CWI

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

W.H. Hundsdorfer, B.I. Steininger

Convergence of linear multistep and one-leg methods
for stiff nonlinear initial value problems

Convergence of linear multistep and one-leg methods

for stiff nonlinear initial value problems

W.H. Hundsdorfer

Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

B.I. Steininger

Gehrenbergstrasse 26, D-7778 Markdorf,
Federal Republic Germany.

To prove convergence of numerical methods for stiff initial value problems, stability is needed but also estimates for the local errors which are not affected by stiffness. In this paper global error bounds are derived for one-leg and linear multistep methods applied to classes of arbitrarily stiff, nonlinear initial value problems. It will be shown that under suitable stability assumptions the multistep methods are convergent for stiff problems with the same order of convergence as for nonstiff problems, provided that the stepsize variation is sufficiently regular.

1. INTRODUCTION

In this paper we shall discuss convergence of multistep methods applied to stiff nonlinear initial value problems

$$u'(t) = f(t, u(t)) \ (0 < t \leq T), \ u(0) \ \text{given}, \tag{1.1}$$

with $u(0) \in \mathbb{R}^m$ and $f : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$. Much of our attention will be given to k-step *one-leg methods*, where successive approximations u_{n+k} to the exact solution $u(t)$ at gridpoints $t_{n+k} = t_{n+k-1} + h$ are computed from

$$\sum_{j=0}^{k} \alpha_j u_{n+j} = hf(\sum_{j=0}^{k} \beta_j t_{n+j}, \sum_{j=0}^{k} \beta_j u_{n+j}) \ (n = 0, 1, 2, ...), \tag{1.2}$$

with starting values $u_0, u_1, \ldots, u_{k-1}$. Compared with the corresponding *linear multistep method*

$$\sum_{j=0}^{k} \alpha_j u_{n+j} = h \sum_{j=0}^{k} \beta_j f(t_{n+j}, u_{n+j}) \ (n = 0, 1, 2, ...), \tag{1.3}$$

the one-leg method (1.2) may have stronger nonlinear stability properties, such as G-stability, and a more robust behaviour on nonuniform grids, see [4], [16]. On the other hand, it is known that to obtain a one-leg method of high *order* (i.e. order of consistency for nonstiff problems) the parameters α_j, β_j have to satisfy more constraints than for linear multistep methods, see [5], [7] and also Section 3.

We shall be concerned with bounds for the global errors $u(t_n) - u_n$ that are not affected by stiffness. Such bounds have been studied quite extensively for Runge-Kutta methods, for example in [8], [9] and [10]. Most Runge-Kutta methods suffer from an *order reduction* in the presence of stiffness, i.e., the order of convergence for stiff problems may be considerably lower than for nonstiff problems, even if the solution $u(t)$ is very smooth. As we shall see, such order reduction will not occur with the

multistep methods, provided that the grid is sufficiently regular.

Stiffness independent error bounds can be obtained for various classes of initial value problems. The most general class that will be considered in this paper consists of the problems (1.1) where the function f satisfies the monotonicity condition

$$(f(t,v) - f(t,\tilde{v}), v - \tilde{v}) \leq 0 \quad \text{(for all } t \in [0,T] \text{ and } v, \tilde{v} \in \mathbf{R}^m) \tag{1.4}$$

with respect to some inner product $(\cdot, \cdot)$ on $\mathbf{R}^m$. Although sufficient stability conditions for one-leg methods are known, this does not lead to convergence results in a straightforward way, since the local errors will depend on the stiffness (except for methods like BDF, which are at the same time one-leg and linear multistep methods). For the linear multistep methods it is stability that causes difficulties in proving convergence, and as we will see additional constraints on the starting values have to be imposed. One of the results of this paper is that any A-stable multistep method (1.2), (1.3) applied with exact starting values and smoothly varying stepsizes to a problem satisfying (1.4) will be convergent, independently of the stiffness, with the same order of convergence as for nonstiff problems.

A complete convergence analysis for the implicit midpoint rule and the trapezoidal rule applied to problems of the above type has been obtained by Kraaijevanger [13]. Our approach is closely related to this analysis. For the sake of simplicity, we shall confine ourselves until Section 5 to uniform grids $t_n = nh$ $(n = 0, 1, 2, ...)$.

After some preliminaries in Section 2, convergence of one-leg methods is discussed in Section 3. It will be shown that the local discretization error, defined as the error introduced in one single step of the integration process, may slightly suffer from an order reduction: as a rule, one order is lost due to stiffness. For stable methods, however, this reduction will not be present in the global discretization error, due to damping and cancellation effects. As we found out after completing this section, Hairer and Wanner derived some convergence results for one-leg methods in a preliminary version of [12] by considering the same local error but without taking into account the cancellation effects (and this leads to a too pessimistic estimate for the order of convergence). In [6], Dahlquist indicated a different way to obtain stiffness independent global error bounds for one-leg methods; this approach however is more complicated and it seems difficult in certain cases (if there is no strong damping) to get the correct order of convergence.

In Section 4, convergence results for linear multistep methods are derived, by using the well known equivalence relations of Dahlquist [3] between the linear multistep method (1.3) and its one-leg counterpart (1.2). Rather general error bounds for linear multistep methods applied to stiff nonlinear problems were given already by Nevanlinna and Odeh [16]; they applied the equivalence relation in a slightly different way, which requires the assumption of strong stability at infinity (and this excludes the trapezoidal rule, for instance). We shall prove convergence for linear multistep methods with appropriate starting values under the sole assumption that the corresponding one-leg method is stable for the class of problems under consideration.

Under certain restrictions on the stepsize variation, the convergence results will carry over to nonuniform grids. This will be demonstrated in Section 5.

Finally, in Section 6, convergence of interpolated one-leg approximations $\beta_k u_{n+k} + ... + \beta_0 u_n$ will be considered.

2. PRELIMINARIES

Consider the polynomials ρ and σ containing the coefficients of the method,

$$\rho(z) = \sum_{j=0}^{k} \alpha_j z^j, \quad \sigma(z) = \sum_{j=0}^{k} \beta_j z^j.$$

Let E stand for the forward shift operator and $\bar{t}_n = \sigma(E)t_n$ for $n = 0, 1, ..., N$, with N being the number of steps needed to cover the interval $[0, T]$ at a given stepsize h. For the one-leg methods it will be assumed that $t_n \leq \bar{t}_n \leq t_{n+k}$ to guarantee that all $\bar{t}_n$-values are inside $[0, T]$; otherwise some modifications would be required. Further it will be assumed throughout the paper that the

polynomials ρ,σ have no common zeros and $\alpha_k\neq 0$ (irreducibility), and that $\rho(1)=0$, $\sigma(1)=\rho'(1)=1$ (consistency). We will consider initial value problems (1.1) such that all derivatives $u^{(j)}(t)$ of the exact solution needed in the analysis exist and $f:[0,T]\times\mathbb{R}^m\to\mathbb{R}^m$ is assumed to be continuously differentiable. The Jacobian matrix $[\partial f(t,v)/\partial v]$ will be denoted by $J(t,v)$. Besides, it will be assumed that the systems of algebraic equations arising from implicitness of the methods if $\beta_k\neq 0$ have unique solutions; cf. [3] for a sufficient condition.

The one-leg scheme (1.2) can be written as

$$\rho(E)u_n = hf(\bar{t}_n,\bar{u}_n), \quad \bar{u}_n = \sigma(E)u_n. \tag{2.1}$$

For the exact solution u of (1.1) we have

$$\rho(E)u(t_n) = hf(\bar{t}_n, u(\bar{t}_n)) + r_n, \quad u(\bar{t}_n) = \sigma(E)u(t_n) + q_n \tag{2.2}$$

where r_n,q_n are discretization errors due to differentiation and interpolation, respectively. These errors, which only depend on the smoothness of u, will be considered more closely in Section 3 (cf. also [5], [6]).

Let $\epsilon_n = u(t_n)-u_n$ denote the global discretization errors of the one-leg scheme, and put $\bar{\epsilon}_n = u(\bar{t}_n)-\bar{u}_n$. By subtraction of (2.1) from (2.2) it follows that

$$\rho(E)\epsilon_n = Z_n\bar{\epsilon}_n + r_n, \quad \bar{\epsilon}_n = \sigma(E)\epsilon_n + q_n \tag{2.3}$$

where the $m\times m$ matrix Z_n is given by $Z_n=hJ_n$ with

$$J_n = \int_0^1 J(\bar{t}_n, \theta u(\bar{t}_n) + (1-\theta)\bar{u}_n)d\theta . \tag{2.4}$$

Elimination of $\bar{\epsilon}_n$ in (2.3) leads to

$$\rho(E)\epsilon_n = Z_n\sigma(E)\epsilon_n + r_n + Z_nq_n. \tag{2.5}$$

This recursion for the global errors can be written in the somewhat more transparant form

$$\epsilon_{n+k} = \sum_{j=1}^{k} \psi_j(Z_n)\epsilon_{n+k-j} + \delta_n \tag{2.6}$$

where $\delta_n = (\alpha_kI-\beta_kZ_n)^{-1}(r_n+Z_nq_n)$ with I the $m\times m$ identity matrix and the ψ_j are rational functions given by $\psi_j(z)= -(\alpha_k-\beta_kz)^{-1}(\alpha_{k-j}-\beta_{k-j}z)$. In order to facilitate the analysis, (2.6) will be written as a one-step recursion in $\mathbb{R}^{km}$

$$e_{n+1} = R_ne_n + d_n \tag{2.7}$$

where $R_n = R(Z_n)$ with

$$R(Z)=\begin{bmatrix} \psi_1(Z) & \cdots & \psi_k(Z) \\ I & 0 & \\ & \ddots & \\ & & I & 0 \end{bmatrix} \quad \text{and} \quad e_n = \begin{bmatrix} \epsilon_{n+k-1} \\ \epsilon_{n+k-2} \\ \cdot \\ \cdot \\ \cdot \\ \epsilon_n \end{bmatrix}, \quad d_n = \begin{bmatrix} \delta_n \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}.$$

To ensure stability of the one-leg scheme - which is, as we see from the above, governed by the matrices $R(Z_n)$ - appropriate assumptions on the method and the class of stiff initial value problems are needed. Consider a class of initial value problems $\mathscr{P}$ with a suitable maximal stepsize $H(\mathscr{P})$. This determines a class $\mathscr{Z}$ of possible matrices $Z_n = hJ_n$ with $0<h\leqslant H(\mathscr{P})$ and J_n given by (2.4). For example, if $\mathscr{P}$ consists of all problems (1.1) with f satisfying (1.4) in arbitrary dimension m, then $\mathscr{Z}$ will contain all $m\times m$ matrices Z with $m\in\mathbb{N}$ such that $(v,Zv)\leqslant 0$ for all $v\in\mathbb{R}^m$; this is irrespective of the value of $H(\mathscr{P})$, which can be chosen as ∞. Let $|\cdot|$ be a given norm on $\mathbb{R}^m$ and define the norm $\|\cdot\|$ on

$\mathbf{R}^{km}$ as

$$\|w\| = \max_{1 \leq j \leq k} |w_j| \quad \text{for } w = (w_1^T, w_2^T, \ldots, w_k^T)^T \in \mathbf{R}^{km}.$$

We shall use $|\cdot|$ and $\|\cdot\|$ also to denote the induced matrix norms for $m \times m$ and $km \times km$ matrices, respectively. A basic assumption in this paper is the existence of a constant $S > 0$ such that

$$\|R_n R_{n-1} \cdots R_l\| \leq S \tag{2.8}$$

for all possible $R_j = R(Z_j)$ with $Z_j \in \mathscr{Z}$ and $1 \leq l \leq n \leq N$. This stability assumption does not depend on our special choice of norm $\|\cdot\|$ on R^{km}, which is merely taken for convenience. Equivalent norms $\|\cdot\|'$ as given in [18] could be used as well and this would only alter the stability constant S. Sufficient conditions for (2.8) have been extensively studied; here we shall give some examples were the stepsizes are allowed to be arbitrarily large.

EXAMPLE 2.1. Suppose the norm $|\cdot|$ on $\mathbf{R}^m$ is generated by an inner product and f satisfies the monotonicity condition (1.4). It was proved in [4] (cf. also [14]) that for any A-stable method a norm $\|\cdot\|'$ can be found such that $\|R(Z_n)\|' \leq 1$ whenever $Z_n = hJ_n$ is given by (2.4) with $h > 0$. The norm $\|\cdot\|'$ (the G-norm) is completely determined by the coefficients α_j, β_j of the method and it is equivalent with $\|\cdot\|$ uniformly in m. Hence, A-stability and (1.4) together are sufficient for (2.8) with a stability constant S determined by the method (and thus independent of the specific problem (1.1), its stiffness or dimension).

For $A(\alpha)$-stable methods (2.8) will hold under more restrictive conditions on f as given in [16]. It should be noted that the results in [16], especially Theorem 4.1, have been formulated in terms of linear multistep methods, but along the same lines stability of one-leg methods follows even in a more simple way. $\square$

EXAMPLE 2.2. Assume again f satisfies (1.4) and $|\cdot|$ is generated by an inner product. A one-leg method is said to be A-contractive in the max-norm [15] if $\|R(z)\| \leq 1$ for all $z \in \mathbb{C}$ with Re $z \leq 0$. It was shown in [14] that for such methods we have $\|R(Z_n)\| \leq 1$ whenever $Z_n = hJ_n$ is given by (2.4) with $h > 0$, and thus (2.8) then holds with $S = 1$.

In [15] also larger classes of methods were considered, for instance methods which are A_0-contractive in the max-norm. For these methods (2.8) holds with $S = 1$ for real, scalar problems where $Z_n \in \mathbb{R}$, $Z_n \leq 0$. $\square$

More general conditions, under the assumption that f is circularly bounded, can be found in [4], [18], [19], for example. The latter two references also deal with norms on $\mathbf{R}^m$ which are not generated by an inner product; the results in [19] have been derived for linear problems, but with one-leg methods they are also applicable to nonlinear problems.

The stability condition (2.8) implies in particular that $\|R(Z_n)\| \leq S$ for any n, and therefore also $|\psi_j(Z_n)| \leq S$ $(1 \leq j \leq k)$. Due to the consistency conditions $\rho(1) = 0$, $\sigma(1) = 1$, we have

$$\sum_{j=1}^{k} \psi_j(Z_n) = I + (\alpha_k I - \beta_k Z_n)^{-1} Z_n. \tag{2.9}$$

It follows that there are constants $S_1, S_2 > 0$, determined by α_k, β_k and S, such that for all n

$$|(\alpha_k I - \beta_k Z_n)^{-1}| \leq S_1, \quad |(\alpha_k I - \beta_k Z_n)^{-1} Z_n| \leq S_2. \tag{2.10}$$

The following lemma shows that slightly perturbed coefficients of a stable scheme will cause only a small perturbation of $\psi_j(Z_n)$ - and consequently of $R(Z_n)$ - even if $|Z_n|$ is very large. Consider $\tilde{\alpha}_j, \tilde{\beta}_j$ $(0 \leq j \leq k)$ such that

$$|\tilde{\alpha}_j - \alpha_j| \leq \gamma h, \quad |\tilde{\beta}_j - \beta_j| \leq \gamma h \quad (0 \leq j \leq k) \tag{2.11}$$

with $\gamma > 0$, and let $\tilde{\psi}_j(z) = -(\tilde{\alpha}_k - \tilde{\beta}_k z)^{-1}(\tilde{\alpha}_{k-j} - \tilde{\beta}_{k-j} z)$ $(1 \leq j \leq k)$.

LEMMA 2.3. *Suppose (2.8), (2.10) and (2.11) hold with constants $S, S_1, S_2, \gamma > 0$. Let $\gamma' = \gamma(S_1 + S_2)$ and assume $\gamma' h < 1$. Then*

$$|\tilde{\psi}_j(Z_n) - \psi_j(Z_n)| \leqslant (1 + S)(1 - \gamma' h)^{-1} \gamma' h \quad (1 \leqslant j \leqslant k).$$

PROOF. For any i, $0 \leqslant i \leqslant k$, we have

$$|(\alpha_k I - \beta_k Z_n)^{-1} [(\tilde{\alpha}_i - \alpha_i)I - (\tilde{\beta}_i - \beta_i)Z_n]| \leqslant S_1 \gamma h + S_2 \gamma h = \gamma' h,$$

and in particular, for $i = k$,

$$|(\alpha_k I - \beta_k Z_n)^{-1}(\tilde{\alpha}_k I - \tilde{\beta}_k Z_n) - I| \leqslant \gamma' h.$$

By some calculations it follows that

$$|(\tilde{\alpha}_k I - \tilde{\beta}_k Z_n)^{-1}(\alpha_k I - \beta_k Z_n) - I| \leqslant (1 - \gamma' h)^{-1} \gamma' h,$$

$$|(\tilde{\alpha}_k I - \tilde{\beta}_k Z_n)^{-1}(\alpha_k I - \beta_k Z_n)| \leqslant (1 - \gamma' h)^{-1}.$$

The proof can now easily be deduced from the relation

$$\tilde{\psi}_j(Z_n) - \psi_j(Z_n) = [(\tilde{\alpha}_k I - \tilde{\beta}_k Z_n)^{-1}(\alpha_k I - \beta_k Z_n) - I] \psi_j(Z_n) +$$

$$+ [(\tilde{\alpha}_k I - \tilde{\beta}_k Z_n)^{-1}(\alpha_k I - \beta_k Z_n)] [(\alpha_k I - \beta_k Z_n)^{-1}((\tilde{\alpha}_{k-j} - \alpha_{k-j})I - (\tilde{\beta}_{k-j} - \beta_{k-j})Z_n)]. \quad \square$$

3. CONVERGENCE OF ONE-LEG METHODS

3.1. Local error bounds

Considering (2.6) we see that $\delta_n = (\alpha_k I - \beta_k Z_n)^{-1} (r_n + Z_n q_n)$ is a local discretization error, in the sense that if $\epsilon_n = \epsilon_{n+1} = \ldots = \epsilon_{n+k-1} = 0$ then $\epsilon_{n+k} = \delta_n$. (In other words, δ_n is the error, due to discretization, which is introduced in one single step of the integration process.) For a q-th order method we expect that $|\delta_n| = O(h^{q+1})$.

Note that δ_n is different from the usual discretization error obtained by substituting the exact solution directly into (1.2). This error, which is approximately given by $r_n + Z_n q_n$, is not bounded uniformly in the stiffness and therefore inappropriate for stiff systems. This was observerd already in [6], where an alternative error was proposed for stepsize control (for our purpose, proving convergence for stiff problems, the δ_n seem more suitable).

Let the numbers ν_i, μ_i be defined by

$$\nu_i = \sum_{j=0}^{k} \alpha_j j^i, \quad \mu_i = \sum_{j=0}^{k} \beta_j j^i \quad (i = 1, 2, \ldots) \tag{3.1}$$

By a Taylor series expansion of u and the consistency conditions $\rho(1) = 0$, $\sigma(1) = \rho'(1) = 1$, it follows that

$$r_n = A_2 h^2 u''(t_n) + A_3 h^3 u'''(t_n) + \cdots \quad , A_i = \frac{1}{i!}(\nu_i - i\mu_1^{i-1}), \tag{3.2a}$$

$$q_n = B_2 h^2 u''(t_n) + B_3 h^3 u'''(t_n) + \cdots \quad , B_i = \frac{1}{i!}(\mu_1^i - \mu_i). \tag{3.2b}$$

Now, for *nonstiff problems*, where f satisfies a Lipschitz condition with a moderate constant, we have $|Z_n| = O(h)$, and consequently

$$\delta_n = (\alpha_k^{-1} + O(h))(r_n + O(h)q_n).$$

In order to have $|\delta_n| = O(h^{q+1})$ it is then necessary and sufficient that

$$A_i = 0 \, (2 \leqslant i \leqslant q) \quad \text{and} \quad B_i = 0 \, (2 \leqslant i \leqslant q - 1), \tag{3.3}$$

6

which are the usual order conditions for one-leg methods (cf. [5], [6]). The more familiar conditions for the linear multistep method (1.3) to be of order q can be written as $A_i + B_{i-1} = 0$ ($2 \leqslant i \leqslant q$), see Section 4 for example. These conditions are the same if $q = 2$, but for $q \geqslant 3$ the one-leg method has to satisfy more constraints.

For *stiff problems* the order conditions (3.3) are not sufficient to have $|\delta_n| \leqslant Ch^{q+1}$ for some moderately sized $C > 0$. From (2.10) it can be concluded that

$$|\delta_n| \leqslant S_1 |r_n| + S_2 |q_n|$$

but since $S_2 \neq O(h)$ in general (for instance if $Z_n = h\lambda, \lambda \to \infty$), (3.3) merely implies $|S_n| = O(h^q)$ for stiff problems. This local order reduction occurs with the implicit midpoint rule [13] and most other one-leg methods.

EXAMPLE 3.1. Consider the second order method CA2, introduced in [15],

$$u_{n+2} - u_{n+1} = hf(t_n + \frac{3}{2}h, \frac{3}{4}u_{n+2} + \frac{1}{4}u_n).$$

This method is A-stable and A_0-contractive in the maximum norm. Consider further the model problem

$$u'(t) = \lambda[u(t) - g(t)] + g'(t), \quad u(0) = g(0), \tag{3.4}$$

with solution $u(t) = g(t)$ for any $\lambda \leqslant 0$. Take $g(t) = \frac{1}{2}t^2$. Application of the above method with exact starting values $u_0 = u(0)$, $u_1 = u(t_1)$, gives

$$u(t_2) - u_2 = -\frac{3}{8}h^2(1 - \frac{3}{4}z)^{-1}z, \quad z = h\lambda.$$

If the problem is nonstiff, $|\lambda| \leqslant 1$ say, then $|u(t_2) - u_2| \leqslant \frac{3}{8}h^3 + O(h^4)$, as we would expect after one step with a second order method. For $\lambda \to -\infty$, however, we only have $|u(t_2) - u_2| = \frac{1}{2}h^2$. Thus, due to stiffness one order of h is lost. $\square$

The local order reduction is absent with methods that satisfy, in addition to (3.3), the extra order condition

$$B_q = 0. \tag{3.5}$$

This is fulfilled, for instance, by the BDF methods (where all B_j are zero). It can also be shown that the local order reduction will not occur in case the initial value problem (1.1) is such that all partial derivatives $[\partial^{i+j}f(t,v)/\partial t^i \partial v^j]$ with $i,j \geqslant 0, (i,j) \neq (0,1)$ are bounded by a moderate constant (cf. [9] for a related result with Runge-Kutta methods). Note that the partial derivative with $(i,j) = (0,1)$, the Jacobian, is always large for stiff systems, since its norm is proportional to the Lipschitz constant. For general stiff systems, where other partial derivatives may be large as well, there will be a local order reduction if (3.5) is not satisfied, as can be seen by considering problems of the type (3.4) with $\lambda \ll 0$ and g a smooth function.

3.2. Global error bounds

For nonstiff problems the local condition $|\delta_n| = O(h^{q+1})$ is necessary to have $|\epsilon_n| = 0(h^q)$, global convergence of order q. For stiff problems there may be damping or cancellation of local errors, as a result of which there can be convergence of order q while $|\delta_n| = O(h^q)$ only. This was shown in [13] to be the case for the implicit midpoint rule. Related results for Runge-Kutta methods can be found in [1].

LEMMA 3.2. *Consider recursion (2.7). Assume the stability condition (2.8) holds with a constant $S>0$. Assume further that a constant $D>0$ and vectors $x_n, y_n \in \mathbb{R}^{km}$ exist such that*

$$d_n = (I - R_n)x_n + y_n,$$
(3.6a)

$$\|x_n\| \leqslant Dh^p, \quad \sum_{j=0}^{n-1} \|x_{j+1} - x_j\| \leqslant Dh^p, \quad \sum_{j=0}^{n-1} \|y_j\| \leqslant Dh^p$$
(3.6b)

for all $n = 0,1,...,N$. Then

$$\|e_n\| \leqslant S\|e_0\| + (3S+1)Dh^p \quad (n = 1,2,...,N).$$

PROOF. Let $\tilde{e}_n = e_n - x_n$. These perturbed errors satisfy

$$\tilde{e}_n = R_{n-1}\tilde{e}_{n-1} + \tilde{d}_{n-1}, \quad \tilde{d}_{n-1} = y_{n-1} + x_{n-1} - x_n \quad (n = 1,2,...,N).$$

By writing out $\tilde{e}_n$ fully in terms of $R_j, \tilde{d}_j (0 \leqslant j \leqslant n-1)$ and $\tilde{e}_0$, it easily follows from (2.8) that

$$\|\tilde{e}_n\| \leqslant S\|\tilde{e}_0\| + S\sum_{j=0}^{n-1} \|\tilde{d}_j\|.$$

Hence

$$\|e_n\| \leqslant S\|e_0\| + S\sum_{j=0}^{n-1} \|y_j\| + S\sum_{j=0}^{n-1} \|x_j - x_{j+1}\| + S\|x_0\| + \|x_n\|,$$

which yields the proof of the lemma. □

THEOREM 3.3. *Consider a one-leg method (1.2) having order q, and assume (2.8) holds with stability constant $S>0$. Then there is a constant $C>0$, only depending on S,T and bounds for derivatives of $u(t)$, such that*

$$|u(t_n) - u_n| \leqslant S \max_{0 \leqslant j < k} |u(t_j) - u_j| + Ch^q \quad \text{for all } n \geqslant k, nh \leqslant T.$$

PROOF. In order to apply Lemma 3.2 it has to be determined whether the vector d_n in (2.7) can be decomposed as indicated. Let $x_n = (x_{1n}^T, \ldots, x_{kn}^T)^T$ and $y_n = (y_{1n}^T, \ldots, y_{kn}^T)^T$ with $x_{jn}, y_{jn} \in \mathbb{R}^m$. These vectors should satisfy

$$\delta_n = (I - \psi_1(Z_n))x_{1n} - \psi_2(Z_n)x_{2n} - \cdots - \psi_k(Z_n)x_{kn} + y_{1n},$$

$$0 = -x_{j-1,n} + x_{jn} + y_{jn} \quad (j = 2,3,...,k).$$

Taking $x_{1n} = x_{2n} = \cdots = x_{kn} = -q_n$, $y_{1n} = (\alpha_k I - \beta_k Z_n)^{-1}r_n$ and $y_{2n} = \cdots = y_{kn} = 0$, it is easily seen from (2.9) that (3.6a) is fulfilled and

$$\|x_n\| \leqslant D_1 h^q, \quad \|x_{n+1} - x_n\| \leqslant D_2 h^{q+1}, \quad \|y_n\| \leqslant D_3 h^{q+1}$$

for all n, with D_1, D_2, D_3 determined by S and the solution u (see Section 3.1). Hence, Lemma 3.2 can be applied with $D = \max\{D_1, D_2 T, D_3 T\}$ which leads to the error bound of the theorem. □

This convergence result shows that the order reduction is annihilated in the transition from local to global error. For stable one-leg schemes the order of convergence for stiff problems will be the same as in the nonstiff case.

REMARK 3.4. A somewhat shorter proof of the convergence result can be given by considering directly, instead of (2.6), a recursion for $\epsilon_n - q_n$. We have not followed this approach since it obscures

the underlying reason for the annihilation of the local order reduction as given in Lemma 3.2. □

4. Convergence of linear multistep methods

In this section convergence of linear multistep methods on uniform grids will be proved under the assumption that the corresponding one-leg method is stable, in the sense of (2.8), for the class of problems under consideration. This will be done by using the equivalence relations of [3] between one-leg and linear multistep schemes for constant stepsizes. This approach is not new, for example Theorem 4.1 in [16] can be viewed as being derived this way. In fact, the one-leg methods were introduced originally only to simplify the analysis of linear multistep methods, see [3]. The new feature in the proof below is that the equivalence relations will not be applied directly to the numerical approximations u_n or the global errors ϵ_n, but will only be used to show a kind of stability of the linear multistep error recursion. In this way, the order of the one-leg method will not be involved and use of the inverse operator $\sigma(E)^{-1}$ can be avoided (otherwise all roots of the σ-polynomial should have modulus strictly less than 1, which corresponds with strong stability at ∞). Moreover, the proof can be formulated in such a way that extension to smoothly varying stepsizes is rather straightforward, see Section 5.

Consider the linear multistep method (1.3). Using the generating polynomials ρ and σ it can be written as

$$\rho(E)u_n = h\sigma(E)f(t_n, u_n).$$ (4.1)

Inserting the exact solution $u(t)$ into this scheme, we obtain

$$\rho(E)u(t_n) = h\sigma(E)f(t_n, u(t_n)) + p_n$$ (4.2)

where p_n is a local discretization error. Taylor expansion of u and $u' = f(t, u)$ around the point t_n leads to

$$p_n = C_2 h^2 u''(t_n) + C_3 h^3 u'''(t_n) + \cdots, \quad C_i = \frac{1}{i!}(\nu_i - i\mu_{i-1})$$ (4.3)

with ν_i, μ_i defined by (3.1). The method has order p if the usual order conditions

$$C_i = 0 \quad (2 \leq i \leq p)$$ (4.4)

are satisfied, and then $p_n = O(h^{p+1})$ independently of the stiffness.

Let $Z_n = hJ_n$ with J_n the $m \times m$ matrix given by

$$J_n = \int_0^1 J(t_n, \theta u(t_n) + (1-\theta)u_n)d\theta.$$ (4.5)

For the global errors we then have the recursion

$$\rho(E)\epsilon_n = \sigma(E)Z_n\epsilon_n + p_n.$$ (4.6)

An important difference with the error recursion (2.5) for one-leg methods is that for each $n \geq 0$ several matrices Z_{n+j} are involved in (4.6). For $n = 0$, for example, we get

$$\epsilon_k = (\alpha_k I - \beta_k Z_k)^{-1}[-\sum_{j=0}^{k-1}(\alpha_j I - \beta_j Z_j)\epsilon_j + p_0].$$ (4.7)

If the Z_n are allowed to vary strongly we may have a situation where $|Z_k| \ll |Z_j| \, (0 \leq j < k)$, and to obtain an estimate for $|\epsilon_k|$ we will need not only bounds for the $|\epsilon_j|$ but also for $|Z_j\epsilon_j| \, (0 \leq j < k)$. Therefore, instead of the ϵ_n, we shall consider the transformed errors

$$\hat{\epsilon}_n = (\alpha_k I - \beta_k Z_n)\epsilon_n.$$ (4.8)

These transformed errors can be seen to satisfy

$$\hat{\epsilon}_{n+k} = \sum_{j=1}^{k} \psi_j(Z_{n+k-j})\hat{\epsilon}_{n+k-j} + p_n. \tag{4.9}$$

In one-step form this can be written as

$$\hat{e}_{n+1} = \hat{R}_n\hat{e}_n + \hat{d}_n \tag{4.10}$$

where $\hat{e}_n = (\hat{\epsilon}_{n+k-1}^T, \ldots, \hat{\epsilon}_n^T)^T$, $\hat{d}_n = (p_n^T, 0, \ldots, 0)^T \in \mathbb{R}^{km}$ and $\hat{R}_n$ is the $km \times km$ block matrix with $m \times m$ blocks $\psi_1(Z_{n+k-1}), \ldots, \psi_k(Z_n)$ on the first row, identity blocks I on the first lower subdiagonal and zero blocks on the other positions. For this recursion to be stable we now need uniform boundedness of products of $\hat{R}_n$.

LEMMA 4.1. *Assume (2.8) holds with stability constant $S > 0$. Then there exists an $\hat{S} > 0$, determined by S and the coefficients α_j, β_j, such that*

$$\|\hat{R}_n\hat{R}_{n-1} \cdots \hat{R}_l\| \leqslant \hat{S} \quad \text{for all } l,n \text{ with } 0 \leqslant l \leqslant n \leqslant N. \tag{4.11}$$

PROOF. Take $l = 0$, for convenience of notation. Consider vector sequences $\{v_n\}$ and $\{w_n\}$ such that

$$\rho(E)v_n = \sigma(E)Z_n v_n, \quad \rho(E)w_n = Z_n\sigma(E)w_n \quad (n \geqslant 0),$$

and let $\hat{v}_n = (\alpha_k I - \beta_k Z_n)v_n$. An equivalent formulation is

$$\hat{V}_{n+1} = \hat{R}_n\hat{V}_n, \quad W_{n+1} = R_n W_n \quad (n \geqslant 0)$$

with $\hat{V}_n = (\hat{v}_{n+k-1}^T, \ldots, \hat{v}_n^T)^T$ and $W_n = (w_{n+k-1}^T, \ldots, w_n^T)$. So, in order to prove the lemma it has to be shown that $|\hat{v}_n|$ can be uniformly bounded in terms of $|\hat{v}_0|, \ldots, |\hat{v}_{k-1}|$, starting from the stability estimate for the $|w_n|$. This will be done by choosing, for arbitrary given $v_0, v_1, \ldots, v_{k-1}$, suitable $w_0, w_1, \ldots, w_{k-1}$.

Since $\rho(E)$ and $\sigma(E)$ commute, the vectors $\bar{w}_n = \sigma(E)w_n$ satisfy

$$\rho(E)\bar{w}_n = \sigma(E)Z_n\bar{w}_n \quad (n \geqslant 0).$$

Hence, we have $\bar{w}_n = v_n$ for all $n \geqslant 0$ provided that $\bar{w}_j = v_j$ for $0 \leqslant j < k$. The latter condition will hold iff $w_0, w_1, \ldots, w_{2k-1}$ are such that

$$\sigma(E)w_j = v_j, \quad \rho(E)w_j = Z_j v_j \quad (0 \leqslant j < k).$$

This determines $w_0, w_1, \ldots, w_{2k-1}$ as a linear combination of v_j and $Z_j v_j (0 \leqslant j < k)$, as can be seen from relation (1.11) of [3] (for this relation to hold we have assumed in Section 2 that ρ and σ have no common zeros). It follows that

$$|w_i| \leqslant D \max_{0 \leqslant j < k} |\alpha_k v_j - \beta_k Z_j v_j| \leqslant D \max_{0 \leqslant j < k} |\hat{v}_j| \quad (0 \leqslant i < 2k).$$

for some $D > 0$ which only depends on the coefficients α_j, β_j, and the stability assumption (2.8) thus implies

$$|w_n| \leqslant SD \max_{0 \leqslant j < k} |\hat{v}_j| \quad (\text{for all } n \geqslant 0).$$

Finally, since

$$\hat{v}_n = \alpha_k\bar{w}_n - \beta_k Z_n\bar{w}_n = \alpha_k\sigma(E)w_n - \beta_k\rho(E)w_n,$$

we arrive at the bound

$$|\hat{v}_n| \leqslant \hat{S} \max_{0 \leqslant j < k} |\hat{v}_j| \quad (\text{for all } n \geqslant 0),$$

with $\hat{S} = SD \max_{0 \leqslant j \leqslant k} |\alpha_k\beta_j - \beta_k\alpha_j|$. $\square$

Returning to (4.9) it now follows that

$$|\hat{\epsilon}_n| \leqslant \hat{S}\|\hat{e}_0\| + \hat{S}\sum_{j=0}^{n-k}\|\hat{d}_j\|$$ (4.12)

for all $n \geqslant 0$. The sum over $\|\hat{d}_j\|$ ($= |p_j|$) will give an $O(h^p)$ contribution. Further, note that

$$\|\hat{e}_0\| = \max_{0 \leqslant j < k} |\alpha_k \epsilon_j - \beta_k Z_j \epsilon_j| \leqslant |\alpha_k| \max_{0 \leqslant j < k} |u(t_j) - u_j| + h|\beta_k| \max_{0 \leqslant j < k} |\hat{u}(t_j) - f(t_j, u_j)|.$$

Therefore, to have $\|\hat{e}_0\| = O(h^p)$ it will be required that

$$|u(t_j) - u_j| \leqslant C_0 h^p, \quad |\hat{u}(t_j) - f(t_j, u_j)| \leqslant C_0 h^{p-1} \quad (0 \leqslant j < k).$$ (4.13)

Finally, by noting that $|\epsilon_n| \leqslant S_1 |\hat{\epsilon}_n|$ (see (2.10)), the following result for the global errors ϵ_n is obtained.

THEOREM 4.2. *Consider a linear multistep method (1.3) with order p. Let $S, C_0 > 0$ be such that (2.8) and (4.13) hold. Then there is a $C > 0$, only depending on C_0, S, T and bounds for derivatives of u, such that*

$$|u(t_n) - u_n| \leqslant Ch^p \quad \text{for all } n \geqslant 0, nh \leqslant T. \quad \square$$

For stiff systems the condition (4.13) is more difficult to fulfil than only $|u(t_j) - u_j| \leqslant C_0 h^p$ ($0 \leqslant j < k$), which is sufficient for convergence of one-leg approximations (see Theorem 3.3). By considering model problems of the type

$$u'(t) = \lambda(t)[u(t) - g(t)] + g'(t), \quad u(0) = g(0)$$ (4.14)

with strongly varying $\lambda(t) \leqslant 0$ and smooth $g(t)$, it can be seen from (4.7) that (4.13) is necessary in the above theorem.

On the other hand, under the assumptions of Theorem 4.2 we also obtain

$$|u'(t_n) - f(t_n, u_n)| \leqslant C'h^{p-1} \quad \text{for all } n \geqslant 0, nh \leqslant T$$ (4.15)

with a $C' > 0$ independent of the stiffness, as it follows from

$$h|u'(t_n) - f(t_n, u_n)| = |Z_n \epsilon_n| \leqslant S_2 |\hat{\epsilon}_n| \quad \text{(see (2.10))}.$$

For one-leg methods such bounds (4.15) will not hold (this can be seen from (2.6) with $n = 0$, $J(t, v) = \lambda \to -\infty$ and $\epsilon_0 = \cdots = \epsilon_{k-1} = 0$).

It should also be noted that the approximation property (4.15) can be used to apply Theorem 4.2 in situations where subintervals of $[0, T]$ are integrated with different linear multistep methods.

REMARK 4.3. If we consider only linear differential equations

$$u'(t) = Au(t) + g(t),$$

then the Z_n are constant and there is no need to introduce the transformed errors $\hat{\epsilon}_n$ or to impose (4.13). Convergence can then be proved directly from (4.6) as for one-leg methods, see [2]. In the same way one can deal with semi-linear problems with an inhomogeneous term $g(t, u(t))$ satisfying a Lipschitz condition with respect to u; stability can then be proved from the linear case by a perturbation argument, similar as in Lemma 2.3 (cf. also [1] for a related result on Runge-Kutta methods). Therefore, in many problems of practical interest the condition (4.13) will not be needed. $\square$

5. VARIABLE STEPSIZES

5.1. Assumptions on the stepsize variations

In this section it will be discussed to what extent the convergence results for the multistep methods will be affected by variable stepsizes.

Consider a nonuniform grid $\{t_n\}$ covering $[0, T]$ and let $h_n = t_n - t_{n-1}$ be the stepsizes with a maximal stepsize h. Further, let $\omega_n = h_{n+1}/h_n$ denote the stepsize ratios. In the variable stepsize formulation of the multistep schemes (1.2) and (1.3), h is replaced by h_{n+k} and the coefficients α_j, β_j are allowed to vary with n,

$$\alpha_{jn} = a_j(\omega_{n+1}, \omega_{n+2}, \ldots, \omega_{n+k-1}), \quad \beta_{jn} = b_j(\omega_{n+1}, \omega_{n+2}, \ldots, \omega_{n+k-1}) \tag{5.1}$$

It will be assumed that these functions a_j, b_j, which describe the method on nonuniform grids, are Lipschitz continuous in a neighbourhood of $(1, 1, \ldots, 1)$, and that there is a positive constant M such that $M^{-1} \leqslant \omega_n \leqslant M$ and $|\alpha_{jn}|, |\beta_{jn}| \leqslant M$. Moreover it will be assumed that the fixed stepsize methods (1.2), (1.3) are extended in such a way that the orders q, p, respectively, are not affected by variable stepsizes. (The order conditions are for any given grid of the same complexity as for uniform grids, but it is difficult in general to find suitable closed expressions (5.1) such that these order conditions will be satisfied independently of the stepsize ratios.)

As we shall see, the convergence results will remain unchanged if the stepsizes are assumed to vary smoothly, in the sense that

$$|\omega_n - 1| \leqslant Kh_n \quad (n \geqslant 1) \tag{5.2}$$

with a moderate constant $K > 0$. This assumption will hold in case the stepsizes are chosen by accuracy considerations only (see [11], pp. 356, 357). Although in any practical numerical code local accuracy is not the only selection principle, (5.2) with often hold in substantial regions of the integration interval and thus has a certain practical relevance. If (5.2) is valid and α_j, β_j are the coefficients of the fixed stepsize method, we will have, for some $\gamma > 0$,

$$|\alpha_{jn} - \alpha_j| \leqslant \gamma h_n, \quad |\beta_{jn} - \beta_j| \leqslant \gamma h_n \quad (n \geqslant 0, 1 \leqslant j \leqslant k). \tag{5.3}$$

5.2. Variable stepsize one-leg methods

Let ρ_n, σ_n be the generating polynomials containing the coefficients α_{jn}, β_{jn}, respectively. Then the one-leg scheme is given by

$$\rho_n(E)u_n = h_{n+k}f(\sigma_n(E)t_n, \sigma_n(E)u_n) \tag{5.4}$$

and the global errors $\epsilon_n = u(t_n) - u_n$ satisfy

$$\epsilon_{n+k} = \sum_{j=1}^{k} \psi_{jn}(Z_n)\epsilon_{n+k-j} + \delta_n. \tag{5.5}$$

Here $Z_n = h_{n+k}J_n$ with J_n as in (2.4), $\psi_{jn}(z) = -(\alpha_{kn} - \beta_{kn}z)^{-1}(\alpha_{k-j,n} - \beta_{k-j,n}z)$ and $\delta_n = (\alpha_{kn}I - \beta_{kn}Z_n)^{-1}(r_n + Z_nq_n)$ with

$$r_n = \rho_n(E)u(t_n) - h_{n+k}u'(\sigma_n(E)t_n), \quad q_n = u(\sigma_n(E)t_n) - \sigma_n(E)u(t_n).$$

Again, these errors are independent of the stiffness. If the one-leg method has order q, we have as in Section 3 (cf. also [5])

$$r_n = O(h_n^{q+1}), \quad q_n = h_n^q B_{qn}u^{(q)}(t_n) + O(h_n^{q+1}) \tag{5.6a}$$

where

$$B_{qn} = \frac{1}{q!}\{[\sum_{j=0}^{k} \beta_{jn}h_n^{-1}(t_{n+j} - t_n)]^q - \sum_{j=0}^{k} \beta_{jn}[h_n^{-1}(t_{n+j} - t_n)]^q\}. \tag{5.6b}$$

In order to prove convergence we need a stability condition like (2.8) where $R_n = R_n(Z_n)$, $R_n(\cdot)$ being formed by the rational functions ψ_{jn}. On arbitrary grids, sufficient conditions are known for some relatively small classes of schemes, see [7], [16]. If we have a smooth stepsize variation (5.2), then Lemma 2.3 and (5.3) show that $\|R_n(Z_n) - R(Z_n)\| = O(h_n)$ where $R(Z_n)$ is the companion matrix of the fixed stepsize scheme. Consequently, stability for variable steps then already follows from fixed stepsize stability by a well-known perturbation argument, see, for example, [17], p. 58 (the proof in this reference for constant h can be easily modified to fit into our situation by noting that (5.2) will cause h/h_n to be uniformly bounded in n).

In the remainder of this section it will simply be assumed that the method, the grid and the class of problems are such that (5.4) is stable. As before, we obtain for stiff systems the local error estimate $|\delta_n| = O(h_n^q)$, and in the standard way stability leads to the global, stiffness independent bound $|\epsilon_n| = O(h^{q-1})$ for all $n \geqslant 0$.

If the grid is sufficiently regular, convergence with order q can also be proved. Proceeding as in the proof of Theorem 3.3 it follows that $|\epsilon_n| = O(h^q)$ (independently of the stiffness) under the assumption

$$\sum_{n=0}^{N-1} |q_{n+1} - q_n| \leqslant Dh^q$$

for some $D > 0$. This will hold for arbitrary, smooth solutions u iff there is a $D' > 0$ such that

$$\sum_{n=0}^{N-1} |Q_{n+1} - Q_n| \leqslant D'h^q, \quad Q_n = h_n^q B_{qn}. \tag{5.7}$$

As a consequence, it can be shown that (5.2) is sufficient for order q convergence. For specific methods this can be proved under assumptions on the stepsizes less restrictive than (5.2).

EXAMPLE 5.1. For the implicit midpoint rule

$$u_{n+1} - u_n = h_{n+1} f(t_n + \tfrac{1}{2}h_{n+1}, \tfrac{1}{2}u_n + \tfrac{1}{2}u_n),$$

we have $q = 2$ and $Q_n = -\tfrac{1}{8}h_{n+1}^2$. The method is stable for arbitrary problems (1.1) where f satisfies (1.4). From (5.7) a result of [13] is reobtained: the method is convergent of order 2, independently of the stiffness, provided that

$$\sum_{n=0}^{N-1} |h_{n+1}^2 - h_n^2| = O(h^2). \tag{5.8}$$

This condition is satisfied if the number of sign changes in the sequence $\{h_{n+1} - h_n\}$ is bounded by a fixed, finite number. This seems a reasonable assumption for numerical codes, where h_n will be somehow related to the smoothness of solutions near t_n. It was also shown in [13] that (5.8) is necessary to guarantee second order convergence; for stepsize sequences like $h_n = \tfrac{1}{2}h$ (for n odd), $h_n = h$ (for n even) the order will reduce to 1. □

EXAMPLE 5.2. The variable stepsize formulation of the Adams-type method CA2 of [15] (cf. Example 3.1) reads

$$u_{n+2} - u_{n+1} = h_{n+2} f(t_{n+1} + \tfrac{1}{2}h_{n+2}, \tfrac{1}{2}\frac{2 + \omega_{n+1}}{1 + \omega_{n+1}}u_{n+2} + \tfrac{1}{2}\frac{\omega_{n+1}}{1 + \omega_{n+1}}u_n).$$

This method is A_0-contractive in the maximum norm for arbitrary stepsize sequences, and we have order $q = 2$ and $Q_n = -\tfrac{1}{8}h_{n+2}^2 - \tfrac{1}{4}h_{n+1}h_{n+2}$. Hence

$$|Q_{n+1} - Q_n| \leqslant \frac{1}{8}|h_{n+2}^2 - h_{n+1}^2| + \frac{1}{4}h_{n+1}|h_{n+2} - h_n| \leqslant$$

$$\leqslant \frac{3}{8}|h_{n+2}^2 - h_{n+1}^2| + \frac{1}{4}|h_{n+1}^2 - h_n^2|.$$

Thus we see that this method is also convergent of order 2, independently of the stiffness, if the grid refinement is such that (5.8) is satisfied. $\square$

5.3. Variable stepsize linear multistep methods

The linear multistep method for variable stepsizes is given by

$$\rho_n(E)u_n = h_{n+k}\sigma_n(E)f(t_n, u_n). \tag{5.9}$$

In the same way as for constant stepsizes we obtain an error recursion

$$\sum_{j=0}^{k}(\alpha_{jn}I - \beta_{jn}h_{n+k}J_{n+j})\epsilon_{n+j} = p_n \tag{5.10}$$

with J_n defined by (4.5) and

$$p_n = \rho_n(E)u(t_n) - h_{n+k}\sigma_n(E)u'(t_n) = O(h_n^{p+1}) \tag{5.11}$$

where p is the order of the method.

It is well known, see for instance [20], pp. 181, 182, and [15], that unfavourable combinations of strongly varying h_n and J_n may cause instability for linear multistep methods. Therefore, convergence can only be proved if either the variation of the stepsizes or the variation of the J_n is limited.

If $J_n = J$ is constant, we obtain from (5.10) a recursion of the type (5.5) with $Z_n = h_{n+k}J$ and $\delta_n = (\alpha_{kn}I - \beta_{kn}Z_n)^{-1}p_n$, leading to convergence with order p if the corresponding one-leg scheme is stable for variable stepsizes. For mildly varying J_n a similar result can be derived, again by using a perturbation argument to prove stability from the constant J case.

If, on the other hand, the stepsize variation is restricted we consider the transformed errors $\hat{\epsilon}_n = (\alpha_{k,n-k}I - \beta_{k,n-k}Z_n)\epsilon_n$ with $Z_n = h_nJ_n$. To make this meaningful for $n < k$ we may put $h_0 = h_1$ and $\alpha_{kl} = \alpha_{k0}, \beta_{kl} = \beta_{k0}$ if $l < 0$. Then it follows from (5.10) that

$$\hat{\epsilon}_{n+k} = \sum_{j=1}^{k}\tilde{\psi}_{jn}(Z_{n+k-j})\hat{\epsilon}_{n+k-j} + p_n \tag{5.12}$$

with rational functions $\tilde{\psi}_{jn}(z) = -(\alpha_{k,n-j} - z\beta_{k,n-j})^{-1}(\alpha_{k-j,n} - z\beta_{k-j,n}h_{n+k}/h_{n+k-j})$. Assuming (5.2), we obtain as in Lemma 2.3

$$|\tilde{\psi}_{jn}(Z_{n+k-j}) - \psi_j(Z_{n+k-j})| = O(h_n)$$

and thus stability, and consequently convergence with order p, again follow from stability for the constant stepsize case, treated in Section 4. As for the one-leg methods, this can be proved for specific linear multistep methods under assumptions on the stepsize variation somewhat less restrictive than (5.2), but in contrast to the one-leg methods we may now have divergence for irregular grids, see for example the results of [13] for the trapezoidal rule.

6. INTERPOLATED ONE-LEG APPROXIMATIONS

Consider a linear multistep method (1.3) with order p and its one-leg counterpart (1.2) with order q, and suppose that $p > q$. As we saw in the previous sections, the one-leg method will be more robust w.r.t. variable step sizes and initial errors, but on the other hand the linear multistep method will often be the more accurate one. As a compromise, one might consider interpolated one-leg approximations $\bar{u}_n = \sigma(E)u_n$ and errors $\bar{\epsilon}_n = u(t_n) - \bar{u}_n$, with the vectors u_n computed from the one-leg scheme. The approximations $\bar{u}_n$ will be stable whenever the u_n are so, while for constant stepsizes the

$\bar{\epsilon}_n$ can be shown to satisfy an error recursion with local error $\bar{p}_n = O(h^{p+1})$ (and thus, as said in [6], from the equivalence point of view the quantity $\bar{\epsilon}_n$ might be a more adequate measure of the global error than the customary one).

In this section it will be proved that even very smooth stepsize variations will cause already a decrease in the order of $\bar{p}_n$ to $O(h^{q+1})$, resulting in convergence of $\bar{u}_n$ to $u(t_n)$ with $O(h^q)$ only, the same as for the original one-leg approximations. A closely related fact was claimed in [6] for a class of high order methods with $q = k$; this was based on a somewhat different local error and on more complicated arguments. Besides, it will be shown in this section that also for constant stepsizes the approximations $\bar{u}_n$ will in general not converge with order p, due to the fact that the initial errors $\bar{\epsilon}_0, \bar{\epsilon}_1, \ldots, \bar{\epsilon}_{k-1}$ will be of $O(h^q)$ at best.

Hence, as far as the order of convergence is concerned the interpolated values have no general advantage over the one-leg approximations. It should be noted that this does not exclude the possibility that the errors $\bar{\epsilon}_n$ are substantially smaller than the ϵ_n, due to the global error constants involved. This is to be expected for (nearly) constant stepsizes and damping of initial errors.

Throughout this section we will assume that $p > q$. Comparing the order conditions (3.3), (4.4) it easily follows that (3.5) cannot hold under this assumption. Consequently, the errors q_n will be of $O(h^q)$ exactly, provided that $u^{(q)}(t_n) \neq 0$ (see (3.2)).

In the following we consider a geometrical stepsize sequence h_n, with constant stepsize ratios $\omega_n = \omega$. Then the coefficients of the multistep methods will not vary with n, but they still may depend on ω. Consider again the error relations for the one-leg method

$$\bar{\epsilon}_n = \sigma(E)\epsilon_n + q_n, \quad \rho(E)\epsilon_n = Z_n\bar{\epsilon}_n + r_n \tag{6.1}$$

(cf. (2.3)) were now $q_n = u(\bar{t}_n) - \sigma(E)u(t_n)$, $r_n = \rho(E)u(t_n) - h_{n+k}u'(t_n)$ and $Z_n = h_{n+k}J_n$. By premultiplying the error relations in (6.1) by $\rho(E)$, $\sigma(E)$, respectively, and using the fact that $\rho(E)$ and $\sigma(E)$ commute, it is easy to eliminate the ϵ_n, yielding

$$\rho(E)\bar{\epsilon}_n = \sigma(E)Z_n\bar{\epsilon}_n + \bar{p}_n \tag{6.2}$$

with

$$\bar{p}_n = \rho(E)q_n + \sigma(E)r_n = \rho(E)u(\bar{t}_n) - \sigma(E)[h_{n+k}u'(t_n)]. \tag{6.3}$$

If the stepsizes are constant, $\omega = 1$, then

$$\bar{p}_n = \rho(E)u(\bar{t}_n) - h\sigma(E)u'(t_n)$$

which is just the local error of the linear multistep scheme on the shifted grid $\{\bar{t}_n\}$. It follows then that $|\bar{p}_n| = O(h^{p+1})$ and from (6.2) we can conclude that $|\bar{\epsilon}_n| = O(h^p)$ for all n provided that the initial errors are sufficiently small. However, the initial errors involved here are $\bar{\epsilon}_0, \bar{\epsilon}_1, \ldots, \bar{\epsilon}_{k-1}$. Suppose that the starting values of the one-leg scheme are exact, $\epsilon_0 = \epsilon_1 = \cdots = \epsilon_{k-1} = 0$. Then (6.1) implies

$$\bar{\epsilon}_0 = \sigma(E)\epsilon_0 + q_0 = \beta_k\epsilon_k + q_0,$$

and from (2.6) we know that

$$\epsilon_k = (\alpha_k I - \beta_k Z_0)^{-1}(r_0 + Z_0 q_0).$$

Hence

$$\bar{\epsilon}_0 = (\alpha_k I - \beta_k Z_0)^{-1}(\alpha_k q_0 + \beta_k r_0).$$

From (2.10) it follows that

$$|\bar{\epsilon}_0| \leq S_1(|\alpha_k q_0| + |\beta_k r_0|) = O(h^q)$$

and if $|Z_0| = O(1)$ this estimate is sharp. Note that since convergence of the one-leg scheme (see

Section 3) implies $|\epsilon_n| = O(h^q)$ we already know

$$|\bar{\epsilon}_n| \leq |\sigma(E)\epsilon_n| + |q_n| = O(h^q)$$

for all n. In view of the above this error bound is optimal.

In this discussion we have not considered the fact that there may be damping for $h > 0$ bounded away from zero. Therefore, in an actual computation the conclusion that $|\epsilon_n|$ and $|\bar{\epsilon}_n|$ will be of similar size might be too pessimistic for the $|\bar{\epsilon}_n|$. On the other hand, in actual computations constant stepsizes are hardly ever used. It will be shown in the following that a slight deviation from the constant stepsize case leads to $|\bar{p}_n| = O(h^{q+1})$, instead of $|\bar{p}_n| = O(h^{p+1})$ as for constant h.

Consider a stepsize sequence $h_n = (1+h_0)^n h_0$, i.e.,

$$\omega = 1+h_0. \tag{6.4}$$

Note that on $[0,T]$ the ratio of maximal and minimal stepsizes $h/h_0 = O(1)$. Let $\bar{h}_n = \bar{t}_n - \bar{t}_{n-1}$ for $n \geq 1$. We have

$$\bar{p}_n = \{\rho(E)u(\bar{t}_n) - \bar{h}_{n+k}\sigma(E)u'(\bar{t}_n)\} + \{\bar{h}_{n+k}\sigma(E)u'(\bar{t}_n) - \sigma(E)[h_{n+k}u'(\bar{t}_n)]\}.$$

The first term on the right hand side is the local error of the linear multistep scheme on the grid $\{\bar{t}_n\}$, and thus it is of $O(\bar{h}_n^{p+1})$. Hence

$$\bar{p}_n = \sum_{j=0}^{k} \beta_j (\bar{h}_{n+k} - h_{n+k+j})u'(\bar{t}_{n+j}) + O(\bar{h}_n^{p+1}).$$

Taylor expansion of $u'(\bar{t}_{n+j})$ around $\bar{t}_n$ for $j = 1,2,...,k$ yields

$$\bar{p}_n = \sum_{i=1}^{p} \bar{C}_i \bar{h}_{n+1}{}^i u^{(i)}(\bar{t}_n) + O(\bar{h}_n^{p+1}) \tag{6.5}$$

were

$$\bar{C}_i = \frac{1}{(i-1)!} \sum_{j=0}^{k} \beta_j (\bar{h}_{n+k} - h_{n+k+j})(\bar{t}_{n+j} - \bar{t}_n)^{i-1} \bar{h}_{n+1}^{-i} \quad (1 \leq i \leq p). \tag{6.6}$$

Since $\bar{h}_n = \sigma(E)h_n$ and $h_{n+j} = \omega^j h_n$, it follows that $\bar{h}_n = \sigma(\omega)h_n$. Therefore $\bar{h}_{n+j} = \omega^j \bar{h}_n$ and

$$\bar{C}_i = \frac{1}{(i-1)!} \sum_{j=0}^{k} \beta_j (\bar{h}_{n+k} - \frac{1}{\sigma(\omega)}\bar{h}_{n+k+j})(\bar{h}_{n+1} + ... + \bar{h}_{n+j})^{i-1} \bar{h}_{n+1}^{-i} =$$

$$= \frac{1}{(i-1)!\sigma(\omega)} \omega^{k-1} \sum_{j=0}^{k} \beta_j (\sigma(\omega) - \omega^j)(1 + \omega + ... + \omega^{j-1})^{i-1}.$$

Let

$$\theta_j = 1 + \omega + ... + \omega^{j-1} \quad \text{and} \quad m_i(\omega) = \sum_{j=0}^{k} \beta_j \theta_j^i.$$

Using $\sigma(1) = 1$, $\theta_j = (\omega-1)^{-1}(\omega^j - 1)$ for $\omega \neq 1$, it follows by some simple calculations that

$$\sigma(\omega) = 1 + (\omega-1)m_1(\omega), \quad \sigma(\omega) - \omega^j = (\omega-1)[m_1(\omega) - \theta_j],$$

and consequently

$$\bar{C}_i = \frac{1}{(i-1)!\sigma(\omega)} \omega^{k-1}(\omega-1)[m_1(\omega)m_{i-1}(\omega) - m_i(\omega)] \quad (1 \leq i \leq p). \tag{6.7}$$

Note that $m_i(1) = \mu_i$, with μ_i being defined by (3.1). The order conditions (3.3) imply $m_1(1)^i = m_i(1)$ $(0 \leq i \leq q-1)$, and since it is assumed that the one-leg method remains of order q for nonconstant stepsizes, it follows in the same way that

$$m_1(\omega)^i = m_i(\omega) \quad (0 \leq i \leq q-1). \tag{6.8}$$

As said before, $p > q$ implies that (3.5) does not hold, i.e.,

$$m_1(1)^q \neq m_q(1). \tag{6.9}$$

Hence we have $\overline{C}_1 = \cdots = \overline{C}_{q-1} = 0$, but

$$\overline{C}_q = \frac{1}{(q-1)!\sigma(\omega)} \omega^{k-1}(\omega-1)[m_1(\omega)^q - m_q(\omega)] = \tag{6.10}$$

$$= \frac{h_0}{(q-1)!}[m_1(1)^q - m_q(1)] + O(h_0^2).$$

Returning to (6.5) we thus see that for our geometrical stepsize sequence we have $|\overline{p}_n| = O(h^{q+1})$ only. From the recursion (6.2) it can be concluded that the erros $\overline{\epsilon}_n$ will be of order q at most (there is no cancellation of local errors in general, for example if $Z_n \simeq 0$ for all n).

REFERENCES

[1] K. BURRAGE, W.H. HUNDSDORFER and J.G. VERWER, *A study of B-convergence of Runge-Kutta methods*, Computing 36 (1986), 17-34.

[2] M. CROUZEIX and P.A. RAVIART, *Approximation d'équations d'évolution linéaires par des méthodes multipas*, In: *Etude des grandes systèmes*, J.L. Lions and G.I. Marchouk (eds.), Dunod, Paris, 1978.

[3] G. DAHLQUIST, *Error analysis for a class of methods for stiff nonlinear initial value problems*, Proc. Num. Anal. Dundee 1975, G.A. Watson (ed.), Lecture Notes in Mathematics 506, Springer Verlag, Berlin, 1976.

[4] G. DAHLQUIST, *G-stability is equivalent to A-stability*, BIT 18 (1978), 384-401.

[5] G. DAHLQUIST, *Some properties of linear multistep and one-leg methods for ODEs*, Proc. 1979 SIGNUM meeting on numerical ODEs, R.D. Skeel (ed.), Report R79/963, Dept. of Comp. Sc., Univ. of Illinois at Urbana-Champaign (1979).

[6] G. DAHLQUIST, *On one-leg multistep methods*, SIAM J. Numer. Anal. 20 (1983), 1130-1138.

[7] G. DAHLQUIST, W. LININGER and O. NEVANLINNA, *Stability of two-step methods for variable integration steps*, SIAM J. Numer. Anal. 20 (1983), 1071-1085.

[8] K. DEKKER and J.G. VERWER, *Stability of Runge-Kutta methods for stiff nonlinear differential equations*, North-Holland, Amsterdam, 1984.

[9] R. FRANK, J. SCHNEID and C.W. UEBERHUBER, *Order results for implicit Runge-Kutta methods applied to stiff systems*, SIAM J. Numer. Anal. 22 (1985), 497-514.

[10] E. HAIRER, CH. LUBICH and M. ROCHE, *Error of Runge-Kutta methods for stiff problems studied via differential algebraic equations*, BIT 28 (1988), 678-700.

[11] E. HAIRER, S.P. NØRSETT and G. WANNER, *Solving ordinary differential equations I - nonstiff problems*, Springer Verlag, Berlin, 1987.

[12] E. HAIRER and G. WANNER, *Solving ordinary differential equations II - stiff and differential algebraic problems*, in preparation.

[13] J.F.B.M. KRAAIJEVANGER, *B-convergence of the implicit midpoint rule and the trapezoidal rule*, BIT 25 (1985), 652-666.

[14] O. NEVANLINNA, *Matrix valued versions of a result of von Neumann with an application to time discretization*, J. Comp. Appl. Math. 12 & 13 (1985), 475-489.

[15] O. NEVANLINNA and W. LININGER, *Contractive methods for stiff differential equations*, Part I, BIT 18 (1978), 457-474; Part II, BIT 19 (1979), 53-72.

[16] O. NEVANLINNA and F. ODEH, *Multiplier techniques for linear multistep methods*, Numer. Funct. Anal. Optimiz. 3 (1981), 377-423.

[17] R.D. RICHTMEYER and K.W. MORTON, *Difference methods for initial-value problems*, Interscience Publ., New York, 1967.

[18] J. SAND, *Choices in contractivity theory*, Appl. Num. Math. 5 (1989), 105-116.

[19] M.N. SPIJKER, *Contractivity in the numerical solution of initial value problems*, Numer. Math. 42

(1983), 271-290.

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