stichting
mathematisch
centrum

MC

AFDELING NUMERIEKE WISKUNDE
(DEPARTMENT OF NUMERICAL MATHEMATICS)

NW 147/83 JANUARI

A.O.H. AXELSSON & J.G. VERWER

BOUNDARY VALUE TECHNIQUES FOR INITIAL VALUE PROBLEMS IN ORDINARY DIFFERENTIAL EQUATIONS

Preprint

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, Kruislaan 413, Amsterdam, The Netherlands.

The Mathematical Centre, founded 11th February 1946, is a non-profit institution for the promotion of pure and applied mathematics and computer science. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

1980 Mathematical subject clasification: 65L05, 65L10

ACM Computer review categories: 5.17

Boundary value techniques for initial value problems in ordinary differential equations *)

by

A.O.H. Axelsson **) & J.G. Verwer

ABSTRACT

The numerical solution of initial value problems in ordinary differential equations by means of boundary value techniques is considered. We discuss a finite difference method which was already investigated by Fox in 1954 and Fox and Mitchell in 1957. Hereby we concentrate on explaining the fundamentals of the method because for initial value problems the boundary value method seems to be fairly unknown. We further propose and discuss new Galerkin methods for initial value problems along the lines of the boundary value approach.

KEY WORDS & PHRASES: *Numerical analysis, Initial value problems for ordinary differential equations, Stiffness, Boundary value techniques*

---

*) This report will be submitted for publication elsewhere.

**) Mathematical Institute, Catholic University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

1. INTRODUCTION

Traditionally, methods used for the numerical integration of initial value problems in ordinary differential equations

$$\dot{y}(x) = f(x,y(x)), \quad a \le x \le b, \quad y(a) \text{ given}, \tag{1.1}$$

are *step-by-step methods*. Familiar step-by-step methods, which are also called *forward step methods*, are the Runge-Kutta and linear multistep methods (see e.g. Henrici [12], Lambert [17], Stetter [24]). The latter, in its most simple form, is defined by the so-called k-step formula

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f(x_{n+j},y_{n+j}), \quad \alpha_j,\beta_j,\ h \in \mathbb{R},\ h > 0,\ k \in \mathbb{N}^+, \tag{1.2}$$

where y_{n+j} represents the approximation to the exact solution value $y(x_{n+j})$ defined by (1.1). The positive real h is called the stepsize. Assuming that h is constant, it is given by $h = (b-a)/N$, N being some positive integer. The points x_{n+j} are called grid points and belong to the uniform grid

$$G_h = \{x_j : x_j = a + jh,\ j = 0(1)N\}. \tag{1.3}$$

In the forward step approach the numerical solution is obtained by stepping through this grid in the direction from a to b, i.e. given approximations y_{n+j} for some integer n and $j = 0(1)k-1$, the approximation y_{n+k} at the next grid point x_{n+k} is computed by solving (1.2) for y_{n+k}. In fact, all results on convergence and numerical stability which emanate from the pioneering work of Dahlquist [5] are based on this forward step application. Nowadays there exist very efficient linear multistep computer codes which automatically start the integration process and vary the stepsize and the order of accuracy of the multistep method according to the smoothness of the problem (see e.g. Hindmarsh [13], Shampine & Gordon [23]).

In this paper we will tackle the numerical solution of (1.1) in a completely different way then in the step-by-step approach. For its numerical solution we will consider (1.1) as a two-point boundary value problem with a given value at the left endpoint and an implicitly defined value, by the

equation $\dot{y}(x) = f(x,y(x))$, at the right endpoint. In this approach formula (1.2) ought to be considered as a *finite difference formula* as is the practice in the numerical solution of genuine two-point boundary value problems for systems of first order differential equations (see Keller [15,16]). One of the aims of this *boundary value approach* is to circumvent the known *Dahlquist-barriers* on convergence and stability which are a direct consequence of the step-by-step application of (1.2). In this respect boundary value methods for (1.1) bear a relationship with the iterative algorithms of Cash [4] for the stable solution of recurrence relations and with Olver's algorithm [19,20].

Up to now boundary value methods for initial value problems have hardly been discussed in the numerical literature. Perhaps because the step-by-step application of formulas of type (1.2) is invariably easier to perform. As far as we know the first contributions have been delivered by Fox [8] in 1954 and Fox & Mitchell [9] in 1957. They discuss a simple finite difference formula for (1.1) and for the derived second order equation

$$\ddot{y}(x) = g(x,y(x)) = \frac{\partial f}{\partial x}(x,y(x)) + \frac{\partial f}{\partial y}(x,y(x))f(x,y(x)). \tag{1.4}$$

A feature of the boundary value method is that all approximations on the grid G_h are generated simultaneously. In 1964 Axelsson [1] proposed a quadrature type method for the integrated form of (1.1) which also computes all approximations over the interval [a,b] simultaneously. This method has been called a global integration method. It is best characterized as a huge implicit Runge-Kutta method which performs just one step with stepsize b-a. A special feature of this global method is that the global errors at the end of the interval are particularly small, even when the problem is mathematically unstable. On the other hand, the errors of step-by-step methods have a tendency to grow due to accumulation at every step, especially when the problem itself is unstable.

Two recent contributions on boundary value methods for initial value problems emanate from Rolfes [21] and Rolfes & Snyman [22]. They consider a finite difference method which has also been proposed by Fox [8] and apply it to stiff equations. Rolfes and Snyman report that the finite diffe-

rence method performs satisfactorily on stiff problems. Fox considered non-stiff equations, but was not satisfied with the method because of an oscillating error behaviour which prevents the application of difference correction for improving the accuracy.

The present contribution consists of two parts. The first part deals with *finite difference methods* while the second one is devoted to *Galerkin methods*. When discussing boundary value techniques for initial value problems it is of course obvious to consider Galerkin methods because of their use in the numerical solution of genuine two point boundary value problems. We shall comment on a relation between the two approaches. To a certain extent our paper has an expository nature. We have concentrated on describing the fundamentals of the boundary value method because for initial value problems this method seems to be fairly unknown.

2. A FINITE DIFFERENCE BOUNDARY VALUE METHOD

In this section we will describe the general principles of the boundary value approach from simple examples of finite difference formulas. We will concentrate on a combination of the explicit midpoint rule with Backward Euler.

2.1. *Outline of the method*

Consider the initial value problem (1.1). Let us discretize the differential equation $\dot{y} = f(x,y)$ on the grid (1.3) by means of the *explicit midpoint rule*

$$y_{n+1} - y_{n-1} - 2hf(x_n,y_n) = 0. \tag{2.1}$$

When we apply (2.1) as a step-by-step method we need two initial values, one at the left endpoint $x = a$ and one at $x = a + h$. The first initial value is known from the problem, while the second one has to be computed by another method. When we apply (2.1) as a boundary value method it is applied at each of the points $x_n \in G_h$ for $n = 1(1)N-1$. In addition to the initial value at the left endpoint $x = a$, we now need a boundary condition at the right endpoint $x = b$. For that purpose one can use the most simple *backward difference formula* (Backward Euler)

$$y_N - y_{N-1} - hf(x_N,y_N) = 0. \tag{2.2}$$

Thus we arrive at the discrete boundary value problem

$$\begin{aligned} &y_0 \text{ given,} \\ &y_{n+1} - y_{n-1} - 2hf(x_n,y_n) = 0, \quad n = 1(1)N-1, \\ &y_N - y_{N-1} - hf(x_N,y_N) = 0, \end{aligned} \tag{2.3}$$

whose solution values $y_1,\ldots,y_N$ must be generated simultaneously. Since f may be non-linear in y, the discrete problem (2.3) must be solved by iteration. A Newton-type iteration is feasible because of the *tridiagonal structure* (block-tridiagonal for systems).

As an alternative for formula (2.2), we mention the more accurate *trapezoidal rule*

$$y_N - y_{N-1} - \tfrac{1}{2}hf(x_{N-1},y_{N-1}) - \tfrac{1}{2}hf(x_N,y_N) = 0, \tag{2.4}$$

or the *second order backward difference formula*

$$y_N - \tfrac{4}{3}y_{N-1} + \tfrac{1}{3}y_{N-2} - \tfrac{2}{3}hf(x_N,y_N) = 0. \tag{2.5}$$

The use of (2.4) or (2.5) instead of (2.2) does not increase the order of accuracy of the method. Both combinations are of order two. Normally, method (2.3) will be somewhat less accurate. Convergence questions are further discussed in section 2.3.

Combination (2.1), (2.5) has already been proposed by Fox [8] and Fox & Mitchell [9]. Rolfes [21] and Rolfes & Snyman [22] have applied this combination to stiff problems. A slight disadvantage is that by using (2.5) the tridiagonal coupling is lost. This might be overcome, however, by eliminating y_{N-2} from (2.5) and the particular equation

$$y_N - y_{N-2} - 2hf(x_{N-1},y_{N-1}) = 0. \tag{2.6}$$

This yields

$$
\begin{aligned}
& y_0 \text{ given,} \\
(2.7) \qquad & y_{n+1} - y_{n-1} - 2hf(x_n,y_n) = 0, \quad n = 1(1)N-1, \\
& \tfrac{4}{3}(y_N - y_{N-1}) - \tfrac{2}{3}hf(x_{N-1},y_{N-1}) - \tfrac{2}{3}hf(x_N,y_N) = 0,
\end{aligned}
$$

which is just method (2.1), (2.4).

Finally we observe that methods like (2.3) can be directly applied to problems with periodic solutions. The last line of (2.3) then should read $y_N = y_0$. In what follows we concentrate on the pure initial value problem.

2.2. *The test model*

In this section we consider the standard test model

$$
(2.8) \qquad \dot{y} = \delta y, \ \delta \in \mathbb{C}, \ a \le x \le b, \ y(a) \text{ given.}
$$

We observe that this model plays an important role in the stability of step-by-step integration methods. The notion of absolute stability (see e.g. [17]) is based on this simple problem which is also very suitable for becoming acquainted with the boundary value approach and for comparison with the step-by-step approach. In section 2.3 the model is linked with a constant coefficient linear system. We will concentrate on method (2.3), i.e. explicit midpoint combined with Backward Euler.

Our discrete boundary value problem (2.3) now reads

$$
\begin{aligned}
& y_0 = y(a) \\
(2.9) \qquad & y_{n+1} - y_{n-1} - 2zy_n = 0, \quad z = h\delta, \quad n = 1,\ldots,N-1, \\
& y_N - y_{N-1} - zy_N = 0,
\end{aligned}
$$

i.e. we have to solve the linear algebraic system

$$
(2.10) \qquad A(z)\, Y = R,
$$

where $Y = [y_1,\ldots,y_N]^T$, $R = [y(a),0,\ldots,0]^T$ and $A(z)$ is given by

$$A(z) = \begin{pmatrix} -2z & 1 & & & \\ -1 & -2z & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & -2z & 1 \\ & & & -1 & 1-z \end{pmatrix} . \tag{2.11}$$

The first question which arises is, for which z-values is Y a well-defined vector of approximations y_n to $e^{nz}y(a)$, n = 1,...,N, i.e. for which z-values is A(z) regular. In what follows we call z *a regular point for* A(z) if A(z) is regular. Otherwise z is called a *singular point*.

Define $\tilde{A}(z)$ = diag (1,...,1,2) A(z), and write $\tilde{A}(z) = E - 2zI$, i.e.

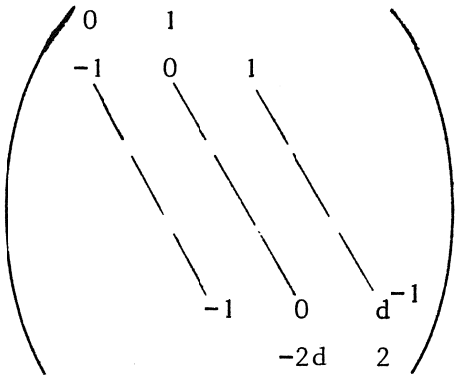
$$E = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -2 & 2 \end{pmatrix} . \tag{2.12}$$

A(z) is singular, iff $\tilde{A}(z)$ is singular. Hence we can use $\tilde{A}(z)$, and in turn the constant matrix E to find the singular points for A(z). Obviously, the location of the eigenvalues λ_j of E is decisive, since z is a singular point, iff $z = \lambda_j/2$.

LEMMA 1. *All eigenvalues* λ_j *of* E *satisfy* $0 < \mathrm{Re}(\lambda_j) \leq 2$, $-2 \leq \mathrm{Im}(\lambda_j) \leq 2$.

PROOF. The inequality $-2 \leq \mathrm{Im}(\lambda_j) \leq 2$ is a direct consequence of Gersgorin's cirle theorem. To prove the inequality for the real part we first perform the similarity transformation

$$\tilde{E} = \mathrm{diag}(1,\ldots,1,d)E\ \mathrm{diag}(1,\ldots,1,d^{-1}) =$$



which leaves the spectrum invariant. Let λ and μ be the real and imaginary parts of an arbitrary eigenvalue and let u and v be the real and imaginary parts of the corresponding eigenvector. Then we easily derive

$$(2.13) \qquad \tfrac{1}{2}[u^T(\tilde{E} + \tilde{E}^T)u + v^T(\tilde{E} + \tilde{E}^T)v] = \lambda(u^Tu+v^Tv).$$

Now we take $d = 1/\sqrt{2}$ for which $\frac{1}{2}(\tilde{E}+\tilde{E}^T) = \mathrm{diag}(0,\ldots,0,2)$. Hence

$$0 \le \tfrac{1}{2}u^T(\tilde{E}+\tilde{E}^T)u \le 2u^Tu, \text{ all } u \in \mathbb{R}^N,$$

so that $0 \le \lambda \le 2$. Finally, assume $\lambda = 0$ and let u_i, v_i denote the i-th component of u and v, respectively. From (2.13) it then follows that $u_N = v_N = 0$. By using the relations $\tilde{E}u = -\mu v$, $\tilde{E}v = \mu u$ and the specific form of $\tilde{E}$ it is now easy to verify that $u_i = v_i = 0$, all i = 1(1)N. This leads to a contradiction showing that $\lambda \neq 0$. □

We thus have the following result:

THEOREM 2. *All singular points* z *for* A(z) *satisfy* $0 < \mathrm{Re}(z) \le 1$, $-1 \le \mathrm{Im}(z) \le 1$. □

We cannot determine the eigenvalues of E explicitly. Note that if in E the last row elements are replaced by -1 and 0, respectively, the eigenvalues become $2i\cos(j\pi/(N+1))$, j = 1(1)N. Figure 2 shows all numerically computed eigenvalues of E and E/2h for some values of $h = N^{-1}$. The eigenvalues of E/2h play an important role in the convergence analysis (cf. section 2.3). We see that when N increases two eigenvalues of E approach $\pm 2i$. This

means that for N large, the points $\pm i$ will act numerically as singular points for A(z).

The second question we now wish to discuss is, how well are the decaying exponentials e^{nz} approximated. From diagonal dominance properties it easily follows that for $Re(z) << 0$ (stiff eigenvalues) $|y_n|$ is an excellent approximation to $|e^{nz}y(a)|$. More precisely, if $z \neq 0$ is a regular point, then (2.10) can be rewritten as

$$Y = -(2z)^{-1}(I-(2z)^{-1}E)^{-1}R,$$

which implies

$$y_1 = -(2z)^{-1}y(a) + O(|z|^{-2}),\ y_n = O(|z|^{-2}),\ n = 2(1)N,\ |z| \to \infty.$$

Observe that the method cannot approximate positive exponentials if $Re(z) >> 0$. Roughly spoken, for $|Re(z)|$ large, the approximations for the negative and positive exponential e^{nz} are of the same magnitude.

To get more insight in the question how decaying exponentials are approximated, we now proceed with the analytical solution of the recurrence equation $y_{n+1} - y_{n-1} - 2zy_n = 0$ defined by the explicit midpoint rule when applied to test-model (2.8):

$$y_n = C_1\mu_1^n + C_2\mu_2^n,\ n = 1,2,\ldots,N, \tag{2.14}$$

where $\mu_1 = z + \sqrt{z^2+1}$, $\mu_2 = z - \sqrt{z^2+1}$ and C_1, C_2 are constants to be determined by boundary conditions. Note that $\mu_1 = e^z + O(z^3)$, $z \to 0$, whereas μ_2 has no relationship with e^z, i.e. μ_2 is the parasitic root.

Solution (2.14) can be adapted to our discrete problem (2.9) via C_1 and C_2 by requiring

$$\begin{aligned} &C_1 + C_2 = y(a), \\ &(1-z)(C_1\mu_1^N + C_2\mu_2^N) = C_1\mu_1^{N-1} + C_2\mu_2^{N-1}. \end{aligned} \tag{2.15}$$

Solving for C_1 and C_2 yields $C_2 = \delta C_1$, $C_1 = y(a)/(1+\delta)$, where

$$(2.16) \qquad \delta = \left(\frac{-1}{\mu_2^2}\right)^{N-1} \eta, \quad \eta = \frac{1-\mu_1(1-z)}{(1-z)\mu_2 - 1},$$

and where it is assumed that $\mathrm{Re}(\delta) \neq -1$. $\mathrm{Re}(\delta) = -1$ means singularity of the 2×2 system (2.15). Like for system (2.10), one thus must distinguish singular and regular points z. We emphasize that the set of singular points for (2.15) is not identical to that of A(z). For example, $z = \pm i$ is a singular point for (2.15) for all N, but not for A(z) according to Theorem 3. Though, as observed before, for numerical computations the points $z = \pm i$ must be regarded also as singular points for A(z). Of course, if z is a singular point for A(z) and not for (2.15), (2.14) defines a particular solution for system (2.10).

Let us consider the behaviour of the principal solution component $C_1\mu_1^n$ and the parasitic component $C_2\mu_2^n$ for varying n and z, where we restrict ourselves to $z \leq 0$ and N even. We observe that for N even, $z \in \mathbb{R}$, the quantity $\delta \geq 0$, since $\eta(0) = 0$ and $\eta(z) < 0$ if $z \neq 0$. Hence, for $z \leq 0$ and N even, solution (2.14) is well-defined and is just the unique solution of system (2.10).

We distinguish between $z = 0$ and $z < 0$. The case $z = 0$ corresponds with $\dot{y}(x) = 0$, i.e. $y(x) = y(a)$, $a \leq x \leq b$. It is readily seen that for $z = 0$, $y_n = y(a)$ for all $n = 1,\ldots,N$. Hence the constant solution is computed without error. For $z < 0$, decaying exponentials, we have $0 < \mu_1 < 1$, $\mu_2 < -1$ and the limit behaviour

$$\mu_1 \sim 1 + z, \quad \mu_2 \sim -1 + z, \quad \eta \sim -\frac{1}{4}z^2 \quad \text{as} \quad z \uparrow 0,$$

$$\mu_1 \to 0, \quad \mu_2 \to -\infty, \quad \eta \sim -\frac{1}{4}z^{-2}, \quad \text{as } z \to -\infty.$$

Taking this into consideration the behaviour of $C_1\mu_1^n$ and $C_2\mu_2^n$ is best described as follows. $C_1\mu_1^n$ approximates the decaying solution for z close to zero and vanishes if $z \to -\infty$. This is true for all $1 \leq n \leq N$. For z close to zero, the parasitic component $C_2\mu_2^n$ is negligibly small (up to the discretization order in z). For n fixed, $C_2\mu_2^n$ increases with n. However, for all $z < 0$, its contribution to y_n is negligible for all n, $1 \leq n \leq N$. We once more note that for $\mathrm{Re}(z) << 0$ (stiff eigenvalues) the strongly decaying

exponential e^{nz} is well approximated . A similar description can be given for $z > 0$.

At this point it is appropriate to make a comparison with the standard step-by-step approach. Suppose that the explicit midpoint rule is applied that way. Consider the general solution (2.14). In order to obtain *absolute stability* μ_1 and μ_2 now must satisfy *the root condition*, i.e. none of the characteristic roots has modulus greater than one and every root with modulus one is simple. The root condition is satisfied, iff z is purely imaginary and $|z| < 1$. Hence, as is well known, the step-by-step explicit midpoint rule has no real interval of absolute stability, which shows that with respect to stability the boundary value method is just opposite the step-by-step method. In fact, from the investigation of equations (2.14)-(2.16) it can be seen that the boundary value method can be applied for $Re(z) < 0$, just because there $|\mu_1| < 1$ and $|\mu_2| > 1$. This conclusion, which is valid for other difference schemes as well, has been drawn before by Rolfes [21]. She considers the tridiagonal infinite Toeplitz matrix with rows (-1 0 1) and shows that the forward-backward substitution of the LU-decomposed Toeplitz matrix can be interpreted as a stable forward recursion ($|\mu_1|<1$) followed by a stable backward recursion ($|\mu_2|>1$) (see also [19,20]).

Finally, to complete our treatment of test model (1.1), in Figure 1 we give some illustrative plots showing the error

$$\varepsilon(r,\phi) = \max_{n=1(1)N} |y_n - e^{nhr\ \exp(i\phi)}|, \quad r \geq 0, \quad h = N^{-1}, \quad N = 20$$

for some values of ϕ; $\varepsilon(r,\phi)$ is the maximum error, taken over all grid-points, of y_n when approximating $y(x_n)=\exp(\delta x_n)$, $x_n = nh$, $\delta = re^{i\phi}$. Observe that for $\phi = \pi/2$, i.e. δ purely imaginary, the value $r = 20$ which corresponds with $z = i$, yields a meaningless result.

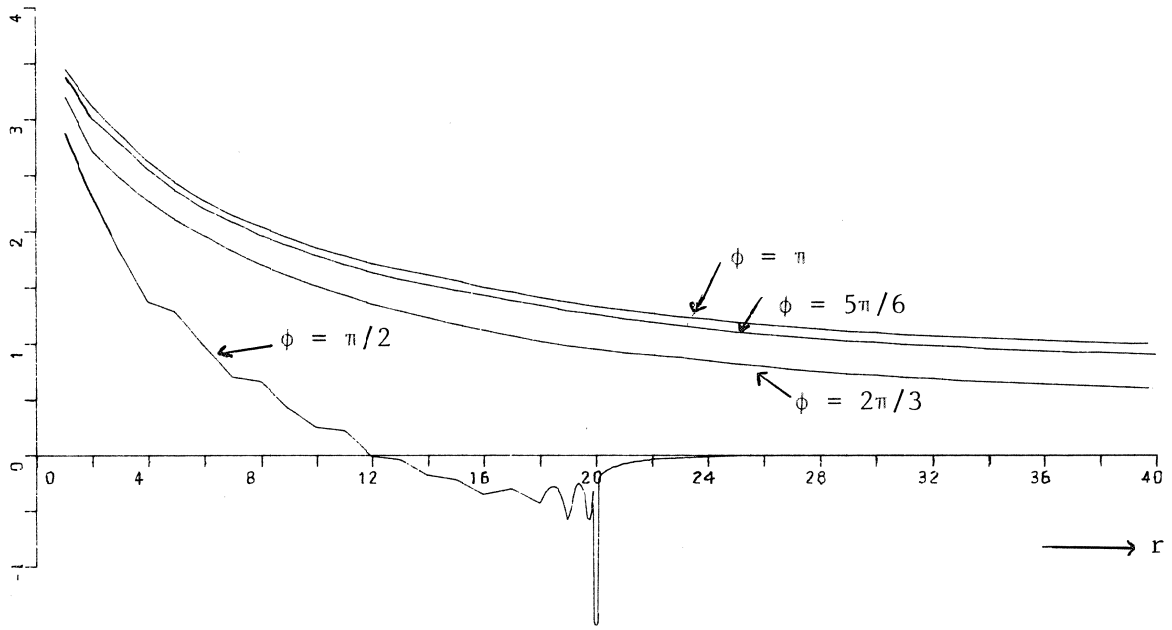


Fig. 1 Some plots of $-{}^{10}\log \varepsilon(r,\phi)$

2.3. *Convergence properties*

This section is devoted to convergence properties of the finite difference boundary value method. As in the preceding section we concentrate on method (2.3). It will be assumed that the vector function $f: [a,b] \times \mathbb{R}^s \to \mathbb{R}^s$ is as smooth as our analysis requires.

We introduce the conventional operators N and N_h (see e.g. [15,16]):

$$Ny \equiv \dot{y}(x) - f(x,y(x)) = 0,\ a \le x \le b,\ y(a) \text{ given},$$

$$N_h y_n \equiv \frac{y_{n+1} - y_{n-1}}{2h} - f(x_n,y_n) = 0,\ n = 1,\ldots,N-1,\ y_0 = y(a),$$

$$N_h y_N \equiv \frac{y_N - y_{N-1}}{h} - f(x_N,y_N) = 0.$$

Next, for any sufficiently smooth function $v(x)$, we define the *local truncation errors* $\tau_n[v] \equiv N_h v(x_n) - Nv(x_n)$, $n = 1(1)N$, and observe that

$$\tau_n[v] = \frac{1}{6}h^2 \dddot{v}(x_n) + O(h^3),\ n = 1(1)N-1,$$

$$\tau_N[v] = -\frac{1}{2}h\ddot{v}(x_N) + O(h^2).$$

Let e_n be the *global error vector* at x_n, i.e. $e_n \equiv y_n - y(x_n)$, $n = 1(1)N$. By substracting $N_h y(x_n)$ from $N_h y_n$ and by using the mean value equation

$$f(x_n, y(x_n) + e_n) - f(x_n, y(x_n)) = M(x_n)\, e_n,$$

$$M(x_n) \equiv \int_0^1 f'(x_n, y(x_n) + \theta e_n)\, d\theta, \quad f'(x,u) \equiv \frac{\partial f}{\partial u}(x,u),$$

it can be seen that e_n satisfies the difference scheme

$$L_h e_n \equiv \frac{e_{n+1} - e_{n-1}}{2h} - M(x_n)e_n = -\tau_n[y], \quad n = 1(1)N-1,$$

$$L_h e_N \equiv \frac{e_N - e_{N-1}}{h} - M(x_N)e_N = -\tau_N[y], \tag{2.17}$$

where e_0 is the zero vector and $y = y(x)$ denotes the exact solution of initial value problem (1.1). Hence method (2.3) is convergent, for a given vector function f, if for this function L_h is a *stable* difference operator (cf. [15,16]).

Let us reformulate (2.17) in the block matrix form

$$\begin{pmatrix} -2hM(x_1) & I & & & \\ -I & -2hM(x_2) & I & & \\ & \ddots & \ddots & \ddots & \\ & & -I & -2hM(x_{N-1}) & I \\ & & & -I & I-hM(x_N) \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_{N-1} \\ e_N \end{pmatrix} = \begin{pmatrix} -2h\tau_1[y] \\ -2h\tau_2[y] \\ \vdots \\ -2h\tau_{N-1}[y] \\ -h\tau_N[y] \end{pmatrix} \tag{2.18}$$

which we denote by

$$A_h \vec{e} \equiv (E_1 \otimes I - 2hM)\vec{e} = -2h\vec{\tau}, \tag{2.19}$$

where E_1 is given by (2.12) with the last row divided by two, and where $\otimes$ denotes the direct matrix product. The definitions of M, $\vec{e}$, and $\vec{\tau}$ are obvious. Stability of L_h is equivalent to the existence and *uniform boundedness* of the inverses of the family of matrices $h^{-1}A_h$.

EXAMPLE 3. To gain some feeling how the local errors τ_n accumulate in the global error we now first consider the scalar equation $\dot{y}(x) = f(x)$, i.e. f does not depend on y. Then $\vec{e}$ satisfies

$$(2.20) \qquad (2h)^{-1}E_1\vec{e} = -\vec{\tau}.$$

E_1 can be decomposed as $E_1 = UL$, where

$$U = \begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix}.$$

From the computation of E_1^{-1} one then finds the global errors

$$(2.21) \qquad \begin{aligned} e_n &= \sum_{j=1}^{n/2} - 2h\tau_{2j-1}, \quad n \text{ even}, \\ e_n &= e_{n-1} + \sum_{j=n}^{N} - 2h(-1)^{j-1}\tau_j + h(-1)^{N-1}\tau_N, \quad n \text{ odd}, \ e_0 = 0. \end{aligned}$$

These expressions reveal a clear distinction between *even and odd numbered gridpoints*. For n even, e_n is the sum of local errors at all odd points which lie behind. Hence, if τ_j does not change sign, $|e_n|$ increases with n. For n odd, $e_n = e_{n-1}$ + an alternating sum of all local errors at points at the right of x_n, including τ_N. Normally the alternating sum will hardly contribute, except for the last element $(-1)^{N-1}h\tau_N$ which contains $\ddot{y}(x_N)$ instead of $\dddot{y}(x_N)$.

To sum up, there will be a *lack of smoothness* in e_n when considering all gridpoints. However, the errors e_{2n-1}, or e_{2n}, for n = 1,2,..., will behave smooth. Finally, it immediately follows that for all n, $e_n = O(h^2)$. Note that $\tau_N = -\frac{1}{2}h\ddot{y}(x_N) + O(h^2)$ occurs only once in each e_n, n odd, and not in e_n if n is even. □

REMARK 4. Fox [8] describes the lack of smoothness for the testmodel (2.8) from equation (2.14). He points out that this behaviour hinders the application of difference correction for improving accuracy. For Richardson extrapolation, however, it does not cause any trouble if in the extrapolation process only even numbered grid points are used. □

In example 3 we considered an over-simplified problem. It nicely illustrates, however, the role of the matrix E, or E_1, in the convergence process, which, as we will show below, plays a similar role for the general problem.

Let us proceed with equation (2.19). Since E_1 is non-singular, we can write

$$(2.22) \qquad (I - 2h(E_1^{-1}\otimes I)M)\vec{e} = \vec{\gamma} \equiv -2h(E_1^{-1}\otimes I)\vec{\tau}.$$

Note that we use I to denote the s × s unit matrix, as well as the sN × sN unit matrix. The sN-vectors $\vec{\tau}$ and $\vec{\gamma}$ consist of N blocks, each of length s. Let $\vec{\tau}_j$ and $\vec{\gamma}_j$ denote the N-vector composed of the j-th element from each block. These vectors are associated to the j-th component of the solution vector y(x). Then, for j = 1(1)s, we have $(2h)^{-1}E_1\vec{\gamma}_j = -\vec{\tau}_j$ as in equation (2.20), implying that each n-th element of $\vec{\gamma}_j$ satisfies relation (2.21). This in turn implies that each element of the whole vector $\vec{\gamma}$ is $O(h^2)$, or, equivalently,

$$(2.23) \qquad \|\vec{\gamma}\|_\infty \le Ch^2, \text{ C a constant not depending on } h \le h_0.$$

THEOREM 5. *Let* $\|M(x)\|_\infty < \frac{1}{2}$ *for all* x ∈ [a,b], *method* (2.3) *is then convergent in the maximum norm with order two.*

PROOF. Consider equation (2.22) and observe that $hE_1^{-1}\otimes I$ is uniformly bounded.

In fact, from the equation for e_1 in (2.21) it follows that $\|hE_1^{-1}\otimes I\|_\infty = 1$. The proof is now easily completed by applying the perturbation lemma to the left hand side matrix of equation (2.22) and by using inequality (2.23). □

This result covers only a rather narrow class of problems due to the norm inequality on M(x). For example, stiff problems do not satisfy this inequality. The above derivation indicates, however, through the introduction of $\vec{\gamma}$, that for the general problem $\dot{y} = f(x,y)$ the global errors show a similar behaviour as described in Example 3. In fact, we observed this behaviour in all our numerical experiments, with non-stiff, as well as stiff problems. In the next theorem we will prove convergence in the spectral norm for a much broader class of problems:

THEOREM 6. *Define*

$$\mu_2(f'(x,u)) \equiv \max_i \lambda_i\left(\frac{f'(x,u)+f'^{T}(x,u)}{2}\right),$$

where λ_i *denotes the* i-th *eigenvalue and assume that* $\mu_2(f'(x,u)) \le \nu < 0$ *for all* $(x,u) \in [a,b] \times \mathbb{R}^s$. *Method* (2.3) *is then convergent in the spectral norm.*

PROOF. We consider the matrix $A_h^* = (-2h)^{-1}A_h$ (cf. (2.19)). By definition

$$\mu_2(A_h^*) = \max_i \lambda_i\left(\operatorname{diag}\left(\frac{M_1+M_1^T}{2},\ldots,\frac{M_{N-1}+M_{N-1}^T}{2},\frac{-2I+h(M_N+M_N^T)}{4h}\right)\right),$$

where $M_n = M(x_n)$. For all $h > 0$ we have

$$\mu_2(A_h^*) \le \max_n \mu_2(M_n) \le \nu.$$

The first inequality is trivial, while the second is a direct consequence of the definition of M_n and of a result given by Dahlquist [5], p. 11. Since $\nu < 0$ does not depend on h, but only on the problem, and since

$$\max_i \operatorname{Re} \lambda_i(A_h^*) \le \mu_2(A_h^*),$$

it is immediate that A_h^{*-1} exists and is uniformly bounded in $\|.\|_2$. More precisely, $\|A_h^{*-1}\|_2 \le -\nu^{-1}$, so that

(2.24) $\quad \|\vec{e}\|_2 \leq -\nu^{-1}\|\vec{\tau}\|_2.$ □

We observe that the method of proof of this theorem cannot be used to deal with equation (2.22). This prevents us from proving order two convergence in the spectral norm. In section 3, however, we are able to prove second order convergence in the spectral norm by considering method (2.3) as a particular Galerkin method.

The inequality $\mu_2(f'(x,u)) \leq \nu < 0$ is satisfied by all differential equations which possess strictly contractive solutions in the Euclidean vector norm (see Dahlquist [5], p. 13 and Frank et al. [10]). Hence Theorem 6 covers a broad and interesting class of problems, including many stiff ones. Furthermore, for these problems the stiffness, i.e. the magnitude of the stiff eigenvalues of $f'(x,u)$, does not enter into the one-sided Lipschitz constant ν. This constant ν is related to the smooth, non-stiff solutions components (see Frank et al. [10] for a clarifying discussion). Inequality (2.24) thus shows that if the solution to be computed is smooth, the global error will not suffer from the stiffness of the problem. Rolfes and Snyman [21,22] observed this in their experiments.

If ν is very close to zero, inequality (2.24) is meaningless. We emphasize, however, that the algorithm then still may perform quite satisfactorily, even if ν is larger than zero. We will explain this from the constant coefficient linear model system

(2.25) $\quad \dot{y}(x) = My(x) + g(x)$, M a normal matrix, $M = XDX^{-1}$.

Consider for (2.25) the matrix A_h given by (2.18), but with the last row again multiplied by two. We then can write $A_h^* \equiv (2h)^{-1}A_h$ in the form

$$A_h^* = (I\otimes X)\left(\frac{E\otimes I}{2h} - I\otimes D\right)(I\otimes X^{-1}),$$

E as in (2.12). The eigenvalues of A_h^* are the sN numbers (cf. [18], p. 259)

(2.26) $\quad \lambda_j/2h - \delta_k$, $j = 1(1)N$, $k = 1(1)s$,

where λ_j and δ_k are the eigenvalues of E and D, respectively (each eigenvalue δ_k of D plays the role of δ of test-model (2.8)). Hence method (2.3) will perform satisfactorily on problem (2.25), for a certain h, if the eigenvalues (2.26) stay away from zero. Figure 2 shows all, numerically computed, eigenvalues of E/2h for some values of the stepsize h. Note that some of the eigenvalues remain close to the imaginary axis if h decreases. Further, max $\mathrm{Re}(\lambda_j/2h)$ slowly increases as h decreases. Figure 2 is useful to ascertain for which spectra of M the method will converge. For example, if M has positive eigenvalues δ_k, i.e. the problem is unstable, the method will perform satisfactorily for $h \leq h_0$ if max $\delta_k <$ max $\mathrm{Re}(\lambda_j/2h_0)$. See also Fox & Mitchell [9], where it is pointed out that boundary value methods may give advantage over step-by-step methods if the problem to be integrated is unstable.

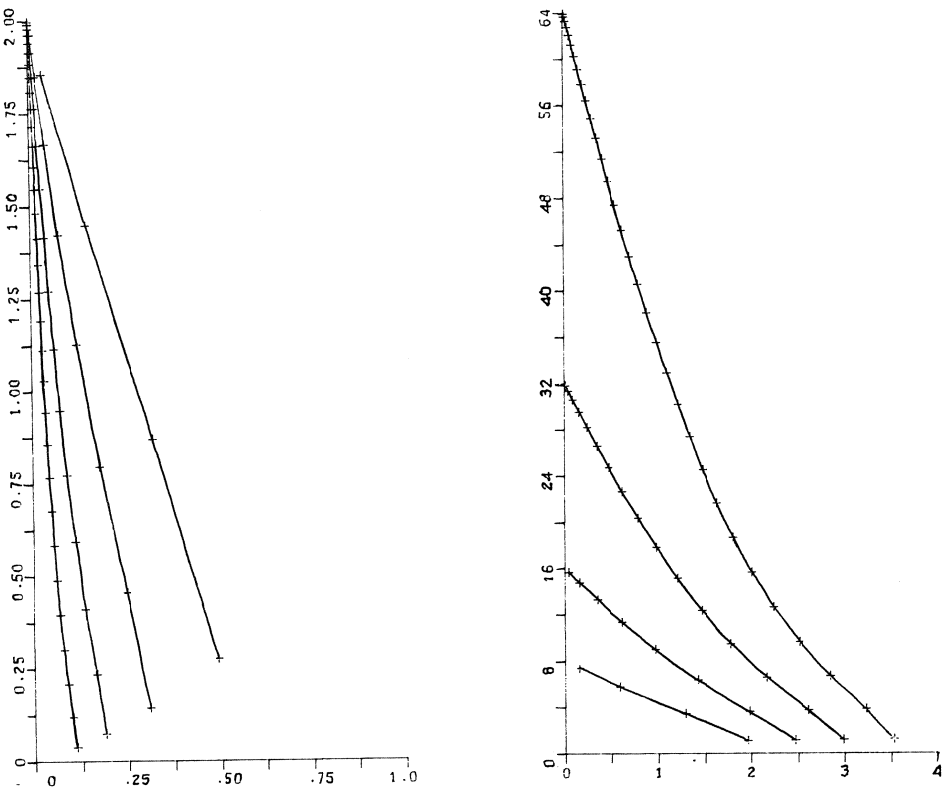


Fig. 2 Eigenvalues of E (left plot) and E/2h (right plot) for $h = \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}$. We have only plotted eigenvalues with non-negative imaginary part.

2.4. *A numerical illustration*

This section deals with a numerical example which serves to illustrate the convergence results derived in the previous section. For that purpose we selected the simple scalar problem

$$\dot{y}(x) = \delta\left(y(x) - \frac{1}{x+1}\right) - \frac{1}{(x+1)^2},\ 0 \le x \le 1,\ y(0) = 1,\ \delta \in \mathbb{R}, \tag{2.27}$$

whose general solution is given by $y(x) = e^{\delta x}(y(0)-1) + 1/(x+1)$. Since $y(0) = 1$, only the smooth solution component $1/(x+1)$ has to be computed. If $\delta << -1$, (2.27) is an example of a stiff problem where $e^{\delta x}y(0)$ represents the strongly varying solution component. In order to give sufficient insight in the error behaviour which has been predicted in Example 3, results will be shown for various choices of h and δ. We wish to emphasize that these results do not stand on their own. On the contrary, in a qualitative sense they are valid for systems as well. We refer to [21,22] for extensive experiments with a known collection of stiff problems.

Table 3 contains results of method (2.3) for $h = 1/4, 1/8, 1/16$, and $\delta = -1, -5, -10, -100$. Table 4 shows results for $\delta = 1, 5, 10, 100$. The following observations are relevant. The lack of smoothness over the grid is clearly observable. However, when we consider either even gridpoints, or odd ones, the error behaves smooth. Recall that we only have to compute the smooth solution of (2.27). For $\delta < 0$ the algorithm nicely shows its order two convergence at even numbered gridpoints. Observe that after halving h the absolute error should decrease with a factor 4 because the method is of order two and that $-^{10}\log(\frac{1}{4}) \simeq 0.6$. At odd gridpoints the order behaviour is much less pronounced as expected from Example 3. For $\delta > 0$ the algorithm yields more or less comparable results, though the second order not always shows up. This is because δ comes too close to the spectrum of E/2h (cf. Fig. 2).

Table 3. Results of method (2.3) for problem (2.27) with $\delta < 0$.
The table contains the value $-{}^{10}\log$ (absolute error).

δ	-1			-5			-10			-100		
x_n \ h	1/4	1/8	1/16	1/4	1/8	1/16	1/4	1/8	1/16	1/4	1/8	1/16
1/16			4.56			3.76			3.83			4.54
2/16		3.18	3.42		3.01	3.56		3.15	3.70		4.02	4.60
3/16			3.54			3.51			3.69			4.69
4/16	2.33	2.64	3.23	2.41	2.91	3.48	2.63	3.15	3.72	3.58	4.18	4.78
5/16			3.39			3.52			3.78			4.87
6/16		2.77	3.15		2.97	3.52		3.26	3.84		4.34	4.95
7/16			3.34			3.60			3.91			5.03
8/16	1.96	2.53	3.12	2.44	3.00	3.58	2.78	3.37	3.97	3.88	4.50	5.10
9/16			3.34			3.73			4.07			5.17
10/16		2.76	3.10		3.24	3.65		3.57	4.09		4.64	5.24
11/16			3.36			3.94			4.28			5.31
12/16	2.25	2.51	3.10	2.98	3.07	3.67	3.33	3.50	4.12	4.20	4.76	5.37
13/16			3.41			4.49			5.03			5.43
14/16		2.86	3.10		4.20	3.62		4.00	3.96		5.14	5.44
15/16			3.48			4.10			4.05			5.61
16/16	1.94	2.51	3.11	2.35	2.91	3.49	2.57	3.05	3.59	3.46	3.81	4.16

Table 4. Result of method (2.3) for problem (2.27) with $\delta > 0$.
The table contains the values $-{}^{10}\log$ (absolute error).

δ	1			5			10			100		
x_n \ h	1/4	1/8	1/16	1/4	1/8	1/16	1/4	1/8	1/16	1/4	1/8	1/16
1/16			2.96			3.07			3.42			4.49
2/16		2.43	3.61		2.62	3.83		2.95	4.02		3.99	4.63
3/16			3.11			3.20			3.70			4.72
4/16	1.95	2.81	3.40	2.20	3.27	3.52	2.54	3.45	4.01	3.57	4.21	4.81
5/16			3.28			3.21			3.92			4.89
6/16		2.76	3.31		2.84	3.28		3.39	4.07		4.37	4.97
7/16			3.49			3.09			4.04			5.05
8/16	2.11	2.66	3.24	2.64	2.91	3.04	3.00	3.56	4.08	3.91	4.52	5.12
9/16			3.81			2.89			4.01			5.19
10/16		3.38	3.20		2.65	2.78		3.48	3.91		4.66	5.26
11/16			5.20			2.65			3.74			5.33
12/16	3.06	2.58	3.16	2.25	2.46	2.52	2.83	3.27	3.54	4.14	4.78	5.39
13/16			3.88			2.39			3.31			5.44
14/16		3.18	3.12		2.21	2.26		2.90	3.06		4.72	5.43
15/16			3.58			2.13			2.81			5.00
16/16	1.97	2.50	3.09	1.88	1.97	2.00	2.30	2.48	2.56	3.43	3.75	4.03

3. A VARIATIONAL APPROACH

3.1. *Preliminaries*

We consider nonlinear systems of ode's

(3.1) $\dot{U} = \tilde{F}(t,U)$, $0 < t \leq T$, $U(.) \in \mathbb{R}^m$, $U(0)$ prescribed.

We make at first a transformation of this equation to a more suitable form. In problems we shall consider, there may exist positive stiffness parameters ε_i, such that parts of $\tilde{F}$ and the corresponding parts of the Jacobian matrix $\frac{\partial \tilde{F}}{\partial U}$ are unbounded as $O(\varepsilon_i^{-1})$, $\varepsilon_i \to 0$. We then multiply the corresponding equations by this parameter to get

(3.2) $\varepsilon \dot{U} = F(t,U)$, $0 < t \leq T$,

where ε is a diagonal matrix with entries ε_i, $0 < \tilde{\varepsilon} \leq \varepsilon_i \leq 1$, and F and $\frac{\partial F}{\partial U}$ are bounded with respect to ε. A typical example is given by $\tilde{F}(t,U) = \tilde{A}U + \tilde{C}$

$$\tilde{A} = \begin{pmatrix} -1800 & 900 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1000 & -2000 \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1000 \end{pmatrix}$$

found in Enright et al. [7]. Here $\varepsilon = \mathrm{diag}(\frac{1}{900}, 1, \ldots, 1, \frac{1}{1000})$ is an obvious choice. In more general problems we may have to multiply by a more general positive definite matrix ε, in order to get a bounded F and $\frac{\partial F}{\partial U}$. We further assume that F satisfies

(3.3) $([F(t,U) - F(t,V)], U - V) \leq \rho(t)\|U-V\|^2, \forall U,V \in \mathbb{R}^m$, $t > 0$

where ρ: $[0,T] \to \mathbb{R}$ is at least piecewise continuous and independent of ε and $\rho(t) \le -\rho_0$, $t \ge t_0 \ge 0$, $\rho_0 > 0$. Further $\|V\| = (V,V)^{\frac{1}{2}}$ where $(.,.)$ is the inner product in $\mathbb{R}^m$. As is well known and easily seen, this means that, if U,V are two solutions of (3.2) corresponding to different initial values, then

$$\frac{1}{2}\frac{d}{dt}(\varepsilon(U-V),U-V) = ([F(t,U) - F(t,V)],U-V)$$

$$\le \rho(t)\|U-V\|^2 \le \rho(t)(\varepsilon(U-V),U-V), \quad t \ge t_0,$$

so

$$\|U(t)-V(t)\|_\varepsilon^2 \le \exp\left(\int_{t_0}^{t} 2\rho(s)ds\right)\|U(t_0)-V(t_0)\|_\varepsilon^2 \le \|U(t_0)-V(t_0)\|_\varepsilon^2, \quad t_0 \le t \le T,$$

where

$$\|V\|_\varepsilon = (\varepsilon V,V)^{\frac{1}{2}}.$$

This means that the system is contractive for $t \ge t_0$ if condition (3.3) holds. We further assume that F is Lipschitz continuous, i.e. there exists a constant C such that

(3.4) $$\|F(t,U) - F(t,V)\| \le C\|U-V\|, \forall U,\ V \in \mathbb{R}^m.$$

In the initial phase $(0,t_0)$, the system does not have to be contractive, i.e. the eigenvalues of the Jacobian may have positive real parts. In this interval we may consider to use a step-by-step method with very small step-sizes, if it is of importance to follow the transients.

3.2. *The Galerkin method*

At first we shall describe the global Galerkin method to be used in the interval (t_0,T). We divide this interval in a number of subintervals (t_{i-1},t_i), $i = 1,2,\ldots,N$ where $t_N = T$. The length of the intervals (t_i-t_{i-1}) may vary smoothly with some function $h(t_i)$, but for ease of presentation, we

assume that the intervals have equal length, i.e. $t_i - t_{i-1} = h$, $i = 1,2,\ldots,N$. We consider each interval as an element on which we place some nodal points, $t_{i,j}$, $j = 0,1,\ldots,p$, and $t_{i,j} = t_i + \xi_j h$, where ξ_j are the Lobatto quadrature points which satisfy $0 = \xi_0 < \xi_1 < \ldots < \xi_p = 1$ and $\xi_j + \xi_{p-j} = 1$. Hence the endpoints of the interval are always nodal points and (if $p>1$) we choose also $(p-1)$ disjoint nodal points in the interior of each element.

To each nodal point we associate a basisfunction $\phi_{i,j}$. The basisfunctions may be exponential or trigonometric functions and may also be discontinuous but in this paper we only consider the most common choice, where they are piecewise continuous and polynomials over each element. Basisfunctions corresponding to interior nodes have support only in the element to which it belongs and those corresponding to endpoints have support over the two adjacent elements (except those at t_0 and at t_N). The number of nodal points at each closed interval then equals the degree p of the polynomial plus one.

Let $\mathring{S}_h$ be the subspace of test functions which are zero at t_0, i.e. $\mathring{S}_h = \text{SPAN}\{\phi_{i,j},\ i=0,1,\ldots,N-1,\ j=1,2,\ldots,p\}$. Let

$$a(U;V) \equiv \int_{t_0}^{T} ([\varepsilon\dot{U}-F(t,U)],V)\ dt,\quad U,V \in [H^1(t_0,T)]^m,$$

where $H^1(t_0,T)$ is the first order Sobolev space of functions with square integrable derivatives. To get an approximation of the solution of (3.2), we take a testfunction (vectorial function) $V = \phi_{i,j}$, and multiply the equation with V to get after integration,

$$(3.5a)\qquad a(\tilde{U};\phi_{i,j}) = \int_{t_{i-1}}^{t_{i+1}} ([\varepsilon\dot{\tilde{U}}-F(t,\tilde{U})],\phi_{i,j}^{(r)})dt = 0,\ j = 0,\ i = 1,2,\ldots,N-1$$

$$(3.5b)\qquad a(\tilde{U};\phi_{i,j}) = \int_{t_i}^{t_{i+1}} ([\varepsilon\dot{\tilde{U}}-F(t,\tilde{U})],\phi_{i,j}^{(r)})dt = 0,\ j = 1,2,\ldots,p-1,\quad i = 0,1,\ldots,N-1$$

At $t_N = T$ we get

$$(3.5c) \qquad a(\tilde{U};\phi_{N,0}) = \int_{t_{N-1}}^{t_N} ([\varepsilon\ddot{\tilde{U}}-F(t,\tilde{U})],\phi_{N,0}^{(r)})dt = 0.$$

Here we choose in turn

$$\phi_{i,j}^{(r)} = \begin{bmatrix} 0 \\ \vdots \\ \phi_{i,j} \\ \vdots \\ 0 \end{bmatrix} (r), \quad r = 1,2,\ldots,m$$

where $\phi_{i,j}$ is the corresponding scalar basisfunction. This defines the Galerkin approximation $\tilde{U}$ corresponding to $\mathring{S}_h$, where

$$\tilde{U} = U(t_0)\phi_{0,0} + \sum_{i=0}^{N-1}\sum_{j=1}^{p} d_{i,j}\phi_{i,j},$$

i.e. we have imposed strongly the essential boundary condition at t_0. Clearly

$$a(U;V) = 0, \forall\ V \in H^1(t_0,T).$$

We get then from (3.5a)

$$(3.6) \qquad a(U;V) - a(\tilde{U};V) = \int_{t_{i-1}}^{t_{i+1}} (\varepsilon[\dot{U}-\ddot{\tilde{U}}]-[F(t,U)-F(t,\tilde{U})],V)dt = 0,$$
$$V = \phi_{i,j},\ j = 0,\ i = 1,2,\ldots,N-1,$$

and similarly for (3.5b,c).

To estimate the Galerkin discretization error $U - \tilde{U}$, we let $U_I \in S_h$ be the interpolant to U on $\{t_{i,j}\}$, $j = 0,1,2,\ldots,p$, $i = 0,1,\ldots,N-1$ and we write

$$U - \tilde{U} = \eta - \theta,$$

where $\eta = U - U_I$ is the interpolation error and

$$\theta = -U + \tilde{U} + \eta = \tilde{U} - U_I.$$

Note that $\theta \in \mathring{S}_h$. Assuming that the solution U is sufficiently smooth, from the interpolation error expansion on integral form we get the usual Sobolev norm estimates

$$\int_{t_0}^{T} \|U-U_I\|^2 dt \le C_0 h^{2(p+1)} \int_{t_0}^{T} \|U\|^2_{p+1} dt,$$

(3.7)

$$\int_{t_0}^{T} \|\dot{U}-\dot{U}_I\|^2 dt \le C_1 h^{2p} \int_{t_0}^{T} \|\dot{U}\|^2_{p+1} dt,$$

for the interpolation error. Here

$$\|U\|^2_{p+1} = \int_{t_0}^{T} \sum_{k=0}^{p+1} \left(\frac{\partial^k U}{\partial t^k}, \frac{\partial^k U}{\partial t^k}\right) dt$$

is the norm in the Sobolev space $H^{p+1}(t_0,T)$.

THEOREM 7. *Let* U *be the solution of* (3.2) *where* (3.3), (3.4) *are satisfied. Then the Galerkin solution* $\tilde{U}$, *in the space of piecewise polynomial continuous functions of degree* p, *defined by* (3.5a,b,c) *satisfies*

$$|||U-\tilde{U}||| = O(h^{p+\nu})\{\|\varepsilon U\|^2_{p+2} + \|U\|^2_{p+1}\}^{\frac{1}{2}}, \quad h \to 0,$$

where $\nu = 1$ *if* $p = 1$, $1 \ge \nu \ge \frac{1}{2}$ *if* $p = 3,5,\ldots$ *and* $\nu = 0$ *if* p is even, *and*

$$|||V|||^2 = \tfrac{1}{2}(\varepsilon V(T),V(T)) - \int_{t_0}^{T} \rho(t)\|V(t)\|^2 dt.$$

(Note that this estimate implies both a least square estimate as well as a pointwise estimate at the endpoint of the interval.)

PROOF. Our first objective is to derive an estimate of θ. We have $\theta(t_0) = \eta(t_0) = 0$ and

$$a(\tilde{U};\theta) - a(U_I;\theta) = \int_{t_0}^{T} \{(\varepsilon\dot{\theta},\theta) - ([F(t,\tilde{U})-F(t,U_I)],\theta)\}dt. \tag{3.8}$$

From

$$\int_{t_0}^{T} (\varepsilon\dot{\theta},\theta)dt = -\int_{t_0}^{T} (\dot{\theta},\varepsilon\theta)dt + [(\varepsilon\theta,\theta)]_{t_0}^{T},$$

if follows that

$$\int_{t_0}^{T} (\varepsilon\dot{\theta},\theta)dt = \tfrac{1}{2}(\varepsilon\theta(T),\theta(T)).$$

Hence by (3.3),

$$(3.9)\qquad a(\tilde{U};\theta) - a(U_I;\theta) = \tfrac{1}{2}(\varepsilon\theta(T),\theta(T)) - \int_{t_0}^{T} ([F(t,\tilde{U})-F(t,U_I)],\theta)dt$$

$$\geq \tfrac{1}{2}(\varepsilon\theta(T),\theta(T)) - \int_{t_0}^{T} \rho(t)\|\theta(t)\|^2 dt := |||\theta|||^2.$$

Further we get from (3.6) and the Lipschitz continuity (3.4),

$$(3.10)\qquad a(\tilde{U};\theta) - a(U_I;\theta) = a(\tilde{U};\theta) - a(U;\theta) + a(U;\theta) - a(U_I;\theta) =$$

$$= a(U;\theta) - a(U_I;\theta) \leq \left|\int_{t_0}^{T} (\varepsilon\dot{\eta},\theta)dt\right| + C\int_{t_0}^{T} \|\eta\|\|\theta\|dt.$$

First we shall estimate the term $\int_{t_0}^{T} (\varepsilon\dot{\eta},\theta)dt$. Since ε is a diagonal matrix, we may consider each component individually, i.e. $\int_{t_0}^{T} \dot{\eta}_k\theta_k dt$, where η_k,θ_k are scalar functions. Since $\theta_k \in \mathring{S}_h$, we have

$$\theta_k = \sum_{i=0}^{N-1}\sum_{j=1}^{p} \gamma_{i,j}\tilde{\phi}_{i,j}.$$

where $\{\tilde{\phi}_{i,j}\}$, $i = 0,1,\ldots,N-1$, $j = 1,2,\ldots,p$ is a set of basisfunctions (testfunctions) spanning $\mathring{S}_h$, but not necessarily equal to $\phi_{i,j}$. We write $\eta_k = \tilde{\eta}_k + \eta_R$, where $\tilde{\eta}_k$ is the leading (polynomial) term in an expansion of the interpolation error (for more details, see e.g. Axelsson and Gustafsson [2]). For instance, if we use piecewise linear basisfunctions, then at t_i, $i = 1,3,\ldots,N-1$ we have

$$\tilde{\eta}_k(t) = \begin{cases} \frac{1}{2}\ddot{U}_k(t_i)(t-t_{i-1})(t-t_i), & t_{i-1} \le t \le t_i \\ \frac{1}{2}\ddot{U}_k(t_i)(t-t_i)(t-t_{i+1}), & t_i < t \le t_{i+1} \end{cases}.$$

(For simplicity, we may assume that N is even.) With the piecewise polynomial basisfunctions of degree p, an easy calculation shows that

$$(3.11) \qquad \left\{ \int_{t_0}^{T} \tilde{\dot{\eta}}_k^2 dt \right\}^{\frac{1}{2}} = O(h^p)\|U_k\|_{p+1}, \left\{ \int_{t_0}^{T} \dot{\eta}_R^2 dt \right\}^{\frac{1}{2}} = O(h^{p+1})\|U_k\|_{p+2}, \ h \to 0.$$

We shall prove that due to cancellation

$$(3.12a) \qquad \left| \int_{t_0}^{T} \dot{\eta}_k \theta_k dt \right| \le \rho_0 \int_{t_0}^{T} \theta_k^2 dt + \frac{1}{2}\theta_k^2(T) + O(h^{2(p+\nu_k)})\|U_k\|_{p+2}^2, \begin{cases} \nu_k = 1, & p=1 \\ \nu_k \ge \frac{1}{2}, & p=3,5,\ldots \end{cases}$$

i.e.

$$(3.12b) \qquad \left| \int_{t_0}^{T} (\varepsilon\dot{\eta},\theta) dt \right| \le \rho_0 \int_{t_0}^{T} \|\theta(t)\|^2 dt + \frac{1}{2}(\varepsilon\theta(T),\theta(T)) + O(h^{2(p+\nu_k)})\|\varepsilon U\|_{p+2}^2.$$

We have

$$(3.13) \qquad \int_{t_0}^{T} \dot{\eta}_k \theta_k dt = \int_{t_0}^{T} \tilde{\dot{\eta}}_k \theta_k dt + \int_{t_0}^{T} \dot{\eta}_R \theta_k dt.$$

From (3.11) it follows

$$\left| \int_{t_0}^{T} \dot{\eta}_R \theta_k dt \right| \le O(h^{p+1}) \left\{ \int_{t_0}^{T} \theta_k^2 dt \right\}^{\frac{1}{2}} \|U_k\|_{p+2} \le \rho_0 \int_{t_0}^{T} \theta_k^2 dt + O(h^{2(p+1)})\|U_k\|_{p+2}^2,$$

and it remains to consider the first term in (3.13). We have

$$\int_{t_0}^{T} \tilde{\dot{\eta}}_k \theta_k dt = - \int_{t_0}^{T} \tilde{\eta}_k \dot{\theta}_k dt.$$

As testfunction $\tilde{\phi}_{i,j}$ we choose the basisfunctions of degree q, q = 1,2,...,p, at each t_i, i = 1,2,...,N, with support on (t_{i-1},t_{i+1}), see Fig. 3.1.

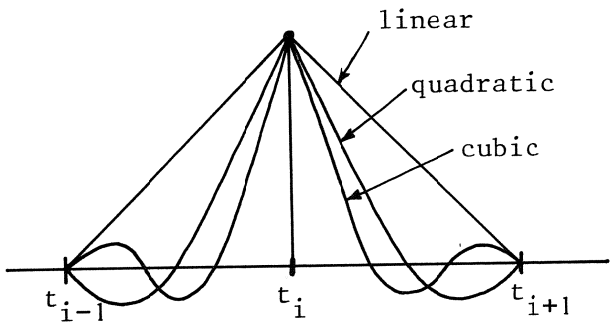


Figure 3.1 Testfunctions $\tilde{\phi}_{i,j}$, j = 1,2,...,p (p=3).

We have

$$(3.14)\qquad \int_{t_0}^{T} \tilde{\eta}_k \dot{\theta}_k dt = \sum_{i,j} \int_{t_0}^{T} \gamma_{i,j} \tilde{\eta}_k \dot{\tilde{\phi}}_{i,j} dt =$$

$$= \sum_{i=1}^{N-1} \sum_{j=1}^{p} \gamma_{i-1,j} \int_{t_{i-1}}^{t_{i+1}} \tilde{\eta}_k \dot{\tilde{\phi}}_{i-1,j} dt + \sum_{j=1}^{p} \gamma_{N-1,j} \int_{t_{N-1}}^{t_N} \tilde{\eta}_k \dot{\tilde{\phi}}_{N-1,j} dt.$$

Note that $\tilde{\eta}_k$ and $\dot{\theta}_k$ are polynomials of degree p + 1 and p - 1, respectively on each subinterval. Also note that $\tilde{\eta}_k$ is zero at the (p+1) Lobatto points. Hence, if we apply Lobatto quadrature we get (from the error term of the quadrature),

$$\int_{t_{i-1}}^{t_i} \tilde{\eta}_k \dot{\tilde{\phi}}_{i-1,j} dt = O\left(h^{2p+1}\right) \frac{\partial^{2p}}{\partial t^{2p}} \left(\tilde{\eta}_k \dot{\tilde{\phi}}_{i-1,j}\right)$$

$$= O\left(h^{2p+1}\right)\binom{2p}{p+1} \frac{\partial^{p+1}}{\partial t^{p+1}} \left(\tilde{\eta}_k\right) \frac{\partial^{p}}{\partial t^{p}} \left(\tilde{\phi}_{i-1,j}\right) = O\left(h^{p+1}\right).$$

The loss of p in the exponent is due to that

$$\frac{\partial^p}{\partial t^p}\left(\tilde{\phi}_{i-1,j}\right) = O\left(h^{-p}\right).$$

Note however that for p odd this derivative, which is piecewise constant, appears with the same numerical value but with opposite signs in the two adjacent intervals (t_{i-1},t_i) and (t_i,t_{i+1}). Hence

$$\int_{t_{i-1}}^{t_{i+1}} \tilde{\eta}_k \dot{\tilde{\phi}}_{i-1,j} dt = 0, \quad i = 1,2,\ldots,N-1$$

and by (3,14),

$$\int_{t_0}^{T} \tilde{\eta}_k \dot{\theta}_k dt = \sum_{j=1}^{p} \gamma_{N-1,j} \int_{t_{N-1}}^{t_N} \tilde{\eta}_k \dot{\tilde{\phi}}_{N-1,j} dt. \tag{3.15}$$

Consider first the case of piecewise linear basisfunctions (p=1). Then we only have one testfunction at t_N. With $\tilde{\phi}_{N-1,1} = \phi_N$, $\gamma_{N-1,1} = \gamma_N = \theta_k(T)$, we get

$$\left| \int_{t_0}^{T} \tilde{\eta}_k \dot{\theta}_k dt \right| = \left| \gamma_N \int_{t_{N-1}}^{t_N} \tilde{\eta}_k \dot{\phi}_N dt \right| \le |\theta_k(T)| \max_{t_{N-1}\le t\le t_N} |\tilde{\eta}_k(t)|$$

$$= |\theta_k(T)| \left| O\left(h^{p+1}\right) \right| \| U_k \|_{p+2}.$$

Here we have used the Sobolev inequality,

$$\max_{t_0\le t\le T} |U_k^{(p+1)}| \le C \| U_k \|_{p+1}, \quad U_k \in H^{p+2}(t_0,T).$$

Hence (3.12a) follows for p = 1. For the general case, p > 1, we choose $\tilde{\phi}_{N-1,j} = \phi_{N-1,j}$ and we then get

$$\left| \int_{t_0}^{T} \tilde{\eta}_k \dot{\theta}_k dt \right| = \left| \sum_{j=1}^{p} \gamma_{N-1,j} \int_{t_{N-1}}^{t_N} \tilde{\eta}_k \dot{\tilde{\phi}}_{N-1,j} dt \right| = \tag{3.16}$$

$$\le C \sum_{j=1}^{p} |\gamma_{N-1,j}| \max_{t_{N-1}\le t\le t_N} |\tilde{\eta}_k| \le C' \left\{ h^{-1} \int_{t_{N-1}}^{t_N} \theta_k^2 dt \right\}^{\frac{1}{2}} \max_{t_0\le t\le t_N} |\tilde{\eta}_k|.$$

Here we have used the fact that

$$\sum_{j=1}^{p} |\gamma_{N-1,j}| \text{ and } \left\{h^{-1}\int_{t_{N-1}}^{t_N} \theta_k^2 dt\right\}^{\frac{1}{2}}$$

are both norms in a finite dimensional space ($\mathbb{R}^p$), and hence equivalent (uniformly in h). Finally we get from (3.16),

$$\left|\int_{t_0}^{T} \tilde{\eta}_k \dot{\theta}_k dt\right| = C''\left\{\int_{t_0}^{T} \theta_k^2 dt\right\}^{\frac{1}{2}} O\left(h^{p+\nu_k}\right)\|U_k\|_{p+1}$$

where

$$(3.17)\qquad \nu_k = \min\left\{1, \frac{1}{2}\left[1 + \left[\log\left(\int_{t_0}^{T} \theta_k^2 dt \Big/ \int_{t_{N-1}}^{t_N} \theta_k^2 dt\right)\right]/\log h^{-1}\right]\right\}.$$

Together with (3.14) and (3.15) this implies (3.12a) for p = 3,5,... , where

$$(3.18)\qquad 1 \geq \nu = \min_{1\leq k\leq m} \nu_k \geq \frac{1}{2}.$$

Finally, from (3.9), (3.10) and (3.12a) we get

$$|||\theta|||^2 \leq O(h^{2(p+\nu)})\|\varepsilon U\|_{p+2}^2 + |||\eta|||^2$$

$$\leq O(h^{2(p+\nu)})\left\{\|\varepsilon U\|_{p+2}^2 + \|U\|_{p+1}^2\right\}^{\frac{1}{2}},\ p = 1,3,\ldots .$$

Clearly, this estimate with $\nu = 0$ is also valid for p even. We now get the Galerkin error

$$|||U-\tilde{U}||| \leq |||U-U_I||| + |||U_I-\tilde{U}||| = |||\eta||| + |||\theta||| \leq$$

$$O(h^{p+\nu})\left\{\|\varepsilon U\|_{p+2}^2 + \|U\|_{p+1}^2\right\}^{\frac{1}{2}},\ h \to 0. \qquad \square$$

Note that if U is smooth, $\tilde{U}$ and U_I will be smooth, so

$$\int_{t_0}^{T} \theta_k^2 dt \Big/ \int_{t_{N-1}}^{T} \theta_k^2 dt = O(h^{-1}).$$

Hence it follows from (3.17), (3.18) that ν will be close to 1.

3.3. *A weighted Galerkin method.*

Consider now the more general case where

$$(F(t,U) - F(t,V), U-V) \le \sum_{k=1}^{m} \rho_k(t)(U_k - V_k)^2$$

and

$$\rho_k(t) \le \begin{cases} -\rho_0, \ \rho_0 > 0 & \text{if } \varepsilon_k << 1, \\ \rho_1, & \text{if } \varepsilon_k = 1. \end{cases}$$

Here ρ_1 may be positive. In this case we use a weighted variational formulation with a positive weight function g,

$$a(U,V) = \int_{t_0}^{T} ([\varepsilon\dot{U} - F(t,U)], V) g(t) dt.$$

Then the analysis goes through as before. Note only that

$$\int_{t_0}^{T} (\varepsilon_k \dot{\theta}_k, \theta_k) g(t) dt = -\frac{1}{2} \int_{t_0}^{T} \varepsilon_k \theta_k^2 \dot{g} dt + \frac{1}{2} \varepsilon_k \theta_k^2(T) g(T).$$

Hence, if we let $\dot{g} < 0$ satisfy

$$-\frac{1}{2}\dot{g} - \rho_1 g \ge c_0 > 0$$

(i.e. for instance $g(t) = e^{-4\rho_1 t}$, if $\rho_1 > 0$) then we have still coercivity as in (3.9) in norm

$$|\!|\!| \theta |\!|\!|^2 = c_0 \int_{t_0}^{T} (\varepsilon\theta, \theta) dt + \frac{1}{2}(\varepsilon\theta(T), \theta(T)) g(T).$$

3.4. *Difference schemes*

In order to get a fully discretized scheme we have to use numerical quadrature, which results in various difference schemes. We shall consider

this only for the case p = 1. Then $\phi_{i,p} = \phi_i$ are the usual hatfunctions and there are no interior nodes. With

$$\tilde{U} = U(t_0)\,\phi_0 + \sum_{i=1}^{N} U_i\phi_i\,,$$

(3.5) and (3.5c) imply

$$(3.19)\qquad \begin{cases} \varepsilon(\tilde{U}_{i+1}-\tilde{U}_{i-1}) = 2\displaystyle\int_{t_{i-1}}^{t_{i+1}} F(t,\tilde{U}_{i-1}\phi_{i-1}+\tilde{U}_i\phi_i+\tilde{U}_{i+1}\phi_{i+1})\phi_i\,dt, \\ i = 1,2,\ldots,N-1, \\ \varepsilon(\tilde{U}_N-\tilde{U}_{N-1}) = \displaystyle\int_{t_{N-1}}^{t_N} F(t,\tilde{U}_{N-1}\phi_{N-1}+\tilde{U}_N\phi_N)\phi_N dt\,. \end{cases}$$

We call this the *generalized midpoint rule difference scheme*. Let $F_i = F(t,\tilde{U})\big|_{t=t_i}$. If we use numerical integration by the trapezoidal rule, i.e.

$$\int_{t_{i-1}}^{t_i} F\,\phi_i\,dt = \frac{1}{2}h\,[F_{i-1}\phi_i(t_{i-1}) + F_i\phi_i(t_i)] = \frac{1}{2}hF_i\,,$$

we recover the difference method (2.1), (2.2). As we know, this scheme is of $O(h^2)$, see Section 2.3. We consider now a more accurate difference scheme which we may derive from (3.19). For this purpose let

$$F(t) \simeq \frac{1}{2}[F_{i-1}+F_i] + (t-t_i+\frac{h}{2})\frac{1}{h}(F_i-F_{i-1}),\ t_{i-1} \le t \le t_i,$$

except that for the last formula in (3.19) we use

$$F(t) \simeq \frac{1}{2}[F_{N-1}+F_N],\ t_{N-1} \le t \le t_N.$$

Then

$$\int_{t_{i-1}}^{t_i} F(t)\phi_i dt = \frac{h}{4}(F_{i-1}+F_i) + \frac{h}{12}(F_i-F_{i-1}) = \frac{h}{6}(F_{i-1}+2F_i),$$

$i = 1,2,\ldots,N-1$, and similary

$$\int_{t_i}^{t_{i+1}} F(t)\phi_i dt = \frac{h}{6}(F_{i+1}+2F_i).$$

Hence the generalized midpoint rule (3.19) takes the form

$$(3.20) \qquad \begin{cases} \varepsilon(\tilde{U}_{i+1}-\tilde{U}_{i-1}) = \frac{h}{3}(F_{i-1}+4F_i+F_{i+1}), \ i = 1,2,\ldots,N-1 \\ \varepsilon(\tilde{U}_N-\tilde{U}_{N-1}) = \frac{h}{2}(F_{N-1}+F_N). \end{cases}$$

We notice that this is a combination of the Simpson and trapezoidal rules.

For this combination numerical tests (see Table 5,6) indicate very accurate results. Note that already on a very coarse mesh ($h=\frac{1}{4}$) the accuracy is high. For $\delta < 0$ (Table 5), the order of convergence seems to be $\simeq 3.5$.

Finally some remarks about methods for the solution of the algebraic systems. These are on block tridiagonal form. If we use a special starting scheme for the calculation of $\tilde{U}_1$, we may use a "shooting method" for the solution of (3.1) i.e.

$$\tilde{U}_{i+1} = \tilde{U}_{i-1} + 2hF(t_i,\tilde{U}_i), \ i = 1,2,\ldots .$$

This is of course nothing but the two-step midpoint rule, which, as well known, is unstable for stiff problems (and of order $O(h^2)$ for non-stiff problems). If the order of the systems (3.1, 3.2) is large and $\partial F/\partial U$ is sparse we may however apply an iterative method, which would preserve sparsity. There exists methods, such as preconditioned generalized conjugate gradient methods, for which the rate of convergence of the iterations is

fast, see for instance Axelsson [3] and Hageman and Young [11].

Hence the large size of the matrices which arise should not be detrimental for the application of the methods described in this papar.

From the analyses and the numerical experiments it is concluded that the global method is a robust reliable method for both stiff systems and systems with increasing fundamental solutions. It is particularly efficient when moderate accuracy is desired. If does not seem to be very sensitive to stiffness.

In the case of high accuracy, large or non-linear systems, the efficiency depends on the availability of good algebraic solvers.

Table 5. Results of method (3.20) for problem (2.27) with $\delta < 0$. The table contains the value $-^{10}\log$ (absolute error).

δ	-1			-5			-10			-100		
x_n \ h	1/4	1/8	1/16	1/4	1/8	1/16	1/4	1/8	1/16	1/4	1/8	1/16
1/16			5.28			5.98			7.08			7.05
2/16		4.44	5.89		5.52	5.92		5.37	6.17		5.95	7.30
3/16			5.34			6.14			6.96			7.42
4/16	3.73	4.59	5.69	4.69	4.72	5.78	4.34	5.02	6.13	5.03	6.42	7.56
5/16			5.38			6.04			6.78			7.69
6/16		4.57	5.59		5.21	5.72		5.74	6.09		6.69	7.80
7/16			5.39			5.89			6.29			7.94
8/16	3.40	4.47	5.53	3.62	4.66	5.66	3.39	4.98	6.00	5.19	6.54	7.98
9/16			5.40			5.75			6.03			8.41
10/16		4.58	5.49		4.79	5.60		4.98	5.86		6.44	7.84
11/16			5.39			5.63			5.82			7.75
12/16	3.77	4.43	5.46	3.73	4.56	5.53	3.84	4.74	5.70	4.76	5.85	7.22
13/16			5.39			5.53			5.64			6.85
14/16		4.54	5.43		4.55	5.45		4.61	5.53		5.39	6.42
15/16			5.38			5.43			5.45			6.02
16/16	3.36	4.40	5.41	3.46	4.42	5.37	3.52	4.43	5.36	4.21	4.90	5.61

Table 6. Results of method (3.20) for problem (2.27) with $\delta > 0$.
The table contains the value $-^{10}\log$ (absolute error).

δ	1			5			10			100		
x_n \ h	1/4	1/8	1/16	1/4	1/8	1/16	1/4	1/8	1/16	1/4	1/8	1/16
1/16			4.96			4.59			5.98			6.99
2/16		4.08	7.70		4.08	4.96		4.93	7.57		5.92	7.49
3/16			4.97			4.44			6.12			7.40
4/16	3.23	5.16	6.33	3.45	4.21	4.53	3.96	5.85	6.73	4.99	6.63	7.64
5/16			4.96			4.24			6.03			7.71
6/16		4.10	5.94		3.75	4.20		5.21	6.02		6.57	7.86
7/16			4.95			4.00			5.65			7.95
8/16	3.72	6.84	5.69	3.43	3.56	3.91	4.72	5.40	5.44	5.36	6.83	8.07
9/16			4.93			3.74			5.15			8.17
10/16		4.09	5.52		3.26	3.62		4.69	4.89		7.20	8.27
11/16			4.90			3.48			4.62			8.40
12/16	3.31	4.93	5.38	2.77	3.00	3.35	3.98	4.19	4.35	4.79	6.10	8.28
13/16			4.88			3.21			4.07			7.97
14/16		4.05	5.27		2.72	3.08		3.64	3.80		5.44	7.00
15/16			4.84			2.94			3.53			6.15
16/16	4.40	4.56	5.16	2.24	2.45	2.80	2.79	3.09	3.26	4.13	4.74	5.28

REFERENCES

[1] AXELSSON, A.O.H., *Global integration of differential equations through Lobatto quadrature*, BIT 4, 69-86, 1964.

[2] AXELSSON, A.O.H., I. GUSTAFSSON, *Quasioptimal finite element approximation of first order hyperbolic and of convection-dominated convection diffusion equations*, in: *Analytical and numerical approaches to asymptotics in analysis*, eds. O. Axelsson, L.S. Frank and A. van der Sluis, North-Holland Mathematics Studies 47, North-Holland Publishing Company, Amsterdam-New York-Oxford. 1981.

[3] AXELSSON, A.O.H., *Conjugate gradient type methods for unsymmetric and inconsistent systems of linear equations*, Linear Algebra and its Applications 29, 1-16, 1980.

[4] CASH, J.R., *Stable recursions*, Academic Press, London, 1979.

[5] DAHLQUIST, G., *Stability and error bounds in the numerical integration of ordinary differential equations*, Trans. Roy. Inst. Tech. No 130, Stockholm, 1959.

[6] DELFOUR, M., W. HAGER and F. TROCHU, *Discontinuous Galerkin methods for ordinary differential equations*, Mathematics of Computation 36, 455-473, 1981.

[7] ENRIGHT, W.H., T.E. HULL and B. LINDBERG, *Comparing numerical methods for stiff systems of* ODEs, BIT 15, 10-48, 1975.

[8] FOX, L., *A note on the numerical integration of first order differential equations*, Quart. J. Mech. and Applied Math. 7, 367-378, 1954.

[9] FOX, L., A.R. MITCHELL, *Boundary value techniques for the numerical solution of initial value problems in ordinary differential equations*, Quart. J. Mech. and Applied Math. 10, 232-243, 1957.

[10] FRANK, R., J. SCHNEID, C.W. UEBERHUBER, *Einseitige Lipschitzbedingungen für gewöhnliche differentialgleichungen*, Bericht No 33/78, institut für Numerische Mathematik, TU-Wien, 1978.

[11] HAGEMAN, L.A., D.M. YOUNG, *Applied iterative methods*, Academic Press, New York, 1981.

[12] HENRICI, P., *Discrete variable methods for ordinary differential equations*, Wiley, New York, 1962.

[13] HINDMARSH, A.C., *Gear, ordinary differential equation system solver*, Lawrence Livermore Laboratory Report UCID-30001, Rev. 3, 1974.

[14] HULME, B.L., *Discrete Galerkin and related one-step methods for ordinary differential equations*, Mathematics of Computation 26, 881-891, 1972.

[15] KELLER, H.B., *Numerical methods for two-point boundary-value problems*, Ginn-Blaisdell, Waltham, Mass., 1968.

[16] KELLER, H.B., *Numerical solution of two-point boundary value problems*, SIAM Regional Conference Series in Applied Mathematics 24, 1976.

[17] LAMBERT, J.D. *Computational methods in ordinary differential equations*,

Wiley, London, 1973.

[18] LANCASTER, P., *Theory of matrices*, Academic Press, New York and London, 1969.

[19] OLVER, F.W.J., *Numerical solution of second order linear difference relations*, J. Res. NBS 71B, 111-129, 1967.

[20] OLVER, F.W.J., D.J. SOOKNE, *Note on backward recurrence algorithms*, Mathematics of Computation 26, 941-947, 1972.

[21] ROLFES, L., *A global method for solving stiff differential equations*, NRIMS Special Report, TWISK 228, CSIR, NRIMS, Pretoria, 1981.

[22] ROLFES L., J.A. SNYMAN, *An evalution of a global method applied to stiff ordinary differential equations*, preprint University of Pretoria, 1982.

[23] SHAMPINE, L.F., M.K. GORDON, *Computer solution of ordinary differential equations*, W.H. Freeman & Co, San Francisco, 1975.

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