Czechoslovak Mathematical Journal

Karel Višňák

A-stable block implicit methods containing second derivatives

Czechoslovak Mathematical Journal, Vol. 27 (1977), No. 1, 14–42

Persistent URL: http://dml.cz/dmlcz/101443

Czechoslovak Mathematical Journal, 27 (102) 1977, Praha

A-STABLE BLOCK IMPLICIT METHODS CONTAINING SECOND DERIVATIVES

KAREL VIŠŇÁK, Praha

(Received September 13, 1974)

1. INTRODUCTION

Many technical problems leading to initial-value problems for ordinary differential equations (O.D.E.) are stiff. Using "basic" numerical methods (such as the linear multistep methods or the Runge-Kutta methods) we must integrate such equations using extremely small steps. The need for very small steps is not caused by the accuracy required but it results from the requirements on the stability of the numerical solution. Therefore, an efficient method for solving stiff systems of O.D.E. should not only have a high speed of convergence, but it should satisfy other requirements among which the A-stability has often proved reasonable. It is well-known that, in the class of "basic" methods, A-stable methods of order higher than 2 do not exist, and the methods exhibiting the A-stability are implicit.

PRÁGER, TAUFER, VITÁSEK have introduced in [2] a class of overimplicit multistep methods including A-stable methods of arbitrarily high orders. The implicit character of these methods is emphasized by the fact that, instead of computing the approximate solution at one point from the known approximate solutions at k preceding points, the approximate solutions at r successive points are calculated simultaneously from the known approximate solutions at k preceding points. A special set of these formulae is a class of one-step methods studied in [1], [2]. These methods calculate the approximate solutions at r successive points simultaneously from the known approximate solution at one preceding point. Only the r-th value is used again as a new starting value. At each step, the new r values are calculated from a certain (generally nonlinear) system of equations. It has been shown in [1] that an A-stable method calculating r new values simultaneously can be constructed for every positive integer r so that the speed of convergence of this method is $O(h^{r+1})$ for $h \to 0$.

A new class of implicit one-step methods including A-stable methods of high orders is introduced in this paper. The system of r (generally nonlinear) equations must be also solved at each step in order to obtain the new r approximate solutions

at r successive points. But on the contrary to the methods in [1], [2], we are able to construct an A-stable method converging with the speed $O(h^{2r+1})$ or $O(h^{2r+2})$ for $h \to 0$ for a fixed number r. This is the contents of §§ 2–5 of the present paper. Our formulae contain the second derivatives of the solution. Thix can be considered a disadvantage of these methods. In § 6, we show how to avoid this disadvantage. We shall give an efficient algorithm for solving the nonlinear systems arising at each step of our method. This algorithm uses only the Jacobian matrix of the right-hand side of the original system of O.D.E. Using this algorithm we are nearly in the same situation as when we solve a system of equations arising by applying the methods quoted in [1], [2] to the original system of O.D.E. by the Newton method. In addition, the algorithm uses an A-stable one-step explicit nonlinear formula quoted in [3] for obtaining a good initial approximation for solving the nonlinear system by our Newton-like iterative procedure. So we are able to integrate very stiff systems of O.D.E. with extremely large steps guaranteeing at the same time a satisfactory accuracy of the approximate solution. This is illustrated in § 7 where two A-stable methods are tested on a very stiff problem arising in reactor kinetics.

2. THE *BIM2*$_r$ METHODS

Now we introduce a class of methods for solving the initial-value problems for systems of O.D.E. For the sake of simplicity, all statements will be formulated and proved for a single differential equation of the form

$$y' = f(x, y(x)), \quad y(x_0) = y_0, \quad x \in \langle x_0, b \rangle . \tag{2.1}$$

The changes that are to be made for systems of O.D.E. are given in remarks.

Notation 2.1. In what follows, h denotes a real positive number, the step size. Further, r denotes a positive integer, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ are real column vectors of the length r. The symbol $\mathbf{e}$ denotes the column vector of the length r, the elements of which are 1; $\mathbf{B}$, $\mathbf{C}$ are real $r \times r$ matrices. The symbol f denotes the right-hand side of (2.1) and the symbol f' denotes the function $\mathrm{d}f/\mathrm{d}x = \partial f/\partial x + (\partial f/\partial y) f$. Further, we suppose that the equation (2.1) has exactly one solution, denoted by $y(x)$.

Definition 2.1. Let the equation (2.1) be given and let h and r be fixed numbers. Let $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ be some chosen vectors of the length r and let $\mathbf{B}$, $\mathbf{C}$ be some given $r \times r$ matrices. Consider the following relation:

$$\begin{bmatrix} y_{n+1} \\ \vdots \\ y_{n+r} \end{bmatrix} = \mathbf{e} \,.\, y_n + h \,.\, \boldsymbol{\beta} \,.\, f_n + h^2 \,.\, \boldsymbol{\gamma} \,.\, f'_n + h \,.\, \mathbf{B} \begin{bmatrix} f_{n+1} \\ \vdots \\ f_{n+r} \end{bmatrix} + h^2 \,.\, \mathbf{C} \begin{bmatrix} f'_{n+1} \\ \vdots \\ f'_{n+r} \end{bmatrix}, \tag{2.2}$$

where $f_{n+j} = f(x_{n+j}, y_{n+j})$, $f'_{n+j} = f'(x_{n+j}, y_{n+j})$ and $x_{n+j} = x_n + jh$, $j = 1, \ldots, r$.

Let us put $n = r \,.\, k$ for $k = 0, 1, \ldots$ an let us determine $y_{n+1}, \ldots, y_{n+r}$ from (2.2).

Then we say that *we solve the differential equation* (2.1) *by the Block Implicit Method* 2 (***BIM2***$_r$ method). The values y_{kr} will be called *the approximate solutions of the equation* (2.1) *at the points* x_{kr}, $k = 0, 1, \ldots$ and the number h will be called *the step* (*or step size*) *of the method.*

So every ***BIM2***$_r$ method is associated with a fixed positive integer r which says how many new values are to be calculated from the known one at each step. Further, the ***BIM2***$_r$ method is described by the matrices $\mathbf{B}$, $\mathbf{C}$ and by the vectors $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$.

Remark 2.1. When we have to solve a system of q O.D.E. of the form (2.1), then $\mathbf{f}$ is assumed to be a q-dimensional vector-valued function of $q + 1$ variables. The symbol $\mathbf{y}_{n+j}$ denotes the q-dimensional vector of apprpximate values of the solution of the system (2.1) at the point x_{n+j}, $j = 1, \ldots, r$. Using the ***BIM2***$_r$ method, we now compute r vectors $\mathbf{y}_{n+j}$ at each step. The formula (2.2) in Definition 2.1 is to be replaced by the following one:

$$(2.3)\qquad \begin{bmatrix} \mathbf{y}_{n+1} \\ \vdots \\ \mathbf{y}_{n+r} \end{bmatrix} - {}'\mathbf{e}\mathbf{y}_n - h\,{}'\boldsymbol{\beta}\mathbf{f}_n - h^2\,{}'\boldsymbol{\gamma}\mathbf{f}'_n - h\,{}'\mathbf{B}\begin{bmatrix} \mathbf{f}_{n+1} \\ \vdots \\ \mathbf{f}_{n+r} \end{bmatrix} - h^2\,{}'\mathbf{C}\begin{bmatrix} \mathbf{f}'_{n+1} \\ \vdots \\ \mathbf{f}'_{n+r} \end{bmatrix} = 0\,,$$

where

$${}'\mathbf{e} = \begin{bmatrix} \mathbf{I}_q \\ \vdots \\ \mathbf{I}_q \end{bmatrix}, \qquad {}'\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \mathbf{I}_q \\ \vdots \\ \beta_r \mathbf{I}_q \end{bmatrix}, \qquad {}'\boldsymbol{\gamma} = \begin{bmatrix} \gamma_1 \mathbf{I}_q \\ \vdots \\ \gamma_r \mathbf{I}_q \end{bmatrix},$$

$${}'\mathbf{B} = \begin{bmatrix} b_{1,1}\mathbf{I}_q, \ldots, b_{1,r}\mathbf{I}_q \\ \vdots \qquad\qquad \vdots \\ b_{r,1}\mathbf{I}_q, \ldots, b_{r,r}\mathbf{I}_q \end{bmatrix}, \qquad {}'\mathbf{C} = \begin{bmatrix} c_{1,1}\mathbf{I}_q, \ldots, c_{1,r}\mathbf{I}_q \\ \vdots \qquad\qquad \vdots \\ c_{r,1}\mathbf{I}_q, \ldots, c_{r,r}\mathbf{I}_q \end{bmatrix}$$

and where $\mathbf{I}_q$ is the $q \times q$ unit matrix and $\mathbf{f}' = \partial\mathbf{f}/\partial x + \mathbf{J} \,.\, \mathbf{f}$, where $\mathbf{J}$ is the Jacobian matrix of the vector-valued function $\mathbf{f}$.

Now we shall examine the properties of the matrices $\mathbf{B}$, $\mathbf{C}$, vectors $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and of the right-hand side f in (2.1) that will guarantee that the approximate solutions obtained by the ***BIM2***$_r$ method converge with a certain speed to the exact solution of (2.1) for $h \to 0$.

Definition 2.2. We say that the approximate solutions y_{kr}, $k = 0, 1, \ldots$ obtained by means of the ***BIM2***$_r$ method *converge to the exact solution* $y(x)$ *of the equation* (2.1) *with the speed* $O(h^p)$ for $h \to 0$, $p \geqq 1$, if

$$(2.4)\qquad |y_{kr} - y(x_{kr})| = O(h^p) \quad \text{for} \quad h \to 0\,, \quad h = (x - x_0)/(kr)$$

at every fixed point $x = x_{kr}$ from $\langle x_0, b\rangle$.

Now we introduce the local error and the order of the ***BIM2***$_r$ method. These notions will be used later.

Definition 2.3. The following vector is said to be *the local error of the* ***BIM2***$_r$ *method applied to the equation* (2.1):

$$\mathbf{L}(h, y(x)) = \begin{bmatrix} y(x+h) \\ \vdots \\ y(x+rh) \end{bmatrix} - \mathbf{e}\, y(x) - h\boldsymbol{\beta} f_0 - h^2 \gamma f_0' - h\mathbf{B} \begin{bmatrix} f_1 \\ \vdots \\ f_r \end{bmatrix} - h^2 \mathbf{C} \begin{bmatrix} f_1' \\ \vdots \\ f_r' \end{bmatrix} \tag{2.5}$$

where $f_j = f(x + jh, y(x + jh))$, $f_j' = f'(x + jh, y(x + jh))$, $j = 0, \ldots, r$.

Remark 2.2. For the system of O.D.E. the symbols $\mathbf{B}$, $\mathbf{C}$, $\mathbf{e}$, $\boldsymbol{\beta}$, γ are to be replaced by $'\mathbf{B}$, $'\mathbf{C}$, ... in the sense of (2.3).

Definition 2.4. The ***BIM2***$_r$ method is said to be of *order at least* p, $p \geqq 1$, if the following relations are fulfilled for the components of $\boldsymbol{\beta}$, γ, $\mathbf{B}$, $\mathbf{C}$:

$$\beta_j + \sum_{k=1}^{r} b_{j,k} = j\,, \tag{2.6}$$

$$\gamma_j + \sum_{k=1}^{r} b_{j,k} \cdot k + \sum_{k=1}^{r} c_{j,k} = \frac{j^2}{2!}\,, \quad j = 1, \ldots, r\,,$$

$$\frac{1}{(i-1)!} \cdot \sum_{k=1}^{r} b_{j,k} \cdot k^{i-1} + \frac{1}{(i-2)!} \cdot \sum_{k=1}^{r} c_{j,k} \cdot k^{i-2} = \frac{j^i}{i!}\,, \tag{2.7}$$

$$i = 3, \ldots, p\,, \quad j = 1, \ldots, r\,.$$

Now we establish the mutual relation between the order of the ***BIM2***$_r$ method and the local error of the ***BIM2***$_r$ method applied to the equation (2.1).

Lemma 2.1. *Let p be a positive integer and let $y(x) \in C^{p+1}(\langle x_0, b\rangle)$. Let L_i, $i = 1, \ldots, r$ be the components of the vector of the local error. Then there exists a constant K (independent of x) so that*

$$|L_i(h, y(x))| \leqq K \,.\, h^{p+1}\,, \quad i = 1, \ldots, r\,, \quad x \in \langle x_0, b\rangle \tag{2.8}$$

if and only if the ***BIM2***$_r$ *method is of order at least p. Moreover, if $y(x) \in C^{p+2}(\langle x_0, b\rangle)$ then there exist constants K_1, K_2 (independent of x) so that*

$$|L_i(h, y(x))| \leqq K_1 \,.\, h^{p+1}\,, \quad i = 1, \ldots, r-1\,, \tag{2.9}$$

$$|L_r(h, y(x))| \leqq K_2 \,.\, h^{p+2}\,, \quad x \in \langle x_0, b\rangle$$

if and only if the ***BIM2***$_r$ *method is of order at least p and the relation* (2.7) *is also fulfilled for $i = p + 1$ with $j = r$.*

Proof follows immediately from Taylor's formula applied to the terms in (2.5).

Remark 2.3. For systems of O.D.E., we assume the vector of the local error to be divided into r blocks $\mathbf{L}_i$ of the length q and therefore the symbols $|\ |$ mean the vector norm.

Now we are able to formulate the main theorem of this section.

Theorem 2.1. *Let p be a positive integer. Let N, M be real constants such that*

$$|f(x, y) - f(x, z)| \leqq N|y - z|\,, \quad |f'(x, y) - f'(x, z)| \leqq M|y - z|\,,$$
$$\text{for} \quad x \in \langle x_0, b\rangle\,, \quad y, z \in E_1\,.$$

Let $y(x) \in C^{p+1}(\langle x_0, b\rangle)$ and let the **BIM2$_r$** *method of the order at least p be given. Then*

1) *the approximate solutions y_{kr} obtained by the* **BIM2$_r$** *method applied to the equation (2.1) converge to the exact solution $y(x)$ with the speed $O(h^p)$ for $h \to 0$.*

2) *Moreover, if $y(x) \in C^{p+2}(\langle x_0, b\rangle)$ and the relation (2.7) is fulfilled also for $i = p + 1$ with $j = r$, the speed of convergence is $O(h^{p+1})$ for $h \to 0$.*

Proof. We prove the second statement only. The proof of the first statement is analogous, but shorter and easier.

Let us introduce the following notation:

$$e_{n+i} = y(x_{n+i}) - y_{n+i}\,,$$

$$s_{n+i} = \begin{cases} \dfrac{f(x_{n+i}, y(x_{n+i})) - f(x_{n+i}, y_{n+i})}{e_{n+i}} & \text{for} \quad e_{n+i} \neq 0 \\ 0 & \text{for} \quad e_{n+i} = 0\,, \end{cases}$$

$$t_{n+i} = \begin{cases} \dfrac{f'(x_{n+i}, y(x_{n+i})) - f'(x_{n+i}, y_{n+i})}{e_{n+i}} & \text{for} \quad e_{n+i} \neq 0 \\ 0 & \text{for} \quad e_{n+i} = 0 \end{cases}$$

for $i = 1, \ldots, r$.

Let us further denote by $\mathbf{S}$ the diagonal matrix

$$\mathbf{S} = \operatorname{diag}(s_{n+1}, \ldots, s_{n+r})$$

and by $\mathbf{T}$ the diagonal matrix

$$\mathbf{T} = \operatorname{diag}(t_{n+1}, \ldots, t_{n+r})\,.$$

According to the assumptions of the theorem, it holds

$$|s_{n+i}| \leqq N\,, \quad i = 1, \ldots, r\,, \tag{2.10}$$

$$|t_{n+i}| \leqq M\,, \quad i = 1, \ldots, r\,. \tag{2.11}$$

Let us put $x = x_n$ in (2.5) and let us subtract (2.2) from (2.5). Further, let us substitute the terms introduced above into this new expression. We have

$$(\mathbf{I} - h\mathbf{BS} - h^2\mathbf{CT}) \begin{bmatrix} e_{n+1} \\ \vdots \\ e_{n+r} \end{bmatrix} = (\mathbf{e} + h\beta s_n + h^2\gamma t_n)\, e_n + \mathbf{L}(h, y(x_n))\,. \tag{2.12}$$

In the following, we shall use the norms of vectors and matrices. For the sake of simplicity, we shall assume the norm of the vector $\mathbf{u} = (u_i)$ to be defined by $\|\mathbf{u}\| = \max_i |u_i|$ and the matrix norm to be induced in the usual sense by this vector norm.

First of all, we show that there exists a constant $h_2 > 0$ so that $(\mathbf{I} - h\mathbf{BS} - h^2\mathbf{CT})^{-1}$ exists for $0 < h < h_2$ and the following estimate holds:

$$\|(\mathbf{I} - h\mathbf{BS} - h^2\mathbf{CT})^{-1}\| \leqq \frac{1}{1 - h\|\mathbf{B}\| N - h^2\|\mathbf{C}\| M}\,. \tag{2.13}$$

Obviously $\|h\mathbf{BS} + h^2\mathbf{CT}\| \leqq h\|\mathbf{B}\| N + h^2\|\mathbf{C}\| M$, and the inequality $h^2\|\mathbf{C}\| M + h\|\mathbf{B}\| N - 1 < 0$ holds for $0 < h < h_2$ where h_2 is given by

$$h_2 = \begin{cases} \dfrac{-\|\mathbf{B}\| N + (\|\mathbf{B}\|^2 N^2 + 4\|\mathbf{C}\| M)^{0.5}}{2\|\mathbf{C}\| M}\,, & \|\mathbf{C}\| \neq 0 \\ \dfrac{1}{\|\mathbf{B}\| N}\,, & \|\mathbf{C}\| = 0\,,\ \|\mathbf{B}\| \neq 0\,. \end{cases} \tag{2.14}$$

Hence we conclude that $(\mathbf{I} - h\mathbf{BS} - h^2\mathbf{CT})^{-1}$ exists for $0 < h < h_2$ and (2.13) holds. Let us confine ourselves to $0 < h < h_2$. We multiply (2.12) by the matrix $(\mathbf{I} - h\mathbf{BS} - h^2\mathbf{CT})^{-1}$ and take the norms on both sides. Using the assumptions of the theorem and Lemma 2.1 for the estimate of $\|\mathbf{L}\|$ and considering (2.10), (2.11), (2.13) we have

$$|e_{n+i}| \leqq G(h)\,|e_n| + H(h)\,, \quad i = 1, \ldots, r\,, \tag{2.15}$$

where

$$G(h) = \frac{1 + h\|\beta\| N + h^2\|\gamma\| M}{1 - h\|\mathbf{B}\| N - h^2\|\mathbf{C}\| M}\,, \qquad H(h) = \frac{K_1 h^{p+1}}{1 - h\|\mathbf{B}\| N - h^2\|\mathbf{C}\| M}\,.$$

Let us take only the last, i.e., the r-th equation in (2.12). After an arrangement we obtain

$$(2.16)\qquad (1 - hb_{r,r}s_{n+r} - h^2c_{r,r}t_{n+r})\, e_{n+r} =$$
$$= (1 + h\beta_r s_n + h^2\gamma_r t_n)\, e_n + h\sum_{i=1}^{r-1} u_i e_{n+i} + L_r\,,$$

where

$$u_i = b_{r,i}s_{n+i} + hc_{r,i}t_{n+i}\,, \quad i = 1, \ldots, r-1$$

and L_r is the r-th component of the vector of the local error. In a similar way as before, we conclude that there exists $h_1 > 0$ so that $(1 - hb_{r,r}s_{n+r} - h^2c_{r,r}t_{n+r}) \neq 0$ for $0 < h < h_1$ and

$$(2.17)\qquad \left|\frac{1}{1 - hb_{r,r}s_{n+r} - h^2c_{r,r}t_{n+r}}\right| \leqq \frac{1}{1 - h|b_{r,r}|\, N - h^2|c_{r,r}|\, M}\,.$$

In the following, we confine ourselves to $h < \min(h_1, h_2)$. In the relation (2.16), we divide by $(1 - hb_{r,r}s_{n+r} - h^2c_{r,r}t_{n+r})$ and take the modulus on both sides. Using (2.10), (2.11), (2.17) and Lemma 2.1, we obtain

$$(2.18)\qquad |e_{n+r}| \leqq a_0(h)\,|e_n| + h\sum_{i=1}^{r-1} a_i(h)\,|e_{n+i}| + a_r(h)\,,$$

where

$$a_0(h) = \frac{1 + h|\beta_r|\, N + h^2|\gamma_r|\, M}{1 - h|b_{r,r}|\, N - h^2|c_{r,r}|\, M}\,, \quad a_r(h) = \frac{K_2 h^{p+2}}{1 - h|b_{r,r}|\, N - h^2|c_{r,r}|\, M}\,,$$

$$a_i(h) = \frac{|b_{r,i}|\, N + h|c_{r,i}|\, M}{1 - h|b_{r,r}|\, N - h^2|c_{r,r}|\, M}\,, \quad i = 1, \ldots, r-1\,.$$

Substituting from (2.15) into (2.18), we have

$$(2.19)\qquad |e_{n+r}| \leqq P(h)\,|e_n| + R(h)$$

where

$$P(h) = a_0(h) + h\,G(h)\sum_{i=1}^{r-1} a_i(h)\,, \quad R(h) = a_r(h) + h\,H(h)\sum_{i=1}^{r-1} a_i(h)\,.$$

Now we recall that $n = rk$. Solving the recurrence (2.19), we obtain

$$(2.20)\qquad |e_{rk}| \leqq P^k(h)\,|e_0| + R(h)\,\frac{P^k(h) - 1}{P(h) - 1}\,.$$

Now we shall study the behaviour of the right-hand side of (2.20) at an arbitrary

but fixed point $x \in \langle x_0, b\rangle$ for $k \to \infty$ and $h = (x - x_0)/(rk)$. Let us calculate $\lim_{k\to\infty} P^k((x - x_0)/(rk))$. Using the l'Hospital rule, we have

$$\lim_{\substack{k\to\infty \\ h=(x-x_0)/(rk)}} P^k(h) = \exp\left(N \frac{x - x_0}{r} \left(|\beta_r| + \sum_{i=1}^{r} |b_{r,i}|\right)\right). \tag{2.21}$$

Further, we obtain easily

$$\lim_{\substack{k\to\infty \\ h=(x-x_0)/(rk)}} \frac{R(h)}{(P(h) - 1) \cdot h^{p+1}} = \frac{K_1 N \sum_{i=1}^{r-1} |b_{r,i}| + K_2}{N(|\beta_r| + \sum_{i=1}^{r} |b_{r,i}|)}. \tag{2.22}$$

Substituting from (2.21) and (2.22) into (2.20) and taking into account that $e_0 = 0$, we have

$$\lim_{\substack{k\to\infty \\ h=(x-x_0)/(rk)}} |e_{rk}/h^{p+1}| \leqq \left(\exp\left(N \frac{x - x_0}{r} \left(\beta_r + \sum_{i=1}^{r-1} |b_{r,i}|\right)\right) - 1\right) \frac{K_1 N \sum_{i=1}^{r-1} |b_{r,i}| + K_2}{N(|\beta_r| + \sum_{i=1}^{r} |b_{r,i}|)}.$$

This relation proves the second statement of the theorem.

3. A-STABILITY

Definition 3.1. A numerical method for solving initial-value problems for O.D.E. is said to be *A-stable*, if the numerical solutions y_k obtained by applying the method to the equation

$$y' = \alpha y \quad (\alpha - \text{complex constant with } \operatorname{Re}(\alpha) < 0) \tag{3.1}$$

with an arbitrary (but fixed) step size $h > 0$ tend to zero for $k \to \infty$.

In the following, we derive some useful lemmas for studying the A-stability of the ***BIM2***$_r$ methods. We substitute from (3.1) into (2.2) and obtain a system of linear algebraic equations for unknown values of the numerical solution y_{n+k}, $k = 1, \ldots, r$:

$$(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C}) \begin{bmatrix} y_{n+1} \\ \vdots \\ y_{n+r} \end{bmatrix} = (\mathbf{e} + \boldsymbol{\beta} z + \boldsymbol{\gamma} z^2)\, y_n \tag{3.2}$$

where $z = \alpha h$.

Notation 3.1. Let a certain ***BIM2***$_r$ method, i.e., matrices $\mathbf{B}, \mathbf{C}$ and vectors $\boldsymbol{\beta}, \boldsymbol{\gamma}$ be given. We shall denote by $\mathbf{D}_k$, $k = 1, \ldots, r$ the matrices arising from $(\mathbf{I} - z\mathbf{B} -$

$-z^2\mathbf{C}$) by replacing the k-th column by the vector $(\mathbf{e} + \boldsymbol{\beta}z + \gamma z^2)$. Further, we shall denote

$$P_0(z) = \det(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C}) \tag{3.3}$$

and

$$P_k(z) = \det \mathbf{D}_k\,, \quad k = 1, \ldots, r\,. \tag{3.4}$$

Obviously, $P_j(z)$, $j = 0, \ldots, r$ are polynomials in z of degree at most $2r$ and normalized so that $P_j(0) = 1$.

Using the above notation and *Cramer's rule* in (3.2), we find

$$y_{n+k} = \frac{P_k(z)}{P_0(z)}\, y_n\,, \quad k = 1, \ldots, r\,. \tag{3.5}$$

If y_n is the value obtained after j steps of the **BIM2**$_r$ method, it holds

$$y_n = y_{jr} = \left(\frac{P_r(z)}{P_0(z)}\right)^j y_0\,. \tag{3.6}$$

Definition 3.1 together with the relation (3.6) immediately imply

Lemma 3.1. *Let a certain* **BIM2**$_r$ *method be given and let* P_0, P_r *be the polynomials defined by the relations* (3.3) *and* (3.4).

Then the given **BIM2**$_r$ *method is A-stable if and only if*

$$|P_r(z)/P_0(z)| < 1$$

for every z *with* $\mathrm{Re}(z) < 0$.

Further, we shall profit from

Lemma 3.2. *Suppose* $P(z)$ *is a real polynomial such that* $P(z) = P(-z)$ *does not hold identically. Let all the roots of* $P(z)$ *which are not roots of* $P(-z)$ *have positive real parts. Then*

$$|P(-z)/P(z)| < 1 \quad \textit{holds for all } z\,, \quad \mathrm{Re}\,(z) < 0\,. \tag{3.7}$$

Conversely, let there exists z, $\mathrm{Re}\,(z) < 0$ *such that* $P(z) = 0$ *and* $P(-z) \neq 0$. *Then* (3.7) *is not true.*

Proof. The second part of the assertion is obvious. We prove the first part in two steps.

1. First, we suppose that all the roots of the polynomial P have positive real parts. Let us write the ratio $P(-z)/P(z)$ in the form of a product

$$\frac{P(-z)}{P(z)} = \left(\frac{-z - a_1}{z - \bar{a}_1}\right)^{b_1} \left(\frac{-z - \bar{a}_1}{z - a_1}\right)^{b_1} \cdots \tag{3.7a}$$

$$\cdots \left(\frac{-z - a_s}{z - \bar{a}_s}\right)^{b_s} \left(\frac{-z - \bar{a}_s}{z - a_s}\right)^{b_s} \left(\frac{-z - c_1}{z - c_1}\right)^{d_1} \cdots \left(\frac{-z - c_k}{z - c_k}\right)^{d_k}$$

where a_i for $i = 1, \ldots, s$ are complex roots of P with multiplicities b_i; c_i for $i = 1, \ldots, k$ are real roots of P with multiplicities d_i.

According to the relation (3.7a) it remains to show that

$$\left|\frac{-z-a}{z-\bar{a}}\right| < 1 \quad \begin{array}{l} \text{holds for every } z\,, \quad \mathrm{Re}\,(z) < 0 \\ \text{and for every } a\,, \quad \mathrm{Re}\,(a) > 0\,. \end{array} \tag{3.7b}$$

The relation (3.7b) can be readily proved by direct calculation.

2. Now, let $P(z)$ and $P(-z)$ have a certain number of common roots and let all the remaining roots of $P(z)$ have positive real parts. Then $|P(-z)/P(z)| = |R(-z)/R(z)|$ where the polynomial $R(z)$ has all the roots with positive real parts and we can apply the assertion proved above. Q.E.D.

Lemma 3.3. *Let a certain* **BIM2**$_r$ *method of order* $p \geqq 2r$ *be given. Let* $P_0(z) = \sum_{i=0}^{2r} a_i z^i$ *and* $P_k(z) = \sum_{i=0}^{2r} a_{k,i} z^i$ *for* $k = 1, \ldots, r$ *be real polynomials defined for the given method by the relations* (3.3) *and* (3.4). *Then*

1) *the coefficients of the polynomials* P_k *are given in terms of the coefficients of the polynomial* P_0 *as follows*:

$$a_{k,i} = \sum_{j=0}^{i} \frac{k^{i-j}}{(i-j)!}\, a_j\,, \quad i = 0, \ldots, 2r\,, \quad k = 1, \ldots, r\,. \tag{3.8}$$

2) *Moreover, if the order of the given method is* $p = 2r + 2$, *the coefficients of the polynomial* P_0 *fulfil the relations*

$$\sum_{j=0}^{2r} a_{2r-j} \frac{k^{j+1}}{(j+1)!} = 0\,, \quad \sum_{j=0}^{2r} a_{2r-j} \frac{k^{j+2}}{(j+2)!} = 0\,, \quad k = 1, \ldots, r\,. \tag{3.9}$$

3) *The relations* (3.9) *viewed as a system of linear algebraic equations yield exactly one set of numbers* $a_1, \ldots, a_{2r}$ *for fixed* a_0.

Proof. Taking into account that the method is of order at least p, we obtain from (2.5) and (3.1) the following relation

$$(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C}) \begin{bmatrix} y(x_n + h) \\ \vdots \\ y(x_n + rh) \end{bmatrix} = (\mathbf{e} + z\boldsymbol{\beta} + z^2\boldsymbol{\gamma})\, y(x_n) + Q(h^{p+1})\,. \tag{3.10}$$

Applying *Cramer's rule* to (3.10), we have

$$y(x_n + kh) = \frac{P_k(z)}{P_0(z)}\, y(x_n) + q_k \tag{3.11}$$

where P_j, $j = 0, \ldots, r$ are defined by the relations (3.3), (3.4) and q_k is a solution of the system

$$(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C}) \begin{bmatrix} q_1 \\ \vdots \\ q_r \end{bmatrix} = O(h^{p+1}). \tag{3.12}$$

Taking into account that $(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C})^{-1} = O(1)$ for $h \to 0$, we rewrite (3.11) in the following way:

$$y(x_n + kh) = \frac{P_k(z)}{P_0(z)} y(x_n) + Q(h^{p+1}). \tag{3.13}$$

Further, we put $\alpha = 1$ and $y(0) = 1$ in (3.1). Then the exact solution of (3.1) is $y(x) = \exp(x)$ and, by substituting into (3.13), we obtain

$$P_0(h) \exp(kh) - P_k(h) = O(h^{p+1}). \tag{3.14}$$

Let us substitute the relations $P_0(h) = \sum_{i=0}^{2r} a_i h^i$ and $P_k(h) = \sum_{i=0}^{2r} a_{k,i} h^i$ for $k = 1, \ldots, r$ into (3.14) and let us order the new expression according to the powers of h. The right-hand side of (3.14) is of order $O(h^{p+1})$ and, therefore, the coefficients at h^j for $j = 0, \ldots, p$ must be equal to zero. For $p = 2r$ this immediately implies the assertion 1.

For $p = 2r + 2$ the coefficients at h^{2r+1} and h^{2r+2} must be also equal to zero, which implies the relations (3.9).

It remains to prove the third assertion. The relations (3.9) can be viewed as conditions for the polynomial

$$P(x) = \sum_{j=0}^{2r} d_{j+2} x^{j+2}, \quad d_{j+2} = a_{2r-j}/(j+2)! \tag{3.15}$$

of order at most $2r + 2$ to vanish together with its first derivative at the $r + 1$ points $0, 1, \ldots, r$. Thus the desired result follows immediately from the interpolation theory. Q.E.D.

Remark 3.1. We denote by $\mathbf{M}$ the diagonal matrix with elements $a_{i,i} = i$, $i = 1, \ldots, r$ and by $\mathbf{e}$ the column-vector of length r all elements of which are 1. Then we can rewrite (3.9) in the following form:

$$\mathbf{S} \begin{bmatrix} a_{2r} \\ \vdots \\ a_1 \end{bmatrix} = -a_0 \begin{bmatrix} \dfrac{\mathbf{M}^{2r+1}\mathbf{e}}{(2r+1)!} \\ \dfrac{\mathbf{M}^{2r+2}\mathbf{e}}{(2r+2)!} \end{bmatrix} \tag{3.16}$$

where

$$\mathbf{S} = \begin{bmatrix} \frac{\mathbf{M}^1\mathbf{e}}{1!}, & \frac{\mathbf{M}^2\mathbf{e}}{2!}, & \dots, & \frac{\mathbf{M}^{2r}\mathbf{e}}{2r!} \\ \frac{\mathbf{M}^2\mathbf{e}}{2!}, & \frac{\mathbf{M}^3\mathbf{e}}{3!}, & \dots, & \frac{\mathbf{M}^{2r+1}\mathbf{e}}{(2r+1)!} \end{bmatrix}.$$

From Lemma 3.3 we know that the matrix $\mathbf{S}$ is non-singular.

4. METHODS OF THE MAXIMAL ORDER

Definition 4.1. Let r be a fixed positive integer. Every ***BIM2***$_r$ method of order $2r + 2$ is said to be *a method of the maximal order* (the ***BIM2M***$_r$ method).

Lemma 4.1. *For every fixed positive integer r, there exists exactly one* ***BIM2M***$_r$ *method.*

Proof. Using the same notation as in Remark 3.1, we can rewrite the conditions (2.7) in the following form:

$$[\mathbf{C}, \mathbf{B}] \,.\, \mathbf{S} = \left[\frac{\mathbf{M}^3\mathbf{e}}{3!}, \frac{\mathbf{M}^4\mathbf{e}}{4!}, \dots, \frac{\mathbf{M}^{2r+2}\mathbf{e}}{(2r+2)!} \right]. \tag{4.1}$$

According to Remark 3.1 the matrix $\mathbf{S}$ is non-singular and solving (4.1) we obtain exactly one pair of matrices $\mathbf{B}$, $\mathbf{C}$. The relations (2.6) yield exactly one pair of vectors β, γ. Q.E.D.

Remark 4.1. The relations (4.1) give us the possibility to construct the ***BIM2M***$_r$ methods and the relations (3.16) determine the coefficients of the polynomial (3.3) constructed for this method.

Lemma 4.2. *Let a certain* ***BIM2M***$_r$ *method be given and let* P_0, P_r *be real polynomials defined for this method by the relations* (3.3), (3.4).
Then $P_r(-z) = P_0(z)$ *identically.*

Proof. We consider the method of the order $2r + 2$, therefore the coefficients of the polynomial P_0 satisfy the relations (3.9) and hence the polynomial (3.15) has roots of multiplicity 2 at $x = 0, \dots, r$, i.e.,

$$P(x) = \sum_{k=0}^{2r} a_{2r-k} \frac{x^{k+2}}{(k+2)!} = x^2(x-1)^2(x-2)^2 \dots (x-r)^2 \,. \tag{4.2}$$

According to (3.8), the coefficients of the polynomial $P_r(z)$ satisfy the relation

$$a_{r,2r-k} = \sum_{j=k}^{2r} a_{2r-j} \frac{r^{j-k}}{(j-k)!}, \quad k = 0, \dots, 2r \,. \tag{4.3}$$

Hence

$$a_{r,2r-k} = P^{(k+2)}(x)\big|_{x=r} \tag{4.4}$$

where the superscript denotes derivative. Let us construct the following polynomial:

$$S(x) = \sum_{k=0}^{2r} a_{r,2r-k}(-1)^k \frac{x^{k+2}}{(k+2)!} \tag{4.5}$$

and let us substitute from (4.4) into (4.5). We obtain

$$S(x) = \sum_{k=0}^{2r} \frac{P(r)^{(k+2)}}{(k+2)!}(-x)^{k+2}.$$

Consequently $S(x) = P(r-x)$, because $P(r) = 0$ and $P^{(1)}(r) = 0$. The relation (4.2) immediately implies $P(r-x) = P(x)$ and hence $P(x) = S(x)$. Comparing (4.5) and (4.2) we have $a_{r,2r-k} = a_{2r-k}(-1)^k$. Q.E.D.

Remark 4.2. The relation (4.2) can be used for a direct calculation of the coefficients of the polynomial P_0. From (4.2) we can also derive relations enabling us to calculate the coefficients of the polynomial P_0 for the **BIM2M**$_{r+1}$ method if the coefficients of P_0 for the **BIM2M**$_r$ method are known.

Theorem 4.1. *Let r be a fixed positive integer and let the* **BIM2M**$_r$ *method be given by* (4.1) *and* (2.6). *Let* $P_0(z) = \sum_{i=0}^{2r} a_i z^i$, $a_0 = 1$, *be the polynomial with* a_i, $i = 1, \ldots, 2r$ *calculated from* (4.2) *or equivalently from* (3.16).

Then the **BIM2M**$_r$ *method is A-stable if and only if all the roots of the polynomial* $P_0(z)$ *which are not roots of the polynomial* $P_0(-z)$ *have positive real parts.*

Proof. The statement of the theorem follows immediately from Lemma 4.2, Lemma 3.2 and Lemma 3.1 if we realize that $P_r(z) = P_0(z)$ cannot hold identically for a convergent method.

Remark 4.3. It has been shown by direct calculation that all roots of the polynomials P_0 defined for the **BIM2M**$_r$ methods, $r = 1, \ldots, 5$, have positive real parts. For $r = 6$ there exists one root of the polynomial P_0 with a negative real part. This root is not a root of $P_0(-z)$. Applying Theorem 4.1 we have

Theorem 4.2. *The* **BIM2M**$_r$ *methods are A-stable for* $r = 1, \ldots, 5$ (*i.e., up to the order* 12). *The* **BIM2M**$_6$ *method is not A-stable.*

The **BIM2M**$_1$ method is the well-known *Pade formula* of the 4-th order: $b = \beta = 0.5$, $c = -\gamma = -1/12$. The coefficients of the **BIM2M**$_2$ method of the order 6 are shown in the following table:

(4.7)

i	β_i	$b_{i,1}$	$b_{i,2}$	$c_{i,1}$	$c_{i,2}$	γ_i
1	101/240	8/15	11/240	−1/6	−1/80	13/240
2	7/15	16/15	7/15	0	−1/15	1/15

Table 4.1. The coefficients of ***BIM2M***$_2$.

5. A-STABLE METHODS OF HIGH ORDERS

In this section it will be discussed how a ***BIM2***$_r$ method of order at least $2r$ can be constructed to a polynomial P of degree exactly $2r$ so that the polynomial P_0 defined for this method by (3.3) and the given polynomial P are identical. Further, the choice of P which yields A-stable ***BIM2***$_r$ methods will be given.

Lemma 5.1. *Let* $P_0(z) = \sum_{i=0}^{2r} a_i z^i$ *be a real polynomial of degree exactly* $2r$ *with* $a_0 = 1$.

Let $P_k(z) = \sum_{i=0}^{2r} a_{k,i} z^i$, $k = 1, \ldots, r$ *be real polynomials with* $a_{k,i}$ *defined by* (3.8). (*Note that* $a_{k,0} = a_k$.)

Then

$$P_0(z), P_1(z), \ldots, P_r(z), z\,P_0(z), z\,P_1(z), \ldots, z\,P_r(z) \tag{5.1}$$

are linearly independent on every interval if and only if

$$\sum_{j=0}^{2r} a_{2r-j} \sum_{k=0}^{r} \binom{r}{k}^2 \frac{k^j}{j!} \neq 0\,. \tag{5.3}$$

Proof. Suppose that

$$\sum_{k=0}^{r} b_k P_k(z) + \sum_{k=0}^{r} c_k z\, P_k(z) = 0 \quad \text{identically}\,. \tag{5.4}$$

The substitution of the expressions for P_k produces $\sum_{k=0}^{2r} \alpha_i z^i = 0$ where

$$\alpha_0 = \sum_{k=0}^{r} a_{k,0} b_k\,, \tag{5.5}$$

$$\alpha_i = \sum_{k=0}^{r} a_{k,i} b_k + \sum_{k=0}^{r} a_{k,i-1} c_k\,, \quad i = 1, \ldots, 2r\,, \tag{5.6}$$

$$\alpha_{2r+1} = \sum_{k=0}^{r} a_{k,2r} c_k\,. \tag{5.7}$$

But the linear independence of $1, z, \ldots, z^{2r+1}$ implies that all α_i are equal to zero. Using (3.8) we obtain from (5.5)–(5.7) successively

$$\sum_{k=0}^{r} b_k = 0 , \tag{5.9a}$$

$$\frac{1}{i!} \sum_{k=0}^{r} k^i b_k + \frac{1}{(i-1)!} \sum_{k=0}^{r} k^{i-1} c_k = 0 , \quad i = 1, \ldots, 2r , \tag{5.9b}$$

$$\sum_{k=0}^{r} c_k \sum_{j=0}^{2r} \frac{k^j}{j!} a_{2r-j} = 0 . \tag{5.9c}$$

We shall show that the equations (5.9) are linearly independent for every choice of a_j satisfying the assumptions of Lemma 5.1 if and only if (5.3) holds. Hence the system (5.9) has exactly one solution $b_k = c_k = 0$ if and only if (5.3) holds.

First we show that the $2r + 1$ equations (5.9a), (5.9b) are linearly independent. Suppose the contrary. Then there exist real numbers r_j, not all equal to zero, so that

$$\sum_{j=0}^{2r} r_j \frac{k^j}{j!} = 0 , \quad \sum_{j=1}^{2r} r_j \frac{k^{j-1}}{(j-1)!} = 0 , \quad k = 0, 1, \ldots, r . \tag{5.10}$$

But the conditions (5.10) tell us that the polynomial

$$P(x) = \sum_{j=0}^{2r} t_j x^j , \quad t_j = r_j/j! \tag{5.11}$$

of degree at most $2r$ should have roots of multiplicity 2 at $x = 0, \ldots, r$, i.e., $P(x)$ should be a polynomial of degree at least $2r + 2$, and this leads to a contradiction.

It remains to show that (5.9c) is a linear combination of the equations (5.9a), (5.9b) if and only if the condition (5.3) is not true. The equation (5.9c) is a such a combination if and only if there exist real numbers r_i, $i = 0, \ldots, 2r$ not all equal to zero, so that

$$\sum_{i=0}^{2r} r_i \frac{k^i}{i!} = 0 , \quad \sum_{i=1}^{2r} r_i \frac{k^{i-1}}{(i-1)!} = \sum_{j=0}^{2r} a_{2r-j} \frac{k^j}{j!} = a_{k,2r} , \quad k = 0, \ldots, r . \tag{5.11a}$$

The conditions (5.11a) tell us that the polynomial (5.11) satisfies the following conditions:

$$P(k) = 0 , \quad P'(k) = a_{k,2r} , \quad k = 0, \ldots, r . \tag{5.11b}$$

The conditions (5.11b) determine exactly one polynomial of degree at most $2r + 1$ and with $t_1 = a_{2r} \neq 0$. Hence the polynomials (5.1) are linearly dependent if and

only if the coefficient at x^{2r+1} in the polynomial given by (5.11b) vanishes. Using Hermitian interpolation, we obtain that the coefficient at x^{2r+1} is

$$const. \times \sum_{k=0}^{r} \binom{r}{k}^2 a_{k,2r} .$$

This yields immediately the condition (5.3). Q.E.D.

Theorem 5.1. *Let* $P_0(z) = \sum_{i=0}^{2r} a_i z^i$ *be a real polynomial of degree exactly* $2r$ *and with* $a_0 = 1$. *Let* $P_k(z) = \sum_{i=0}^{2r} a_{k,i} z^i$, $k = 1, \ldots, r$ *be polynomials with* $a_{k,i}$ *defined by* (3.8). *Let* (5.3) *hold. Then*

1) *we can calculate exactly one set of numbers*

$$\beta_j, \gamma_j, b_{j,k}, c_{j,k} \tag{5.12}$$

from

$$\beta_j P_0(z) + \sum_{k=1}^{r} b_{j,k} P_k(z) + \gamma_j z P_0(z) + \sum_{k=1}^{r} c_{j,k} z P_k(z) = \frac{P_j(z) - P_0(z)}{z}, \tag{5.13}$$

$$j = 1, \ldots, r .$$

2) *The numbers* (5.12) *define a* $\mathbf{BIM2}_r$ *method of order at least* $2r$.

3) $P_0(z) = \det(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C})$ *and* $P_k(z) = \det \mathbf{D}_k$, *where* $\mathbf{D}_k$, $k = 1, \ldots, r$ *are defined in Notation* 3.1.

Proof. According to Lemma 5.1 the polynomials (5.1) form a basis in the $2r + 2$ dimensional space of polynomials having degree at most $2r + 1$. Because of $a_0 = a_{k,0}$, $k = 1, \ldots, r$, the right-hand term in (5.13) is a polynomial of degree at most $2r - 1$, and hence it must have a unique representation in terms of the basis (5.12), which implies the existence of exactly one set of numbers (5.12). Let us choose $2r + 2$ distinct real numbers $p_1, \ldots, p_{2r+2}$ and let us put successively $z = p_i$, $i = 1, \ldots, 2r + 2$ in (5.13). We obtain r systems of $2r + 2$ linear algebraic equations for the numbers (5.12). Later we shall show a better possibility how the numbers (5.12) can be calculated.

Now let us prove the second statement of Theorem 5.1. The substitution of the expressions for P_k into (5.13) yields

$$\sum_{i=0}^{2r+1} p_i z^i = 0 \quad \text{identically},$$

where

$$p_0 = \beta_j + \sum_{k=1}^{r} b_{j,k} - a_{j,1} + a_1 , \tag{5.15}$$

$$p_i = \beta_j a_i + \sum_{k=1}^{r} b_{j,k} a_{k,i} + \gamma_j a_{i-1} + \sum_{k=1}^{r} c_{j,k} a_{k,i-1} - a_{j,i+1} + a_{i+1}\,, \tag{5.16}$$

$$p_{2r} = \beta_j a_{2r} + \sum_{k=1}^{r} b_{j,k} a_{k,2r} + \gamma_j a_{2r-1} + \sum_{k=1}^{r} c_{j,k} a_{k,2r-1}\,, \tag{5.17}$$

$$p_{2r+1} = \gamma_j a_{2r} + \sum_{k=1}^{r} c_{j,k} a_{k,2r}\,, \quad j = 1, \ldots, r\,. \tag{5.18}$$

Using the fact that $p_i = 0$ and the relations (3.8) we have

$$\beta_j + \sum_{k=1}^{r} b_{j,k} - \frac{j}{1!} = 0\,, \tag{5.19}$$

$$\begin{aligned} &\frac{1}{i!} \sum_{k=1}^{r} b_{j,k} k^i + \frac{1}{(i-1)!} \sum_{k=1}^{r} c_{j,k} k^{i-1} - \frac{j^{i+1}}{(i+1)!} + \\ &+ \sum_{s=1}^{i-2} a_s \left(\frac{1}{(i-s)!} \sum_{k=1}^{r} b_{j,k} k^{i-s} + \frac{1}{(i-s-1)!} \sum_{k=1}^{r} c_{j,k} k^{i-s-1} - \frac{j^{i+1-s}}{(i+1-s)!} \right) + \\ &+ a_{i-1} \left(\gamma_j + \sum_{k=1}^{r} c_{j,k} + \sum_{k=1}^{r} b_{j,k} k - \frac{j^2}{2!} \right) + a_i \left(\beta_j + \sum_{k=1}^{r} b_{j,k} - \frac{j^1}{1!} \right) = 0\,, \\ &i = 1, \ldots, 2r-1\,, \end{aligned} \tag{5.20}$$

$$\begin{aligned} &\frac{1}{2r!} \sum_{k=1}^{r} b_{j,k} k^{2r} + \frac{1}{(2r-1)!} \sum_{k=1}^{r} c_{j,k} k^{2r-1} + \sum_{s=1}^{2r-2} a_s \left(\frac{1}{(2r-s)!} \sum_{k=1}^{r} b_{j,k} k^{2r-s} + \right. \\ &\left. + \frac{1}{(2r-1-s)!} \sum_{k=1}^{r} c_{j,k} k^{2r-1-s} \right) + a_{2r-1} \left(\gamma_j + \sum_{k=1}^{r} c_{j,k} + \sum_{k=1}^{r} b_{j,k} k \right) + a_{2r} \left(\beta_j + \sum_{k=1}^{r} b_{j,k} \right) = 0\,, \end{aligned} \tag{5.21}$$

$$a_{2r} \gamma_j + \sum_{k=1}^{r} c_{j,k} \sum_{s=0}^{2r} a_s \frac{k^{2r-s}}{(2r-s)!} = 0\,. \tag{5.22}$$

Further, we obtain from (5.19)–(5.22) successively the following r systems of $2r + 2$ linear algebraic equations for the unknown components of vectors $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and matrices $\mathbf{B}$, $\mathbf{C}$:

$$\beta_j + \sum_{k=1}^{r} b_{j,k} = \frac{j}{1!}\,, \tag{5.23}$$

$$\sum_{k=1}^{r} b_{j,k} k + \sum_{k=1}^{r} c_{j,k} + \gamma_j = \frac{j^2}{2!}\,,$$

$$\frac{1}{(i-1)!} \sum_{k=1}^{r} b_{j,k} k^{i-1} + \frac{1}{(i-2)!} \sum_{k=1}^{r} c_{j,k} k^{i-2} = \frac{j^i}{i!}\,, \quad i = 3, \ldots, 2r\,,$$

$$\frac{1}{2r!}\sum_{k=1}^{r} b_{j,k}k^{2r} + \frac{1}{(2r-1)!}\sum_{k=1}^{r} c_{j,k}k^{2r-1} = -\sum_{s=0}^{2r-1} a_{2r-s}\frac{j^{s+1}}{(s+1)!}, \tag{5.24}$$

$$\gamma_j a_{2r} + \sum_{k=1}^{r} c_{j,k}\sum_{s=0}^{2r} a_{2r-s}\frac{k^s}{s!} = 0 \quad \text{for} \quad j = 1, \ldots, r. \tag{5.25}$$

We see that the matrix of each of this systems ($j = 1, \ldots, r$) is the same as the matrix of (5.9) and therefore it is non-singular. Moreover, it is obvious that every solution of (5.23)–(5.25) is also a solution of (5.13) and vice versa. Now the second statement is proved by the conditions (5.23). (Cf. (2.7).) It remains to prove the third assertion. The first two statements imply that the numbers (5.12) found from (5.13) or equivalently from (5.23)–(5.25) define a ***BIM2***$_r$ method of order at least $2r$. Let us define the polynomials $P_0^*, P_1^*, \ldots, P_r^*$ by the relations (3.3) and (3.4). According to (3.2) we have

$$(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C})\begin{bmatrix} \frac{P_1^*(z)}{P_0^*(z)} \\ \vdots \\ \frac{P_r^*(z)}{P_0^*(z)} \end{bmatrix} = (\mathbf{e} + z\boldsymbol{\beta} + z^2\boldsymbol{\gamma}) \tag{5.26}$$

for z, $P_0^*(z) \neq 0$. By an easy arrangement of (5.13) we can show that the functions $P_1(z)/P_0(z), \ldots, P_r(z)/P_0(z)$ also fulfil the relations (5.26) for z, $P_0(z) \neq 0$. This implies that $P_j^*(z)/P_0^*(z) = P_j(z)/P_0(z)$ for z, $P_0(z) \neq 0$, $P_0^*(z) \neq 0$. Further, it follows that there exists a non-zero polynomial R such that $P_j(z) = R(z)\,P_j^*(z)$, $j = 0, \ldots, r$. The linear independence of the polynomials (5.1) immediately yields $R(z) = const.$ The coefficients at z^0 are the same for P_0 and P_0^*. Hence $R(z) = 1$ and we obtain $P_j(z) = P_j^*(z)$, $j = 0, \ldots, r$. Q.E.D.

In the proof of Theorem 5.1 we have shown a suitable way of determining the matrices $\mathbf{B}$, $\mathbf{C}$ and the vectors $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$. Hence we can reformulate Theorem 5.1 in the following more convenient form.

Theorem 5.2. *Let* $P_0(z) = \sum_{i=0}^{2r} a_i z^i$ *be a real polynomial of degree exactly* $2r$ *such that* $a_0 = 1$.

Let $P_k(z) = \sum_{i=0}^{2r} a_{k,i} z^i$, $k = 1, \ldots, r$ *be real polynomials with* $a_{k,i}$ *defined by* (3.8). *Let* (5.3) *hold.*

Then there exists exactly one ***BIM2***$_r$ *method of order at least* $2r$ (*i.e., matrices* $\mathbf{B}$, $\mathbf{C}$ *and vectors* $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$) *so that* $P_0(z) = \det(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C})$ *and* $P_k(z) = \det \mathbf{D}_k$, $k = 1, \ldots, r$ (*cf. Notation* 3.1). *The matrices* $\mathbf{B}$, $\mathbf{C}$ *and the vectors* $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ *are given by* (5.23)–(5.25).

Corollary 5.1. *For every polynomial P satisfying*

$$\sum_{j=0}^{2r} a_{2r-j} \frac{k^{j+1}}{(j+1)!} = 0$$

there exists exactly one ***BIM2***$_r$ *method of order at least* $2r+1$ *such that* $P(z) = \det(\mathbf{I} - z\mathbf{B} - z^2\mathbf{C})$ *and* $P_k(z) = \det \mathbf{D}_k$ *where* $\mathbf{D}_k$ *are defined in Notation* 3.1.

Proof. In this case, the condition (5.23) hold for $i = 2r+1$, too. The system (5.23)–(5.25) is equivalent to that with the non-singular matrix $\mathbf{S}$ (see (3.16) for the matrix $\mathbf{S}$).

Remark 5.1. We see from the proof of Theorem 5.1 that if the condition (5.3) is not fulfilled and the system (5.23)–(5.25) has a solution, we obtain a set of methods of order at least $2r$. All the polynomials P_i^* defined for these methods by (3.3) and (3.4) fulfil $P_j^*(z)/P_0^*(z) = P_j(z)/P_0(z)$, but $P_j^* = P_j$ need not be true. It can be guaranteed that this case occurs if (5.3) is not true and the condition of a quadratic form

$$\sum_{j=0}^{2r} \sum_{k=0}^{2r} a_{2r-j} a_{2r-k} p_k q_j + \sum_{j=0}^{2r} a_{2r-j} u_j = 0$$

holds. It is somewhat complicated to obtain p_j, q_j, u_j in a general form (i.e., as functions of r); therefore we omit it here. We shall only illustrate on examples for $r = 1, 2$ that all the possibilities can occur.

For $r = 1$ the condition (5.3) assumes the form

$$a_2 + 0{\cdot}5a_1 + 0{\cdot}25 \neq 0 .$$

For every polynomial not satisfying it, we obtain a set of rational functions

$$\frac{P_1^*(z)}{P_0^*(z)} = \frac{(2\gamma z + 1)(1 + 0{\cdot}5z)}{(2\gamma z + 1)(1 - 0{\cdot}5z)}$$

where γ is an arbitrary real number. On the other hand, for $r = 2$ the condition of singularity assumes the form

$$a_4 + a_3 + 2a_2/3 + a_1/3 + 5/36 = 0 .$$

If this condition is satisfied, we obtain the following condition for the existence of a solution of the system (5.23)–(5.25).

$$6(a_3 + a_2 + 2a_1/3)(a_3 + a_2 + 7a_1/12) + 2\tfrac{43}{60} a_3 + 2\tfrac{91}{120} a_2 + \tfrac{109}{120} a_1 + \tfrac{379}{1140} = 0 .$$

Therefore, for the polynomial $P_0(z) = -5z^4/36 + 1$, no ***BIM2***$_2$ method of order at least 4 exists so that $P_j^*/P_0^* = P_j/P_0$.

For our further investigation it will be essential to show a suitable choice of the polynomial P_0 which yields an A-stable method. First, we recall the definition of the Padé approximation of the exponential function (cf. [5]).

Definition 5.1. Let $P_{j,k}$ be a real polynomial of degree at most k and let $Q_{j,k}$ be a real polynomial of degree at most j. Let

$$\exp(z)\, Q_{j,k}(z) - P_{j,k}(z) = O(|z|^{j+k+1}) \quad \text{for} \quad z \to 0\,.$$

Then the rational function $P_{j,k}(z)/Q_{j,k}(z)$ is said to be the (j, k)-*Padé approximation* of the exponential function.

Further, we shall use

Lemma 5.2. *Let $P_{j+1,j}(z)/Q_{j+1,j}(z)$ be the $(j+1, j)$-Padé approximation of the exponential function. Then*

$$\left|\frac{P_{j+1,j}(z)}{Q_{j+1,j}(z)}\right| < 1 \quad \textit{for every } z\,, \quad \operatorname{Re}(z) < 0\,.$$

Proof can be found in [5].

Remark 5.2. The polynomials $P_{j+1,j}$ and $Q_{j+1,j}$ are explicitly given by

$$P_{j+1,j}(z) = \sum_{k=0}^{j} \frac{(2j+1-k)!\, j!\, z^k}{(2j+1)!\, k!\, (j-k)!} \tag{5.27}$$

$$Q_{j+1,j}(z) = \sum_{k=0}^{j+1} \frac{(-1)^k (2j+1-k)!\, (j+1)!\, z^k}{(2j+1)!\, k!\, (j+1-k)!}\,.$$

Now we are able to formulate the main theorem.

Theorem 5.3. *Let $Q_{2r,2r-1}(rz)$ denote the polynomial (5.27) for $j = 2r-1$ and for the argument rz. Let the relation (5.3) for the coefficients a_i of the polynomial $Q_{2r,2r-1}(rz)$ be fulfilled.*

Then solving (5.23)–(5.25) we obtain the matrices $\mathbf{B}$, $\mathbf{C}$ *and vectors* β, γ *defining a* ***BIM2***$_r$ *method of order at least* $2r$. *The method obtained in this way is A-stable. Moreover, if the solution of (2.1) belongs to $C^{2r+2}(\langle x_0, b\rangle)$, the speed of convergence of this* ***BIM2***$_r$ *method is $O(h^{2r+1})$ for $h \to 0$.*

Proof. We see from Definition 5.1 that

$$P_{2r,2r-1}(rz) = Q_{2r,2r-1}(rz)\exp(rz) + O(|z|^{4r})\,, \quad z \to 0\,. \tag{5.28}$$

Let us write

$$Q_{2r,2r-1}(rz) = \sum_{i=0}^{2r} a_i z^i \quad \text{and} \quad P_{2r,2r-1}(rz) = \sum_{i=0}^{2r} a_{r,i} z^i\,,$$

and let us substitute these relations into (5.28). Ordering this new expression according to the powers of z and using the fact that all the coefficients at z^i, $i = 0, \ldots, 2r + 1$ must be equal to zero, we finally obtain

$$a_{r,i} = \sum_{j=0}^{i} \frac{r^{i-j}}{(i-j)!} a_j , \tag{5.29}$$

$$\sum_{j=0}^{2r} a_{2r-j} \frac{r^{j+1}}{(j+1)!} = 0 . \tag{5.30}$$

We put $P_0(z) = Q_{2r,2r-1}(rz)$ in Theorem 5.2. The relation (5.29) shows that the polynomial $P_r(z)$ constructed in accordance with the assumptions of Theorem 5.2 coincides with $P_{2r,2r-1}(rz)$. Theorem 5.2 and Lemma 5.2 prove the existence of an A-stable ***BIM2***$_r$ method of order at least $2r$ and they also show the way how the method is to be constructed. Considering (5.30) and (5.24) we obtain the rest of the statement from Theorem 2.1. Q.E.D.

As an example, we show the ***BIM2***$_2$ method obtained by means of Theorem 5.3. The speed of convergence of this method is $O(h^5)$ for $h \to 0$.

(5.31)

i	β_i	$b_{i,1}$	$b_{i,2}$	γ_i	$c_{i,1}$	$c_{i,2}$
1	$4463/b$	$59/a$	$689/b$	$447/b$	$-2384/b$	$-169/b$
2	$37/a$	$112/a$	$61/a$	$3/a$	$-16/a$	$-11/a$

$a = 105,\ b = 112 \,.\, a$

Remark 5.3. It is obvious that every ***BIM2***$_r$ method constructed according to Theorem 5.3 has the following property useful for stiff systems:

$$\lim_{z \to \infty} \frac{P_r(z)}{P_0(z)} = 0 . \tag{5.32}$$

Remark 5.4. For Padé polynomials the condition (5.3) assumes the form

$$\sum_{k=0}^{r} \binom{r}{k}^2 \sum_{j=0}^{2r} \binom{2r-1+j}{j} \binom{2r}{j} r^{2r-j} k^j (-1)^j \neq 0 .$$

It has been proved for $r = 1, \ldots, 20$ by direct calculation that the above condition holds.

6. AN APPROPRIATE STRATEGY FOR SOLVING SYSTEMS OF NONLINEAR EQUATIONS

When we have to solve a system of O.D.E. by an implicit method, the efficiency of the computation is essentially influenced by the choice of the numerical method for solving the systems of (generally nonlinear) algebraic equations arising at each step of the method.

The use of an unsuitable method or of an unsuitable initial approximation can cause that the iterations are not convergent for larger values of the integration step. This can be so bad with nonlinear stiff systems that we can loose the advantage of the A-stability and have to integrate with steps so small as if any "basic" method were used. We suggest here a technique to avoid such difficulties. First of all we introduce the following notation.

Notation 6.1. We suppose that the right-hand side $\mathbf{f}$ of (2.1) is a q-dimensional vector-valued function of $q + 1$ variables. The symbol $\mathbf{y}_{n+j}$ denotes a q-dimensional vector of approximate values of the solution of the system (2.1) at the point x_{n+j} $j = 1, \ldots, r$. We shall denote the value of the vector function $\mathbf{f}$ at the $q + 1$-dimensional point $(x_{n+j}, \mathbf{y}_{n+j})$ by $\mathbf{f}_{n+j}$. Further, we shall deal with sequences of q-dimensional vectors ${}^k\mathbf{y}_{n+j}$ where k denotes the k-th member of the sequence. The values of $\mathbf{f}$ at the points $(x_{n+j}, {}^k\mathbf{y}_{n+j})$ will be denoted by ${}^k\mathbf{f}_{n+j}$. Analogously, ${}^k\mathbf{J}_{n+j}$ denotes the value of the Jacobian matrix of the function $\mathbf{f}$ taken at the point $(x_{n+j}, {}^k\mathbf{y}_{n+j})$. Further, ${}^k\mathbf{J}_{x,n+j}$ and ${}^k\mathbf{J}_{y_i,n+j}$ denote the partial derivatives of the Jacobian matrix of the function $\mathbf{f}$ with respect to the variable x or y_i, $i = 1, \ldots, q$ respectively, the value of which is taken at the point $(x_{n+j}, {}^k\mathbf{y}_{n+j})$. Analogously, the symbol ${}^k\mathbf{f}_{x,n+j}$ denotes the partial derivative of $\mathbf{f}$ with respect to the variable x taken at the point $(x_{n+j}, {}^k\mathbf{y}_{n+j})$. Further, for $k = 0, 1, \ldots$ we introduce the following $qr \times qr$ matrices:

$$
\begin{aligned}
{}^k\mathbf{D} &= \operatorname{diag}\left({}^k\mathbf{J}_{n+1}, \ldots, {}^k\mathbf{J}_{n+r}\right), \\
{}^k\mathbf{G} &= \operatorname{diag}\left({}^k\mathbf{J}_{x,n+1}, \ldots, {}^k\mathbf{J}_{x,n+r}\right), \\
{}^k\mathbf{H} &= \operatorname{diag}\left({}^k\mathbf{H}_{n+1}, \ldots, {}^k\mathbf{H}_{n+r}\right)
\end{aligned}
$$

where ${}^k\mathbf{H}_{n+j}$, $j = 1, \ldots, r$ are the $q \times q$ matrices

$$
{}^k\mathbf{H}_{n+j} = \left({}^k\mathbf{J}_{y_1,n+j}\,{}^k\mathbf{f}_{n+j}, \ldots, {}^k\mathbf{J}_{y_r,n+j}\,{}^k\mathbf{f}_{n+j}\right).
$$

The symbol diag denotes the diagonal block matrix, the elements of its diagonal following the symbol diag.

The left-hand side of the system (2.3) will be denoted by $\mathbf{F}$ and its value at the $(q + 1) \times r$−dimensional point $(x_{n+1}, \ldots, x_{n+r}, {}^k\mathbf{y}_{n+1}, \ldots, {}^k\mathbf{y}_{n+r})$ will be denoted by ${}^k\mathbf{F}$. The Jacobian matrix of the function $\mathbf{F}$ with respect to qr variables $\mathbf{y}_{n+j}$, $j = 1, \ldots, r$ will be denoted by $\mathbf{DF}$ and its value at the $(q + 1) \times r$−dimensional

point $(x_{n+1}, \ldots, x_{n+r}, {}^k\mathbf{y}_{n+1}, \ldots, {}^k\mathbf{y}_{n+r})$ by ${}^k\mathbf{DF}$. Finally, ${}^k\Delta$ is the $q \times r$–dimensional vector

$$ {}^k\Delta = \begin{bmatrix} {}^{k+1}\mathbf{y}_{n+1} - {}^k\mathbf{y}_{n+1} \\ \vdots \\ {}^{k+1}\mathbf{y}_{n+r} - {}^k\mathbf{y}_{n+r} \end{bmatrix}. $$

The symbols $'\mathbf{B}$, $'\mathbf{C}$, $'\beta$, $'\gamma$ have the same meaning as in (2.3).

Now we try to solve the nonlinear system (2.3) by Newton's method, i.e., we try to construct a sequence of approximations from the formula

$$ {}^k\mathbf{DF}\,{}^k\Delta = -{}^k\mathbf{F}, \quad k = 0, 1, \ldots \tag{6.1} $$

with a starting vector $({}^0\mathbf{y}_{n+1}, \ldots, {}^0\mathbf{y}_{n+r})$. The Jacobian matrix ${}^k\mathbf{DF}$ has the explicit form

$$ {}^k\mathbf{DF} = {}'\mathbf{I} - h\,{}'\mathbf{B}({}^k\mathbf{D}) - h^2\,{}'\mathbf{C}({}^k\mathbf{G} + ({}^k\mathbf{D})^2 + {}^k\mathbf{H}). \tag{6.2} $$

From (6.2) it is immediately seen that the use of Newton's method is expensive for solving the system (2.3) because the matrices

$$ {}^k\mathbf{J}_{x,n+j}, {}^k\mathbf{J}_{y_1,n+j}, \ldots, {}^k\mathbf{J}_{y_r,n+j}, \quad j = 1, \ldots, r $$

must be also calculated.

We suggest other methods resembling Newton's one. These methods require only that the Jacobian matrix of the function $\mathbf{f}$ be evaluated.

Method I. We construct the sequence of approximations from the relation

$$ \mathbf{T}({}^k\Delta) = -{}^k\mathbf{F}, \quad k = 0, \ldots \tag{6.3} $$

where

$$ \mathbf{T} = {}'\mathbf{I} - h\,{}'\mathbf{B}({}^0\mathbf{D}) - h^2\,{}'\mathbf{C}({}^0\mathbf{D})^2. \tag{6.4} $$

Method II. The sequence of approximations is calculated from

$$ {}^k\mathbf{T}({}^k\Delta) = -{}^k\mathbf{F}, \quad k = 0, \ldots \tag{6.5} $$

where

$$ {}^k\mathbf{T} = {}'\mathbf{I} - h\,{}'\mathbf{B}({}^k\mathbf{D}) - h^2\,{}'\mathbf{C}({}^k\mathbf{D})^2. \tag{6.6} $$

From the relations (6.2), (6.4) and (6.6) we see that

$$ \mathbf{T} - {}^0\mathbf{DF} = h^2\,{}'\mathbf{C}({}^0\mathbf{G} + {}^0\mathbf{H}), \tag{6.7} $$

$$ {}^k\mathbf{T} - {}^k\mathbf{DF} = h^2\,{}'\mathbf{C}({}^k\mathbf{G} + {}^k\mathbf{H}). \tag{6.8} $$

Now we show conditions needed for the convergence of the iterations (6.3). As far as the iterations (6.5) are concerned a similar theorem can be proved. However, we did not succeed in proving that the speed of convergence is better than that of the iterations (6.3) even though the computations showed this phenomenon. Therefore, we present only the theorem on the convergence of the iterations (6.3) as an illustration. The following lemma will be of great importance for it.

Lemma 6.1. *Let* R, S *be Banach spaces. Let* $\mathbf{F}$ *be a mapping from* R *into* S *and let* $\mathbf{Y}_0$ *be a fixed element of the space* R. *Let* Ω_0 *be the following set:* $\Omega_0 = \{\mathbf{Y} \in R, \|\mathbf{Y} - \mathbf{Y}_0\| \leqq p\}$, *where* p *is a certain number. Let there exist a continuous second derivative of the mapping* $\mathbf{F}$ *at each point* $\mathbf{Y} \in \Omega_0$. (*It will be denoted by* $\mathbf{F}''(\mathbf{Y})$.) *Let* $\mathbf{T}$ *be a linear mapping from* R *into* S *such that* $\mathbf{T}^{-1}$ *exists. Further, let us suppose that there exist constants* a_i, $i = 1, \ldots, 4$ *so that*

$$\|\mathbf{T}^{-1}\| \leqq a_1, \quad \|\mathbf{F}(\mathbf{Y}_0)\| \leqq a_2, \quad \|\mathbf{T} - \mathbf{F}'(\mathbf{Y}_0)\| \leqq a_3, \tag{6.9}$$

$$\|\mathbf{F}''(\mathbf{Y})\| \leqq a_4 \quad \text{for} \quad \mathbf{Y} \in \Omega_0 .$$

Let us write $E = a_1 a_2$, $K = a_1 a_4$, $L = a_1 a_3$, $U = KE/(1 - L)^2$ *and let* $U < 0.5$, $L < 1$.

If the number p *fulfils*

$$\frac{(1 - (1 - 2U)^{0.5})\, E}{U(1 - L)} < p < \frac{(1 + (1 - 2U)^{0.5})\, E}{U(1 - L)}$$

then there exists exactly one element $\mathbf{Y}^* \in \Omega_0$ *so that* $\mathbf{F}(\mathbf{Y}^*) = \mathbf{\Theta}$. *The sequence of the approximations* $\mathbf{Y}_k$ *obtained from*

$$\mathbf{Y}_{k+1} = \mathbf{Y}_k - \mathbf{T}^{-1}\, \mathbf{F}(\mathbf{Y}_k), \quad k = 0, \ldots \tag{6.10}$$

converges to $\mathbf{Y}^*$ *in the norm of the space* R *and the following estimate is true*:

$$\|\mathbf{Y}^* - \mathbf{Y}_k\| \leqq U^{-1}(1 - (1 - L)(1 - 2U)^{0.5})^{k+1}\, E/(1 - L)^2 . \tag{6.11}$$

Proof. The conditions (6.9) imply that the assumptions of Theorems 1, 2 in Chapter XVIII, § 2 in [6] are fulfilled. The statement follows from Theorems 1, 2 and from the remark after Theorem 2.

Theorem 6.1. *Let a system of* q *O.D.E. of the form* (2.1) *be given and let the right-hand side* $\mathbf{f}$ *of* (2.1) *have continuous partial derivatives with respect to all variables up to the third order in a certain domain* $\Omega = I \times Q$, *where* $\langle x_0, b\rangle \subset I$ *and the domain* Q *contains the solution of the system* (2.1).

Then there exists $h_0 > 0$ *so that the system* (2.3) *can be solved for every* $0 < h <$

$< h_0$ *by iterations* (6.3) *with the initial approximation* ${}^0\mathbf{y}_{n+j} = \mathbf{y}_n$, $j = 1, \ldots, r$ *and the speed of convergence is given by*

$$\|\mathbf{Y}^* - \mathbf{Y}_k\| = O(h^{k+1}).$$

Proof. Let R be the set of $q \times r-$dimensional real vectors and let the norm of the vector $\mathbf{Y} = (y_1, \ldots, y_{qr})$ be defined by $\|\mathbf{Y}\| = \max_j |y_j|$. The left-hand side $\mathbf{F}$ of (2.3) can be regarded as a mapping from the Banach space R into itself. Let us further denote by $\mathbf{Y}_0$ the column vector of the length qr, composed of the r-dimensional vectors $\mathbf{y}_n$ (approximate values of the solution of (2.1) at the point x_n). Let the norm of any mapping from R into R be defined in the usual sense by the norm in R.

The assumptions of the theorem imply the existence of a sphere Ω_0 with its centre at the point $\mathbf{Y}_0$ such that $\mathbf{F}'(\mathbf{Y})$ and $\mathbf{F}''(\mathbf{Y})$ exist for every $\mathbf{Y} \in \Omega_0$ (The symbols $\mathbf{F}'(Y)$, $\mathbf{F}''(\mathbf{Y})$ denote the Fréchet derivatives of the mapping $\mathbf{F}$.) Moreover, there exist constants a_4 and h_4 so that

$$\|\mathbf{F}''(\mathbf{Y})\| \leqq a_4 \quad \text{for every} \quad \mathbf{Y} \in \Omega_0 \quad \text{and} \quad 0 < h < h_4 . \tag{6.12}$$

The relation (6.4) implies (in a similar way as in the proof of Theorem 2.1) that the matrix $\mathbf{T}$ is non-singular for every $h > 0$ less than a certain h_1 and

$$\|\mathbf{T}^{-1}\| \leqq \frac{1}{1 - h\|\mathbf{B}\| \Lambda - h^2\|\mathbf{C}\| \Lambda^2} = a_1(h), \quad 0 < h < h_1 , \tag{6.13}$$

where $\Lambda = \max_j \|{}^0\mathbf{J}_{n+j}\|$. Further, we see from the relation (2.3) that

$$\|\mathbf{F}(\mathbf{Y}_0)\| \leqq hK_1 + h^2K_2 = a_2(h) \tag{6.14}$$

where K_1, K_2 are constants. Let us recall that $\mathbf{F}'(\mathbf{Y}_0) = {}^0\mathbf{DF}$. Then using (6.7), we have

$$\|\mathbf{F}'(\mathbf{Y}_0) - \mathbf{T}\| \leqq h^2K_3 = a_3(h) . \tag{6.15}$$

Let us put $L = a_1(h)\, a_3(h)$, $E = a_1(h)\, a_2(h)$, $U = a_1(h)^2\, a_2(h)\, a_4/(1 - L)^2$.

In virtue of (6.12)–(6.15) we have $\lim_{h \to 0} U = 0$, which implies the existence of $h_3 > 0$ so that $U < 0{\cdot}5$ for $0 < h < h_3$. In the same way we establish that there exists $h_2 > 0$ so that $L < 1$ for $0 < h < h_2$. Let us put $h_0 = \min(h_1, \ldots, h_4)$. Then for every $0 < h < h_0$, the assumptions of Lemma 6.1 are fulfilled and hence our statement is proved. By substitution from (6.12)–(6.15) into (6.11) we obtain the estimate for the speed of convergence.

Remark 6.1. Solving practical problems we have used both Method I and Method II. The speed of convergence of Method II was in all test problems better than that of Method I. The iterations in Method II were convergent also for those h for which the iterations of Method I did not converge. Method II needed less iterations than

Method I to achieve a given accuracy, so that the total time for computations needed for Method II was somewhat less than that needed for Method I.

Using a better initial approximation than ${}^0\mathbf{y}_{n+j} = \mathbf{y}_n$, $j = 1, \ldots, r$ we have obtained an algorithm working with extremely large step size h.

We suggest to use the following one-step explicit nonlinear A-stable method quoted in [3] for obtaining a good initial approximation ${}^0\mathbf{y}_{n+j}$:

$$(\mathbf{I} - h({}^0\mathbf{J}_n) + 0{\cdot}5h^2({}^0\mathbf{J}_n)^2)(\mathbf{y}_{n+1} - \mathbf{y}_n) = \tag{6.16}$$
$$= h({}^0\mathbf{f}_n) + 0{\cdot}5h^2({}^0\mathbf{f}_{x,n} - ({}^0\mathbf{J}_n)\,{}^0\mathbf{f}_n - h({}^0\mathbf{J}_n)\,{}^0\mathbf{f}_{x,n}).$$

We use the formula (6.16) r-times and the values obtained are used as an initial approximation ${}^0\mathbf{y}_{n+j}$, $j = 1, \ldots, r$ for the iterations (6.3) or (6.5). This algorithm worked on a very stiff system with extremely large steps (see § 7).

The use of Method I and Method II requires the evaluation of the Jacobian matrix of the right-hand side $\mathbf{f}$ of the system (2.1) and the evaluation of the vector $\mathbf{f}_x$ only. So we are nearly in the same situation as if we applied the methods defined in [1] on the system (2.1) and solved the arising systems of nonlinear equations by Newton's method. Nevertheless, the convergence of the ***BIM2***$_r$ method can be twice as fast.

Example. The iterations (6.5) for a ***BIM2***$_2$ method are given by the matrix

$$ {}^k\mathbf{T} = \begin{bmatrix} (\mathbf{I} - hb_{1,1}({}^k\mathbf{J}_{n+1}) - h^2c_{1,1}({}^k\mathbf{J}_{n+1})^2), & -(hb_{1,2}\,{}^k\mathbf{J}_{n+2} + h^2c_{1,2}({}^k\mathbf{J}_{n+2})^2) \\ -(hb_{2,1}({}^k\mathbf{J}_{n+1}) + h^2c_{2,1}({}^k\mathbf{J}_{n+1})^2), & (\mathbf{I} - hb_{2,2}({}^k\mathbf{J}_{n+2}) - h^2c_{2,2}({}^k\mathbf{J}_{n+2})^2) \end{bmatrix} \tag{6.17}$$

where $b_{i,j}$, $c_{i,j}$ are the elements of the matrices $\mathbf{B}$, $\mathbf{C}$ respectively, and $\mathbf{I}$ is the $q \times q$ unit matrix.

7. NUMERICAL EXPERIMENTS

Many successful numerical experiments have been done with the formula (4.7) and with the formula (5.31) on about 15 stiff systems of O.D.E. In this section we present only the results obtained by solving the following system of O.D.E. arising in reactor kinetics (see [7]):

$$\begin{aligned} y_1' &= -0{\cdot}04y_1 + 10^4y_2y_3\,, \\ y_2' &= 0{\cdot}04y_1 - 10^4y_2y_3 - 3\,.\,10^7y_2^2\,, \\ y_3' &= 3\,.\,10^7y_2^2\,, \\ y_1(0) &= 1\,, \quad y_2(0) = y_3(0) = 0\,. \end{aligned} \tag{7.1}$$

The system is nonlinear, therefore its stiffness is described by the behaviour of the eigenvalues of the linearized system. One of these eigenvalues is always equal to zero, the remaining eigenvalues differ sufficiently from each other and vary very rapidly in the range $\langle -0{\cdot}04, -10^{+4}\rangle$ for $x \in \langle 0, 10\rangle$. Thus the system is very stiff.

Both the formula (4.7) and the formula (5.31) are used with the algorithm (6.5) described in § 6. The initial approximation for iterations is obtained by means of the formula (6.16). Computations were performed on an IBM 370/135 computer in REAL*8 precision. The results obtained by the formula (5.31) with different steps are shown in Table 1, where the symbol § denotes a reference solution obtained by the *Runge-Kutta method of order* 4 with the automatic step-size control.

$x = 10$

h	y_1	$10^4 y_2$	y_3	Time in sec.
2·0	0·841863	0·162729	0·158121	1·85
1·0	0·841500	0·162442	0·158484	2·42
0·4	0·841391	0·162356	0·158593	4·20
0·2	0·841375	0·162343	0·158609	5·98
0·1	0·841371	0·162340	0·158613	9·66
0·04	0·841370	0·162339	0·158614	22·86
§	0·841370	0·162339	0·158614	

Table 1. Solutions of the system (7.1) by the formula (5.31).

The relative errors of the solution shown in Table 1 are given in Table 2.

h	err_1	err_2	err_3
2·0	$6 . 10^{-4}$	$3 . 10^{-3}$	$3 . 10^{-3}$
1·0	$2 . 10^{-4}$	$1 . 10^{-3}$	$8 . 10^{-4}$
0·4	$3 . 10^{-5}$	$2 . 10^{-4}$	$1 . 10^{-4}$
0·2	$6 . 10^{-6}$	$3 . 10^{-5}$	$3 . 10^{-5}$
0·1	$1 . 10^{-6}$	$6 . 10^{-6}$	$6 . 10^{-6}$
0·04	0	0	0

Table 2. The relative errors of the solution from Table 1.

The formula (4.7) has been also tested on the system (7.1). The results obtained are shown in Table 3 and the relative errors in Table 4.

$x = 10$

h	y_1	$10^4 y_2$	y_3
0·4	0·842071	0·163715	0·157912
0·2	0·841521	0·162552	0·158463
§	0·841370	0·162339	0·158614

Table 3. Solutions obtained by the formula (4.7).

h	err_1	err_2	erc_3
0·4	$8 . 10^{-4}$	$2 . 10^{-3}$	$4 . 10^{-3}$
0·2	$2 . 10^{-4}$	$1 . 10^{-3}$	$1 . 10^{-3}$

Table 4. Relative errors of the solution in Table 3.

A comparison of Tables 1–4 shows that the formula (5.31) approximates all components of the solution better than the formula (4.7). The computing time was for the formula (5.31) also somewhat shorter. *The formula* (5.31) *fits the exponential function not only for small values of the argument* x, *but for* $x \to -\infty$, *too.* (See (5.32).) The formula (4.7) has not this property. Therefore we recommend rather to use the formula (5.31) for stiff systems, even though (5.31) converges only $O(h^5)$, $h \to 0$.

From Table 1 we see that also for extremely large steps as 1 or 2 we obtained the solution with a sufficient accuracy. The following Table 5 shows the time comparison with the Runge-Kutta method of order 4 with the automatic step-size control and with the controlling constant, giving the requested accuracy at one step, chosen equal to 10^{-5}. We note that for the value 10^{-4} of this constant the method was unstable.

$x = 10$	err_1	err_2	err_3	*time (in sec.)*
formula (5.31) $h = 2$	$6 . 10^{-4}$	$3 . 10^{-3}$	$3 . 10^{-3}$	1·85
R–K method *eps* $= 10^{-5}$	$5 . 10^{-2}$	$2 . 10^{-1}$	$3 . 10^{-2}$	87·24

Table 5. Time and accuracy comparison.

We see that the formula (5.31) used with the iterative procedure (6.5) and with the formula (6.16) has worked about 50 *times* faster and with 10–100 *times better accuracy.*

Acknowledgement. I should like to thank Prof. M. PRÁGER and Mr. M. KUBÍČEK for useful discussions I had with them during the present research.

References

[1] *H. A. Watts, L. F. Shampine:* A-stable block implicit one-step methods, BIT *12* (1972), 252—266.

[2] *M. Práger, J. Taufer, E. Vitásek:* Overimplicit multistep methods, Aplikace matematiky *18* (1973), No. 6, 399—421.

[3] *M. Kubíček, K. Višňák:* A nonlinear explicit algorithm for efficient integration of stiff systems of O.D.E., Chem. eng. com. (1974), 291—296.

[4] *K. Višňák, M. Kubíček:* A class of numerical methods for stiff problems in ordinary differential equations. JIMA, to be published.

[5] *R. D. Grigorieff:* Numerik gewöhnlicher Differentialgleichungen, Teubner Verlag, Stuttgart 1972.

[6] *Л. В. Канторович, Г. П. Акилов:* Функциональный анализ в нормированных пространствах, Гос. издат. физикоматематической лит., Москва 1959.

[7] *L. Lapidus, J. H. Seinfeld:* Numerical solution of O.D.E., Academic Press, N.Y. 1971.

Author's address: 115 67 Praha 1, Žitná 25, ČSSR (Matematický ústav ČSAV).