A Study of (m,k)-Methods for Solving Differential-Algebraic Systems of Index 1

Alexander I. Levykin[1] and Eugeny A. Novikov[2](✉)

[1] Institute of Computational Mathematics and Mathematical Geophysics,
Academy of Sciences, Siberian Branch, pr. Ak. Lavrent'eva 6, Akademgorodok,
630090 Novosibirsk, Russia
lai@osmf.sscc.ru
http://www.sscc.ru/

[2] Institute of Computational Modeling, Academy of Sciences, Siberian Branch,
Akademgorodok 50, Str. 4, 660036 Krasnoyarsk, Russia
novikov@icm.krasn.ru
http://icm.krasn.ru/

Abstract. A class (m,k)-methods is discussed for the numerical solution of the initial value problems for implicit systems of ordinary differential equations. The order conditions and convergence of the numerical solution in the case of implementation of the scheme with the time-lagging of matrices derivatives for systems of index 1 are obtained. At $k \leq 4$ the order conditions are studied and schemes optimal computing costs are obtained.

Keywords: Stiff systems · Differential-algebraic systems of index 1 · Numerical methods

1 Introduction

Many applied problems lead to systems of differential equations given implicitly as [1–4]

$$F(x, x') = 0,\ x(t_0) = x_0,\ t_0 \leq t \leq t_k, \tag{1}$$

where x and F are functions of the same dimension, and F is assumed to have sufficiently many bounded derivatives. Such problems arise in simulation of chemical reaction kinetics [4], electrical networks [5–6], control engineering etc. A non-autonomous systems $F(x, x', t) = 0$ is brought to the form (1) by adding the equation for the independent variable, $t' = 1$.

The modern methods for numerical solution of the initial-value problem for systems of ordinary differential equations (ODE) suppose usually the explicit dependence of the derivative of the solution [7]

$$x' = \varphi(x),\ x(t_0) = x_0,\ t_0 \leq t \leq t_k. \tag{2}$$

Support of RFBR Ander Grants 14-01-00047 and 15-01-00977.

N. Danaev et al. (Eds.): CITech 2015, CCIS 549, pp. 94–107, 2015.
DOI: 10.1007/978-3-319-25058-8_10

However, a reduction of (1) to the form (2) requires a large additional numerical costs at every integration step, because this is connected with the inversion of the matrix $F_y = \partial F/\partial y$ which generally is singular. The numerical problem appeares to be more complicated because of the stiffness of explicit equations systems: in this case it is necessary to apply of special methods with conversion of the Jacobian matrix . A class of the schemes is offered [8], in which the reduction to the form (1) and the calculation of the approximate solution are carried out simultaneously. The given methods were generated by the (m, k)-schemes [9] for solving the explicit ODE systems.

We use classification of implicit systems, based on the concept of the index for such systems [1–2]. We say that system (1) is:

a) of index 0, if $\|F_y^{-1}\| \le c < \infty$ (i.e., when (1) is solvable);

b) of index 1, if (1) can be reduced to

$$x' = f(x, y),\ 0 = g(x, y), \tag{3}$$

where $\|g_y^{-1}\| \le c < \infty$;

c) of index 2, if (1) can be reduced to

$$x' = f(x, y),\ 0 = g(x),$$

where $\|(g_x f_y)^{-1}\| \le c < \infty$.

In addition, it is assumed that functions F, f, and g are Lipschitz bounded, which ensures existence and uniqueness of the solution to problem (1) [10].

Using the notation $x' = y$, problem (1) can be written in the form

$$x' = y,\ F(x, y) = 0,\ x(t_0) = x_0,\ y(t_0) = y_0,\ t_0 \le t \le t_k. \tag{4}$$

The additional condition $y(t_0) = y_0$ can be found, for example, by solving the problem $F(x_0, y) = 0$ and using the stabilization technique.

2 The Numerical Schemes

We define the class of the (m, k)–schemes for solving problem (4). Let m and k, $(m \ge k)$ be given integers and consider the sets

$$\begin{aligned} &M_m = \{1, \dots, m\},\\ &M_k = \{m_i\ |m_1 = 1,\ m_{i-1} < m_i,\ m_i \le m,\ 2 \le i \le k\},\\ &J_i = \{m_j - 1\ |j > 1,\ m_j \in M_k, m_j \le i\}, \quad 2 \le i \le m. \end{aligned} \tag{5}$$

Then (m, k)-methods for the systems of index 0 have the form

$$x_{n+1} = x_n + \sum_{i=1}^{m} \mu_i k_{xi}, \quad y_{n+1} = y_n + \sum_{i=1}^{m} \mu_i k_{yi}, \tag{6}$$

where the internal stages are given by

$$D_n = A_2 + ahA_1,$$

$$D_n k_{xi} = h[A_2(y_n + \sum_{j=1}^{i-1} \beta_{ij} k_{yj}) - F(x_n + \sum_{j=1}^{i-1} \beta_{ij} k_{xj}, y_n + \sum_{j=1}^{i-1} \beta_{ij} k_{yj})] +$$
$$+ \eta A_2 \sum_{j \in J_i} \alpha_{ij} k_{xj} + (1-\eta) h A_1 \sum_{j \in J_i} \gamma_{ij} k_{xj},$$

$$k_{yi} = \frac{1}{ah}[k_{xi} - h(y_n + \sum_{j=1}^{i-1} \beta_{ij} k_{yj}) - \eta \sum_{j \in J_i} \alpha_{ij} k_{xj} - (1-\eta) h \sum_{j \in J_i} \gamma_{ij} k_{yj}],$$

if $i \in M_k$ and

$$D_n k_{xi} = \eta A_2 (k_{x(i-1)} + \sum_{j \in J_i} \alpha_{ij} k_{xj}) + (\eta - 1) h A_1 (k_{x(i-1)} + \sum_{j \in J_i} \gamma_{ij} k_{xj}),$$

$$k_{yi} = \frac{1}{ah}(k_{xi} - \eta(k_{x(i-1)} + \sum_{j \in J_i} \alpha_{ij} k_{xj}) - (1-\eta) h (k_{y(i-1)} + \sum_{j \in J_i} \gamma_{ij} k_{yj}).$$

when $i \in M_m \backslash M_k$. Here, a, μ_i, β_{ij}, α_{ij} and γ_{ij} are parameters defining properties of stability and accuracy (6), h is the integration step, A_1 and A_2 are matrices approximating the derivatives $F_{ny} = \partial F(x_n, y_n)/\partial y$ and $F_{nx} = \partial F(x_n, y_n)/\partial x$. In what follows we use the notation $c_{ij} = \beta_{ij} + \gamma_{ij}$, where $\gamma_{ij} = 0$ if $j \notin J_i$ and $\gamma_{i,i-1} = 1$ if $j \in M_m \backslash M_k$. The matrix D_n is non-singular because $\det F_y \neq 0$. For the systems of index 1 or 2 the stages of the method are given by

$$(E - ahA_1) k_{xi} - ahA_2 k_{yi} = \delta_i h f(x_n + \sum_{j=1}^{i-1} \beta_{ij} k_{xj}, y_n + \sum_{j=1}^{i-1} \beta_{ij} k_{yj}) +$$
$$+ \eta \sum_{j \in J_i} \alpha_{ij} k_{xj} + (1-\eta) h \sum_{j \in J_i} \gamma_{ij} (A_1 k_{xj} + A_2 k_{yj}), \tag{7}$$

$$-aB_1 k_{xi} - aB_2 k_{yi} = \delta_i g(x_n + \sum_{j=1}^{i-1} \beta_{ij} k_{xj}, y_n + \sum_{j=1}^{i-1} \beta_{ij} k_{yj}) +$$
$$+ (1-\eta) h \sum_{j \in J_i} \gamma_{ij} (B_1 k_{xj} + B_2 k_{yj}), \tag{8}$$

where A_1, A_2, B_1, B_2 are matrices approximating the derivatives

$$f_{nx} = \frac{\partial f(x_n, y_n)}{\partial x},\ f_{ny} = \frac{\partial f(x_n, y_n)}{\partial y},\ g_{nx} = \frac{\partial g(x_n, y_n)}{\partial x},\ g_{ny} = \frac{\partial g(x_n, y_n)}{\partial y},$$

and $\delta_i = 1$ if $i \in M_k$, $\delta_i = 0$ if $i \in M_m \backslash M_k$.

Reversibility of the matrix D_n is ensured for systems of index 1 by the reversibility of the matrix g_y, while for systems of index 2 – by the matrix $g_x f_y$.

The parameter η equals to 0 or 1. At $\eta = 0$, the schemes are preferable for computations, since they require less multiplications of a matrix by vector, and at $\eta = 1$ the schemes are more convenient in implementation .

The main feature of the schemes presented when compared to the conventional methods [11–14] is that in (m, k)–schemes the function F is evaluated k times at each step, and the number of stages is equal to m, $m \geq k$. The given schemes can be considered as a special form of ROW-methods, in which the set of definition of the scheme parameters is given more exactly. This simplifies the analysis of the order conditions, and the study of the problem how to use the time-lagged matrix D_n are carried out. The linear system of algebraic equations, arising in calculation of stages, is solved by the LU-decomposition of the matrix D_n. At every step once decomposition of the matrix D_n is evaluated, the function of the right side of a differential problem k times is calculated, backward in the Gauss method m times is executed. For given m and k the cost of one step is completely determined, and numbers $m_1, \ldots, m_k$ do only distribute this costs inside the step.

Two implementations of (6) for the systems of index 1 will be further considered:

a) the matrix D_n is reevaluated at each integration step;

b) the matrix D_n and the integration step h similar to [5] are not changed at several steps, thus $D_n = D_{n+\vartheta}$, $h_n = h_{n+\vartheta}$, $-Q \leq \vartheta \leq 0$ where Q is the maximum number of steps in the time-lagging of matrices derivatives.

3 Convergence and Order Conditions

The local error of the scheme (6) when solving (3) is defined as the difference between the exact and the numerical solution provided the initial values are choosen on the exact solution

$$\delta_x(t) = x_1 - x(t+h),\ \delta_y(t) = y_1 - y(t+h).$$

We recall that order of consistency with respect to x is p and with respect to y is q, if

$$\delta_x(t) = O(h^{p+1}), \quad \delta_y(t) = O(h^{q+1}).$$

The condition for the parameters of a scheme ensuring the required order consistency can be obtained by equating the coefficients of the expansion of the approximate solution x_{n+1}, y_{n+1} to the exact solution

$$x(t_n + h) = x_n + \sum_{r=1}^{\infty} \frac{h^r}{r!} \sum_{\substack{\mathbf{t} \in \mathbf{L}T1X \\ \rho(\mathbf{t})=r}} [F(\mathbf{t})]_n, \tag{9}$$

$$y(t_n+h) = y_n + \sum_{r=1}^{\infty} \frac{h^r}{r!} \sum_{\substack{\mathbf{t}\in \mathbf{L}T1Y \\ \rho(\mathbf{t})=r}} [F(\mathbf{t})]_n, \tag{10}$$

where $[F(\mathbf{t})]_n$ denotes a value of the elementary differential $F(\mathbf{t})$ of the order $\rho(\mathbf{t})$ at a point (x_n, y_n). Expressions (9), (10), trees set definition $\mathbf{T1} = \mathbf{T1X} \cup \mathbf{T1X}$, and the corresponding elementary differentials $F(\mathbf{t})$, $\mathbf{t} \in \mathbf{LT1}$ were introduced in [3].

Now we find an expansion similar to (9), (10) for the numerical solution at $\big(x_{n+1}, y_{n+1}\big)$ for our scheme (6).

Assume that the (m,k)-scheme is implemented with time-lagging of the matrices derivatives. The following proposition gives the derivatives with respect to h at $h=0$ of the entries of the matrix

$$\begin{bmatrix} f_x(x_n+\vartheta h) & f_y(y_n+\vartheta h) \\ g_x(x_n+\vartheta h) & g_y(y_n+\vartheta h) \end{bmatrix}$$

at a point (x_n, y_n).

Proposition 1. *Let $p \equiv f \vee g$ and $r \equiv x \vee y$. Then*

$$p_r^{(q)}(x_{n+\vartheta}, y_{n+\vartheta})|_{h=0} = \sum_{\substack{\mathbf{t}\in\mathbf{LT1X} \\ \rho(\mathbf{t})=q}} \vartheta^q [A_{pr}(\mathbf{t})]_n, \tag{11}$$

where $[A_{pr}(\mathbf{t})]_n$ is a value of the differential

$$\frac{\partial^{k+l+1} p}{\partial r \partial x^k \partial y^l}(F(\mathbf{t}_1), \cdots, F(\mathbf{t}_k), F(\mathbf{u}_1), \cdots, F(\mathbf{u}_l)),$$

$\mathbf{t} = [\mathbf{t}_1, \cdots, \mathbf{t}_k, \mathbf{u}_1, \cdots, \mathbf{u}_l] \in \mathbf{LT1}$ *in the point (x_n, y_n).*

Differentiating p_r with respect to t gives

$$\frac{d^q p_r(x_n, y_n)}{dt^q} = \sum_{\substack{\mathbf{t}\in\mathbf{LT1X} \\ \rho(\mathbf{t})=q}} \frac{\partial^{k+l+1} p}{\partial r \partial x^k \partial y^l}(x^{(\alpha_1)}, \cdots, x^{(\alpha_k)}, y^{(\beta_1)}, \cdots, y^{(\beta_l)}).$$

Substituting of the expression for $x^{(\alpha_i)}, \cdots, y^{(\beta_j)}$ obtained from (9), (10) using the change of variables $dt = \vartheta dh$, gives the stated result as $h \to 0$.

We denote

$$\mathbf{t} = [\mathbf{t}_1{}^{\lambda_1}, \cdots, \mathbf{t}_n{}^{\lambda_n}]_r, \tag{12}$$

for the tree $\mathbf{t} \in \mathbf{T1}$, where the index λ_i is the multiplicity of a inclusion of a corresponding subtree $\mathbf{t}_i \in \mathbf{T1}$, $r \equiv x \vee y$.

The number of a possible labelling $\alpha(\mathbf{t})$ of the tree $\mathbf{t} \in \mathbf{T1}$ is defined recursively by $\alpha(\mathbf{t}) = 1$, if

$$\rho(\mathbf{t}) = 1,\ \alpha(\mathbf{t}) = \bar{\rho}(\mathbf{t}) \prod_{i=1}^{n} \frac{1}{\lambda_i!} \left(\frac{\alpha(\mathbf{t}_i)}{\rho(\mathbf{t}_i!)} \right)^{\lambda_i},$$

where $\bar{\rho} = (\rho(\mathbf{t}) - 1)!$, if $\mathbf{t} \in \mathbf{T1X}$, $\bar{\rho} = \rho(\mathbf{t})!$, if $\mathbf{t} \in \mathbf{T1Y}$.

The integer number $\Gamma(\mathbf{t})$ corresponding to a tree $\mathbf{t} \in \mathbf{T1}$ is defined recursively by

$$\Gamma(\mathbf{t}) = 1, \text{ if } \quad \rho(\mathbf{t}) = 1,$$

$$\Gamma(\mathbf{t}) = \rho(\mathbf{t}) \prod_{i=1}^{n} \Gamma(\mathbf{t}_i)^{\lambda_i}, \text{ if } \quad \mathbf{t} = [\mathbf{t}_1{}^{\lambda_1}, \cdots, \mathbf{t}_n{}^{\lambda_n}]_x,$$

$$\Gamma(\mathbf{t}) = \prod_{i=1}^{n} \Gamma(\mathbf{t}_i)^{\lambda_i}, \text{ if } \quad \mathbf{t} = [\mathbf{t}_1{}^{\lambda_1}, \cdots, \mathbf{t}_n{}^{\lambda_n}]_y.$$

We put $\tilde{c}_{ij} = c_{ij}$, if $i > j$, $\tilde{c}_{ii} = a$, $\tilde{c}_{ij} = 0$, if $i < j \cdot \omega = (\omega_{ij})$ is the inverse of the matrix $(\tilde{c}_{ij})$.

The expression $\phi_i(\mathbf{t}) = \phi_{1i}(\mathbf{t}) + \phi_{2i}(\mathbf{t})/\Gamma(\mathbf{t})$, $\mathbf{t} \in \mathbf{T}1$, $1 \le i \le m$, is defined recursively by

$$\phi_{1i}(\mathbf{t}) = \delta_i,\ \phi_{2i}(\mathbf{t}) = 0,$$

if $\rho(\mathbf{t}) = 1$,

$$\phi_{1i}(\mathbf{t}) = \delta_i \prod_{r=1}^{n} (\sum_{\nu_r=1}^{i-1} \beta_{i\nu_r} \phi_{\nu_r}(\mathbf{t}_r))^{\lambda_r},$$

$$\phi_{2i}(\mathbf{t}) = \rho(\mathbf{t}) \sum_{j=1}^{i} \gamma_{ij} \sum_{r=1}^{n} (\lambda_r \Gamma(\mathbf{t}_r) \vartheta^{(\rho(\mathbf{t}) - \rho(\mathbf{t}_r) - 1)} \phi_j(\mathbf{t}_r)),$$

if $\mathbf{t} = [\mathbf{t}_1{}^{\lambda_1}, \cdots, \mathbf{t}_n{}^{\lambda_n}]_x$,

$$\phi_{1i}(\mathbf{t}) = \sum_{j=1}^{i} \omega_{ij} \delta_j \prod_{r=1}^{n} (\sum_{\nu_r=1}^{j-1} \beta_{j\nu_r} \phi_{\nu_r}(\mathbf{t}_r))^{\lambda_r},$$

$$\phi_{2i}(\mathbf{t}) = \sum_{1 \le v \le j \le i} \omega_{ij} \gamma_{jv} \sum_{r=1}^{n} (\lambda_r \Gamma(\mathbf{t}_r) \vartheta^{(\rho(\mathbf{t}) - \rho(\mathbf{t}_r))} \phi_v(\mathbf{t}_r)),$$

if $\mathbf{t} = [\mathbf{t}_1{}^{\lambda_1}, \cdots, \mathbf{t}_n{}^{\lambda_n}]_y$.

The expansion of the derivatives of the numerical solution is given by the following proposition.

Proposition 2.

$$k_{xi}^{(q)} = \sum_{\substack{\mathbf{t} \in \mathbf{LT1X} \\ \rho(\mathbf{t}) = q}} \Gamma(\mathbf{t}) \phi_i(\mathbf{t}) [F(\mathbf{t})]_n, \tag{13}$$

$$k_{yi}^{(q)} = \sum_{\substack{\mathbf{t}\in\mathbf{LT1Y} \\ \rho(\mathbf{t})=q}} \Gamma(\mathbf{t})\phi_i(\mathbf{t})[F(\mathbf{t})]_n. \tag{14}$$

This proposition generalizes the theorem (4.4) from [3] and , for $q = 1$, (13), (14) coincide with the corresponding expressions from [3].

The order conditions are defined by the following proposition.

Proposition 3.
$\delta_n^x = O(h^{p+1})$, *if the conditions* $\sum_{i=1}^{m} \mu_i\phi_i(\mathbf{t}) = \frac{1}{\Gamma(\mathbf{t})}$,
hold for all trees $\mathbf{t} \in \mathbf{T1X}$ *of order* $\rho(\mathbf{t}) \le p$,
$\delta_n^y = O(h^{q+1})$, *if* $\sum_{i=1}^{m} \mu_i\phi_i(\mathbf{t}) = \frac{1}{\Gamma(\mathbf{t})}$ *hold,*
for all trees $\mathbf{t} \in \mathbf{T1Y}$ *of order* $\rho(\mathbf{t}) \le q$.

A numerical solution converges with order p with respect to x and with order q with respect to y if the global error

$$e_n^x = x_n - x(t_n), \quad e_n^y = y_n - y(t_n)$$

satisfies

$$e_n^x = O(h^p), \quad e_n^y = O(h^q).$$

Applying methods (6) for solving the scalar test equation $x' = \lambda x$ we obtain $x_{n+1} = R(z)x_n$, $z = h\lambda$, where $R(z)$ is called a stability function.

The following theorem answers the question on convergence of the (m, k)-methods (6).

Proposition 4. *Assume that scheme (6) is consistent of order p with respect to x and of order $(q-1)$ with respect to y Suppose that the stability factor is such that $|R(\infty)| < 1$ (stability function at ∞). Then numerical solution converges to the exact solution with the order $\overline{p}$ on variable x and with the order q on variable y, where the value $\overline{p}$ is set by above chosen implementation of the scheme a) or b):*

$$\text{a) } \overline{p} = \min(p, 2q), \text{ b) } \overline{p} = \min(p, q+1).$$

We note, that the given proposition in the case p=q follows from the theorem 1 [3] true for a wider class of the one-step methods.

In Tables 1, 2 the order conditions ensuring convergence of (m, k)-methods up to the fourth order of accuracy are tabulated. We use the notations

$$\gamma_i = \sum \gamma_{ij}\delta_j,\ \tilde{c}_i = \sum \tilde{c}_{ij}\delta_j,\ \beta_i = \sum \beta_{ij}\delta_j.$$

Table 1. Order conditions for the x-component

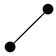
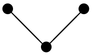
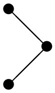
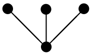
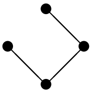
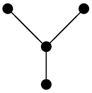
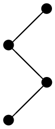
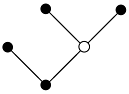
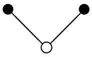
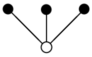
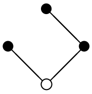
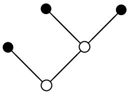
$\rho(\mathbf{t})$	$\mathbf{t}$		
1		$\sum \mu_i \delta_i = 1$	(15.a)
2		$\sum \mu_i \tilde{c}_i = \frac{1}{2}$	(15.b)
3		$\sum \mu_i \beta_i^2 + 2\vartheta \sum \mu_i \gamma_i = \frac{1}{3}$	(15.c)
3		$\sum \mu_i \tilde{c}_{ij} \tilde{c}_j = \frac{1}{6}$	(15.d)
4		$\sum \mu_i \beta_i^3 + 3\vartheta \sum \mu_i \gamma_i = \frac{1}{4}$	(15.e)
4		$\sum \mu_i \beta_i \beta_{ij} \tilde{c}_j + \vartheta \sum \mu_i \gamma_{ij} \tilde{c}_j + \frac{1}{2}\vartheta^2 \sum \mu_i \gamma_i = \frac{1}{8}$	(15.f)
4		$\sum \mu_i \tilde{c}_{ij} \beta_j^2 + 2\vartheta \sum \mu_i \tilde{c}_{ij} \gamma_j = \frac{1}{12}$	(15.g)
4		$\sum \mu_i \tilde{c}_{ij} \tilde{c}_{jk} \tilde{c}_k = \frac{1}{24}$	(15.h)
4		$\sum \mu_i \beta_i \beta_{ij} \omega_{jk} \beta_k^2 + \vartheta \sum \mu_i (2\beta_i \beta_{ij} \omega_{jk} \gamma_k +$ $+\gamma_{ij} \omega_{jk} \beta_k^2) + \vartheta^2 \sum \mu_i (\gamma_i + 2\gamma_{ij} \omega_{jk} \gamma_k) = \frac{1}{4}$	(15.i)

Table 2. Order conditions for the y - component

$\rho(\mathbf{t})$	$\mathbf{t}$		
2		$\sum \mu_i \omega_{ij} \beta_j^2 + 2\vartheta \sum \mu_i \omega_{ij} \gamma_j = 1$	(15.j)
3		$\sum \mu_i \omega_{ij} \beta_j^2 + 3\vartheta^2 \sum \mu_i \omega_{ij} \gamma_j = 1$	(15.k)
3		$\sum \mu_i \omega_{ij} \beta_j \beta_{jk} \tilde{c}_k + \vartheta \sum \mu_i \omega_{ij} \gamma_{jk} \tilde{c}_k +$ $+\frac{1}{2}\vartheta^2 \sum \mu_i \omega_{ij} \gamma_j = \frac{1}{2}$	(20.l)
3		$\sum \mu_i \omega_{ij} \beta_j \beta_{jk} \omega_{ks} \beta_s^2 + \vartheta \sum \mu_i \omega_{ij} (2\beta_j \beta_{jk} \omega_{ks} \gamma_s +$ $+\gamma_{jk} \omega_{ks} \beta_s^2) + \vartheta^2 \sum \mu_i \omega_{ij} (\gamma_j + 2\gamma_{jk} \omega_{ks} \gamma_s) = 1$	(15.m)

4 (m,k)-Schemes of the Optimum Order

We study the utmost achievable order of accuracy by (m, k)-schemes for given $k \leq 4$ for system (1) of index 1. First we consider the case of the implementation a) of the scheme (6).

Let $k = 1$ and let us consider the schemes with one evaluation of the function F at a step. In the case $m = 1$ the stability function takes the form $R(z) = [1 + (\mu_1 - a)z]/(1 - az)$. Under $\mu_1 = 1$, $a = 0.5$ the order conditions of the second order are satisfied. However, unlike ODE systems, the scheme has only the first order of accuracy, as far as $|R(\infty)| = 1$. Under $\mu_1 = a = 1$ we have the L-stable ($R(\infty) = 0$) scheme of the first order, which in [6] is applied to solve the problem of index 0.

In the case $m = 2$ the conditions of the second order yield $\mu_1 = 1$, $\mu_2 = 0.5a$, and

$$R(z) = \frac{1 + (1 - 2a)z + (0.5 - 2a + a^2)z^2}{(1 - az)^2},\ R(\infty) = \frac{0.5 - 2a - a^2}{a^2}.$$

Setting $a = 1 - 0.5\sqrt{2}$ or $a = 1 + 0.5\sqrt{2}$ we obtain the parameters of L-stable $(2, 1)$-scheme of the second order.

Proposition 5. *For all m there exists no $(m, 1)$-method of order higher than 2.*

The given proposition is a consequence an analogous statement from [9].

Let $k = 2$ and we consider the schemes with two evaluation of the function F on a step. Easily to be convinced, that at $m = 2$ the maximum order is equal to 2. In the case $m = 3$, $M_2 = \{1, 2\}$ the conditions of the third order imply

$$\mu_1 = \beta_{21}^{-2}(3\beta_{21}^2 - 1)/3,\ \mu_2 = \beta_{21}^{-2}/3,\ \mu_3 = \beta_{21}^{-2}(a - 3a^2)/3,$$

$$c_{21} = \frac{-6a^2 + 6a - 1}{6a^2 - 2a}\beta_{21}^2,\ c_{31} = \frac{18a^3 - 21a^2 + 9a - 1}{18a^4 - 12a^3 + 2a^2}\beta_{21}^2 - 1,$$

where a and β_{21} are free parameters. Under $1/3 \leq a \leq 1.068579$ [12] a scheme is A-stable, and under $a \approx 0.43587$ (i.e. a is root of the $a^3 - 3a^2 + 2a/3 - 1/6 = 0$) a scheme is L-stable.

Proposition 6. *For all m and for any choice of sets (5) there exists no $(m, 3)$-method of order higher than 3 for the y-component.*

Let $k = 3$, $M_3 = \{1, s, r\}$, $1 < s < r \leq m$. We denote

$$q_s = \sum_{i=s}^{m} \mu_i \omega_{ij},\ q_r = \sum_{i=r}^{m} \mu_i \omega_{ij},\ u_r = \sum_{r>j\geq l}^{m} \beta_{rj}\omega_{jl}\beta_l^2.$$

The conditions of the fourth order (15.c), (15.e), (15.j), (15.k), (15.i), (15.m) yields

$$\mu_s\beta_s^2 + \mu_r\beta_r^2 = \frac{1}{3},\ \mu_s\beta_s^3 + \mu_r\beta_r^3 = \frac{1}{4},\ q_s\beta_s^2 + q_r\beta_r^2 = 1,$$

$$q_s\beta_s^3 + q_r\beta_r^3 = 1,\ \mu_r\beta_r u_r = \frac{1}{4},\ q_r\beta_r u_r = 1.$$

We introduce the matrices

$$A = \left\{ \begin{matrix} \mu_s & \mu_r \\ q_s & q_r, \end{matrix} \right\}, \quad B = \left\{ \begin{matrix} \beta_s^2 & \beta_s^3 \\ \beta_r^2 & \beta_r^3 \end{matrix} \right\}, \ C = \left\{ \begin{matrix} 1/3 & 1/4 \\ 1 & 1 \end{matrix} \right\}, \quad D = \left\{ \begin{matrix} 1 & 0 \\ -4 & 1 \end{matrix} \right\},$$

then the first four equations can be represented in the form of the matrix equality: $AB = C$. We notice that $\beta_s \neq 0$, as far as $\det(C) \neq 0$. The last two equations give $q_r = 4\mu_r$. Multiplying the matrix equality from the right-hand-side by matrix B^{-1} and from the left by D, we have for the right bottom element of the product

$$0 = \beta_s/(3\beta_r^2(\beta_s - \beta_r)).$$

The obtained contradiction proves the proposition.

However for the explicit problem (2) it is possible to obtain the methods of the fourth order, in addition ensuring the third order for the problem (4) of index 1. In the case $m = 4$, $M_2 = \{1, 3\}$ the parameters of the A-stable scheme are

$$a = \frac{1}{2},\ \mu_1 = \frac{11}{27},\ \mu_2 = -\frac{8}{27},\ \mu_3 = \frac{16}{27},\ \mu_4 = -\frac{4}{27},$$
$$\beta_{31} = \frac{3}{4},\ \beta_{32} = -\frac{3}{32},\ c_{32} = -\frac{9}{32},\ c_{42} = -\frac{21}{16}.$$

and parameters of the L-stable scheme at $m = 5$, $M_2 = \{1, 3\}$ are

$$\mu_1 = \frac{11}{27},\ \mu_2 = \frac{-22a + 5}{54},\ \mu_3 = \frac{16}{27},\ \mu_4 = \frac{-16a + 4}{27},$$

$$\mu_5 = \frac{48a^3 - 32a^2 + 4a}{27},\ \beta_{31} = \frac{3}{4},\ \beta_{32} = \frac{-24a + 9}{32},$$

$$c_{32} = \frac{216a^4 - 864a^3 + 648a^2 - 144a + 9}{384a^2 - 256a + 32},$$

$$c_{52} = \frac{-6912a^6 + 16416a^5 - 14832a^4 + 6296a^3 - 1263a^2 + 114a - 4}{6912a^6 - 13824a^5 + 10944a^4 - 4352a^3 + 912a^2 - 96a + 4},$$

$$c_{42} = \big[c_{52}(576a^5 - 768a^4 + 352a^3 - 64a^2 + 4a) -$$
$$-216a^4 + 4a^3 + 159a^2 - 45a - 3\big]/(192a^3 - 176a^2 + 48a - 4),$$

where a is choosen such that $0.2479 < a < 0.67604$ [12].

Note, that the properties of stability of (m, k)-methods depend on the choice of the set M_k. The following proposition in particular holds.

Proposition 7. *There exists a L-stable* $(4, 3)$*-scheme of order 4 with respect to x and of order 3 with respect to y.*

However, the study of methods at $M_3 = \{1, 2, 3\}$ shows that $|R(z)| > 1$. If we consider the case $M_3 = \{1, 2, 4\}$, the parameters of the L-stable scheme are the following:

$$\mu_4 = \frac{4\beta_2 - 3}{12\beta_4^2(\beta_2 - \beta_4)},\ \mu_2 = \frac{1 - 3\mu_4\beta_4}{3\beta_2},\ \mu_1 = 1 - \mu_2 - \mu_4,$$

$$c_{21} = \frac{(-24a^3 + 36a^2 - 12a + 1)\beta_2}{24a^3 - 16a^2 + 2a},\ c_{43} = \frac{12a^3 - 8a^2 + a}{12\mu_4\beta_2^2},$$

$$c_{31} = \frac{(-12\mu_4 c_{43} + 12a^2 - 12a + 2)\beta_2^2 + (4a - 1)c_{21}}{12\mu_4 c_{43}\beta_2^2},$$

$$\beta_{43} = \frac{(-8a^3 + 3a)\beta_2 - 6a^2 c_{21}}{24\mu_4(c_{21} + ac_{31} + a)},$$

$$\beta_{42} = \frac{4\mu_4\beta_4\beta_{43} + a}{4a\mu_4\beta_2\beta_4},\ \beta_{41} = \beta_4 - \beta_{42},$$

$$\mu_3 = \frac{-12\mu_4\beta_{42}\beta_2^2 - 4a + 1}{12\beta_2^2},$$

where β_2 and β_4 are free parameters, $a \approx 0.572816$.

Proposition 8. *There exist embedded* $(5, 4)$*-schemes of order 4 and 3 determined by the set* $M_4 = \{1, 3, 4, 5\}$*. The scheme of order 4 is L-stable and the scheme of order 3 is A-stable.*

Let β_3, β_4, β_5, β_{32}, β_{54}, c_{54} be, in general, free parameters. We use a short notation

$$q_s = \sum_{i=s+1}^{5} \mu_i\omega_{ij},\ u_s = \sum_{s>j\geq l}^{m} \beta_{sj}\omega_{jl}\beta_l^2,\ s = 3, 4, 5.$$

The conditions of the fourth order (15.j), (15.k), (15.i), (15.m) yields

$$q_3\beta_3^2 + q_4\beta_4^2 = 1 - \frac{1}{3a},$$

$$q_3\beta_3^3 + q_4\beta_4^3 = 1 - \frac{1}{4a},$$

$$\mu_4\beta_4 u_4 + \mu_5\beta_5 u_5 = \frac{1}{4},\ q_4\beta_4 u_4 = 1.$$

Having chosen the free parameters, we obtain q_3, q_4 from the first two equations and μ_5 from the expression $q_4 = -a^{-2}c_{54}\mu_5$.

Equations (15.c), (15.e)

$$\mu_3\beta_3^2 + \mu_4\beta_4^2 + \mu_5\beta_5^2 = \frac{1}{3},$$

$$\mu_3\beta_3^3 + \mu_4\beta_4^3 + \mu_5\beta_5^3 = \frac{1}{4}$$

give μ_3, μ_4.

Now u_4, u_5 are obtained from (15.i), (15.m). Using $u_4 = \beta_3^2\beta_{43}$ we get β_{43}.

Parameters c_{43}, c_{53} are obtained from the expression

$$q_3 = a^{-3}\big[\mu_5(c_{43}c_{54} - ac_{53}) - a\mu_4 c_{43}\big]$$

and from the equation (15.g)

$$\mu_4\beta_3^2 c_{43} + \mu_5\big(\beta_3^2 c_{53} + \beta_4^2 c_{54}\big) = \frac{1}{12} - \frac{a}{3}.$$

The expression

$$u_5 = a^{-2}\big[(a\beta_{53} - c_{43}\beta_{54})\beta_3^2 + a\beta_{54}\beta_4^2\big]$$

gives β_{53}.

From the conditions (15.l), (15.f)

$$q_3\beta_3\beta_{32} + q_4\beta_4(\beta_{43}\beta_3 + \beta_{42}) = \frac{5}{6} - \frac{1}{8a},$$

$$\mu_3\beta_3\beta_{32} + \mu_4\beta_4(\beta_{42} + \beta_{31}\beta_{43}) + \mu_5\beta_5(\beta_{52}+$$

$$+\beta_{54}c_{43} + \beta_{41}\beta_{54} + \beta_{31}\beta_{53}) = \frac{1}{8} - \frac{a}{3},$$

the equation

$$\mu_5 c_{32}c_{43}c_{54} = a^5 - 4a^4 + 3a^3 - \frac{2a^2}{3} + \frac{a}{24},$$

ensuring the L-stability, and from the conditions (15.h), (15.d), (15.b), (15.a)

$$\mu_4 c_{32}c_{43} + \mu_5(\beta_3 c_{43}c_{54} + c_{32}c_{53} + c_{42}c_{54}) = \frac{1}{24} - \frac{a}{2} + \frac{3a^2}{2} - a^3,$$

$$\mu_3 c_{32} + \mu_4(\beta_3 c_{43} + c_{42}) + \mu_5(\beta_3 c_{53} + \beta_{41}c_{54} + c_{43}c_{54} + c_{52}) = \frac{1}{6} - a + a^2,$$

$$\mu_2 + \mu_3\beta_{31} + \mu_4(\beta_{41} + c_{43}) + \mu_5(\beta_{51} + c_{53} + c_{54}) = \frac{1}{2} - a,$$

$$\mu_1 + \mu_3 + \mu_4 + \mu_5 = 1$$

we evaluate sequentially the parameters β_{42}, β_{52}, c_{32}, c_{42}, c_{52}, μ_1, μ_2.

Degeneration of the minor

$$\left\{\begin{array}{ccc} \beta_3^2 & \beta_4^2 & \beta_5^2 \\ c_{32} & \beta_3 c_{43} + c_{42} & c_{52} + \beta_3 c_{53} + c_{54}(\beta_{41} + c_{43}) \\ 0 & -a\beta_3^2 c_{43} & -a\beta_3^2 c_{53} + c_{54}(\beta_3^2 c_{43} - a\beta_4^2) \end{array}\right\},$$

corresponding to the parameters $\tilde{\mu}_3$, $\tilde{\mu}_4$, $\tilde{\mu}_5$ in the order conditions (15.c), (15.d), (15.j) of the embedded scheme, ensures the existence of embedded method of the order 3. This condition gives the algebraic equation at parameter a

$$84a^4 - 132a^3 + 72a^2 - 15a + 1 = 0$$

having two of the solutions in $\mathbf{R}:\ a_1 \approx 0.130354,\ a_2 \approx 0.239192$.

Choosing the second value a and setting the free parameter $\tilde{\mu}_5$ we obtain from conditions (15.a) – (15.d) of the embedded scheme other values of the parameters $\tilde{\mu}_i$.

Now we consider the implementation b) with the time-lagging of matrices derivatives. Assume that the coefficients of the scheme are independent of the parameter ϑ. In the case of order 3 accuracy this yields the two additional order conditions

$$\sum \mu_i \beta_i = \frac{1}{2}, \tag{15.c'}$$

$$\sum \mu_i \omega_{ij} \beta_j = 1. \tag{15.j'}$$

Proposition 9. *For all m there exists no $(m, 2)$-method of order 3 satisfying* (15.c'), (15.j').

This follows from the inconsistency of (15.c'), (15.j'), and (15.j).

In the case $m = 3$, $k = 3$ the parameters of the L-stable scheme are

$$\mu_1 = \frac{(6a-1)\beta_3 - 2a}{4(3a-1)\beta_3},\ \mu_3 = -\frac{a}{\beta_3((6a-3)\beta_3 - 6a + 2)},$$

$$\mu_2 = \frac{(6a-3)(1-2\mu_3\beta_3)}{4(3a-1)},\ \beta_2 = \frac{6a-2}{6a-3},\ \beta_{32} = \frac{a(1-2a)}{2\mu_3\beta_2},$$

$$c_{21} = \frac{6a^2 - 6a + 1}{6\mu_3 c_{32}},\ c_{31} = \frac{1 - 2\mu_2 c_{21} - 2\mu_3 c_{32} - 2a}{2\mu_3},$$

where $a \approx 0.43587$, and β_3, β_{32} are free parameters.

In addition in the case of the order 4 accuracy it is necessary to satisfy 7 conditions

$$\sum \mu_i \beta_{ij} c_j = \frac{1}{6} - \frac{a}{2}, \tag{15.f'}$$

$$\sum \mu_i c_{ij} \beta_j = \frac{1}{6} - \frac{a}{2}, \tag{15.g'}$$

$$2\sum \mu_i \beta_i \beta_{ij} \omega_{jk} \beta_k + \sum \mu_i \beta_{ij} \omega_{jk} \beta_k^2 = 1, \tag{15.i'}$$

$$\sum \mu_i \beta_{ij} \omega_{jk} \beta_k = \frac{1}{2}, \tag{15.i''}$$

$$\sum \mu_i \tilde{\omega}_{ij} \beta_{jk} c_k = \frac{1}{2} - a - \frac{1}{6a}, \tag{15.l'}$$

$$2\sum \mu_i \tilde{\omega}_{ij} \beta_j \beta_{jk} \omega_{kl} \beta_l + \sum \mu_i \tilde{\omega}_{ij} \beta_{jk} \omega_{kl} \beta_l^2 = 3 - \frac{1}{a}, \tag{15.m'}$$

$$\sum \mu_i \tilde{\omega}_{ij} \beta_{jk} \omega_{kl} \beta_l = 1 - \frac{1}{2a}. \tag{15.m''}$$

Here is $\tilde{\omega} = \omega - a^{-1}I$, where I is the identity matrix.

The following result for the scheme with the time-lagging of matrices derivatives similar to Proposition 7 holds.

Proposition 10. *There exists an L-stable* $(10, 4)$*-scheme of order 4 accuracy in both variables with the time-lagging of matrices derivatives.*

We present this result without the proof, since the proof is too complicated.

In conclusion we note that at $m \leq 9$ there exists no the $(m, 4)$-scheme of the order 4 accuracy in both variables with the time-lagging of matrices derivatives.

References

1. Boyarincev, Y.A., Danilov, V.A., Loginov, A.A., Chistyakov, V.F.: The Numerical Methods for Singular Systems. Nauka, Novosibirsk (1989)
2. Gear, C.W.: Differential-algebraic equations index transformations. SIAM J. Sci. Stat. Comput. **9**, 39–47 (1988)
3. Roche, M.: Rosenbrock methods for differential algebraic equation. Numer. Math. **52**, 45–63 (1988)
4. Levykin, A.I., Novikov, E.A.: A study of (m, k)-methods of the order 3 for implicit systems of ordinary differential equations. Novosibirsk, Preprint Computing Center SB RAS **882** (1990)
5. Werwer, J.G., Scholz, S., Blom, J.G., Louter-Nool, M.: A class Runge-Kutta-Rosenbrock Methods for solving stiff differential equations. ZAAM **63**, 13–20 (1983)
6. Novikov, E.A., Yumatova, L.A.: Some onestep methods for solving differential algebraic equation. Novosibirsk, Preprint Computing Center SB RAS **661** (1986)
7. Dekker, K., Verver, J.G.: Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations. North-Holland, Amsterdam (1984)
8. Levykin, A.I., Novikov, E.A.: (m, k)-method of the order 2 for implicit systems of ordinary differential equations. Novosibirsk, Preprint Computing Center SB RAS **768** (1987)
9. Novikov, E.A., Shitov, Y.A., Shokin, Y.I.: A class of (m, k)-methods for solving stiff systems. Dokl. Akad. Nauk. **301**(6), 1310–1313 (1988)
10. Elsgolts, L.E.: Differential Equation and Variatsional Calculus. Nauka, Moskow (1969)
11. Deuflhard, P., Hairer, E., Zugck, J.: One-step and extrapolation methods for differential-algebraic systems. Numer. Math. **51**, 501–516 (1987)
12. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems. Springer, Heidelberg (1991)
13. Mao, G., Petzold, L.R.: Efficient integration over discontinuities for differential-algebraic systems. Computers Mathematics with Applications **43**(1), 65–79 (2002)
14. Boscarino, S.: Error Analysis of IMEX RungeKutta Methods Derived from Differential-Algebraic Systems. SIAM J. Numer. Anal. **45**(4), 1600–1621 (2007)