Journal of Computational Mathematics, Vol.18, No.3, 2000, 325–336.

MULTISTEP DISCRETIZATION OF INDEX 3 DAES*[1)]

Yang Cao Qing-yang Li
(*Department of Mathematics, Tsinghua University, Beijing 100084, China*)

Abstract

In the past Index-3 DAEs were solved by BDF methods as multistep methods or implicit Runge-Kutta methods as one-step methods. But if the equations are nonstiff, not only BDF but other multistep methods may be applied. This paper considers four different types of multistep discretization of index 3 DAEs in hessenberg form. The convergence of these methods is proven under the condition that the multistep formula is striculy infinite stable. numerical tests also confirm the results.

Key words: Multistep methods, Adams method, BDF, DAEs, Index 3

1. Introduction

In this paper, we will consider the multistep discrezations of the differential-algebraic equations (DAEs) in Hessenberg form

$$y' = F(y,z), \tag{1.1a}$$
$$z' = K(y,z,u), \tag{1.1b}$$
$$0 = G(y), \tag{1.1c}$$

where $F \in \mathbb{R}^{N+M} \to \mathbb{R}^N$, $K \in \mathbb{R}^{N+M+L} \to \mathbb{R}^M$, $G \in \mathbb{R}^N \to R^L$, the initial value (y_0, z_0, u_0) at x_0 are assumed to be consistent, i.e., they satisfy

$$0 = G(y_0), \tag{1.2a}$$
$$0 = (G_y F)(y_0, z_0), \tag{1.2b}$$
$$0 = (G_{yy}(F,F) + G_y F_y F + G_y F_z K)(y_0, z_0, u_0). \tag{1.2c}$$

We supposes, F, G and K are sufficiently differentiable, and that

$$\|[G_y F_z K_u)(y,z,u)]^{-1}\| \le C. \tag{1.3}$$

in a neighbourhood of the solution. Such problems often appear in the simulation of mechanical systems with constraints and the singularly perturbed problems (see [2,

* Received April 19, 1997.
1) Supported by National Natural Science Foundation of China.

5, 7]). BDF-methods were the first formula for solving DAEs. Convergence of BDF-methods for (1.1) was given in [3]. During the same period, the implicit Runge-Kutta (IRK) methods for solving index 3 DAEs were also considered. In [5] convergence results of IRK-methods were given, which have been sharpened by L. Jay in [8, 9]. These methods are directly derived from those methods for solving stiff ODEs. But if the equations are nonstiff, much more methods can be applied. For example, the Adams_methods, Simposon formula, etc. Then a natural question is: do these formula converge? So we need to consider the general multistep discretizations of DAEs of index 3. In [1], the multistep discretization for DAEs of index one and two were discussed. Since the numerical solution of index 3 DAEs is such more complicate than that of index 2 DAEs, different discretization may be applied to different part of the equations (1.1). There are two different types of discretization: implicit formula and explicit formula, which can be applied to (1.1a) and (1.1b) independently. So we distinguish four types of discretizations as following:

Type I: Implicit-Implicit type

$$y_{n+k} = \sum_{i=0}^{k-1} \alpha_i y_{n+i} + h \sum_{i=0}^{k} \beta_i F(y_{n+i}, z_{n+i}), \tag{1.4a}$$

$$z_{n+k} = \sum_{i=0}^{k-1} a_i z_{n+i} + h \sum_{i=0}^{k} b_i K(y_{n+i}, z_{n+i}, u_{n+i}), \tag{1.4b}$$

$$0 = G(y_{n+k}), \tag{1.4c}$$

where $y_{n+k}, z_{n+k}, u_{n+k}$ are unknown.

Type II: Implicit-Explicit type

$$y_{n+k} = \sum_{i=0}^{k-1} \alpha_i y_{n+i} + h \sum_{i=0}^{k} \beta_i F(y_{n+i}, z_{n+i}), \tag{1.5a}$$

$$z_{n+k} = \sum_{i=0}^{k-1} a_i z_{n+i} + h \sum_{i=0}^{k-1} b_i K(y_{n+i}, z_{n+i}, u_{n+i}), \tag{1.5b}$$

$$0 = G(y_{n+k}), \tag{1.5c}$$

where $y_{n+k}, z_{n+k}, u_{n+k-1}$ are unknown.

Type III: Explicit-Implicit type

$$y_{n+k+1} = \sum_{i=0}^{k-1} \alpha_i y_{n+i+1} + h \sum_{i=0}^{k-1} \beta_i F(y_{n+i+1}, z_{n+i+1}), \tag{1.6a}$$

$$z_{n+k} = \sum_{i=0}^{k-1} a_i z_{n+i} + h \sum_{i=0}^{k} b_i K(y_{n+i}, z_{n+i}, u_{n+i}), \tag{1.6b}$$

$$0 = G(y_{n+k}), \tag{1.6c}$$

where $y_{n+k+1}, z_{n+k}, u_{n+k}$ are unknown.

Type IV: Explicit-Explicit type

$$y_{n+k+1} = \sum_{i=0}^{k-1} \alpha_i y_{n+i+1} + h \sum_{i=0}^{k-1} \beta_i F(y_{n+i+1}, z_{n+i+1}), \tag{1.7a}$$

$$z_{n+k} = \sum_{i=0}^{k-1} a_i z_{n+i} + h \sum_{i=0}^{k-1} b_i K(y_{n+i}, z_{n+i}, u_{n+i}), \tag{1.7b}$$

$$0 = G(y_{n+k}), \tag{1.7c}$$

where $y_{n+k+1}, z_{n+k}, u_{n+k-1}$ are unknown.

The BDF methods in [3] belong to the type I, and the half-explicit methods (see [7]) belong to the type IV. The following figures can illustrate the difference between these four types of discretizations.

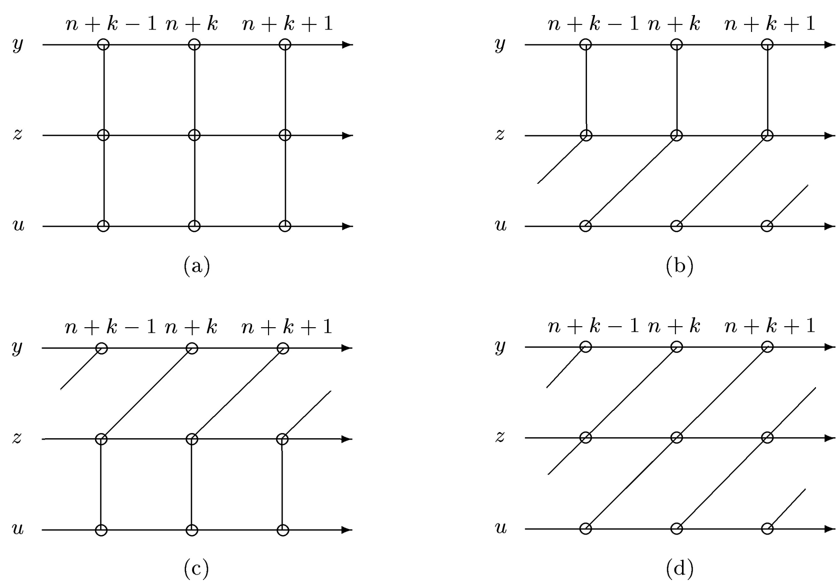

Figure (a) shows the process of solving (1.1) by the type I discretization (1.4). In each step we get y_{n+k}, z_{n+k}, u_{n+k} from the former values. Figure (b)–(d) show the process by type II–IV respectively.

In the following paragraphs, §2 gives the existence, uniqueness and the influence of perturbations. §3 proves the convergence. Because the proof for each type is similar, we focus on the proof to type I. These results are confirmed in §4 by some numerical tests.

2. Existence, Uniqueness and Influence of Perturbations

This section is devoted to analyseing the solution of the nonlinear equations (1.4). the analysis to other types are similar. [5, p75] gives an example to which the implicit Euler formula, which fall into the type I, has no solution. So some limits must be added to the equations. Here we assume that $K(y, z, u)$ depends linearly on u. The important class of Euler-Lagrange equations describing a constrained mechanical system just satisfies this requirement. Moreover, numerical tests show that even if K is nonlinear on u, the solution may still exist, so long as (1.3) is satisfied.

Theorem 2.1. *Suppose that for a solution $y(x)$, $z(x)$, $u(x)$ of* (1.1) *the starting values (y_j, z_j, u_j) satisfy for $j = 0, \cdots, k-1$ and $x_j = x_0 + jh$:*

$$\begin{aligned} y_j - y(x_j) &= O(h), \quad G(y_j) = O(h^2), \\ z_j - z(x_j) &= O(h), \quad (G_y, F)(y_j, z_j) = O(h), \end{aligned} \tag{2.1}$$

and (1.3) *holds in a neighbourhood of this solution, K depends linearly on u and if β_k, $b_k \neq 0$, then the nonlinear system* (1.4) *has a solution for $h \leq h_0$. This solution is locally unique and satisfies*

$$y_k - y(x_k) = O(h), \quad z_k - z(x_k) = O(h), \quad u_k - u(x_k) = O(1). \tag{2.2}$$

Proof. Homotopy technique is used in our proof just like that in [5], Theorem 6.1. We first reformulate (1.4) as following:

$$y_k = \eta + h\beta_k F(y_k, z_k), \tag{2.3a}$$

$$z_k = \xi + hb_k K(y_k, z_k, u_k), \tag{2.3b}$$

$$0 = G(y_k). \tag{2.3c}$$

where $\eta = \sum_{i=0}^{k-1} (\alpha_i y_i + h\beta_i F(y_i, z_i))$ and $\xi = \sum_{i=0}^{k-1} (a_i z_i + hb_i K(y_i, z_i, u_i))$ have been known. Select ν close to $u(x_k)$ such that $(G_{yy}(F, F) + G_y F_y F + G_y F_z k)\ (\eta, \xi, \nu) = 0$. From the initial conditons (2.1), it is easy to verify that:

$$\eta - y(x_k) = O(h), \quad \xi - z(x_k) = O(h), \quad G(\eta) = O(h^2), \quad (G_y F)(\eta, \xi) = O(h). \tag{2.4}$$

Consider the homotopy

$$y(r) = \eta + h\beta_k F(y(r), z(r)) - (1-r)(h\beta_k F(\eta, \xi)), \tag{2.5a}$$

$$z(r) = \xi + hb_k K(y(r), z(r), u(r)) - (1-r)(hb_k K(\eta, \xi, \nu)), \tag{2.5b}$$

$$0 = G(y(r)) - (1-r)G(\eta), \tag{2.5c}$$

Obviously, when $r = 0$, (η, ξ, ν) is the solution, and when $r = 1$, the solution (if it exist) is just y_k, z_k, u_k. Differentiate (2.5) with respect to r, we get

$$\dot{y} = h\beta_k F_y \dot{y} + h\beta_k F_z \dot{z} + h\beta_k F(\eta, \xi), \tag{2.6a}$$

$$\dot{z} = hb_k K_y \dot{y} + hb_k K_z \dot{z} + hb_k K_u \dot{u} + hb_k K(\eta, \xi, \nu)), \tag{2.6b}$$

$$0 = G_y \dot{y} + G(\eta), \tag{2.6c}$$

Inserting (2.6a) into (2.6c) and substituting (2.6a) and (2.6b) into the resulting equation, we get

$$0 = (\beta_k G_y F_y^2 + b_k G_y F_z K_y)\dot{y} + (\beta_k G_y F_y F_z + b_k G_y F_z K_z)\dot{z} + b_k G_y F_z K_u \dot{u}$$

$$+ \beta_k G_y F_y F(\eta,\xi) + b_k G_y F_z K(\eta,\xi,\nu) + \frac{1}{h} G_y F(\eta,\xi) + \frac{1}{h^2\beta_k} G(\eta).$$

With (2.6a) and (2.6b), we have

$$M \begin{pmatrix} \dot{y} \\ \dot{z} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} h\beta_k F(\eta,\xi) \\ hb_k K(\eta,\xi,\nu) \\ \omega \end{pmatrix} \tag{2.7}$$

where

$$M = \begin{pmatrix} I - h\beta_k F_y & -h\beta_k F_z & 0 \\ -hb_k K_y & I - hb_k K_z & -hb_k K_u \\ \beta_k G_y F_y^2 + b_k G_y F_z K_y & \beta_k G_y F_y F_z + b_k G_y F_z K_z & b_k G_y F_z K_u \end{pmatrix}$$

and

$$\omega = \beta_k G_y F_y F(\eta,\xi) + b_k G_y F_z K(\eta,\xi,\nu) + \frac{1}{h} G_y F(\eta,\xi) + \frac{1}{h^2\beta_k} G(\eta).$$

M has a bounded inverse

$$M^{-1} = \begin{pmatrix} O(1) & O(h) & O(h^2) \\ O(h) & O(1) & O(h) \\ O(1) & O(1) & O(1) \end{pmatrix} \tag{2.8}$$

provided that (1.3) holds at $(y(r), z(r), u(r))$. Since $b_k, \beta_k \neq 0$ and K is linear on u, this condition need only that $(y(r), z(r))$ lie in a neighbourhood V of $(y(x_k), z(x_k))$. But η, ξ do, so we have $y(r) - y(x_k) = O(h)$, $z(r) - z(x_k) = O(h)$ for $r \leq 1$ and sufficiently small h. When $r = 1$, this give the existence of the solution (y_k, z_k, y_k) and (2.2).

Uniqueness can be gotten from the following result about perturbations. □

Theorem 2.2. *Let (y_k, z_k, u_k) be given by (2.3) and consider perturbed values $(\hat{y}_k, \hat{z}_k, \hat{u}_k)$ satisfying*

$$\hat{y}_k = \hat{\eta} + h\beta_k F(\hat{y}_k, \hat{z}_k) + h\delta, \tag{2.9a}$$

$$\hat{z}_k = \hat{\xi} + hb_k K(\hat{y}_k, \hat{z}_k, \hat{u}_k) + h\theta, \tag{2.9b}$$

$$0 = G(\hat{y}_k) + \mu. \tag{2.9c}$$

Suppose that

$$\hat{\eta} - \eta = O(h^2), \quad \hat{\xi} - \xi = O(h), \ \delta = O(h), \ \theta = O(1), \ \mu = O(h^3). \tag{2.10}$$

Then for $h \leq h_0$ we have the estimate:

$$\begin{aligned} \|\hat{y}_k - y_k\| \leq & C(\|\hat{\eta} - \eta\| + h\|\hat{\xi} - \xi\| + h\|\delta\| + h^2\|\theta\| + \|\mu\|) \\ \|\hat{z}_k - z_k\| \leq & \frac{C}{h}(\|G_y(\eta)(\hat{\eta} - \eta)\| + h\|\hat{\eta} - \eta\| + h\|\hat{\xi} - \xi\| \\ & + h\|\delta\| + h^2\|\theta\| + \|\mu\|) \\ \|\hat{u}_k - u_k\| \leq & \frac{C}{h^2}(\|G_y(\eta)(\hat{\eta} - \eta)\| + h\|\hat{\eta} - \eta\| + h\|G_y(\eta)F_z(\eta,\xi)(\hat{\xi} - \xi)\| \end{aligned} \tag{2.11}$$

$$+ h^2\|\hat{\xi} - \xi\| + h\|\delta\| + h^2\|\theta\| + \|\mu\|)$$

Proof. Consider the homotopy

$$y(r) = \eta + h\beta_k F(y(r), z(r)) + (1-r)(\hat{\eta} - \eta + h\delta), \tag{2.12a}$$
$$z(r) = \xi + hb_k K(y(r), z(r), u(r)) + (1-r)(\hat{\xi} - \xi + h\theta), \tag{2.12b}$$
$$0 = G(y(r)) + (1-r)\mu, \tag{2.12c}$$

then as in the proof in Theorem 2.1, we obtain the differential equations

$$M \begin{pmatrix} \dot{y} \\ \dot{z} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} -(\hat{\eta} - \eta + h\delta) \\ -(\hat{\xi} - \xi + h\theta) \\ v \end{pmatrix}, \tag{2.13}$$

where

$$v = \frac{1}{\beta h^2}(G_y(\hat{\eta} - \eta + h\delta) + \mu) + \frac{1}{h}(G_y F_y(\hat{\eta} - \eta + h\delta) + G_y F_z(\hat{\xi} - \xi + h\theta))$$

From (2.8), we get the desired estimate by integration. □

For the multistep methods, the following corollary is more useful.

Corollary 2.1. *Let* (y_k, z_k, u_k) *be given by*

$$y_k = \sum_{i=0}^{k-1} \alpha_i y_i + h \sum_{i=0}^{k} \beta_i F(y_i, z_i), \tag{2.14a}$$
$$z_k = \sum_{i=0}^{k-1} a_i z_i + h \sum_{i=0}^{k} b_i K(y_i, z_i, u_i), \tag{2.14b}$$
$$0 = G(y_k), \tag{2.14c}$$

and perturbed values $(\hat{y}_k, \hat{z}_k, \hat{u}_k)$ *satisfy*

$$\hat{y}_k = \sum_{i=0}^{k-1} \alpha_i \hat{y}_i + h \sum_{i=0}^{k} \beta_i F(\hat{y}_i, \hat{z}_i) + h\delta, \tag{2.15a}$$
$$\hat{z}_k = \sum_{i=0}^{k-1} a_i \hat{z}_i + h \sum_{i=0}^{k} b_i K(\hat{y}_i, \hat{z}_i, \hat{u}_i) + h\theta, \tag{2.15b}$$
$$0 = G(\hat{y}_k) + \mu, \tag{2.15c}$$

Suppose for $j = 0, \cdots, k-1$

$$\hat{y}_j - y_j = O(h^2),\ \hat{z}_j - z_j = O(h),\ \delta = O(h),\ \theta = O(1),\ \mu = O(h^3).$$

The for $h \le h_0$ *we have the estimates*

$$\|\hat{y}_k - y_k\| \le C(\|\hat{Y}_0 - Y_0\| + h\|\hat{Z}_0 - Z_0\| + h\|\delta\| + h^2\|\theta\| + \|\mu\|)$$
$$\|\hat{z}_k - z_k\| \le \frac{C}{h}\Big(\Big\|\sum_{i=0}^{k-1} G_y(y_k)(\hat{y}_j - y_j)\Big\| + h\|\hat{Y}_0 - Y_0\| + h\|\hat{Z}_0 - Z_0\|$$

$$+ h\|\delta\| + h^2\|\theta\| + \|\mu\|\Big) \tag{2.16}$$

$$\|\hat{u}_k - u_k\| \leq \frac{C}{h^2}\Big(\Big\|\sum_{i=0}^{k-1} G_y(y_k)(\hat{y}_j - y_j)\Big\| + h^2\|\hat{Y}_0 - Y_0\| + h\|\hat{Z}_0 - Z_0\|$$

$$+ h\Big\|\sum_{i=0}^{k-2} G_y(y_k)F_z(y_k, z_k)(\hat{z}_j - z_j)\Big\| + h\|\delta\| + h^2\|\theta\| + \|\mu\|\Big)$$

where $\hat{Y}_0 - Y_0 = (\hat{y}_{k-1} - y_{k-1}, \cdots, \hat{y}_0 - y_0)^T$, $\|\hat{Y}_0 - Y_0)\| = \max_{0\leq j\leq k-1} \|\hat{y}_j - y_j\|$ *and likewise for the* z*-component.*

3. Local and Global Convergence

When initial values $y_j = y(x_j)$, $z_j = z(x_j)$, $u_j = u(x_j)$ $(j = 0, \cdots, k-1)$, applying the multistep formula once, the difference $y_k - y(x_k), z_k - z(x_k)$, $u_k - u(x_k)$ are called the local errors of the method.

In the following paragraph, denote

$$\rho_1(\xi) = \xi^k - \sum_{i=0}^{k-1} \alpha_i \xi^i, \quad \sigma_1 = \sum_{i=0}^{k} \beta_i \xi^i, \quad \rho_2(\xi) = \xi^k - \sum_{i=0}^{k-1} a_i \xi^i, \quad \sigma_2 = \sum_{i=0}^{k} b_i \xi^i.$$

If the multistep discretizaiton (ρ_1, σ_1) has order p, (ρ_2, σ_2) has order q for ODEs, set $\hat{y}_j = y(x_j)$, $\hat{z}_j = z(x_j)$, $\hat{u}_j = u(x_j)$, $\delta = O(h^p)$, $\theta = O(h^q)$, $\mu = 0$ in Corollary 2.1. We get immediately.

Lemma 3.1. *If the multistep discretization* (α, β) *has order* p, (a, b) *has order* q, *the local errors satisfy*

$$\begin{aligned} &y_k - y(x_k) = O(h^{p+1} + h^{q+2}),\ z_k - z(x_k) = O(h^p + h^{q-1}), \\ &u_k - u(x_k) = O(hp - 1 + h^q). \end{aligned} \tag{3.1}$$

Theorem 3.1. *If the multistep discretization* (ρ_1, σ_1) *has order* p, (ρ_2, σ_2) *has order* q, *and both these discretizations are stable and strictly stable at infinity, which implies that the zeros of* σ_1, σ_2 *lie inside the unit disc* $\|\xi\| < 1$, *then for* $x_n = nh \leq$ *Const the global errors satisfy*

$$\begin{aligned} &y_n - y(x_n) = O(h^p + h^q),\ z_n - z(x_n) = O(h^p + h^q), \\ &u_n - u(x_n) = O(h^{p-1} + h^q), \end{aligned} \tag{3.2}$$

whenever the initial values satisfy (for $j = 0, \cdots, k-1$*)*

$$\begin{aligned} &y_j - y(x_j) = O(h^{p+1} + h^{q+1}),\ z_j - z(x_j) = O(^{p+1} + h^{q+1}), \\ &u_j - u(x_j) = O(h^{p-1} + h^q). \end{aligned}$$

Proof. This proof is inspired by Lady Windermere's Fan (see [5, Figure 4.1, p.36]). We need just give thenpropagation of the errors, and use the standard technique in [5]. As in [5], let

$$S = K_u(G_y F_z K_u)^{-1} G_y,\ Q_y = F_z S,\ P_y = I - Q_y,\ Q_z = S F_z,\ P_z = I - Q_z$$

and for any $1 \le i \le n$,

$$\begin{aligned}
&\Delta Y_n = (\Delta y_{n+k-1}, \cdots, \Delta y_n)^T, && \Delta Z_n = (\Delta z_{n+k-1}, \cdots, \Delta z_n)^T,\\
&\Delta Y_n^1 = (P_{y,n+k-1}\Delta y_{n+k-1}, \cdots, P_{y,n}\Delta y_n)^T, && \Delta Y_n^2 = (Q_{y,n+k-1}\Delta y_{n+k-1}, \cdots, Q_{y,n}\Delta y_n)^T,\\
&\Delta Z_n^1 = (P_{z,n+k-1}\Delta z_{n+k-1}, \cdots, P_{z,n}\Delta z_n)^T, && \Delta Z_n^2 = (Q_{z,n+k-1}\Delta z_{n+k-1}, \cdots, Q_{z,n}\Delta z_n)^T,\\
&\Delta U_n = (\Delta u_{n+k-1}, \cdots, \Delta u_n)^T.
\end{aligned}$$

Consider two neighbouring multistep solutions $\hat{y}_n, \hat{z}_n, \hat{u}_n$ and $\tilde{y}_n, \tilde{z}_n, \tilde{u}_n$, from (1.4), we have

$$\begin{aligned}
\Delta y_{n+k} =& \sum_{i=0}^{k-1} \alpha_i \Delta y_{n+i} + h \sum_{i=0}^{k} \beta_i F_z(\hat{y}_{n+i}, \hat{z}_{n+i}) \Delta z_{n+i}\\
&+ O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\|), \qquad (3.3a)\\
\Delta z_{n+k} =& \sum_{i=0}^{k-1} \alpha_i \Delta z_{n+i} + h \sum_{i=0}^{k} b_i K_u(\hat{y}_{n+i}, \hat{z}_{n+i}, \hat{u}_{n+i}) \Delta u_{n+i}\\
&+ O(h\|\Delta Y_n\| + h\|\Delta Z_n\| + h^2\|\Delta U_n\|), \qquad (3.3b)\\
0 =& G_y(\hat{y}_{n+k}) \Delta y_{n+k} + O(h\|\Delta Y_n\|), \qquad (3.3c)
\end{aligned}$$

where $\Delta y_{n+k} = \hat{y}_{n+k} - \tilde{y}_{n+k}$, $\Delta z_{n+k} = \hat{z}_{n+k} - \tilde{z}_{n+k}$, $\Delta u_{n+k} = \hat{u}_{n+k} - \tilde{u}_{n+k}$. From (3.3b),

$$\begin{aligned}
h \sum_{i=0}^{k} b_i \Delta u_{n+i} =& (G_y F_z K_u)^{-1} G_y F_z \Big(\Delta z_{n+k} - \sum_{i=0}^{k-1} a_i \Delta z_{n+i} \Big)\\
&+ O(h\|\Delta Y_n\| + h\|\Delta Z_n\| + h^2\|\Delta U_n\|)
\end{aligned}$$

and so

$$P_z \Delta z_{n+k} = \sum_{i=0}^{k-1} a_i \Delta z_{n+i} + O(h\|\Delta Y_n\| + h\|\Delta Z_n\| + h^2\|\Delta U_n\|) \qquad (3.4)$$

Inserting (3.3a) into (3.3c) and Substituting (3.3b) into the result formula lead to

$$\begin{aligned}
h^2 \beta_k b_k \Delta u_{n+k} =& - \sum_{i=0}^{k-1} (h^2 \beta_k b_i \Delta u_{n+i} + (G_y F_z K_u)^{-1} G_y (h \beta_k a_i F_z \Delta z_{n+i}\\
&+ h \beta_i F_z \Delta z_{n+i} + \alpha_i \Delta y_{n+i}))\\
&+ O(h\|\Delta Y_n\| + h^2\|\Delta Z_n \ + h^3\|\Delta U_n\|). \qquad (3.5)
\end{aligned}$$

And then

$$\begin{aligned}
h \beta_k \Delta z_{n+k} =& \sum_{i=0}^{k-1} (h \beta_k a_i P_z \Delta z_{n+i} - h \beta_i Q_z \Delta z_{n+i} - \alpha_i S Q_y \Delta y_{n+i})\\
&+ O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|), \qquad (3.6)\\
\Delta y_{n+k} =& \sum_{i=0}^{k-1} (\alpha_i P_y \Delta y_{n+i} + h(\beta_k a_i + \beta_i) F_z P_z \Delta z_{n+i}
\end{aligned}$$

$$+ O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|). \tag{3.7}$$

Multiplying it by $P_y(\hat{y}_{n+k}, \hat{z}_{n+k}, \hat{u}_{n+k})$, $Q_y(\hat{y}_{n+k}, \hat{z}_{n+k}, \hat{u}_{n+k})$, $Q_z(\hat{y}_{n+k}, \hat{z}_{n+k}, \hat{u}_{n+k})$ respectly, we get

$$\begin{aligned} P_y\Delta y_{n+k} =& \sum_{i=0}^{k-1}(\alpha_i P_y \Delta y_{n+i} + h(\beta_k a_i + \beta_i) F_z P_z \Delta z_{n+i}) \\ &+ O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\| \end{aligned} \tag{3.8a}$$

$$Q_y\Delta y_{n+k} = O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|, \tag{3.8b}$$

$$\begin{aligned} h\beta_k Q_z \Delta z_{n+k} =& -\sum_{i=0}^{k-1}(h\beta_i Q_z \Delta z_{n+i} - \alpha_i S Q_y \Delta y_{n+i}) \\ &+ O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|),) \end{aligned} \tag{3.8c}$$

In the upper equations, P_y, Q_y, P_z, Q_z are at $(\hat{y}_{n+k}, \hat{z}_{n+k}, \hat{u}_{n+k})$. But because of smoothness, $P_y(\hat{y}_{n+k}) - P_y(\hat{y}_{n_i}) = O(h)$ for $0 \le i \le k-1$, and likewise for Q_y, P_z, Q_z. Denote

$$A_1 = \begin{pmatrix} \alpha_{k-1} & \cdots & \alpha_1 & \alpha_0 \\ 1 & & 0 & 0 \\ & \ddots & \vdots & \vdots \\ & & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} \alpha_{k-1} & \cdots & \alpha_1 & \alpha_0 \\ 1 & & 0 & 0 \\ & \ddots & \vdots & \vdots \\ & & 1 & 0 \end{pmatrix},$$

$$N = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ 1 & & 0 & 0 \\ & \ddots & \vdots & \vdots \\ & & 1 & 0 \end{pmatrix},$$

and

$$B_1 = \begin{pmatrix} -\beta'_{k-1} & \cdots & -\beta'_1 & -\beta'_0 \\ 1 & & 0 & 0 \\ & \ddots & \vdots & \vdots \\ & & 1 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -b_{k-1} & \cdots & -b'_1 & -b'_0 \\ 1 & & 0 & 0 \\ & \ddots & \vdots & \vdots \\ & & 1 & 0 \end{pmatrix},$$

where $\beta'_i = \dfrac{\beta_i}{\beta_k}$, $b'_i = \dfrac{b_i}{b_k}$. Then from (3.5), (3.8) and (3.4), we have

$$\begin{aligned} \Delta Y^1_{n+1} =& (A_1 \otimes I)\Delta Y^1_n + O(1)h\Delta Z^1_n \\ &+ O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|), \end{aligned} \tag{3.9a}$$

$$\Delta Y^2_{n+1} = (N \otimes I)\Delta Y^2_n + O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|), \tag{3.9b}$$

$$h\Delta Z^1_{n+1} = (A_2 \otimes I)h\Delta Z^1_n + O(h^2\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|), \tag{3.9c}$$

$$\begin{aligned} h\Delta Z^2_{n+1} =& (B_1 \otimes I)h\Delta Z^2_n + O(1)\Delta Y^2_n \\ &+ O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|), \end{aligned} \tag{3.9d}$$

$$h^2\Delta U_{n+1} =(B_2\otimes I)h^2\Delta U_n + O(1)\Delta Y_n^2 + O(1)h\Delta Z_n^2 + O(h\|\Delta Y_n\| + h^2\|\Delta Z_n\| + h^3\|\Delta U_n\|), \tag{3.9e}$$

According to [6, Lemma III.4.4, p.343], we choose norms such that $\|A_1\otimes I\|_{A_1}\le 1$, $\|A_2\otimes I\|_{A_2}\le 1$. And because $9\rho_1,\sigma_1)$, (ρ_2,σ_2) are both strictly stable, we can also choose norms such that $\|B_1\otimes I\|_{B_1}\le \kappa_1<1$, $\|B_2\otimes I\|_{B_2}\le\kappa_2<1$. The norm $\|\cdot\|_N$ is also chosen to satisfy $\|N\otimes I\|_N\le\varrho<1$. Then

$$\begin{pmatrix} h\|\Delta Z_{n+1}^1\|_{A_2}\\ \|\Delta Y_{n+1}^1\|_{A_1}\\ \|\Delta Y_{n+1}^2\|_N\\ h\|\Delta Z_{n+1}^2\|_{B_1}\\ h^2\|\Delta U_{n+1}\|_{B_2}\end{pmatrix} = \begin{pmatrix} 1+O(h) & O(h^2) & O(h) & O(h) & O(h)\\ O(1) & 1+O(h) & O(h) & O(h) & O(h)\\ O(h) & O(h) & \varrho+O(h) & O(h) & O(h)\\ O(h) & O(h) & O(1) & \kappa_1+O(h) & O(h)\\ O(h) & O(h) & O(1) & O(1) & \kappa_2+O(h)\end{pmatrix} \cdot\begin{pmatrix} h\|\Delta Z_{n+1}^1\|_{A_2}\\ \|\Delta Y_n^1\|_{A_1}\\ \|\Delta Y_n^2\|_N\\ h\|\Delta Z_n^2\|_{B_1}\\ h^2\|\Delta U_n\|_{B_2}\end{pmatrix}. \tag{3.10}$$

Denote the iterative matrix by W. Then W is power-bounded by

$$W^n = \begin{pmatrix} O(1) & O(h) & O(h) & O(h) & O(h)\\ O(n) & O(1) & O(1) & O(1) & O(1)\\ O(1) & O(h) & O(1) & O(h) & O(h)\\ O(1) & O(h) & O(1) & O(1) & O(h)\\ O(1) & O(h) & O(1) & O(1) & O(1)\end{pmatrix}, \tag{3.11}$$

and now if $nh\le Const$, from the initial conditions, we get

$$\|\Delta y_n\| = O(h^{p+1}+h^{q+1}),\ \|G_y(\hat y_{n+1})\Delta y_{n+j}\| = O(h^{p+1}+h^{q+1}),$$
$$\|\Delta z_n\| = O(h^{p+1}+h^{q+1}),\ \|G_yF_z(\hat y_{n+k},\hat z_{n+k})\Delta z_{n+j}\| = O(h^p+h^q)$$

Summing up the propagated errors as in [7, p.496], we get

$$\|y_n - y(x_n)\| \le \sum_{i=0}^{n-k+1}\|y_n^l - y_n^{l=1}\| = O(h^p+h^q)$$
$$\|z_n - z(x_n)\| \le \sum_{i=0}^{n-k+1}\|z_n^l - z_n^{l+1}\| = O(h^p+h^q)$$

And applying Corollary 2.1 with $\hat y_i = y(x_i)$, $\hat z_i = z(x_i)$, $\hat u_i = u(x_i)$, $\delta = O(h^p)$, $\theta = O(h^q)$ and $\mu = 0$ yields

$$\|u_n - u(x_n)\| \le \frac{C}{h^2}\sum_{j=1}^{k}(G_y(y(x_n))(y_{n-j}-y(x_{n-j})) + hG_yF_z(y(x_n),z(x_n))(z_{n-j}-z(x_{n-j})) + O(h^{p+1}) + O(h^{q+2}))$$

$$=O(h^{p-1}+h^q)$$

Remark 3.1. The ∞-stable requirement enable the application of Adams-Bashforth methods and BDF methods. But unfortunately, all Adams-Moulton methods except the trapeziodal rule are excluded.

Remark 3.2. The estimation of W in (3.11) is different from that in [5]. In fact, there is an error in [5, Lemma 6.5, p.82], where

$$V=\begin{pmatrix} 1+O(h) & O(1) \\ O(h) & 1+O(h) \end{pmatrix}$$

was considered be power-bounded by

$$V^n=\begin{pmatrix} O(1) & O(n) \\ O(1) & O(1) \end{pmatrix}.$$

It can be proved wrong. Another estimation like (3.4) must be added so that $O(h^2)$ appears in

$$V=\begin{pmatrix} 1+O(h) & O(1) \\ O(h^2) & 1+O(h) \end{pmatrix}.$$

Then the desired power-bound can be derived.

4. Numerical Tests

Numerical tests are carried out with the following two index-3 problems:

Problem 4.1.

$$\begin{aligned} &y_1'=2y_1y_2z_1z_2, \quad y_2'=-y_1y_2z_2^2, \\ &z_1'=(y_1y_2+z_1z_2)u, \quad z_2'=-y_1y_2^2z_2^2u, \\ &0=y_1y_2^2-1. \end{aligned} \tag{4.1}$$

where K is linear in u. For the consistent initial values $y_0=(1,1)^T$, $z_0=(1,1)^T$ and $u_0=1$ the exact solution is

$$y_1(x)=z_1(x)=e^{2x},\ y_2(x)=z_2(x)=e^{-x},\ u(x)=e^x. \tag{4.2}$$

Problem 4.2.

$$\begin{aligned} &y_1'=2y_1y_2z_1z_2, \quad y_2'=-y_1y_2z_2^2, \\ &z_1'=(y_1y_2+z_1z_2)u, \quad z_2'=-y_1y_2^2z_2^3u^2, \\ &0=y_1y_2^2-1. \end{aligned} \tag{4.3}$$

wher K is nonlinear in u. The solution is the same as in Problem 4.1.

Different formula-pairs are used to discrete the above two problems. We integrate the equations with stepsize $h=0.1$ initially. And change it to $\frac{h}{2},\frac{h}{4},\cdots$, then the corresponding convergence order can be approximated by $\frac{\log \|error\|}{\log h}$. these results are listed in Table 1, which confirm our convergence results in §3. To our surprise, the convergence order of u is not $(p-1,q)$ as in Theorem 3.1 but (p,q). This difference implies that there should be deeper estimation for the convergence order of u. To Problem 4.2, the convergence order remains the same as that to Problem 4.1. But it is still an open question that what conditions should be added so that when K is nonlinear in u, the existence of a solution can be assured.

Table 1 Convergence order for different formula-pairs

Type	Methods	Problem 4.1			Problem 4.2		
		y	z	u	y	z	u
I	BDF-3/BDF-3	3	3	3	3	3	3
I	ADM-3/ADM-3	*	*	*	*	*	*
II	BDF-4/ADB-2	2	2	2	2	2	2
II	ADM-3/ADB-3	*	*	*	*	*	*
III	ADB-2/BDF-4	2	2	2	2	2	2
IV	ADB-3/ADB-3	3	3	3	3	3	3

"*": divergence, "BDF-k": k-step BDF method,
"ADB-k": k-step A dams-Bashforth method,
"ADM-k": k-step A dams-Moulton method.

References

[1] C. Arévalo, G. Söderlind, Convergence of multistep discretizations of DAEs, *BIT,* **35** (1995), 143–168.

[2] K.E. Brenan, S.L. Campbell, L.R. Petzold, Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations, North-Holland, 1989.

[3] K.E. Brenan, B.E. Engquist, Backward differentiation approximations of nonlinear differential-algebraic systems, *Mathematics of Computation,* **51**:184 (1988), 659–676.

[4] Y. Cao, Q. Li, Highest order of multistep formula for solving index-2 DAEs, *BIT,* 38 (1998), 663–673.

[5] E. Hairer, C. Lubich, M. Roche, The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods, Lecture Notes in mathematics, 1409, Spring-Verlag, 1989.

[6] E. Hairer, G. Wanner, Solving Ordinary Differential Equations I. Notiff problems, Springer series in Computational Mathematics, 8, Springer-Verlag, New York, 1987.

[7] E. hairer, G. Wanner, Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems, Springer series in Computational Mathematics, 14, Springer-Verlag, New York, 1991.

[8] L. jay, Collocation methods for differential-algebraic equations of index 3, *Numer. Math.,* **65** (1993),, 407–421.

[9] L. Jay, Convergence of Runge-Kutta methods for differential-algebraic systems of index 3, *Applied Numer. Math.,* **17** (1995), 97–118.