626461_{10}-2$
1	0	$-0.19949026507992_{10}-2$	$0.26009035761455_{10}-2$
1	0	$-0.36024229851479_{10}-2$	$0.46983584120506_{10}-2$
1	0	$-0.59607171394383_{10}-2$	$0.77759742631620_{10}-2$
1	0	$-0.96319035551034_{10}-2$	$0.12566498098988_{10}-2$
1	0	$-0.15827347046527_{10}-1$	$0.20647876976121_{10}-1$
1	0	$-0.27575393221043_{10}-1$	$0.35961884124349_{10}-1$
1	0	$-0.54358937105922_{10}-1$	$0.70842630567026_{10}-1$
1.26196439161229	-0.26196439161229	-0.17691526753511	0.23032252201367
1.75	-0.75	-0.60527159061348	0.85527159061345

TABLE 4.2. Parameter matrix corresponding to parameters (4.5), m = 10.

5. NUMERICAL EXAMPLES

The integration formula defined by the parameter matrix given in table 4.2, will be applied to two parabolic equations. We shall concentrate on the experimental verification of the theoretically derived stability condition. To begin with we have chosen a non-linear diffusion problem which proceeds from FEHLBERG [3]:

$$
\begin{aligned}
&\frac{\partial u}{\partial t} = \frac{\exp(2-u)}{4(2+x^2)} \frac{\partial^2 u}{\partial x^2}, \qquad 0 \le x \le 1, \quad t \ge 0,\\
&u(x,0) = 2[1 - \ln(2-x^2)], \qquad 0 \le x \le 1,\\
&\frac{\partial u}{\partial x} = 0, \qquad x = 0, \quad t \ge 0,\\
&u(1,t) = 2 + \ln(1+t), \qquad t \ge 0.
\end{aligned}
\tag{5.1}
$$

The analytical solution of this problem is given by

$$
u(x,t) = 2 + \ln(1+t) - 2\ln(2-x^2). \tag{5.2}
$$

By using the method of lines, i.e. by discretizing with respect to x, we can replace (5.1) by an initial value problem for a system of ordinary differential equations of type (1.1). We divide the x-interval into N equal intervals of length $\Delta x = 1/N$. Let $u_j(t)$ denote an approximation to the exact solution $u(x,t)$ at $x_j = j\Delta x$, $j = 0,\ldots,N-1$. At the internal grid points x_j, $j = 2,\ldots,N-3$, we approximate the partial derivatives $\partial^2 u/\partial x^2$ by means of the 5-point central difference formula, i.e.

$$
\frac{\partial^2 u_j}{\partial x^2} \simeq \frac{1}{12(\Delta x)^2} (-u_{j-2} + 16u_{j-1} - 30u_j + 16u_{j+1} - u_{j+2}), \qquad j = 2,\ldots,N-3. \tag{5.3}
$$

At the grid point x_{N-2} we can also use the 5-point central difference formula. Let $u_N(t) \equiv t$. Then we have

$$(5.4)\qquad \frac{\partial^2 u_{N-2}}{\partial x^2} \simeq \frac{1}{12(\Delta x)^2}\,(-u_{N-4} + 16u_{N-3} - 30u_{N-2} + 16u_{N-1} - 2 - \ln(1+u_N)).$$

At the grid point x_{N-1} we apply the 6-point difference approximation

$$(5.5)\qquad \frac{\partial^2 u_{N-1}}{\partial x^2} \simeq \frac{1}{12(\Delta x)^2}\,(u_{N-5} - 6u_{N-4} + 14u_{N-3} - 4u_{N-2} - 15u_{N-1} + 10(2+\ln(1+u_N))).$$

Because of the symmetry at the left boundary, the partial derivatives at the grid points $x_j, j = 0,1$, can be approximated by

$$(5.6)\qquad \frac{\partial^2 u_0}{\partial x^2} \simeq \frac{1}{12(\Delta x)^2}\,(-30u_0 + 32u_1 - 2u_2),$$

$$\frac{\partial^2 u_1}{\partial x^2} \simeq \frac{1}{12(\Delta x)^2}\,(16u_0 - 31u_1 + 16u_2 - u_3).$$

The approximations (5.3) - (5.6) are all third order exact.

By substituting (5.3) - (5.6) into (5.1), we arrive at the following initial value problem:

$$(5.7)\qquad \frac{du_0}{dt} = \frac{d_0}{12(\Delta x)^2}\,(-30u_0 + 32u_1 - 2u_2),$$

$$\frac{du_1}{dt} = \frac{d_1}{12(\Delta x)^2}\,(16u_0 - 31u_1 + 16u_2 - u_3),$$

$$\frac{du_j}{dt} = \frac{d_j}{12(\Delta x)^2}\,(-u_{j-2} + 16u_{j-1} - 30u_j + 16u_{j+1} - u_{j+2}), \quad j = 2,\ldots,N-3,$$

$$\frac{du_{N-2}}{dt} = \frac{d_{N-2}}{12(\Delta x)^2}\,(-u_{N-4} + 16u_{N-3} - 30u_{N-2} + 16u_{N-1} - 2 - \ln(1+u_N)),$$

(5.7)

$$\frac{du_{N-1}}{dt} = \frac{d_{N-1}}{12(\Delta x)^2}\,(u_{N-5} - 6u_{N-4} + 14u_{N-3} - 4u_{N-2} - 15u_{N-1} + 10(2 + \ln(1+u_N))),$$

$$\frac{du_N}{dt} = 1\ ;$$

$$u_j(0) = 2(1 - \ln(2-j^2\Delta^2 x)), \quad j = 0,\ldots,N-1,$$

$$u_N(0) = 0,$$

where d_j, $j = 0,\ldots,N-1$, is given by

$$d_j = \frac{\exp(2-u_j)}{4(2+j^2(\Delta x)^2)}. \tag{5.8}$$

The Jacobian matrix, say J, of (5.7) can be expressed as

$$J = \frac{1}{12(\Delta x)^2}\,DM, \tag{5.9}$$

$$D = \begin{pmatrix} d_0 & & & 0 \\ & \ddots & & \\ & & d_{N-1} & \\ 0 & & & 1 \end{pmatrix}, \tag{5.10}$$

(5.11)

$$M = \begin{pmatrix} a_0 & 32 & -2 & & & & & \\ 16 & a_1 & 16 & -1 & & & & \\ -1 & 16 & a_2 & 16 & -1 & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & -1 & 16 & a_{N-3} & 16 & -1 & \\ & & & -1 & 16 & a_{N-2} & 16 & b_{N-2} \\ & & 1 & -6 & 14 & -4 & a_{N-1} & b_{N-1} \\ & & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix};$$

The entries a_j are defined by

$$\begin{aligned} a_0 &= -(30-30u_0+32u_1-2u_2), \\ a_1 &= -(31+16u_0-31u_1+16u_2-u_3), \\ a_j &= -(30-u_{j-2}+16u_{j-1}-30u_j+16u_{j+1}-u_{j+2}), \quad j = 2,\ldots,N-3, \\ a_{N-2} &= -(30-u_{N-4}+16u_{N-3}-30u_{N-2}+16u_{N-1}-2-\ln(1+u_N)), \\ a_{N-1} &= -(15+u_{N-5}-6u_{N-4}+14u_{N-3}-4u_{N-2}-15u_{N-1}+10(2+\ln(1+u_N))); \end{aligned} \tag{5.12}$$

the entries b_{N-2} and b_{N-1} are defined by

$$\begin{aligned} b_{N-2} &= \frac{1}{1+u_N}, \\ b_{N-1} &= \frac{-10}{1+u_N}. \end{aligned} \tag{5.13}$$

For an experimental verification of the theoretically derived stability condition we need the spectral radius $\sigma(J)$. However, the eigenvalues of M are not so easily found. Therefore, as an estimate of $\sigma(M)$, we shall use the spectral radius of the matrix which only represents the central differences (5.3), that is, we neglect the boundary conditions. Moreover, we approximate the diagonal entries a_j with the constant -30. The approximating matrix for

M is a well-known symmetric difference matrix of which the eigenvalues are situated in the interval [-64,0]. Thus we approximately have real eigenvalues for J, and

$$\sigma(J) \simeq \frac{16 \max(d_j)}{3 (\Delta x)^2} . \tag{5.14}$$

The corresponding stability condition is

$$h \leq \frac{3 \beta (\Delta x)^2}{16 \max(d_j)} , \quad \beta = 181.1. \tag{5.15}$$

Problem (5.7) shall be integrated for three values of N. For the additional starting values we use the analytical solution (5.2). As it is our aim to verify the theoretically derived stability condition, we shall neglect any accuracy condition by integrating with approximately the maximal step allowed by condition (5.15). From solution (5.2) we know that d_j should be monotonically decreasing for increasing t. This means that, as the integration proceeds, the steplength h must be increased. This will be done by step doubling. Thus, the stepsize strategy is: as soon as 2h satisfies (5.15), the steplength is doubled. The integration is stopped as soon as $t \geq 100$. Results are listed in table 5.1. In this table we give the number of integration steps, denoted with steps, and the maximal relative error

$$\max_j \left| \frac{u_j - u(j\Delta x,t)}{u(j\Delta x,t)} \right| , \tag{5.16}$$

denoted with error.

TABLE 5.1.

N	16	32	64
error	$2.5_{10}{}^{-2}$	$1.0_{10}{}^{-3}$	$5.5_{10}{}^{-5}$
steps	28	101	397

The results of table 5.1 indicate that to a certain extent the stability conditions, which are derived for linear equations, also apply in case of non-linear equations.

Because of the fact that $p_0 = -\frac{3}{4}$ (remember $p_0 = -1$ violates the convergence condition), we expect that the given two-step methods are not so accurate. In order to get some insight in the accuracy behaviour we compare the given 10-point two-step method with a stabilized one-step method of second order which also uses 10 function evaluations per step. This method proceeds from VAN DER HOUWEN [9]. For the one-step method the real stability boundary $\beta = 81.11$.

Again we integrate problem (5.7), while for both methods the stepsize strategy is applied as described above. However, condition (5.15) is replaced by

$$(5.17) \qquad h \leq \frac{3\,\beta\,(\Delta x)^2}{16\,\max(d_j)}, \qquad \beta = 81.11$$

The integration is stopped as soon $t \geq 100$. Results are given in table 5.2.

TABLE 5.2

	one-step		two-step	
N	16	32	16	32
error	$4.1_{10}{}^{-4}$	$3.0_{10}{}^{-5}$	$3.2_{10}{}^{-3}$	$1.6_{10}{}^{-4}$
steps	56	223	58	223

The results of the one-step method indeed are more accurate than the results of the two-step method. The ratios between the given errors are approximately 7.8 and 5.3 for N = 16 and N = 32, respectively. These results thus indicate that it is of interest to investigate two-step schemes of class (3.2) with $p_0 > -\frac{3}{4}$. On the other hand, the two step method has a much larger boundary of absolute stability than the one-step method has. As a consequence, for problems where the steplength of the time integration is completely determined by stability conditions, and not by accuracy conditions, two-step methods shall be more efficient than one-step methods.

To illustrate this we integrate the following linear problem (VAN DER HOUWEN [12]):

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + e^{-t}(x^{10}+90x^8-x),\quad 0 \le x \le 1,\quad t \ge 0,$$

$$u = 1 + x(1-x^9),\quad 0 \le x \le 1,\quad t = 0, \tag{5.18}$$

$$u = 1,\quad x = 0,\quad x = 1,\quad t \ge 0.$$

Problem (5.18) is solved by the function

$$u(x,t) = 1 + e^{-t}x(1-x^9). \tag{5.19}$$

Again we divide the x-interval into N equal intervals of length $\Delta x = 1/N$. Let $u_j(t)$ denote an approximation to $u(j\Delta x,t)$, $j = 1,\ldots,N-1$. Proceeding in the same way as in the first example we arrive at the following initial value problem.

$$\frac{du_1}{dt} = (-\frac{5}{4}u_1 - \frac{1}{3}u_2 + \frac{7}{6}u_3 - \frac{1}{2}u_4 + \frac{1}{12}u_5)/(\Delta x)^2 +$$
$$+ ((\Delta x)^{10}+90(\Delta x)^8-\Delta x)e^{-u_N} + 5/6(\Delta x)^2,$$

$$\frac{du_2}{dt} = (\frac{4}{3}u_1 - \frac{5}{2}u_2 + \frac{4}{3}u_3 - \frac{1}{12}u_4)/(\Delta x)^2 +$$
$$+ ((2\Delta x)^{10}+90(2\Delta x)^8-2\Delta x)e^{-u_N} - 1/12(\Delta x)^2,$$

$$\frac{du_j}{dt} = (-\frac{1}{12}u_{j-2} + \frac{4}{3}u_{j-1} - \frac{5}{2}u_j + \frac{4}{3}u_{j+1} - \frac{1}{12}u_{j+2})/(\Delta x)^2 + \tag{5.20}$$
$$+ ((j\Delta x)^{10}+90(j\Delta x)^8-j\Delta x)e^{-u_N},\quad j = 3,\ldots,N-3,$$

$$\frac{du_{N-2}}{dt} = (-\frac{1}{12}u_{N-4} + \frac{4}{3}u_{N-3} - \frac{5}{2}u_{N-2} + \frac{4}{3}u_{N-1})/(\Delta x)^2 +$$
$$+ (((N-2)\Delta x)^{10}+90((N-2)\Delta x)^8-(N-2)\Delta x)e^{-u_N} - 1/12(\Delta x)^2,$$

$$\frac{du_{N-1}}{dt} = (\frac{1}{12}u_{N-5} - \frac{1}{2}u_{N-4} + \frac{7}{6}u_{N-3} - \frac{1}{3}u_{N-2} - \frac{5}{4}u_{N-1})/(\Delta x)^2 +$$
$$+ (((N-1)\Delta x)^{10}+90((N-1)\Delta x)^8-(N-1)\Delta x)e^{-u_N} + 5/6(\Delta x)^2,$$

$$\frac{du_N}{dt} = 1,$$

and

$$u_j(0) = 1 + j\Delta x(1-(j\Delta x)^9), \quad j = 1,\ldots,N-1,$$

(5.21)

$$u_N(0) = 0.$$

By using the same argument as in the preceding example, we have the stability condition

$$h \leq \frac{3\beta(\Delta x)^2}{16} . \tag{5.22}$$

Problem (5.20) - (5.21) shall be solved for N = 32 with the 10-point two-step method given above, and with a strongly stable 10-point one-step method given by VAN DER HOUWEN [11], section 2.6.6. We observe that the one-step method used for the first example is weakly stable. Van der Houwen calls a method strongly stable if its amplification factors are inside the unit circle. For the strongly stable one-step method there holds: $\beta = 79.70$.

The integration is performed with the maximal constant steplength allowed by (5.22) and is stopped as soon as $t \geq 5$. The results are given in table 5.3. These results clearly illustrate that in cases where the error due to the space discretization dominates the two-step method is more effective than the one-step method.

TABLE 5.3.

	one-step	two-step
error	$4.9_{10}-3$	$4.9_{10}-3$
steps	342	150

REFERENCES

[1] ABRAMOWITZ, M. & I.A. STEGUN, *Handbook of mathematical functions*, National Bureau of Standards Applied Mathematics Series 55, U.S. Government Printing Office, Washington, 1964.

[2] BUTCHER, J.C., *On the convergence of numerical solutions to ordinary differential equations*, Math. Comp. 20, pp. 1-10, 1966.

[3] FEHLBERG, E., *Klassische Runge-Kutta-Formeln vierter und niedriger Ordnung mit Schrittweiten-Kontrolle und ihre Anwendung auf Wärmeleitungsprobleme*, Computing 6, pp. 61-71, 1970.

[4] GEAR, C.W., *Hybrid methods for initial value problems in ordinary differential equations*, SIAM J. Numer. Anal. 2, pp. 69-86, 1965.

[5] HENRICI, P., *Discrete variable methods in ordinary differential equations*, John Wiley & Sons, New York, 1962.

[6] HENRICI, P., *Error propagation for difference methods*, The SIAM Series in Applied Mathematics, John Wiley & Sons, New York, 1963.

[7] RICHTMYER, R.D. & K.W. MORTON, *Difference methods for initial-value problems*, Interscience, New York, 1967.

[8] STETTER, H.J., *Analysis of discretization methods for ordinary differential equations*, Springer-Verlag, Berlin, 1973.

[9] VAN DER HOUWEN, P.J., *Explicit Runge-Kutta formulas with increased stability boundaries*, Numer. Math. 20, pp. 149-164, 1972.

[10] VAN DER HOUWEN, P.J., *A note on two-step Runge-Kutta methods*, Report TN 61/71, Mathematisch Centrum, Amsterdam, 1971.

[11] VAN DER HOUWEN, P.J., *Construction of integration formulas for initial value problems*, North-Holland Publishing Company, Amsterdam, (to be published).

[12] VAN DER HOUWEN, P.J., *One-step methods for linear initial value problems*, ZAMM 51, T58-T59, 1971.

[13] WATT, J.M., *The asymptotic discretization error of a class of methods for solving ordinary differential equations*, Proc. Camb. Phil. Soc. 63, pp. 461-472, 1967.