

STABILITY AND B -CONVERGENCE PROPERTIES OF MULTISTEP RUNGE-KUTTA METHODS

SHOUFU LI

ABSTRACT. This paper continues earlier work by the same author concerning the stability and B -convergence properties of multistep Runge-Kutta methods for the numerical solution of nonlinear stiff initial-value problems in a Hilbert space. A series of sufficient conditions and necessary conditions for a multistep Runge-Kutta method to be algebraically stable, diagonally stable, B - or optimally B -convergent are established, by means of which six classes of high order algebraically stable and B -convergent multistep Runge-Kutta methods are constructed in a unified pattern. These methods include the class constructed by Burrage in 1987 as special case, and most of them can be regarded as extension of the Gauss, RadauIA, RadauIIA and LobattoIIIC Runge-Kutta methods. We find that the classes of multistep Runge-Kutta methods constructed in the present paper are superior in many respects to the corresponding existing one-step Runge-Kutta schemes.

1. INTRODUCTION

During the past twenty years and more, considerable progress has taken place in the stability and convergence theory for discretizations of nonlinear stiff initial value problems. In 1975, Dahlquist [9] was among the first to introduce the concept of one-sided Lipschitz continuity into the analysis of numerical methods for stiff systems and the concept of G -stability for one-leg and linear multistep methods. In the same year, Butcher [7] developed the theory of B -stability for Runge-Kutta methods. To unify and extend these results, Burrage and Butcher [6] presented the concept of monotonicity and established the algebraic stability criterion for general linear methods in 1980. Using one-sided Lipschitz continuity and B -stability theory as bases, Frank, Schneid and Ueberhuber [10, 11, 12] introduced the concept of B -convergence in 1981, and established B -theory for Runge-Kutta methods in 1984. The author of the present paper [14, 15, 16, 17] developed B -theory for general linear methods in 1988, also based on the one-sided Lipschitz continuity, and established much more extensive B -theory for nonlinear general multivalued methods for stiff problems in Banach spaces in 1990, based on the characteristic vectors of the problems considered. Note that the characteristic vector is a new concept much more general than that of one-sided Lipschitz constant (cf. [16, 17]); however, it is beyond the scope of this paper and we shall not explain it in detail. B -theory based

Received by the editor September 21, 1995 and, in revised form, May 12, 1998 and November 4, 1998.

1991 *Mathematics Subject Classification.* Primary 65L05; Secondary 65J99.

Key words and phrases. Numerical analysis, nonlinear stability, B -convergence, multistep Runge-Kutta methods.

The project supported by National Natural Science Foundation of China.

on one-sided Lipschitz continuity still suffers from considerable restrictions, since it is only suitable for stiff problems satisfying a one-sided Lipschitz condition with a one-sided Lipschitz constant not strongly positive. Since 1990, Auzinger, Frank and Kirlinger have published a series of papers [1, 2, 3, 4] to this topic and have given an important extension of B -convergence theory for Runge-Kutta methods. It turns out that B -theory based on one-sided Lipschitz continuity is only a special case of modern quantitative convergence theory (cf. [3]), such as B -theory based on the characteristic vector mentioned above; however, it is still of great importance as a guideline to select methods for the solution of many nonlinear or nonautonomous stiff problems, such as stiff dissipative systems arising after initial/boundary value problems in certain partial differential equations have been discretized in space.

Multistep Runge-Kutta methods are an important subclass of general linear methods which includes many commonly used methods, such as linear multistep methods, one-leg methods, Runge-Kutta methods, hybrid methods and multistep collocation methods, and also many new classes of methods which are still not investigated. It had been thought that the algebraic stability, as well as B -stability and B -convergence, of general linear methods, especially of multistep Runge-Kutta methods, was too difficult to study in a rigorous way because of the necessity of finding a stability matrix which should be nonnegative definite. However, in 1987, Burrage [5] constructed a class of high order multistep Runge-Kutta methods and proved it to be algebraically stable. After that the present author [20, 21] studied this class of methods further and proved that most of them are also B -convergent, and some other classes of algebraically stable multistep Runge-Kutta methods of high order were also constructed. In the present paper, we continue these works. In Section 2, a series of sufficient and necessary conditions for a multistep Runge-Kutta method to be algebraically stable, diagonally stable, B - or optimally B -convergent are established, by means of which, six classes of high order multistep Runge-Kutta methods are presented and proved to be all algebraically stable and B -convergent. They include the aforementioned classes as special cases, and most of them can be regarded as extension of the Gauss, RadauIA, RadauIIA and LobattoIIIC Runge-Kutta methods. In Section 3, we give a general approach for the construction of the six classes of multistep Runge-Kutta methods mentioned above, and a series of examples and related results. We point out at the end of this paper that the classes of multistep Runge-Kutta methods constructed here are superior in many respects to the corresponding existing one-step Runge-Kutta schemes.

Consider the multistep Runge-Kutta method

$$(1.1) \quad \begin{cases} Y^{(n)} = hC_{11}F(t_n, Y^{(n)}) + C_{12}y^{(n-1)}, \\ y^{(n)} = hC_{21}F(t_n, Y^{(n)}) + C_{22}y^{(n-1)}, \xi_n = \beta y^{(n)} \end{cases}$$

for solving the initial value problem

$$(1.2) \quad \begin{cases} y'(t) = f(t, y(t)), & t \in [0, T], \\ y(0) = y_0, & y_0 \in X, \end{cases}$$

where X is a real or complex Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and the corresponding norm $\| \cdot \|$, $f : [0, T] \times X \rightarrow X$ is a sufficiently smooth mapping satisfying a one-sided Lipschitz condition

$$\operatorname{Re}\langle u - v, f(t, u) - f(t, v) \rangle \leq \nu \| u - v \|^2, \quad u, v \in X, t \in [0, T],$$

and the problem (1.2) is assumed to have a unique solution $y(t)$ on the interval $[0, T]$. For the method (1.1) we assume that $h > 0$ is the fixed stepsize,

$$Y^{(n)} := (Y_1^{(n)}, Y_2^{(n)}, \dots, Y_s^{(n)}) \in X^s, \quad y^{(n)} := (y_1^{(n)}, y_2^{(n)}, \dots, y_r^{(n)}) \in X^r, \quad \xi_n \in X$$

are approximations to

$$Y(t_n) := (y(t_n + \mu_1 h), y(t_n + \mu_2 h), \dots, y(t_n + \mu_s h)) \in X^s,$$

$$H(t_n) := (y(t_n + h), y(t_n + 2h), \dots, y(t_n + rh)) \in X^r, \quad y(t_n + rh) \in X$$

respectively,

$$F(t_n, Y^{(n)}) := (f(t_n + \mu_1 h, Y_1^{(n)}), f(t_n + \mu_2 h, Y_2^{(n)}), \dots, f(t_n + \mu_s h, Y_s^{(n)})) \in X^s,$$

$C_{11}, C_{12}, C_{21}, C_{22}$ and β are linear mappings corresponding to the real matrices

$$(1.3) \quad \begin{cases} C_{11} = [C_{ij}^{11}] \in \mathbf{R}^{s \times s}, & C_{12} = [C_{ij}^{12}] \in \mathbf{R}^{s \times r}, & C_{21} = \begin{bmatrix} 0 \\ \gamma^T \end{bmatrix} \in \mathbf{R}^{r \times s}, \\ C_{22} = \begin{bmatrix} 0 & I_{r-1} \\ \alpha^T & \end{bmatrix} \in \mathbf{R}^{r \times r}, & \beta = [0, \dots, 0, 1] \in \mathbf{R}^{1 \times r}, \end{cases}$$

respectively (cf. [15]), where $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_r]^T$, $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_s]^T$, I_m ($m \geq 1$) denotes the $m \times m$ identity matrix, $t_n = t_0 + nh$, t_0 and μ_i ($i = 1, 2, \dots, s$) are real constants chosen appropriately. Furthermore, throughout this paper we always assume that

$$\begin{cases} (1.4a) & \sum_{j=1}^r \alpha_j = 1, \quad \sum_{j=1}^r C_{ij}^{12} = 1, i = 1, 2, \dots, s; \\ (1.4b) & \alpha_1 > 0, \quad \alpha_j \geq 0, j = 2, 3, \dots, r; \\ (1.4c) & \mu_i \neq \mu_j \quad \text{whenever } i \neq j, \end{cases}$$

where the relation (1.4a) is called the preconsistency condition. For simplicity, write $\mu = [\mu_1, \mu_2, \dots, \mu_s]^T$, $\zeta = [0, 1, \dots, r - 1]^T$, introduce the simplifying conditions (cf. [5])

$$\begin{aligned} B(\tau) : b_i &:= i\gamma^T \mu^{i-1} - r^i + \alpha^T \zeta^i = 0, & i = 1, 2, \dots, \tau; \\ C(\tau) : c_i &:= iC_{11} \mu^{i-1} - \mu^i + C_{12} \zeta^i = 0, & i = 1, 2, \dots, \tau; \\ D(\tau) : d_i &:= iC_{11}^T \text{diag}(\gamma) \mu^{i-1} - r^i \gamma + \text{diag}(\gamma) \mu^i = 0, & i = 1, 2, \dots, \tau; \\ E(\tau) : e_i &:= iC_{12}^T \text{diag}(\gamma) \mu^{i-1} - r^i \alpha + \text{diag}(\alpha) \zeta^i = 0, & i = 1, 2, \dots, \tau, \end{aligned}$$

and make the following notational conventions.

- (1) For any vector $x = [x_1, x_2, \dots, x_N]^T$, $x^i = [x_1^i, x_2^i, \dots, x_N^i]^T$.
- (2) For a real vector $x = [x_1, x_2, \dots, x_N]^T$, $x \geq 0$ (resp. > 0) means that all the elements $x_i \geq 0$ (resp. > 0), $i = 1, 2, \dots, N$.
- (3) For a real symmetric matrix M , $M \geq 0$ (resp. $M > 0$) means that M is nonnegative definite (resp. positive definite).
- (4) For $\tau = 0$ the simplifying conditions $B(\tau), C(\tau), D(\tau)$ and $E(\tau)$ do not, by definition, impose any restriction on the method.
- (5) We shall frequently make use of the following quantities in the remainder of this paper:

$$\begin{aligned} G &:= \text{diag}(\alpha_1, \alpha_1 + \alpha_2, \dots, \sum_{j=1}^r \alpha_j), & Q &:= \text{diag}(\gamma), \\ C &:= [c_1, c_2, \dots, c_s], & D &:= [d_1, d_2, \dots, d_s], \\ U &:= [u_1, u_2, \dots, u_s], & V &:= [v_1, v_2, \dots, v_s], \\ u_i &:= \zeta^i, & v_i &:= i\mu^{i-1}, & i &= 1, 2, 3, \dots, \\ \hat{e}_m &:= [1, 1, \dots, 1]^T \in \mathbf{R}^m \quad (m \geq 1), \end{aligned}$$

$$C_p := \begin{cases} [c_1, c_2, \dots, c_p] \in \mathbf{R}^{s \times p} & \text{for } p \geq 0, \\ 0 \in \mathbf{R}^{s \times 1} & \text{for } p=0, \end{cases}$$

$$\widehat{C}_p := \begin{cases} [c_{p+1}, c_{p+2}, \dots, c_s] \in \mathbf{R}^{s \times (s-p)} & \text{for } 0 \leq p < s, \\ 0 \in \mathbf{R}^{s \times 1} & \text{for } p = s, \end{cases}$$

D_p, E_p, U_p and V_p are defined in a similar way as C_p , and $\widehat{D}_p, \widehat{U}_p$ and \widehat{V}_p in a similar way as \widehat{C}_p .

For the convenience of the reader, we now recall some definitions and results which have been presented in [15, 17, 20] and will be used later.

Definition 1.1. Let W denote a real symmetric, positive definite $r \times r$ matrix, B a nonnegative diagonal $s \times s$ matrix. The method (1.1) is said to be *algebraically stable* for the matrices W, B if the corresponding matrix

$$\begin{bmatrix} W - C_{22}^T W C_{22} & C_{12}^T B - C_{22}^T W C_{21} \\ BC_{12} - C_{21}^T W C_{22} & C_{11}^T B + BC_{11} - C_{21}^T W C_{21} \end{bmatrix} \geq 0.$$

With Definition 1.1 we see that the method (1.1) is algebraically stable for the matrices G, Q if $G > 0, Q \geq 0$ and the matrix

(1.5)

$$M := \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \geq 0,$$

where

$$(1.6) \quad \begin{aligned} M_{11} &= G - C_{22}^T G C_{22}, \quad M_{12} = M_{21}^T = C_{12}^T Q - C_{22}^T G C_{21}, \\ M_{22} &= C_{11}^T Q + Q C_{11} - C_{21}^T G C_{21}. \end{aligned}$$

Definition 1.2. The method (1.1) is said to be *diagonally stable* if there exists an $s \times s$ diagonal matrix $B > 0$ such that $C_{11}^T B + BC_{11} > 0$.

Definition 1.3. The method (1.1) is said to be *B-stable* if there exist a real number $\delta > 0$, a nonnegative bounded function $\phi : (0, \delta] \rightarrow \mathbf{R}$, and a real symmetric, positive definite $r \times r$ matrix $W = [w_{ij}]$, where δ, ϕ and W depend only on the method, such that for any given problem (1.2) and any two parallel calculating steps $(t_{n-1}, y^{(n-1)}) \rightarrow (t_n, Y^{(n)}, y^{(n)}, \xi_n)$ and $(t_{n-1}, z^{(n-1)}) \rightarrow (t_n, Z^{(n)}, z^{(n)}, \zeta_n)$ we have

$$\| y^{(n)} - z^{(n)} \|_W \leq (1 + h\nu\phi(h\nu))^{\frac{1}{2}} \| y^{(n-1)} - z^{(n-1)} \|_W, \quad 0 < h\nu \leq \delta,$$

where $\| \cdot \|_W$ denotes a norm on X^r defined by

$$\| U \|_W = \langle U, WU \rangle^{\frac{1}{2}} = \left(\sum_{i,j=1}^r w_{i,j} \langle u_i, u_j \rangle \right)^{\frac{1}{2}}, \quad U = (u_1, u_2, \dots, u_r) \in X^r,$$

where each $u_j \in X$.

Definition 1.4. The method (1.1) is said to be *optimally B-convergent of order p* if the approximation sequences $\{y^{(n)}\}$ and $\{\xi_n\}$ (produced by the method (1.1) applied to any given problem (1.2) with starting value $y^{(0)}$) satisfy

$$\| \xi_n - y(t_n + rh) \| \leq \tilde{C}_0(t_n) \| y^{(0)} - H(t_0) \| + \tilde{C}_1(t_n) h^p, \quad 0 < h \leq h_0,$$

and

$$\| y^{(n)} - H(t_n) \| \leq \tilde{C}_0(t_n) \| y^{(0)} - H(t_0) \| + \tilde{C}_1(t_n) h^p, \quad 0 < h \leq h_0,$$

where the functions $\bar{C}_i(t)$ and $\tilde{C}_i(t), i = 0, 1$, depend only on the method, the one-sided Lipschitz constant ν of the right-hand side function $f(t, y)$ and some bounds M_i for the true solution $y(t)$:

$$\left\| \frac{d^i y(t)}{dt^i} \right\| \leq M_i, \quad t \in [0, T];$$

the maximum stepsize h_0 depends only on ν and the method. Here and later, the norm $\| \cdot \|$ on $X^N (N \geq 1)$ is defined by

$$\| U \| = \left(\sum_{i=1}^N \| u_i \|^2 \right)^{\frac{1}{2}} \quad \forall U = (u_1, u_2, \dots, u_N) \in X^N.$$

Furthermore, the method (1.1) is said to be *B-convergent of order p* if $\bar{C}_i(t), \tilde{C}_i(t)$ and h_0 are allowed to depend, in addition to the quantities mentioned above, on bounds κ_{ij} for certain derivatives of the right-hand side $f(t, y)$ (but not on $\kappa_{0,1}$):

$$\left\| \frac{\partial^{i+j} f(t, y)}{\partial t^i \partial y^j} \right\| \leq \kappa_{ij}, \quad t \in [0, T], \quad y \in X.$$

It should be pointed out that in most early published papers, such as [10, 12, 14, 15], the starting perturbation $y^{(0)} - H(t_0)$ was not considered in the definition of *B-convergence*.

Definition 1.5. The method (1.1) is said to have *generalized stage order p* if p is the largest nonnegative integer which possesses the following properties:

For any given initial value problem (1.2) and stepsize $h \in (0, h_0]$, there exist mappings Y^h and H^h (from some subinterval of $[0, T]$ into X^s and X^r respectively),

$$\begin{aligned} Y^h(t) &= (Y_1^h(t), Y_2^h(t), \dots, Y_s^h(t)) \in X^s, \\ H^h(t) &= (H_1^h(t), H_2^h(t), \dots, H_r^h(t)) \in X^r, \end{aligned}$$

such that

$$\begin{aligned} \| H^h(t) - H(t) \| &\leq d_0 h^p, & \| \Delta^h(t) \| &\leq d_1 h^{p+1}, \\ \| \delta^h(t) \| &\leq d_2 h^{p+1}, & \| \sigma^h(t) \| &\leq d_3 h^p, \end{aligned}$$

where $h_0 > 0$ is only required to be so small that for $h \in (0, h_0]$ all the time nodes belong to the integration interval $[0, T]$; each $d_i (i = 0, 1, 2, 3)$ depends only on the method and bounds M_i for certain derivatives of the exact solution $y(t)$; $\Delta^h(t), \delta^h(t)$ and $\sigma^h(t)$ are determined by the equations

$$\begin{cases} Y^h(t) = hC_{11}F(t, Y^h(t)) + C_{12}H^h(t - h) + \Delta^h(t), \\ H^h(t) = hC_{21}F(t, Y^h(t)) + C_{22}H^h(t - h) + \delta^h(t), \\ y(t + rh) = \beta H^h(t) + \sigma^h(t). \end{cases}$$

Furthermore, if the quantities $d_i (i = 0, 1, 2, 3)$ are also allowed to depend on bounds κ_{ij} for certain partial derivatives of the mapping f (but not on $\kappa_{0,1}$), then the aforementioned integer p is known as *generalized weak stage order* of the method. For the special case where $H^h(t) = H(t)$, the generalized stage order and generalized weak stage order are simply called *stage order* and *weak stage order*, respectively.

Theorem 1.1. *Suppose a multistep Runge-Kutta method of the form (1.1) is algebraically stable and diagonally stable. Then this method is B-stable, and furthermore, this method is optimally B-convergent (resp. B-convergent) of order p provided that it has generalized stage order p (resp. generalized weak stage order p).*

For the proof of Theorem 1.1 we refer to [15, 17].

2. MAIN RESULTS AND THEIR PROOFS

Lemma 2.1. *Let*

$$(2.1) \quad S = \begin{bmatrix} I_r & 0 \\ 0 & V \end{bmatrix}, \quad Y = [I_r, -U], \quad Z = [z_{ij}] = (D + \text{diag}(\gamma)C)^T V.$$

Then $B(s)$ and $E(s)$ yield

$$(2.2a) \quad V^T M_{22} V = U^T M_{11} U + Z,$$

$$(2.2b) \quad S^T M S = Y^T M_{11} Y + \begin{bmatrix} 0 & 0 \\ 0 & Z \end{bmatrix},$$

where the matrices M , M_{11} and M_{22} are defined by (1.5) and (1.6).

Proof. From (1.3), (1.5), (1.6) and (2.1) we get

$$(2.3) \quad \begin{cases} M_{11} = \text{diag}(\alpha) - \alpha\alpha^T, \\ M_{12} = M_{21}^T = C_{12}^T \text{diag}(\gamma) - \alpha\gamma^T, \\ M_{22} = C_{11}^T \text{diag}(\gamma) + \text{diag}(\gamma)C_{11} - \gamma\gamma^T, \\ S^T M S = \begin{bmatrix} M_{11} & W \\ W^T & R \end{bmatrix} \end{cases}$$

with

$$(2.4) \quad \begin{cases} W = [w_1, w_2, \dots, w_s] = M_{12} V, \\ R = [r_1, r_2, \dots, r_s] = V^T M_{22} V. \end{cases}$$

Because of $B(s)$ and $E(s)$, (2.4) and (2.3) lead to

$$\begin{aligned} w_i &= r^i \alpha - \text{diag}(\alpha)\zeta^i - \alpha(r^i - \alpha^T \zeta^i) \\ &= -(\text{diag}(\alpha) - \alpha\alpha^T)\zeta^i = -M_{11}\zeta^i, \end{aligned}$$

$$\begin{aligned} r_i &= V^T [iC_{11}^T \text{diag}(\gamma)\mu^{i-1} + i\text{diag}(\gamma)C_{11}\mu^{i-1} - \gamma(r^i - \alpha^T \zeta^i)] \\ &= V^T (d_i + \text{diag}(\gamma)c_i - M_{12}^T \zeta^i), \end{aligned}$$

and therefore

$$(2.5) \quad \begin{cases} W = -M_{11}U, \\ R = V^T (D + \text{diag}(\gamma)C - M_{12}^T U) = -U^T W + Z = U^T M_{11}U + Z. \end{cases}$$

Thus the conclusion follows from (2.4), (2.5) and substitution of (2.5) into the last equality of (2.3). \square

Lemma 2.2. *The (i, j) -elements of the matrix Z defined by (2.1) can be expressed by*

$$(2.6) \quad \begin{aligned} z_{ij} &= (d_i^T + c_i^T \text{diag}(\gamma))j\mu^{j-1} \\ &= c_i^T \text{diag}(\gamma)j\mu^{j-1} + c_j^T \text{diag}(\gamma)i\mu^{i-1} + b_{i+j} - b_j r^i - e_i^T \zeta^j \\ &= d_i^T j\mu^{j-1} + d_j^T i\mu^{i-1} - b_{i+j} + b_i r^j + e_j^T \zeta^i, \quad i, j = 1, 2, \dots, s. \end{aligned}$$

In particular,

$$(2.7) \quad \begin{aligned} z_{ii} &= (d_i^T + c_i^T \text{diag}(\gamma))i\mu^{i-1} = 2c_i^T \text{diag}(\gamma)i\mu^{i-1} + b_{2i} - b_i r^i - e_i^T \zeta^i \\ &= 2d_i^T i\mu^{i-1} - b_{2i} + b_i r^i + e_i^T \zeta^i, \quad i = 1, 2, \dots, s. \end{aligned}$$

Furthermore, if $B(s)$ and $E(s)$ hold, then

$$(2.8) \quad \begin{aligned} z_{ij} &= c_i^T \text{diag}(\gamma) j \mu^{j-1} + c_j^T \text{diag}(\gamma) i \mu^{i-1} + b_{i+j} \\ &= d_i^T j \mu^{j-1} + d_j^T i \mu^{i-1} - b_{i+j}, \quad i, j = 1, 2, \dots, s, \end{aligned}$$

$$(2.9) \quad z_{ii} = 2c_i^T \text{diag}(\gamma) i \mu^{i-1} + b_{2i} = 2d_i^T i \mu^{i-1} - b_{2i}, \quad i = 1, 2, \dots, s.$$

Proof. Since

$$\begin{aligned} d_j^T i \mu^{i-1} &= (j \mu^{j-1})^T \text{diag}(\gamma) C_{11} i \mu^{i-1} - r^j \gamma^T i \mu^{i-1} + \gamma^T i \mu^{i+j-1} \\ &= (j \mu^{j-1})^T \text{diag}(\gamma) (c_i + \mu^i - C_{12} \zeta^i) - r^j (b_i + r^i - \alpha^T \zeta^i) \\ &\quad + i \gamma^T \mu^{i+j-1} \\ &= (i+j) \gamma^T \mu^{i+j-1} - (j C_{12}^T \text{diag}(\gamma) \mu^{j-1})^T \zeta^i - r^{i+j} + r^j \alpha^T \zeta^i \\ &\quad + (j \mu^{j-1})^T \text{diag}(\gamma) c_i - r^j b_i \\ &= b_{i+j} - r^j b_i + c_i^T \text{diag}(\gamma) j \mu^{j-1} - e_j^T \zeta^i, \end{aligned}$$

we have

$$(2.10) \quad b_{i+j} = r^j b_i - c_i^T \text{diag}(\gamma) j \mu^{j-1} + d_j^T i \mu^{i-1} + e_j^T \zeta^i, \quad i, j = 1, 2, 3, \dots$$

Thus the conclusion follows directly from (2.1) and (2.10). \square

Lemma 2.3. *Let p, q be positive integers. Then the following implications hold:*

$$\begin{aligned} &B(p), C(p), E(p) \text{ with } p \leq s \implies z_{ij} = b_{i+j}, \quad i, j = 1, 2, \dots, p; \\ &B(p), D(p), E(p) \text{ with } p \leq s \implies z_{ij} = -b_{i+j}, \quad i, j = 1, 2, \dots, p; \\ &C(p), D(p) \text{ with } p \leq s \implies z_{ij} = 0 \text{ for } i \leq p; \\ &B(s), C(p), D(p), E(s) \text{ with } p \leq s \implies z_{ij} = 0 \text{ for either } i \leq p \text{ or } j \leq p; \\ &B(p), C(p), D(q), E(q) \implies B(p+q); \\ &B(p+q), C(p), E(q), D_q^T \widehat{V}_p = 0 \text{ with } p \leq s \implies D(q); \\ &B(p+q), D(q), E(q), C_p^T \text{diag}(\gamma) \widehat{V}_q = 0 \text{ with } q \leq s \text{ and } \prod_{i=1}^s \gamma_i \neq 0 \implies C(p); \\ &B(p+q), C(p), D(q), E_q^T \widehat{U}_p = 0 \text{ with } p \leq s \text{ and } r \leq s+1 \implies E(q). \end{aligned}$$

Proof. All the results in Lemma 2.3, except the last one, can be derived directly from (2.6) and (2.10). To prove the last statement, we note that the equality (2.10) together with $B(p+q)$, $C(p)$, $D(q)$, $E_q^T \widehat{U}_p = 0$ and $p \leq s$ leads to

$$U^T e_j = 0, \quad j = 1, 2, \dots, q,$$

the preconsistency condition (1.4a) together with $B(q)$ leads to

$$\widehat{e}_r^T e_j = 0, \quad j = 1, 2, \dots, q,$$

and therefore

$$(2.11) \quad [U, \widehat{e}_r]^T e_j = 0, \quad j = 1, 2, \dots, q.$$

Since $r \leq s+1$, $\text{rank}([U, \widehat{e}_r]) = r$. Thus (2.11) implies that $E(q)$ holds. \square

Remark 1. The following results can be directly derived from the last three implications in Lemma 2.3:

$$(2.12) \quad \begin{cases} B(s+q), C(s), E(q) \implies D(q); \\ B(s+p), D(s), E(s) \text{ with } \prod_{i=1}^s \gamma_i \neq 0 \implies C(p); \\ B(s+q), C(s), D(q) \text{ with } r \leq s+1 \implies E(q). \end{cases}$$

Theorem 2.1. *Suppose $B(s)$ and $E(s)$ hold. Then the method (1.1) is algebraically stable for the matrices G, Q if and only if $\gamma \geq 0$ and $Z \geq 0$. Here the $s \times s$ symmetric matrix Z is defined by (2.1).*

Proof. Since $G > 0$ is guaranteed by (1.4b), algebraic stability of the method (1.1) is equivalent to $\gamma \geq 0$ and $M \geq 0$ with M defined by (1.5) and (1.6), and we only need to prove that $M \geq 0$ is equivalent to $Z \geq 0$. According to Lemma 2.1, $B(s)$ and $E(s)$ yield the equality (2.2b). Since $\mu_1, \mu_2, \dots, \mu_s$ are distinct, the matrices V and S in (2.2a), (2.2b) are obviously nonsingular. Since $\sum_{j=1}^r \alpha_j = 1$ and $\alpha_j \geq 0, j = 1, 2, \dots, r$, it is easy to verify that $M_{11} \geq 0$. Thus, in view of (2.2b), $Z \geq 0$ leads to $S^T M S \geq 0$ and $M \geq 0$. For the converse, we assume that $M \geq 0$. For any given $\xi \in \mathbf{R}^s$, write $\eta = [(U\xi)^T, \xi^T]^T$; we thus have

$$\eta^T S^T M S \eta \geq 0, \quad Y \eta = 0,$$

and therefore together with (2.2b)

$$\xi^T Z \xi \geq 0.$$

This means that $Z \geq 0$ and completes the proof of Theorem 2.1. □

Theorem 2.2. *Suppose $B(s), E(s)$ and $C(p)$ with $1 \leq p \leq s$ hold. Then the following statements are equivalent:*

- (a) *the method (1.1) is algebraically stable for the matrices G, Q ;*
 - (b) *$D(p-1)$ holds, $\widehat{Z}_{p-1} \geq 0, \gamma \geq 0$;*
 - (c) *$B(2p-1)$ holds, $D_{p-1}^T \widehat{V}_p = 0, \widehat{Z}_{p-1} \geq 0, \gamma \geq 0$.*
- Here and later, we define for $0 \leq j < s$*

$$(2.13) \quad \widehat{Z}_j = \begin{bmatrix} z_{j+1,j+1} & z_{j+1,j+2} & \cdots & z_{j+1,s} \\ z_{j+2,j+1} & z_{j+2,j+2} & \cdots & z_{j+2,s} \\ \cdots & \cdots & \cdots & \cdots \\ z_{s,j+1} & z_{s,j+2} & \cdots & z_{s,s} \end{bmatrix}.$$

Proof. Since $B(s)$ and $E(s)$ hold, according to Theorem 2.1 algebraic stability of the method (1.1) (for G, Q) is equivalent to $\gamma \geq 0$ and $Z \geq 0$ with the symmetric matrix Z defined by (2.1). Thus for $p = 1$ the conclusion follows trivially, and without loss of generality we can now assume that $2 \leq p \leq s$, so that $B(s)$ and $C(p)$ yield

$$(2.14) \quad b_1 = b_2 = 0, \quad c_1 = c_2 = \cdots = c_p = 0.$$

In view of Lemma 2.3, $B(s)$, $E(s)$ and $C(p)$ lead to

$$(2.15) \quad Z = \begin{bmatrix} b_2 & b_3 & \dots & b_p & b_{p+1} & z_{1,p+1} & \dots & z_{1,s} \\ b_3 & b_4 & \dots & b_{p+1} & b_{p+2} & z_{2,p+1} & \dots & z_{2,s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_p & b_{p+1} & \dots & b_{2p-2} & b_{2p-1} & z_{p-1,p+1} & \dots & z_{p-1,s} \\ b_{p+1} & b_{p+2} & \dots & b_{2p-1} & b_{2p} & z_{p,p+1} & \dots & z_{p,s} \\ z_{p+1,1} & z_{p+1,2} & \dots & z_{p+1,p-1} & z_{p+1,p} & z_{p+1,p+1} & \dots & z_{p+1,s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ z_{s,1} & z_{s,2} & \dots & z_{s,p-1} & z_{s,p} & z_{s,p+1} & \dots & z_{s,s} \end{bmatrix},$$

and therefore, together with (2.14), $Z \geq 0$ is equivalent to

$$(2.16) \quad \begin{cases} \widehat{Z}_{p-1} \geq 0, \\ D_{p-1}^T V = \begin{bmatrix} b_2 & b_3 & \dots & b_{p+1} & z_{1,p+1} & \dots & z_{1,s} \\ b_3 & b_4 & \dots & b_{p+2} & z_{2,p+1} & \dots & z_{2,s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_p & b_{p+1} & \dots & b_{2p-1} & z_{p-1,p+1} & \dots & z_{p-1,s} \end{bmatrix} = 0. \end{cases}$$

This leads to the required conclusion, since it is easily seen that the statements (b), (c) in Theorem 2.2 are equivalent to (2.16) together with $\gamma \geq 0$, respectively. \square

Theorem 2.3. *Suppose $B(s)$, $E(s)$ and $D(p)$ with $1 \leq p \leq s$ hold. Then the following statements are equivalent:*

- (a) *the method (1.1) is algebraically stable for the matrices G , Q ;*
- (b) *$\text{diag}(\gamma)C_{p-1} = 0$, $\widehat{Z}_{p-1} \geq 0$, $\gamma \geq 0$;*
- (c) *$B(2p - 1)$ holds, $C_{p-1}^T \text{diag}(\gamma)\widehat{V}_p = 0$, $\widehat{Z}_{p-1} \geq 0$, $\gamma \geq 0$.*

Proof. In view of Theorem 2.1, the statement (a) is equivalent to $Z \geq 0$ and $\gamma \geq 0$ because of $B(s)$ and $E(s)$. Thus for $p = 1$ the conclusion follows trivially, and without loss of generality we can now assume that $2 \leq p \leq s$, so that $B(s)$ and $D(p)$ yield

$$(2.17) \quad b_1 = b_2 = 0, \quad d_1 = d_2 = \dots = d_p = 0.$$

By Lemma 2.3, $B(s)$, $E(s)$ and $D(p)$ lead to

$$(2.18) \quad Z = \begin{bmatrix} -b_2 & -b_3 & \dots & -b_p & -b_{p+1} & z_{1,p+1} & \dots & z_{1,s} \\ -b_3 & -b_4 & \dots & -b_{p+1} & -b_{p+2} & z_{2,p+1} & \dots & z_{2,s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -b_p & -b_{p+1} & \dots & -b_{2p-2} & -b_{2p-1} & z_{p-1,p+1} & \dots & z_{p-1,s} \\ -b_{p+1} & -b_{p+2} & \dots & -b_{2p-1} & -b_{2p} & z_{p,p+1} & \dots & z_{p,s} \\ z_{p+1,1} & z_{p+1,2} & \dots & z_{p+1,p-1} & z_{p+1,p} & z_{p+1,p+1} & \dots & z_{p+1,s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ z_{s,1} & z_{s,2} & \dots & z_{s,p-1} & z_{s,p} & z_{s,p+1} & \dots & z_{s,s} \end{bmatrix},$$

and therefore, together with (2.17), $Z \geq 0$ is equivalent to

$$(2.19) \quad \begin{cases} \widehat{Z}_{p-1} \geq 0, \\ C_{p-1}^T \text{diag}(\gamma)V \\ = \begin{bmatrix} -b_2 & -b_3 & \dots & -b_{p+1} & z_{1,p+1} & \dots & z_{1,s} \\ -b_3 & -b_4 & \dots & -b_{p+2} & z_{2,p+1} & \dots & z_{2,s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -b_p & -b_{p+1} & \dots & -b_{2p-1} & z_{p-1,p+1} & \dots & z_{p-1,s} \end{bmatrix} \end{cases} = 0.$$

This leads to the required conclusion, since it is easily seen that the statements (b), (c) in Theorem 2.3 are equivalent to (2.19) together with $\gamma \geq 0$, respectively. \square

For algebraically stable methods satisfying the assumptions of Theorem 2.2, we have $b_{2p} = z_{p,p} \geq 0$. Thus we can separate these methods into two classes, depending on whether $b_{2p} > 0$ or $b_{2p} = 0$. Similarly, algebraically stable methods satisfying the assumptions of Theorem 2.3 can also be separated into two classes, depending on whether $b_{2p} < 0$ or $b_{2p} = 0$. For the methods satisfying $b_{2p} = 0$ we have the following results:

Theorem 2.4. *Suppose $B(s)$, $E(s)$ and $C(p)$ with $1 \leq p \leq s$ hold. Then the following statements are equivalent:*

- (a) $b_{2p} = 0$, and the method (1.1) is algebraically stable for G, Q ;
- (b) $D(p)$ holds, $\gamma \geq 0$, and either $p = s$ or $\widehat{Z}_p \geq 0$;
- (c) $B(2p)$ holds, $D_p^T \widehat{V}_p = 0$, $\gamma \geq 0$, and either $p = s$ or $\widehat{Z}_p \geq 0$.

Proof. Following a similar line as in the proof of Theorem 2.2, we find that in this case the relation (2.15) also holds, and algebraic stability of the method (1.1) (for G, Q) is equivalent to $\gamma \geq 0$ and $Z \geq 0$. Therefore together with $B(s)$ and $C(p)$ it follows that the statement (a) in Theorem 2.4 is equivalent to

$$(2.20) \quad \begin{cases} D_p^T V = \begin{bmatrix} b_2 & b_3 & \dots & b_{p+1} & z_{1,p+1} & \dots & z_{1,s} \\ b_3 & b_4 & \dots & b_{p+2} & z_{2,p+1} & \dots & z_{2,s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ b_{p+1} & b_{p+2} & \dots & b_{2p} & z_{p,p+1} & \dots & z_{p,s} \end{bmatrix} = 0, \\ \gamma \geq 0, \\ \text{either } p = s \text{ or } \widehat{Z}_p \geq 0. \end{cases}$$

This leads to the required conclusion, since it is easily seen that statements (b), (c) in Theorem 2.4 are also equivalent to (2.20), respectively. \square

Theorem 2.5. *Suppose $B(s)$, $E(s)$ and $D(p)$ with $1 \leq p \leq s$ hold. Then the following statements are equivalent:*

- (a) $b_{2p} = 0$, and the method (1.1) is algebraically stable for G, Q ;
- (b) $\text{diag}(\gamma)C_p = 0$, $\gamma \geq 0$, and either $p = s$ or $\widehat{Z}_p \geq 0$;
- (c) $B(2p)$ holds, $C_p^T \text{diag}(\gamma)\widehat{V}_p = 0$, $\gamma \geq 0$, and either $p = s$ or $\widehat{Z}_p \geq 0$.

Proof. Following a similar line as in the proof of Theorem 2.3, we find that in this case the relation (2.18) also holds, and algebraic stability of the method (1.1) (for

G, Q) is equivalent to $\gamma \geq 0$ and $Z \geq 0$. Therefore together with $B(s)$ and $D(p)$ it follows that statement (a) in Theorem 2.5 is equivalent to

$$(2.21) \quad \begin{cases} C_p^T \text{diag}(\gamma)V \\ = \begin{bmatrix} -b_2 & -b_3 & \dots & -b_{p+1} & z_{1,p+1} & \dots & z_{1,s} \\ -b_3 & -b_4 & \dots & -b_{p+2} & z_{2,p+1} & \dots & z_{2,s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -b_{p+1} & -b_{p+2} & \dots & -b_{2p} & z_{p,p+1} & \dots & z_{p,s} \end{bmatrix} = 0, \\ \gamma \geq 0, \\ \text{either } p = s \text{ or } \widehat{Z}_p \geq 0. \end{cases}$$

This leads to the required conclusion, since it is easily seen that the statements (b), (c) in Theorem 2.5 are also equivalent to (2.21), respectively. \square

For the application of Theorems 2.1–2.5, we are now interested in investigating the following special classes of multistep Runge-Kutta methods given by (1.1) and (1.3) with coefficients satisfying (1.4a), (1.4b) and the following conditions:

Class 1:

$$(2.22) \quad B(s), E(s), C(s) \text{ and } D(s);$$

Class 2:

$$(2.23) \quad B(s), E(s), C(s) \text{ and } D(s - 1) \text{ with } b_{2s} > 0;$$

Class 3:

$$(2.24) \quad B(s), E(s), D(s) \text{ and } C(s - 1) \text{ with } b_{2s} < 0;$$

Class 4:

$$(2.25) \quad B(s), E(s), C(s - 1), D(s - 1), \gamma \geq 0, z_{s,s} \geq 0, |c_s||d_s| \neq 0;$$

Class 5:

$$(2.26) \quad B(s), E(s), C(s - 1), D(s - 2), \gamma \geq 0, \widehat{Z}_{s-2} \geq 0, b_{2s-2} \neq 0;$$

Class 6:

$$(2.27) \quad \begin{cases} B(s), E(s), D(s - 1), \text{diag}(\gamma)C_{s-2} = 0, \\ \gamma \geq 0, \widehat{Z}_{s-2} \geq 0, b_{2s-2} \neq 0. \end{cases}$$

Note that in view of Lemma 2.3 and Theorems 2.2–2.5, the condition (2.22) is equivalent to either of the following:

$$(2.28a) \quad B(2s), E(s) \text{ and } C(s),$$

$$(2.28b) \quad B(2s), E(s) \text{ and } D(s),$$

and if $r \leq s + 1$, also to

$$(2.28c) \quad B(2s), C(s) \text{ and } D(s),$$

the condition (2.23) is equivalent to

$$(2.29) \quad B(2s - 1), E(s) \text{ and } C(s) \text{ with } b_{2s} > 0,$$

(2.24) to

$$(2.30) \quad B(2s - 1), E(s) \text{ and } D(s) \text{ with } b_{2s} < 0,$$

(2.25) to

(2.31a)

$$\begin{cases} B(\max\{2s - 2, s\}), E(s), C(s - 1), D_{s-1}^T \widehat{V}_{s-1} = 0, \\ \gamma \geq 0, z_{s,s} \geq 0, |c_s| |d_s| \neq 0, \end{cases}$$

and if $\gamma > 0$, also to

$$(2.31b) \quad \begin{cases} B(\max\{2s - 2, s\}), E(s), D(s - 1), C_{s-1}^T \text{diag}(\gamma) \widehat{V}_{s-1} = 0, \\ \gamma \geq 0, z_{s,s} \geq 0, |c_s| |d_s| \neq 0, \end{cases}$$

(2.26) to

$$(2.32) \quad \begin{cases} B(\max\{2s - 3, s\}), E(s), C(s - 1), D_{s-2}^T \widehat{V}_{s-1} = 0, \\ \gamma \geq 0, \widehat{Z}_{s-2} \geq 0, b_{2s-2} \neq 0, \end{cases}$$

and (2.27) to

$$(2.33) \quad \begin{cases} B(\max\{2s - 3, s\}), E(s), D(s - 1), C_{s-2}^T \text{diag}(\gamma) \widehat{V}_{s-1} = 0, \\ \gamma \geq 0, \widehat{Z}_{s-2} \geq 0, b_{2s-2} \neq 0. \end{cases}$$

The methods of classes 1, 2, 3 and 4 can be regarded as extension of the Gauss, RadauIIA, RadauIA and LobattoIIIC Runge-Kutta methods, respectively, since for the special case $r = 1$ it is easy to verify that $E(s)$ is equivalent to $B(s)$.

Note also that in view of Lemma 2.3 we have $z_{s,s} = b_{2s}$ for class 2, and $z_{s,s} = -b_{2s}$ for class 3, and that For the classes 1–3, $\gamma > 0$ is guaranteed by (1.4a), (1.4b) and $B(2s - 1)$ (cf. [5]). Thus, specializing Theorems 2.1–2.5 to the cases of $p = s, s - 1$, we obtain immediately

Theorem 2.6. *The methods of classes 1–6 are all algebraically stable for the matrices G, Q .*

It should be pointed out that algebraic stability of the methods of class 1 has also been presented by Burrage [5] and Li [21]; in [21] the results for classes 2 and 3 have also been obtained.

Theorem 2.7. *Suppose the method (1.1) satisfying (1.4a), (1.4b) is algebraically stable for the matrices G, Q , and satisfies $B(s), E(s)$ with $\gamma > 0$ and α having at least m positive elements. Then this method is diagonally stable provided $s \leq m$. Furthermore, if the additional condition*

$$(2.34) \quad Z = \begin{bmatrix} 0 & 0 \\ 0 & \widehat{Z}_p \end{bmatrix}, \quad \widehat{Z}_p > 0,$$

is also satisfied, then this method is diagonally stable provided $p \leq m$, where $0 \leq p < s$, and where Z and \widehat{Z}_p are defined by (2.1) and (2.13), respectively.

Proof. Since $\gamma > 0$ and $\mu_1, \mu_2, \dots, \mu_s$ are distinct, for the diagonal stability of the method (1.1) we only need to prove that

$$(2.35) \quad H_\eta := \eta^T V^T (C_{11}^T \text{diag}(\gamma) + \text{diag}(\gamma) C_{11}) V \eta > 0 \\ \forall \eta = [\eta_1, \eta_2, \dots, \eta_s]^T \in \mathbf{R}^s, \quad \eta \neq 0.$$

Let

$$(2.36) \quad \begin{cases} \rho_\eta(x) = \sum_{j=1}^s \eta_j x^{j-1}, & \xi = [\xi_1, \xi_2, \dots, \xi_r]^T = U\eta, \\ \delta = [\sqrt{\alpha_1}, \sqrt{\alpha_2}, \dots, \sqrt{\alpha_r}]^T, & \sigma = \text{diag}(\delta)\xi. \end{cases}$$

In view of Lemma 2.1, $B(s)$ and $E(s)$ lead to (2.2a), (2.2b), and therefore together with (2.3), (2.35) and (2.36)

$$(2.37) \quad \begin{aligned} H_\eta &= (\gamma^T V \eta)^2 + \xi^T M_{11} \xi + \eta^T Z \eta \\ &= (r\rho_\eta(r) - \alpha^T \xi)^2 + \xi^T M_{11} \xi + \eta^T Z \eta. \end{aligned}$$

Here

$$\eta^T Z \eta \geq 0$$

due to algebraic stability of the method (cf. Theorem 2.1). From (1.4a), (1.4b) (2.3), (2.36) and application of Schwarz inequality we get

$$\xi^T M_{11} \xi = \|\delta\|_0^2 \|\sigma\|_0^2 - (\delta^T \sigma)^2 \geq 0,$$

which vanishes only for δ and σ linearly dependent, or equivalently, for

$$(2.38) \quad \xi_i = \text{constant}, \quad i \in \Omega_\alpha := \{i : i = 1, 2, \dots, r; \alpha_i \neq 0\}.$$

Hence, the three terms in the right hand side of (2.37) are all nonnegative, and the only remaining work is to prove that for any given $\eta \neq 0$ there exists at least one term which is positive in the right hand side of (2.37). Suppose on the contrary that the three terms are all equal to zero for some $\eta \neq 0$. Then (2.38) and

$$r\rho_\eta(r) = \alpha^T \xi$$

hold, and therefore together with (2.36) and (1.4b) we have

$$(2.39) \quad \begin{cases} \rho_\eta(i-1) = 0, & i \in \Omega_\alpha \setminus \{1\}, \\ \rho_\eta(r) = \frac{1}{r} \alpha^T \xi = 0. \end{cases}$$

Since α has at least m positive elements, (2.39) implies that the polynomial $\rho_\eta(x)$ has at least m distinct real roots, and consequently, we have

$$(2.40) \quad m \leq s-1, \quad \widehat{\eta}_m := [\eta_{m+1}, \eta_{m+2}, \dots, \eta_s]^T \neq 0.$$

This contradicts the assumption that $s \leq m$, and shows that the method is diagonally stable for $s \leq m$. For the more restricted case where the additional assumption (2.34) is also satisfied with $p \leq m$, combination of (2.34) and (2.40) leads to

$$\eta^T Z \eta = \widehat{\eta}_p^T \widehat{Z}_p \widehat{\eta}_p > 0,$$

where $\widehat{\eta}_p = [\eta_{p+1}, \eta_{p+2}, \dots, \eta_s]^T \neq 0$ since $p \leq m$ and $\widehat{\eta}_m \neq 0$. This contradicts the assumption that the three terms in the right hand side of (2.37) are all equal to zero, and shows that the method is diagonally stable for $p \leq m$. □

Applying Theorem 2.7 to the methods of classes 1–6, we immediately obtain

Theorem 2.8. *Suppose α has at least m positive elements. Then*

- (a) *a method of class 1 is diagonally stable provided $m \geq s$;*
- (b) *a method of class 2 or 3 is diagonally stable provided $m \geq s-1$;*
- (c) *a method of class 4 is diagonally stable provided either $m \geq s$ and $\gamma > 0$ or $m = s-1$, $z_{s,s} > 0$ and $\gamma > 0$;*
- (d) *a method of class 5 or 6 is diagonally stable provided either $m \geq s$ and $\gamma > 0$ or $s-2 \leq m < s$, $\widehat{Z}_{s-2} > 0$ and $\gamma > 0$.*

Remark 2. For the methods of classes 1–6 the (i, j) -elements z_{ij} of the matrix Z can be computed by (2.8). However, using (2.2a), (2.3) and $B(s)$ we can easily deduce that

$$\begin{aligned}
 z_{ij} &= (i\mu^{i-1})^T (C_{11}^T \text{diag}(\gamma) + \text{diag}(\gamma)C_{11})j\mu^{j-1} \\
 &\quad - [(r^i - \alpha^T \zeta^i)(r^j - \alpha^T \zeta^j) + \alpha^T \zeta^{i+j} - \alpha^T \zeta^i \alpha^T \zeta^j] \\
 (2.41) \quad &= (i\mu^{i-1})^T (C_{11}^T \text{diag}(\gamma) + \text{diag}(\gamma)C_{11})j\mu^{j-1} \\
 &\quad - \sum_{k=1}^r \alpha_k (r^i - (k-1)^i)(r^j - (k-1)^j).
 \end{aligned}$$

The formula (2.41) seems to be simpler than (2.8).

In the following theorem, we recall some results about stage order of the method (1.1), which have been presented in [20].

Theorem 2.9. *The method (1.1) satisfying (1.4a) has stage order at least τ if $B(\tau)$ and $C(\tau)$ hold, has weak stage order at least $\tau + 1$ if $B(\tau + 1)$ and $C(\tau)$ hold, and has generalized stage order at least $\tau + 1$ if $B(\tau + 1)$, $C(\tau)$ hold and there exists a real number ν such that $c_{\tau+1} = \nu \widehat{e}_s$.*

Let m denote the number of positive elements of α . Then a combination of Theorems 1.1, 2.2–2.4 and 2.6–2.9 yields the following results.

Theorem 2.10. *Suppose the method (1.1) satisfying (1.4a), (1.4b) satisfies $B(s)$, $E(s)$, $C(p)$ and one of the statements (a)–(c) in Theorem 2.2, $1 \leq p \leq s$, $\gamma > 0$, and either $m \geq s$ or $m \geq p - 1$ and $\widehat{Z}_{p-1} > 0$. Then*

- (a) *this method is B-stable and optimally B-convergent of order p ;*
- (b) *if $s \geq 2$, then this method is B-convergent of order $p + 1$;*
- (c) *if $s \geq 2$, and there exists a real number ν such that $c_{p+1} = \nu \widehat{e}_s$, then this method is optimally B-convergent of order $p + 1$.*

Theorem 2.11. *Suppose the method (1.1) satisfying (1.4a), (1.4b) satisfies $B(s)$, $E(s)$, $D(p)$ and one of the statements (a)–(c) in Theorem 2.3, $1 \leq p \leq s$, $\gamma > 0$, and either $m \geq s$ or $m \geq p - 1$ and $\widehat{Z}_{p-1} > 0$. Then*

- (a) *this method is B-stable, optimally B-convergent of order $p - 1$ (for $p > 1$) and B-convergent of order p ;*
- (b) *if in addition there exists a real number ν such that $c_p = \nu \widehat{e}_s$, then this method is optimally B-convergent of order p .*

Theorem 2.12. *Suppose the method (1.1) satisfying (1.4a), (1.4b) satisfies $B(s)$, $E(s)$, $C(p)$ and one of the statements (a)–(c) in Theorem 2.4, $1 \leq p \leq s$, $\gamma > 0$, and either $m \geq s$ or $\min(m, s - 1) \geq p$ and $\widehat{Z}_p > 0$. Then*

- (a) *this method is B-stable, optimally B-convergent of order p and B-convergent of order $p + 1$;*
- (b) *if in addition there exists a real number ν such that $c_{p+1} = \nu \widehat{e}_s$, then this method is optimally B-convergent of order $p + 1$.*

Theorem 2.13. *For the methods of classes 1–6 we have the following results.*

- (a) *A method of class 1 is B-stable, optimally B-convergent of order s and B-convergent of order $s + 1$, provided $m \geq s$.*
- (b) *A method of class 2 is B-stable, optimally B-convergent of order s and B-convergent of order $\min(s + 1, 2s - 1)$, provided $m \geq s - 1$.*

(c) A method of class 3 is B-stable, optimally B-convergent of order $s - 1$ (for $s > 1$) and B-convergent of order s , provided $m \geq s - 1$.

(d) A method of class 4 is B-stable, optimally B-convergent of order $s - 1$ (for $s > 1$) and B-convergent of order s , provided that $\gamma > 0$ and either $m \geq s$ or $m = s - 1$ and $z_{s,s} > 0$.

(e) A method of class 5 with $s \geq 3$ is B-stable, optimally B-convergent of order $s - 1$ and B-convergent of order s , provided that $\gamma > 0$ and either $m \geq s$ or $s - 2 \leq m < s$ and $\widehat{Z}_{s-2} > 0$.

(f) A method of class 6 with $s \geq 3$ is B-stable, optimally B-convergent of order $s - 2$ and B-convergent of order $s - 1$, provided that $\gamma > 0$ and either $m \geq s$ or $s - 2 \leq m < s$ and $\widehat{Z}_{s-2} > 0$.

Furthermore, these methods have optimal B-convergence order one higher than that mentioned above if, in addition, $c_l = \nu \widehat{e}_s$ for some real number ν , where $l = s + 1$ for classes 1 and 2, $l = s$ for classes 3, 4 and 5, $l = s - 1$ for class 6, $s \geq 2$ for class 2 and $s \geq 3$ for classes 5 and 6.

Note that similar results for methods of class 1 have also been obtained in [20].

3. CONSTRUCTION OF B-CONVERGENT METHODS

Methods of class 1 have been constructed by Burrage [5]. In this section we examine mainly the construction of methods of classes 2-6. These methods all satisfy the simplifying conditions $B(s)$ and $E(s)$, which are equivalent to

$$(3.1) \quad \gamma^T V = R - \alpha^T U$$

and

$$(3.2) \quad C_{12}^T \text{diag}(\gamma)V = \alpha R - \text{diag}(\alpha)U$$

respectively, where $R = [r, r^2, \dots, r^s]$. Let

$$p(x) = \prod_{k=1}^s (x - \mu_k).$$

Then it is readily shown that for any given integer $q \in [0, s - 1]$, the simplifying condition $B(2s - q)$ holds if and only if, in addition to (3.1), the equation

$$(3.3) \quad \int_0^r x^{k-1} p(x) dx = \sum_{j=2}^r \alpha_j \int_0^{j-1} x^{k-1} p(x) dx, \quad k = 1, 2, \dots, s - q,$$

is satisfied (cf.[5]). For $q = 0$ equation (3.3) together with (1.4a), (1.4b) leads to

$$p(x) = \det \begin{bmatrix} h_1 & h_2 & \dots & h_{s+1} \\ h_2 & h_3 & \dots & h_{s+2} \\ \dots & \dots & \dots & \dots \\ h_s & h_{s+1} & \dots & h_{2s} \\ 1 & x & \dots & x^s \end{bmatrix},$$

and, for $0 < q < s$, to

$$(3.4) \quad p(x) = \det \begin{bmatrix} h_1 & h_2 & \dots & h_{s+1} \\ h_2 & h_3 & \dots & h_{s+2} \\ \dots & \dots & \dots & \dots \\ h_{s-q} & h_{s-q+1} & \dots & h_{2s-q} \\ 1 & \mu_{i_1} & \dots & \mu_{i_1}^s \\ 1 & \mu_{i_2} & \dots & \mu_{i_2}^s \\ \dots & \dots & \dots & \dots \\ 1 & \mu_{i_q} & \dots & \mu_{i_q}^s \\ & x & \dots & x^s \end{bmatrix},$$

where

$$(3.5) \quad h_l = \frac{1}{l} [r^l - (r-1)^l + \sum_{j=1}^{r-1} \alpha_j ((r-1)^l - (j-1)^l)], \quad l = 1, 2, 3, \dots$$

Note that, here and later, some constant factor of $p(x)$ is ignored for simplicity. Note also that for the case of $q > 0$, the real numbers $\mu_{i_1}, \mu_{i_2}, \dots, \mu_{i_q}$, which are q distinct roots of $p(x)$, and $\alpha_1, \alpha_2, \dots, \alpha_r$ should be appropriately chosen in advance so that the polynomial $p(x)$ defined by (3.4) and (3.5) is of degree s and has s distinct real roots.

For simplicity, we may write

$$(3.6) \quad p(x) = \hat{p}(x) \prod_{m=1}^q (x - \mu_{i_m}).$$

Then (3.3) leads to

$$(3.7) \quad \hat{p}(x) = \det \begin{bmatrix} g_1 & g_2 & \dots & g_{s-q+1} \\ g_2 & g_3 & \dots & g_{s-q+2} \\ \dots & \dots & \dots & \dots \\ g_{s-q} & g_{s-q+1} & \dots & g_{2(s-q)} \\ 1 & x & \dots & x^{s-q} \end{bmatrix},$$

where

$$(3.8) \quad g_l = \int_{r-1}^r x^{l-1} \prod_{m=1}^q (x - \mu_{i_m}) dx + \sum_{j=1}^{r-1} \alpha_j \int_{j-1}^{r-1} x^{l-1} \prod_{m=1}^q (x - \mu_{i_m}) dx, \\ l = 1, 2, 3, \dots$$

The polynomial $p(x)$ can also be determined by (3.6), (3.7) and (3.8), which seems to be simpler than by (3.4) and (3.5).

To construct methods of classes 2–6, we thus first choose $\alpha_1, \alpha_2, \dots, \alpha_r$ appropriately to satisfy (1.4a), (1.4b), and $\mu_{i_1}, \mu_{i_2}, \dots, \mu_{i_q}$ which are real and distinct, where $q = 1$ for classes 2 and 3, $q = 2$ for class 4 and $q = 3$ for classes 5 and 6. Then compute the roots $\mu_1, \mu_2, \dots, \mu_s$ of the polynomial $p(x)$ defined by either (3.4) and (3.5) or (3.6), (3.7) and (3.8), and compute γ from (3.1) and C_{12} from (3.2). Finally, the coefficient matrix C_{11} can be determined by other conditions which the methods should satisfy.

Example 1. For $r = 2$ and $s = 1$, the methods of classes 1–4 are

$$(3.9) \quad \begin{cases} Y = hcf(t_n + uh, Y) + \frac{2a}{1+a}y_n + \frac{1-a}{1+a}y_{n+1}, \\ y_{n+2} = h(1+a)f(t_n + uh, Y) + ay_n + (1-a)y_{n+1}, \end{cases}$$

where $0 < a \leq 1$, $c \geq (1 + 3a)/[2(1 + a)]$, and

$$(3.10) \quad \begin{cases} c = \frac{1+3a}{2(1+a)}, & u = 2 - c = c + \frac{1-a}{1+a} & \text{for class 1,} \\ c > \frac{1+3a}{2(1+a)}, & u = c + \frac{1-a}{1+a} & \text{for class 2,} \\ c > \frac{1+3a}{2(1+a)}, & u = 2 - c & \text{for class 3,} \\ c \geq \frac{1+3a}{2(1+a)}, & u \neq 2 - c, \quad u \neq c + \frac{1-a}{1+a} & \text{for class 4.} \end{cases}$$

It follows from Theorems 2.6, 2.8 and 2.13 that these methods are all algebraically stable, diagonally stable, B -stable and B -convergent of order 1, and that the methods of class 2 are optimally B -convergent of order 1; the methods of class 1 are optimally B -convergent of order 2 since for the methods of class 1 the condition $B(2)$ holds and we have $c_2 = \nu \hat{e}_1$ provided choose $\nu = c_2$. Note that for the special case $a = 1$, (3.9) specialized to

$$(a) \ c = u = 1, \quad (b) \ c = u = 2, \quad (c) \ c = 2, \ u = 0$$

leads to the one-stage Gauss, RadauIIA and RadauIA Runge-Kutta methods with stepsize $2h$ respectively. Note also that the simplifying condition $C(1)$ does not hold for the methods of classes 3 and 4.

Example 2. Consider the methods of classes 2 and 3 with $r = s = 2$. Let

$$\begin{cases} (3.11a) & \alpha_1 = a, \quad \alpha_2 = 1 - a, \\ (3.11b) & \mu_1 = u, \quad p(x) = \hat{p}(x)(x - u), \end{cases}$$

with $0 < a \leq 1$ and the real number u to be determined. Then it follows from (3.7) and (3.8) that

$$\hat{p}(x) = g_1x - g_2, \quad g_1 = \frac{3+a}{2} - (1+a)u, \quad g_2 = \frac{7+a}{3} - \frac{3+a}{2}u,$$

and consequently we have

$$(3.12) \quad \mu_1 = u, \quad \mu_2 = \left(\frac{7+a}{3} - \frac{3+a}{2}u\right) / \left(\frac{3+a}{2} - (1+a)u\right).$$

To guarantee $\hat{p}(x)$ having degree 1 and $\mu_1 \neq \mu_2$ it is necessary and sufficient that

$$(3.13) \quad u \neq \frac{3+a}{2(1+a)}, \quad (1+a)u^2 - (3+a)u + \frac{7+a}{3} \neq 0.$$

Thus γ can be uniquely determined by (3.1), i.e.

$$(3.14) \quad \gamma_1 = \frac{1}{\mu_2 - \mu_1} \left((1+a)\mu_2 - \frac{3+a}{2} \right), \quad \gamma_2 = \frac{-1}{\mu_2 - \mu_1} \left((1+a)\mu_1 - \frac{3+a}{2} \right).$$

Since $B(2s - 1)$ holds, we have $\gamma > 0$, and C_{12} can be uniquely determined by (3.2), i.e.

$$(3.15) \quad C_{12} = \begin{bmatrix} \frac{4a(\mu_2-1)}{2(1+a)\mu_2-(3+a)} & \frac{(1-a)(2\mu_2-3)}{2(1+a)\mu_2-(3+a)} \\ \frac{4a(\mu_1-1)}{2(1+a)\mu_1-(3+a)} & \frac{(1-a)(2\mu_1-3)}{2(1+a)\mu_1-(3+a)} \end{bmatrix}.$$

Finally, for methods of class 2 C_{11} can be uniquely determined by $C(2)$, i.e.

$$(3.16) \quad C_{11} = ([\mu, \mu^2] - C_{12}U)V^{-1},$$

and for class 3 by $D(2)$, i.e.

$$(3.17) \quad C_{11} = (\gamma R - \text{diag}(\gamma)[\mu, \mu^2])V^{-1}[\text{diag}(\gamma)]^{-1}.$$

Furthermore, it is easily seen from (3.1) that

$$(3.18) \quad \begin{aligned} b_{2s} = b_4 &= 4(R - \alpha^T U)V^{-1}\mu^3 - 16 + \alpha_2 \\ &= 2(3 + a)(\mu_1^2 + \mu_1\mu_2 + \mu_2^2) - 4(1 + a)\mu_1\mu_2(\mu_1 + \mu_2) - 15 - a \\ &=: \varphi(a, \mu). \end{aligned}$$

In view of Theorems 2.6, 2.8 and 2.13 we can thus state the following propositions:

Proposition 3.1. *For any given $a \in (0, 1]$ and u satisfying (3.13) and $\varphi(a, \mu) > 0$, there exists a two-value two-stage multistep Runge-Kutta method of class 2 given by (1.1), (1.3), (3.11a), (3.12), (3.14), (3.15) and (3.16), which is algebraically stable, diagonally stable, B-stable, optimally B-convergent of order 2 and B-convergent of order 3. For the special case $a = 1$ and $u = 2$, the corresponding method is equivalent to the two-stage RadauIIA Runge-Kutta method with stepsize $2h$.*

Proposition 3.2. *For any given $a \in (0, 1]$ and u satisfying (3.13) and $\varphi(a, \mu) < 0$, there exists a two-value two-stage multistep Runge-Kutta method of class 3 given by (1.1), (1.3), (3.11a), (3.12), (3.14), (3.15) and (3.17), which is algebraically stable, diagonally stable, B-stable, optimally B-convergent of order 1 and B-convergent of order 2. For the special case $a = 1$ and $u = 0$, the corresponding method is equivalent to the two-stage RadauIA Runge-Kutta method with stepsize $2h$.*

Example 3. Consider the methods of class 4 with $r = s = 2$. Let

$$(3.19) \quad \alpha_1 = a, \quad \alpha_2 = 1 - a, \quad 0 < a \leq 1,$$

and let μ_1, μ_2 be real numbers satisfying $\mu_1 < \mu_2$. Then γ and C_{12} can be determined by (3.14) and (3.15) respectively. To guarantee $\gamma > 0$ it is necessary and sufficient that

$$(3.20) \quad \mu_1 < \frac{3 + a}{2(1 + a)}, \quad \mu_2 > \frac{3 + a}{2(1 + a)}.$$

The conditions $C(1)$, $2d_1^T \mu = 0$ and $z_{2,2} \geq 0$ are equivalent to

$$\begin{aligned} C_{11}\hat{e}_2 &= \mu - C_{12}\zeta, & 2\gamma^T C_{11}\mu &= 4\gamma^T \mu - 2\gamma^T \mu^2, \\ 8\gamma^T \text{diag}(\mu)C_{11}\mu &= 16\gamma^T \mu - 16 + \alpha_2 + \varepsilon, & \varepsilon &\geq 0, \end{aligned}$$

and therefore, together with $B(2)$,

$$(3.21) \quad C_{11} = [\delta, \sigma_\varepsilon]V^{-1},$$

where

$$(3.22) \quad \delta = \mu - C_{12}\zeta = \begin{bmatrix} \mu_1 - \frac{(1-a)(2\mu_2-3)}{2(1+a)\mu_2-(3+a)} \\ \mu_2 - \frac{(1-a)(2\mu_1-3)}{2(1+a)\mu_1-(3+a)} \end{bmatrix},$$

$$(3.23) \quad \begin{aligned} \sigma_\varepsilon &= \begin{bmatrix} \gamma_1 & \gamma_2 \\ 4\gamma_1\mu_1 & 4\gamma_2\mu_2 \end{bmatrix}^{-1} \begin{bmatrix} 8 - 2\alpha_2 - 2\gamma^T\mu^2 \\ 16 - 7\alpha_2 + \varepsilon \end{bmatrix} \\ &= \begin{bmatrix} \frac{2\mu_2(6+2a-2\gamma^T\mu^2)-\frac{1}{2}(9+7a+\varepsilon)}{2(1+a)\mu_2-(3+a)} \\ \frac{2\mu_1(6+2a-2\gamma^T\mu^2)-\frac{1}{2}(9+7a+\varepsilon)}{2(1+a)\mu_1-(3+a)} \end{bmatrix} \end{aligned}$$

with $\varepsilon \geq 0$. In order that $|c_2||d_2| \neq 0$, we only need $b_3 \neq 0$, i.e.

$$(3.24) \quad 3(\gamma_1\mu_1^2 + \gamma_2\mu_2^2) \neq 7 + a.$$

In view of Theorems 2.6, 2.8 and 2.13 we thus obtain the next proposition:

Proposition 3.3. *For any given $\mu_1, \mu_2, \varepsilon \geq 0$ and $a \in (0, 1]$ satisfying (3.20) and (3.24) with γ_1 and γ_2 defined by (3.14), there exists a two-value two-stage multistep Runge-Kutta method of class 4 given by (1.1), (1.3), (3.19), (3.14), (3.15), (3.21), (3.22) and (3.23), which is algebraically stable, diagonally stable, B-stable, optimally B-convergent of order 1 and B-convergent of order 2, provided that either $a < 1$ or $\varepsilon > 0$. For the special case $a = 1, \mu_1 = 0, \mu_2 = 2$ and $\varepsilon = 16$, the corresponding method is equivalent to the two-stage LobattoIIIC Runge-Kutta method with stepsize $2h$.*

Example 4. Consider the construction of methods of class 5 with $r = 2$ and $s = 3$. Let

$$(3.25) \quad \alpha = [a, 1 - a]^T, \quad \mu = [\mu_1, \mu_2, \mu_3]^T,$$

where

$$(3.26) \quad 0 < a \leq 1, \quad \mu_1 < \mu_2 < \mu_3.$$

Using $B(3)$, we get

$$(3.27) \quad \gamma = \gamma(a, \mu) = \begin{bmatrix} 1 & 2\mu_1 & 3\mu_1^2 \\ 1 & 2\mu_2 & 3\mu_2^2 \\ 1 & 2\mu_3 & 3\mu_3^2 \end{bmatrix}^{-T} \begin{bmatrix} 1 + a \\ 3 + a \\ 7 + a \end{bmatrix}.$$

We also assume that $\gamma(a, \mu) > 0$, so that $E(3)$ leads to

$$(3.28) \quad C_{12} = [\text{diag}(\gamma)]^{-1}V^{-T}(\alpha R - \text{diag}(\alpha)U)^T,$$

where

$$(3.29) \quad V = [\hat{e}_3, 2\mu, 3\mu^2], \quad R = [2, 4, 8], \quad U = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

The elements of the matrix

$$\hat{Z}_1 = \begin{bmatrix} z_{22} & z_{23} \\ z_{32} & z_{33} \end{bmatrix}$$

are

$$\begin{cases} z_{22} = b_4 = 4\gamma^T\mu^3 - 15 - a, \\ z_{23} = z_{32} = (2C_{11}\mu)^T \text{diag}(\gamma)3\mu^2 + 2\gamma^T \text{diag}(\mu)C_{11}3\mu^2 - 21 - 11a, \\ z_{33} = 6\gamma^T \text{diag}(\mu^2)C_{11}3\mu^2 - 49 - 15a. \end{cases}$$

Thus the conditions $C(2)$, $3d_1^T \mu^2 = 0$, $\widehat{Z}_1 \geq 0$ and $b_4 \neq 0$ lead to

$$(3.30) \quad b_4 \equiv 4\gamma^T \mu^3 - 15 - a > 0$$

and

$$\begin{cases} C_{11}[\widehat{e}_3, 2\mu] = [\mu - C_{12}\zeta, \mu^2 - C_{12}\zeta^2], \\ \gamma^T C_{11}3\mu^2 = 6\gamma^T \mu^2 - 3\gamma^T \mu^3, \\ 6\gamma^T \text{diag}(\mu^2)C_{11}3\mu^2 = 49 + 15a + \varepsilon_1, \\ 2\gamma^T \text{diag}(\mu)C_{11}3\mu^2 = 21 + 11a - 3\gamma^T \text{diag}(\mu^2 - C_{12}\zeta^2)\mu^2 \pm \sqrt{\varepsilon_1 b_4 - \varepsilon_2}, \end{cases}$$

where

$$(3.31) \quad \varepsilon_1 \geq 0, \quad 0 \leq \varepsilon_2 \leq \varepsilon_1 b_4,$$

and therefore, together with $B(3)$ and $E(3)$,

$$(3.32) \quad C_{11} = [\mu - C_{12}\zeta, \mu^2 - C_{12}\zeta^2, \delta]V^{-1},$$

where

$$(3.33) \quad \zeta = [0, 1]^T, \quad \delta = \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 \\ 2\gamma_1\mu_1 & 2\gamma_2\mu_2 & 2\gamma_3\mu_3 \\ 6\gamma_1\mu_1^2 & 6\gamma_2\mu_2^2 & 6\gamma_3\mu_3^2 \end{bmatrix}^{-1} \eta$$

with

$$\eta = \begin{bmatrix} 14 + 2a - 3\gamma^T \mu^3 \\ 21 + 11a - 3\gamma^T \mu^4 + 8\alpha^T \zeta^2 - \alpha^T \zeta^5 \pm \sqrt{\varepsilon_1 b_4 - \varepsilon_2} \\ 49 + 15a + \varepsilon_1 \end{bmatrix}.$$

Thus, in view of Theorems 2.6, 2.8 and 2.13 we obtain

Proposition 3.4. *For any given $\mu_1, \mu_2, \mu_3, a, \varepsilon_1$ and ε_2 satisfying (3.26), (3.30), (3.31) and $\gamma(a, \mu) > 0$, there exists a two-value three-stage multistep Runge-Kutta method of class 5 given by (1.1), (1.3), (3.25), (3.27), (3.28), (3.29), (3.32) and (3.33), which is algebraically stable, diagonally stable, B-stable, optimally B-convergent of order 2 and B-convergent of order 3, provided $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$.*

The methods of class 6 with $r = 2$ and $s = 3$ can be constructed similarly, but in this case we have $z_{2,2} = -b_4$, and the coefficient matrix C_{11} should be determined by the conditions $D(2)$, $3c_1^T \text{diag}(\gamma)\mu^2 = 0$, $\widehat{Z}_1 \geq 0$ and $b_4 \neq 0$.

Example 5. To construct methods of class 4 with $r = 2$ and $s = 3$, we let

$$(3.34) \quad \alpha_1 = a, \quad \alpha_2 = 1 - a,$$

and

$$p(x) = \widehat{p}(x)(x - \mu_1)(x - \mu_2),$$

where

$$(3.35) \quad 0 < a \leq 1, \quad \mu_1 < \mu_2.$$

In order to satisfy $B(4)$, the relations (3.7) and (3.8) with $s = 3$ and $r = q = 2$ must hold, which lead to

$$\begin{cases} \widehat{p}(x) = g_1 x - g_2, \\ g_1 = \int_1^2 (x - \mu_1)(x - \mu_2)dx + a \int_0^1 (x - \mu_1)(x - \mu_2)dx, \\ g_2 = \int_1^2 x(x - \mu_1)(x - \mu_2)dx + a \int_0^1 x(x - \mu_1)(x - \mu_2)dx, \end{cases}$$

and consequently

$$(3.36) \quad \mu_3 = \mu_3(a, \mu_1, \mu_2) = \frac{\int_1^2 x(x - \mu_1)(x - \mu_2)dx + a \int_0^1 x(x - \mu_1)(x - \mu_2)dx}{\int_1^2 (x - \mu_1)(x - \mu_2)dx + a \int_0^1 (x - \mu_1)(x - \mu_2)dx}.$$

Here we assume that

$$(3.37) \quad g_1 \neq 0, \quad \mu_3(a, \mu_1, \mu_2) \neq \mu_1, \quad \mu_3(a, \mu_1, \mu_2) \neq \mu_2.$$

The only remaining task now is to calculate γ , C_{12} and C_{11} ; but this is the same as mentioned in example 4 except that in this case we have $b_4 = 0$ and $\varepsilon_2 = 0$, and require

$$(3.38) \quad b_5 \equiv 5\gamma^T \mu^4 - 31 - a \neq 0$$

instead of (3.30). Thus, in view of Theorems 2.6, 2.8 and 2.13 we obtain

Proposition 3.5. *For any given μ_1, μ_2, a and ε_1 satisfying (3.35), (3.37), (3.38), $\gamma(a, \mu) > 0$ and $\varepsilon_1 \geq 0$, there exists a two-value three-stage multistep Runge-Kutta method of class 4 given by (1.1), (1.3), (3.34), (3.36), (3.27), (3.28), (3.29), (3.32) and (3.33) with $\varepsilon_2 = 0$, which is algebraically stable, diagonally stable, B-stable, optimally B-convergent of order 2 and B-convergent of order 3, provided that $\varepsilon_1 > 0$ and $a < 1$. For the special case $a = 1, \mu_1 = 0, \mu_2 = 1$ and $\varepsilon_1 = 52$, the corresponding method is equivalent to the three-stage LobattoIIIC Runge-Kutta method with stepsize $2h$.*

The nonlinear stability and B-convergence theory for multistep Runge-Kutta methods established in the present paper enrich the theoretical foundations of stiff computation, and without doubt of importance. However, in practice, it is natural to ask whether there exist multistep Runge-Kutta methods of classes 1–6 which are superior in some respects to the well-known one-step Runge-Kutta schemes. Fortunately, the answer is indeed positive, due to the following facts.

(1) Any one-step Gauss Runge-Kutta method has a stability function $R(\bar{h})$ whose value at infinity satisfies $|R(\infty)| = 1$. In contrast, the multistep Runge-Kutta methods of class 1 allow the spectral radius of the stability matrix $S(\bar{h})$ at infinity to be smaller than 1: $\rho(S(\infty)) < 1$, which improves stability at infinity, and ensures that the errors of the extremely stiff components are damped out quickly. For example, for the 2-step 3-stage multistep Runge-Kutta method of class 1 with coefficients

$$\left\{ \begin{array}{l} \alpha = [0.0254294608860966, 0.974570539113903]^T, \\ \gamma = [0.0292020628426463, 0.578611565044865, 0.417615832998585]^T, \\ \mu = [0.388710707597604, 1.27430628101834, 1.82951690035238]^T, \\ C_{11} = \begin{bmatrix} 0.337337979617462 & -0.292009898095809 & 0.108941719003734 \\ 0.0488493196803534 & 0.289675767260957 & -0.0331596021107289 \\ 0.0214005894331623 & 0.624526898213954 & 0.208808290132131 \end{bmatrix}, \\ C_{12} = \begin{bmatrix} 0.765559092927782 & 0.234440907072218 \\ 0.0310592038122419 & 0.968940796187758 \\ 0.0252188774268619 & 0.974781122573138 \end{bmatrix}, \end{array} \right.$$

it is readily checked that $\rho(S(\infty)) \approx 0.34$. Furthermore, Hairer and Wanner [13] have also pointed out that there exists a 2-step 2-stage multistep Runge-Kutta method of class 1 for which $\rho(S(\infty)) = \sqrt{2} - 1 \approx 0.41421$.

(2) The coefficient matrices of the one-step s -stage Gauss, RadauIA and RadauIIA Runge-Kutta methods all have complex eigenvalues for $s > 1$. In contrast, there exist multistep Runge-Kutta methods of classes 1–6 whose coefficient matrices only possess distinct real eigenvalues; these are called *multistep Runge-Kutta methods with real eigenvalues* in [18]. For example, the 2-stage multistep Runge-Kutta method

$$\begin{cases} Y_1 = \frac{1}{14}[5y_{n-3} + 9y_{n-1} + h(15f(t_n - h, Y_1) - 5f(t_n + h, Y_2))], \\ Y_2 = \frac{1}{2}[-y_{n-3} + 3y_{n-1} + h(-f(t_n - h, Y_1) + 3f(t_n + h, Y_2))], \\ y_n = \frac{1}{20}[2y_{n-3} + 18y_{n-1} + h(21f(t_n - h, Y_1) + 3f(t_n + h, Y_2))] \end{cases}$$

constructed in [18] is of class 2 and has a coefficient matrix

$$C_{11} = \begin{bmatrix} \frac{15}{14} & \frac{-5}{14} \\ -\frac{1}{2} & \frac{3}{2} \end{bmatrix}$$

whose eigenvalues $\lambda_{1,2} = \frac{1}{7}(9 \pm \sqrt{11})$ are two distinct real numbers. This can essentially improve the computational efficiency. In fact, for s -stage whole implicit methods applied to m -dimensional stiff problems, nonlinear systems of dimension sm arise. Applying a simplified Newton process leads to linear systems of the same dimension. If we use Gaussian elimination to solve such a linear system, then this would cost $\frac{2}{3}s^3m^3$ arithmetic operations for the LU -decomposition, which can be very expensive for large sm . In order to reduce these costs, we usually bring the Newton iteration matrix to block diagonal form by a Butcher transformation (cf. [8]). For the aforementioned one-step Runge-Kutta methods, each complex eigenvalue pair of the coefficient matrix corresponds with a block of size 2, for which the LU -costs can only be reduced to $\frac{8}{3}m^3$. However, for multistep Runge-Kutta methods with real eigenvalues, all the blocks are of size 1, and for each such block the LU -costs can be reduced to $\frac{2}{3}m^3$. Therefore, if we use an appropriate number of processors to compute a large-scale stiff problem in parallel by a multistep Runge-Kutta method with real eigenvalues, then for each processor the computational cost per step is almost the same as for the serial computation by a widely used BDF method, whereas the stability and convergence properties are far better than that of BDF. On the other hand, it is easily seen from the above discussion that for large scale parallel stiff computation the aforementioned one-step Runge-Kutta methods are about 4 times more expensive than the corresponding multistep Runge-Kutta methods with real eigenvalues.

Recently, we have found a family of multistep Runge-Kutta methods with real eigenvalues, such as 2-stage 3-step third-order methods of class 2, 2-stage 4-step fourth-order methods of class 1, 3-stage 4-step fifth-order methods of class 2, 3-stage 8-step sixth-order methods of class 1 and 4-stage 7-step seventh-order methods of class 2, which will be published in a separate paper [19]. Since these methods not only possess very good stability and convergence properties, but also can be performed in parallel at high speed, it is expected that for large scale parallel stiff

computations a code based on these methods will be superior in many respects to the existing codes based on traditional methods.

Furthermore, although the s -stage Gauss and Radau type one-step Runge-Kutta methods with $s > 1$ cannot be singly implicit, we have also found some 2-stage singly implicit multistep Runge-Kutta methods of classes 1 and 2 which can improve the efficiency for serial computation, and will also be published in [19].

(3) It is well known that the 1-step 3-stage LobattoIIIC Runge-Kutta method is not diagonally stable (cf. [11]) and not B -convergent for problems with an optimal one-sided Lipschitz constant $\nu \geq 0$ (cf. [22]). In contrast, there exist many three-stage multistep Runge-Kutta methods of class 4, determined by Proposition 3.5 of this paper which are all diagonally stable, B -stable and B -convergent of order 3. This means that there do exist multistep Runge-Kutta methods, constructed in the present paper, whose stability and convergence properties are all better than those of the corresponding traditional Runge-Kutta methods.

REFERENCES

- [1] W.Auzinger, R.Frank & G.Kirlinger, A note on convergence concepts for stiff problems, *Computing* **44**(1990), 197–208. MR **91h**:65116
- [2] W.Auzinger, R.Frank & G.Kirlinger, An extension of B -convergence for Runge-Kutta methods, *Appl. Numer. Math.* **9**(1992), 91–109. MR **92k**:65109
- [3] W.Auzinger, R.Frank & G.Kirlinger, Modern convergence theory for stiff initial value problems, *J. Comput. Appl. Math.* **45**(1993), 5–16. MR **93j**:65124
- [4] W.Auzinger, R.Frank & G.Kirlinger, Extending convergence theory for nonlinear stiff problems, part I, *BIT* **4**(1996), 635–652. MR **97i**:34019
- [5] K.Burrage, High order algebraically stable multistep Runge-Kutta methods, *SIAM J. Numer. Anal.* **24**(1987), 106–115. MR **88f**:65132
- [6] K. Burrage, & J.C. Butcher, Nonlinear stability of a general class of differential equation methods, *BIT* **20**(1980), 185–203. MR **82b**:65076
- [7] J.C. Butcher, A stability property of implicit Runge-Kutta methods, *BIT* **15**(1975), 358–361.
- [8] J.C. Butcher, On the implementation of implicit Runge-Kutta methods, *BIT* **16**(1976), 237–240. MR **58**:8263
- [9] G. Dahlquist, Error analysis for a class of methods for stiff nonlinear initial value problems, *Num. Anal. Dundee, 1975. Lect. Notes in Math.* 506, Springer-Verlag, Berlin, 1976, 60–74. MR **56**:7203
- [10] R.Frank, J.Schneid & C.W.Ueberhuber, The concept of B -convergence, *SIAM J. Numer. Anal.* **18**(1981), 753–780. MR **82h**:65054
- [11] R.Frank, J.Schneid & C.W.Ueberhuber, Stability properties of implicit Runge-Kutta methods, *SIAM J. Numer. Anal.* **22**(1985), 497–514. MR **86m**:65076a
- [12] R.Frank, J.Schneid & C.W.Ueberhuber, Order results for implicit Runge-Kutta methods applied to stiff systems, *SIAM J. Numer. Anal.* **22**(1985), 515–534. MR **86m**:65076b
- [13] E.Hairer & G.Wanner, Solving ordinary differential equations II, Springer-Verlag Berlin Heidelberg, 1991. MR **92a**:65016
- [14] Shoufu Li, B -convergence of general linear methods, *Proc. BAIL-V International conference*, Shanghai, 1988, 203–208, Boole Press Conf. Ser. 12, 1988.
- [15] Shoufu Li, Stability and B -convergence of general linear methods, *Proc. 3rd International Congress on Comput. Appl. Math.*, Belgium, 1988. *J. Comput. Appl. Math.* **28**(1989), 281–296.
- [16] Shoufu Li, B -convergence of General Multivalued Methods Applied to Stiff Problems in Banach Spaces, *Proc. National Conf. on Comput. Math.*, Tianjin, 1990. *SCIENCE IN CHINA (Series A)*, Chinese Series: 1992, 5:476–485, English Series: 36(1993), 1:1–13.
- [17] Shoufu Li, Theory of Computational Methods for Stiff Differential Equations, Science and Technology Press, Hunan, China, 1997.
- [18] Shoufu Li, Multistep Runge-Kutta methods with real eigenvalues and its parallel implementation, *Proc. 5-th CSIAM Conference*, Qinghua University Publishing House, 1998.

- [19] Shoufu Li, High order algebraically stable and B -convergent multistep Runge-Kutta methods with real eigenvalues, To appear.
- [20] Shoufu Li, B -convergence properties of multistep Runge-Kutta methods, Research Report, *International Conference on Sci. Comput.*, Hangzhou, 1991. *Math. Comput.* **62**(1994), 565—575.
- [21] Shoufu Li, Algebraic stability of multistep Runge-Kutta methods, *Acta Simulata Systematica Sinica* **5**(1993), 51—56 (in Chinese). Also *Syst. Engin. Electr.*, **6**(1995), 3:76-82 (in English).
- [22] J.Schneid, B -convergence of Lobatto IIIC formulas, *Numer. Math.* **51**(1987), 229–235. MR **88f**:65113

DEPARTMENT OF MATHEMATICS, XIANGTAN UNIVERSITY, XIANGTAN 411105, HUNAN PROVINCE,
PEOPLE'S REPUBLIC OF CHINA

E-mail address: `lisf@xtu.edu.cn`