B-CONVERGENCE PROPERTIES OF
MULTISTEP RUNGE-KUTTA METHODS

SHOUFU LI

Abstract. By using the theory of $B$-convergence for general linear methods to
the special case of multistep Runge-Kutta methods, a series of $B$-convergence
results for multistep Runge-Kutta methods is obtained, and it is proved that
the family of algebraically stable $r$-step $s$-stage multistep Runge-Kutta methods
with parameters $\alpha_1, \alpha_2, \ldots, \alpha_r$ presented by Burrage in 1987 is optimally $B$-
convergent of order at least $s$, and $B$-convergent of order $s+1$, provided that
$r \geq s$ and $\alpha_j > 0$, $j = 1, 2, \ldots, r$. Furthermore, this family of methods is
optimally $B$-convergent of order $s+1$ if some other additional conditions are
satisfied.

1. Introduction

Let $X$ be a real or complex Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and
the corresponding norm $\| \cdot \|$, $f: X \to X$ be a given sufficiently smooth mapping
satisfying a one-sided Lipschitz condition

$$\text{Re}(f(y) - f(z), y - z) \leq m\|y - z\|^2 \quad \forall y, z \in X.$$ 

Consider the initial value problem

$$y'(t) = f(y(t)), \quad 0 \leq t \leq T; \quad y(0) = y_0, \quad y_0 \in X$$

and the multistep Runge-Kutta method for solving (1.1):

$$Y^{(n)} = \tilde{A}y^{(n-1)} + h\tilde{B}F(Y^{(n)}),$$
$$y^{(n)} = \tilde{C}y^{(n-1)} + h\tilde{E}F(Y^{(n)}),$$
$$\xi_n = \tilde{B}y^{(n)}.$$

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

$$Y^{(n)} = (Y_1^{(n)}, Y_2^{(n)}, \ldots, Y_s^{(n)}) \in X^s, \quad y^{(n)} = (y_1^{(n)}, y_2^{(n)}, \ldots, y_r^{(n)}) \in X^r,$$
$$\xi_n \in X, \quad F(Y^{(n)}) = (f(Y_1^{(n)}), f(Y_2^{(n)}), \ldots, f(Y_s^{(n)})) \in X^s.$$
$h > 0$ is the stepsize, $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{E}$, and $\tilde{\beta}$ are linear mappings corresponding respectively to the real matrices

$$
A = [a_{ij}] \in \mathbb{R}^{s \times r}, \quad B = [b_{ij}] \in \mathbb{R}^{s \times s}, \quad C = \begin{bmatrix} 0 & I_{r-1} \\ \alpha T \end{bmatrix} \in \mathbb{R}^{r \times r},
$$

$$
E = \begin{bmatrix} 0 \\ y T \end{bmatrix} \in \mathbb{R}^{r \times s}, \quad \beta = [0, \ldots, 0, 1] \in \mathbb{R}^{1 \times r}
$$

(cf. [11]), where $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_r]^T$, $\gamma = [\gamma_1, \gamma_2, \ldots, \gamma_s]^T$, $I_{r-1}$ is the $(r-1) \times (r-1)$ identity matrix, $Y_i^{(n)}$, $y_i^{(n)}$, and $\xi_n$ are approximations to $y(t_n + c_i h)$, $y(t_n + i h)$, and $y(t_n + rh)$, respectively, where

$$
t_n = t_0 + nh; \quad c_i = \sum_{j=1}^{s} b_{ij} + \sum_{j=1}^{r} (j-1)a_{ij}, \quad i = 1, 2, \ldots, s.
$$

For simplicity, we write $c = [c_1, c_2, \ldots, c_s]^T$, $\zeta = [0, 1, \ldots, r - 1]^T$, $e_N = [1, 1, \ldots, 1]^T \in \mathbb{R}^N$ with $N \geq 1$, $Y(t) = (y(t + c_1 h), y(t + c_2 h), \ldots, y(t + c_r h)) \in X^s$, $H(t) = (y(t + h), y(t + 2h), \ldots, y(t + rh)) \in X^r$, introduce the simplifying conditions (cf. [1])

$$
B(\tau): \quad p\gamma^T c^{p-1} = r^p - \alpha^T \zeta^p, \quad p = 1, 2, \ldots, \tau;
$$

$$
C(\tau): \quad pBc^{p-1} = c^p - A^p \zeta^p, \quad p = 1, 2, \ldots, \tau;
$$

$$
E(\tau): \quad pA^T \text{diag} (\gamma) c^{p-1} = \text{diag} (\alpha) (p^r e_r - \zeta^p), \quad p = 1, 2, \ldots, \tau,
$$

and adopt the notational convention: $M > 0$ (or $\geq 0$) for a real symmetric matrix to mean that $M$ is positive definite (or nonnegative definite).

Note that multistep Runge-Kutta methods are a subclass of the General Linear Methods of Butcher, and it is proved by Lie and Nørsett [13] that multistep collocation methods are a subclass of multistep Runge-Kutta methods.

In 1987, Burrage [1] obtained the following results:

**Theorem 1.1.** Suppose the method (1.2)–(1.3) satisfies the conditions $B(2s)$, $C(s)$, and $E(s)$, $c_i \neq c_j$ whenever $i \neq j$, $\sum_{j=1}^{r} \alpha_j = 1$, $\alpha_1 > 0$, and $\alpha_j \geq 0$, $j = 2, 3, \ldots, r$. Then this method is algebraically stable for the matrices

$$
G = \text{diag} \left( \alpha_1, \alpha_1 + \alpha_2, \ldots, \sum_{j=1}^{r} \alpha_j \right), \quad D = \text{diag} (\gamma_1, \gamma_2, \ldots, \gamma_s),
$$

and necessarily $G > 0$, $D > 0$.

**Theorem 1.2.** Suppose that $\sum_{j=1}^{r} \alpha_j = 1$, $\alpha_1 > 0$, $\alpha_j \geq 0$, $j = 2, 3, \ldots, r$. Then the multistep Runge-Kutta methods defined by (1.2), (1.3) and
\[ \gamma_j = \int_0^r l_j(x) \, dx - \sum_{k=2}^r \alpha_k \int_0^{k-1} l_j(x) \, dx, \quad j = 1, 2, \ldots, s; \]
\[ a_{ij} = \frac{\alpha_j}{\gamma_i} \int_{j-1}^r l_i(x) \, dx, \quad i = 1, 2, \ldots, s, \quad j = 1, 2, \ldots, r; \]
\[ b_{ij} = \int_0^{c_i} l_j(x) \, dx - \sum_{k=2}^r a_{ik} \int_0^{k-1} l_j(x) \, dx, \quad i, j = 1, 2, \ldots, s; \]
(1.5)
\[ l_j(x) = \frac{P(x)}{(x - c_j)P'(c_j)}, \quad j = 1, 2, \ldots, s; \]
\[ P(x) = \prod_{k=1}^s (x - c_k) = \det \begin{bmatrix} h_1 h_2 \cdots h_{s} h_{s+1} \\ h_2 h_3 \cdots h_{s+1} h_{s+2} \\ \vdots \\ h_s h_{s+1} \cdots h_{2s-1} h_{2s} \\ 1 x \cdots x^{s-1} x^s \end{bmatrix}; \]
\[ h_i = \frac{1}{i} (r^i - \alpha^T \xi^i), \quad i = 1, 2, \ldots, 2s, \]
satisfy all the hypotheses of Theorem 1.1, and they are all algebraically stable for the matrices \( G > 0, \ D > 0 \) defined by (1.4).

In 1988, the author of the present paper [10, 11] established the theory of \( B \)-convergence (\( B \)-theory) for general linear methods. We here only recall one of the basic principles:

**Theorem 1.3.** If a general linear method is \( BH \)-stable and \( BH \)- (resp. \( BH^* \)-) consistent of order \( p \), then this method is optimally \( B \)-convergent of order \( p \) (resp. \( B \)-convergent of order \( p \)).

In the present paper, the \( B \)-theory for general linear methods is applied to the special case of multistep Runge-Kutta methods. We first discuss the generalized stage order and diagonal stability of the methods (see Theorems 2.1-2.3); then, in view of \( B \)-theory and Theorems 1.1-1.3, a series of \( B \)-convergence results for multistep Runge-Kutta methods is obtained (see Theorems 2.4-2.7).

### 2. Main results and their proofs

**Definition 2.1.** The method (1.2) is said to be **diagonally stable**, if there exists an \( s \times s \) diagonal matrix \( Q > 0 \) such that \( QB + BTQ > 0 \).

**Definition 2.2.** The method (1.2) 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 (1.1) and stepsize \( h \in (0, h_0] \), there exist abstract functions \( Y^h \) and \( H^h \):

\[ Y^h(t) = (Y_1^h(t), Y_2^h(t), \ldots, Y_s^h(t)) \in X^s, \]
\[ H^h(t) = (H_1^h(t), H_2^h(t), \ldots, H_r^h(t)) \in X^r, \]
such that

\[ \| H^h(t) - H(t) \| \leq d_0 h^p, \quad \| \Delta^h(t) \| \leq d_1 h^{p+1}, \]
\[ \| \delta^h(t) \| \leq d_2 h^{p+1}, \quad \| \sigma^h(t) \| \leq d_3 h^p, \]

where \( h > 0 \) 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 on bounds \( M_i \) of some derivatives of the
exact solution \( y(t) \): \( \| d^i y(t)/dt^i \| \leq M_i, \; t \in [0, T] \); \( \Delta^h(t), \; \delta^h(t), \) and \( \sigma^h(t) \)
are determined by the equations

\[
\begin{aligned}
Y^h(t) &= \tilde{A} H^h(t) + h \tilde{B} F(Y^h(t)) + \Delta^h(t), \\
H^h(t) &= \tilde{C} Y^h(t) + h \tilde{E} F(Y^h(t)) + \delta^h(t), \\
y(t + rh) &= \beta H^h(t) + \sigma^h(t); \\
\end{aligned}
\]

the norm \( \| \cdot \| \) on \( X^N \) \((N \geq 1)\) is defined by

\[
\| U \| = \left( \sum_{i=1}^{N} \| u_i \|^2 \right)^{1/2}, \quad \forall U = (u_1, u_2, \ldots, u_N) \in X^N.
\]

Furthermore, if the quantities \( d_i \) \((i = 0, 1, 2, 3)\) are also allowed to depend
on bounds \( \kappa_i \) for certain derivatives of the mapping \( f \) (but not on \( \kappa_1)\):
\( \| d^i f(y)/dy^i \| \leq \kappa_i, \; y \in X \), then the aforementioned integer \( p \) is known
as generalized weak stage order of the method. For the special case where
\( H^h(t) \equiv H(t) \), the generalized stage order and generalized weak stage order are
simply called stage order and weak stage order, respectively.

Note that these two definitions follow from related previous papers, such as
[2, 5, 6, 7, 11].

**Theorem 2.1.** The method \((1.2)-(1.3)\) has stage order not smaller than \( \tau \) if
\[ \sum_{j=1}^{\tau} \alpha_j = 1, \sum_{j=1}^{\tau} a_{ij} = 1, \; i = 1, 2, \ldots, s, \; \text{and the conditions } B(\tau), C(\tau) \]
hold.

**Proof.** Let \( H^h(t) = H(t), \; Y^h(t) = Y(t) \). Substituting this in \((2.1)\), we get by
Taylor expansion

\[
\begin{aligned}
[\Delta^h(t)]_i &= \sum_{p=1}^{r} \frac{h^p}{p!} \left( c_i^p - \sum_{j=1}^{r} a_{ij} (j-1)^{p-1} \sum_{j=1}^{s} b_{ij} c_j^{p-1} \right) y^{(p)}(t) + R_{\Delta^h}(t), \\
[\delta^h(t)]_i &= \sum_{p=1}^{r} \frac{h^p}{p!} \left( r^p - \sum_{j=1}^{r} \alpha_j (j-1)^{p-1} \sum_{j=1}^{s} \gamma_j c_j^{p-1} \right) y^{(p)}(t) + R_{\delta^h}(t); \\
[\sigma^h(t)]_i &= 0, \; i = 1, 2, \ldots, r-1; \quad \sigma^h(t) = 0; \quad H^h(t) - H(t) = 0,
\end{aligned}
\]

where
(2.3) \[ R_{\tau}(t) = \int_0^t \left[ \frac{(1 - \theta)^{r}}{r!} \left( c_i^{\tau+1} y^{(\tau+1)}(t + \theta c_i h) \right) \right. \]
\[ \left. - \sum_{j=1}^{r} a_{ij} (j - 1)^{\tau+1} y^{(\tau+1)}(t + \theta (j - 1)h) \right] \]
\[ \left. - \frac{(1 - \theta)^{\tau-1}}{(\tau - 1)!} \sum_{j=1}^{s} b_{ij} c_i^{\tau} y^{(\tau+1)}(t + \theta c_j h) \right] h^{\tau+1} d\theta, \]
\[ i = 1, 2, \ldots, s; \]
\[ R_{\tau}(t) = \int_0^t \left[ \frac{(1 - \theta)^{r}}{r!} \left( r^{\tau+1} y^{(\tau+1)}(t + \theta r h) \right) \right. \]
\[ \left. - \sum_{j=1}^{r} \alpha_j (j - 1)^{\tau+1} y^{(\tau+1)}(t + \theta (j - 1)h) \right] \]
\[ \left. - \frac{(1 - \theta)^{\tau-1}}{(\tau - 1)!} \sum_{j=1}^{s} y_j c_i^{\tau} y^{(\tau+1)}(t + \theta c_j h) \right] h^{\tau+1} d\theta, \]
and therefore
(2.4) \[ \| R_{\tau}(t) \| \leq k_1 h^{\tau+1} M_{\tau+1}, \quad \| R_{\tau}(t) \| \leq k_1 h^{\tau+1} M_{\tau+1}, \]
where \( k_{i\tau} \) \((i = 1, 2, \ldots, s)\) and \( k_\tau \) depend only on the method. Thus, using the conditions \( B(\tau) \) and \( C(\tau) \), we get the conclusion from (2.2), (2.4), and Definition 2.2. \( \square \)

**Theorem 2.2.** Suppose the method (1.2)–(1.3) satisfies the conditions \( B(\tau + 1) \), \( C(\tau) \), and \( \sum_{j=1}^{r} \alpha_j = 1, \sum_{j=1}^{r} a_{ij} = 1, \ i = 1, 2, \ldots, s. \) Then
(i) this method has weak stage order not smaller than \( \tau + 1 \);
(ii) if there exists a real number \( \nu \) such that
\[ c^{\tau+1} - A^\tau - (\tau + 1)Bc^\tau = \nu e_\tau, \]
then this method has generalized stage order not smaller than \( \tau + 1 \).

**Proof.** Let
\[ H_i^h(t) = y(t + ih) + \delta h^{\tau+1} y^{(\tau+1)}(t), \quad i = 1, 2, \ldots, r; \]
\[ Y_i^h(t) = y(t + c_i h) + \mu_i h^{\tau+1} y^{(\tau+1)}(t), \quad i = 1, 2, \ldots, s, \]
where \( \mu_i \) and \( \delta \) are constants to be determined. Substituting this in (2.1), expanding into Taylor series, and using the conditions \( B(\tau + 1) \) and \( C(\tau) \), we get
(2.6a) \[ [\Delta^h(t)]_i = \left[ \frac{1}{(\tau + 1)!} (c^{\tau+1} - A^\tau - (\tau + 1)Bc^\tau) + \mu - \delta e_\tau \right] h^{\tau+1} y^{(\tau+1)}(t) \]
\[ + \delta h^{\tau+2} \int_0^t y^{(\tau+2)}(t - \theta h) d\theta + R_{i, \tau+1}(t) \]
\[ + h \sum_{i=1}^{s} b_{ij} Q_j(t; \mu, \tau, h), \quad i = 1, 2, \ldots, s; \]
\[ [\delta^h(t)]_r = \delta h^{r+2} \int_0^1 y^{(r+2)}(t - \theta h) \, d\theta + R_{r+1}(t) \]

\[ + h \sum_{j=1}^s \gamma_j Q_j(t; \mu, \tau, h); \]

\[ [\delta^h(t)]_i = \delta h^{r+2} \int_0^1 y^{(r+2)}(t - \theta h) \, d\theta, \quad i = 1, 2, \ldots, r-1; \]

\[ \sigma^h(t) = -\delta h^{r+1} y^{(r+1)}(t); \]

\[ [H^h(t) - H(t)]_i = \delta h^{r+1} y^{(r+1)}(t), \quad i = 1, 2, \ldots, r, \]

where

\[ \mu = [\mu_1, \mu_2, \ldots, \mu_s]^T, \]

\[ Q_j(t; \mu, \tau, h) = f(y(t + c_j h)) - f(y(t + c_j h) + \mu_j h^{r+1} y^{(r+1)}(t)), \]

and \( R_i, \tau_{i+1}(t), R_{\tau+1}(t) \) are given by (2.3). Therefore, we have

\[ \|H^h(t) - H(t)\| \leq \sqrt{r}|\delta h^{r+1} M_{r+1}|, \quad \|\sigma^h(t)\| \leq |\delta h^{r+1} M_{r+1}|, \]

\[ \|[\delta^h(t)]_i\| \leq |\delta h^{r+2} M_{r+2}|, \quad i = 1, 2, \ldots, r-1, \]

and by Taylor expansion,

\[ Q_j(t; \mu, \tau, h) \]

\[ = -\mu_j h^{r+1} \left\{ f'(y(t)) y^{(r+1)}(t) \right. \]

\[ + \int_0^1 [f''((1 - \theta)y(t) + \theta y(t + c_j h) - y(t)) \]

\[ + (1 - \theta) \mu_j h^{r+1} f''(y(t + c_j h) + \theta \mu_j h^{r+1} y^{(r+1)}(t)) y^{(r+1)}(t)] \]

\[ \times y^{(r+1)}(t) \, d\theta \}. \]

By the technique in [7], we can easily prove that

\[ \|f'(y(t)) y^{(r+1)}(t)\| \leq N_\tau \]

with \( N_\tau \) depending only on some bounds \( M_i \) and \( \kappa_i \) (but not on \( \kappa_1 \)). The relations (2.9) and (2.10) lead to

\[ \|Q_j(t; \mu, \tau, h)\| \leq N_{\mu_\tau} h^{r+1}, \quad 0 < h \leq h_0, \]

where the constant \( h_0 \) only need to satisfy the requirement mentioned in Definition 2.2, and \( N_{\mu_\tau} \) depends only on the method and on some bounds \( M_i \) and \( \kappa_i \) (but not on \( \kappa_1 \)). Now choose

\[ \delta = 0, \quad \mu = -\frac{1}{(\tau + 1)!} (c^{r+1} - A^r_{\tau+1} - (\tau + 1) B c^r). \]
B-CONVERGENCE OF RUNGE-KUTTA METHODS

Then the relations (2.4), (2.6a), (2.6b), and (2.11) lead to

\[
\begin{align*}
\|[\Delta^h(t)]_r \| \leq & \left( k_{i,\tau+1} M_{\tau+2} + N_{\mu_1} \sum_{j=1}^{s} |b_{ij}| \right) h^{r+2}, \quad i = 1, 2, \ldots, s, \\
\|[\delta^h(t)]_r \| \leq & \left( k_{\tau+1} M_{\tau+2} + N_{\mu_1} \sum_{j=1}^{s} |\gamma_j| \right) h^{r+2},
\end{align*}
\]

provided that \( h \in (0, \; h_0] \). Thus, it is easily seen from (2.8), (2.12), and Definition 2.2 that the method (1.2)–(1.3) has weak stage order not smaller than \( \tau + 1 \).

Furthermore, if the additional condition (2.5) is satisfied, then we would instead choose \( \mu = 0 \) and \( \delta = \nu/(\tau + 1)! \). In this case, (2.4), (2.6a), (2.6b), and (2.7) lead to

\[
\begin{align*}
\|[\Delta^h(t)]_r \| \leq & (|\nu|/(\tau + 1)! + k_{i,\tau+1}) h^{r+2} M_{\tau+2}, \quad i = 1, 2, \ldots, s, \\
\|[\delta^h(t)]_r \| \leq & (|\nu|/(\tau + 1)! + k_{\tau+1}) h^{r+2} M_{\tau+2},
\end{align*}
\]

and it follows from (2.8), (2.13), and Definition 2.2 that the method (1.2)–(1.3) has generalized stage order not smaller than \( \tau + 1 \). \( \square \)

**Theorem 2.3.** Suppose the method (1.2)–(1.3) satisfies the conditions \( B(2s) \), \( C(s) \), and \( E(s) \), \( r \geq s \), \( c_i \neq c_j \) whenever \( i \neq j \), \( \sum_{j=1}^{r} \alpha_j = 1 \) and \( \alpha_j > 0 \), \( j = 1, 2, \ldots, r \). Then this method is diagonally stable.

This theorem was first proved in 1989 by the author and his post-graduate student Cao Xuenian in a research report “\( BH \)-algebraic stability of general multivalue methods” at Xiangtan University. In the following we give an alternative proof.

Let \( Q = \text{diag}(\gamma_1, \gamma_2, \ldots, \gamma_s) \). Then it is seen from Theorem 1.1 that \( Q > 0 \).

Thus, we only need to prove \( QB + B^TQ > 0 \). Let \( p_l(x) = \prod_{k=0}^{l-1} (x - k) \), \( l = 1, 2, \ldots, s \). Making a congruence transform based on the transformation matrix

\[
V = \begin{bmatrix}
\rho_1'(c_1) & \rho_2'(c_1) & \cdots & \rho_s'(c_1) \\
\rho_1'(c_2) & \rho_2'(c_2) & \cdots & \rho_s'(c_2) \\
\cdots & \cdots & \cdots & \cdots \\
\rho_1'(c_s) & \rho_2'(c_s) & \cdots & \rho_s'(c_s)
\end{bmatrix},
\]

and using the conditions \( B(2s) \), \( C(s) \), and \( E(s) \), with the technique in [1] we obtain

\[
V^T(QB + B^TQ)V = [\delta_{l,m}],
\]

where
\[\delta_{lm} = \sum_{i=1}^{s} \gamma_i \rho'_i(c_i) \sum_{j=1}^{s} b_{ij} \rho'_m(c_j) + \sum_{i=1}^{s} \gamma_i \rho'_m(c_i) \sum_{j=1}^{s} b_{ij} \rho'_i(c_j)\]
\[= \sum_{i=1}^{s} \gamma_i [\rho_i(x) \rho_m(x)]'_{x=c_i}\]
\[- \sum_{j=1}^{r} \rho_m(j-1) \sum_{i=1}^{s} \gamma_i a_{ij} \rho'_i(c_i) - \sum_{j=1}^{r} \rho_i(j-1) \sum_{i=1}^{s} \gamma_i a_{ij} \rho'_m(c_i)\]
\[= \rho_l(r) \rho_m(r) - \sum_{j=1}^{r} \alpha_j \rho_l(j-1) \rho_m(j-1)\]
\[- \sum_{j=1}^{r} \rho_m(j-1) \alpha_j [\rho_l(r) - \rho_l(j-1)] - \sum_{j=1}^{r} \rho_l(j-1) \alpha_j [\rho_m(r) - \rho_m(j-1)]\]
\[= \rho_l(r) \rho_m(r) - \sum_{j=1}^{r} \alpha_j \rho_l(r) \rho_m(j-1) - \sum_{j=1}^{r} \alpha_j \rho_m(r) \rho_l(j-1)\]
\[+ \sum_{j=1}^{r} \alpha_j \rho_l(j-1) \rho_m(j-1), \quad l, m = 1, 2, \ldots, s.\]

Let
\[R = \begin{bmatrix} \alpha_2 & 0 & \cdots & -\alpha_2 \\ \alpha_3 & -\alpha_3 \\ \vdots \\ -\alpha_2 & -\alpha_3 & \cdots & -\alpha_r \end{bmatrix}, \quad U = \begin{bmatrix} \rho_1(1) & \rho_2(1) & \cdots & \rho_s(1) \\ \rho_1(2) & \rho_2(2) & \cdots & \rho_s(2) \\ \vdots \\ \rho_1(r) & \rho_2(r) & \cdots & \rho_s(r) \end{bmatrix}.\]

It is readily verified that the \((l, m)\)-element of the matrix \(U^T RU\) is also equal to \(\delta_{lm}\), \(l, m = 1, 2, \ldots, s\). Therefore,
\[V^T(QB + B^TQ)V = U^T RU.\]

Since \(\sum_{j=1}^{r} \alpha_j = 1\) and \(\alpha_j > 0\), \(j = 1, 2, \ldots, r\), for any given
\[x = [x_1, x_2, \ldots, x_r]^T \neq 0\]
we have
\[x^T Rx = \sum_{i=1}^{r-1} \alpha_{i+1} x_i^2 + x_r^2 - 2x_r \sum_{i=1}^{r-1} \alpha_{i+1} x_i\]
\[\geq \alpha_1 \sum_{i=1}^{r-1} \alpha_{i+1} x_i^2 + \left(x_r - \sum_{i=1}^{r-1} \alpha_{i+1} x_i\right)^2 > 0.\]

Thus, \(R > 0\). Since \(r \geq s\) and \(c_1, c_2, \ldots, c_2\) are distinct, \(\text{rank}(V) = \text{rank}(U) = s\), and therefore the conclusion \(QB + B^TQ > 0\) follows from (2.14) and \(R > 0\). \(\Box\)
In view of the $B$-theory for general linear methods (cf. [11]), a combination of Theorems 2.1–2.3 and 1.1–1.3 yields the following results:

**Theorem 2.4.** Suppose the method (1.2)–(1.3) is algebraically stable and diagonally stable, and satisfies $B(\tau)$, $C(\tau)$, $\sum_{j=1}^{s} \alpha_j = 1$, and $\sum_{j=1}^{r} a_{ij} = 1$, $i = 1, 2, \ldots, s$. Then this method is optimally $B$-convergent of order at least $\tau$.

**Theorem 2.5.** Suppose the method (1.2)–(1.3) is algebraically stable and diagonally stable, and satisfies $B(\tau + 1)$, $C(\tau)$, $\sum_{j=1}^{s} \alpha_j = 1$, and $\sum_{j=1}^{r} a_{ij} = 1$, $i = 1, 2, \ldots, s$. Then

(i) this method is $B$-convergent of order $\tau + 1$;

(ii) if there exists a real number $v$ such that (2.5) holds, then this method is optimally $B$-convergent of order $\tau + 1$.

**Theorem 2.6.** Suppose the method (1.2)–(1.3) satisfies the conditions $B(w)$, $C(\eta)$, and $E(\xi)$, $r, \eta, \xi \geq s$, $w \geq 2s$, $c_i \neq c_j$ whenever $i \neq j$, $\sum_{j=1}^{s} \alpha_j = 1$, and $\alpha_j > 0$, $j = 1, 2, \ldots, r$. Then

(i) this method is optimally $B$-convergent of order at least $\min\{w, \eta\}$;

(ii) this method is $B$-convergent of order $\min\{w, \eta + 1\}$;

(iii) if there exists a real number $v$ such that (2.5) holds with $\tau = \eta$, then this method is optimally $B$-convergent of order $\min\{w, \eta + 1\}$.

**Theorem 2.7.** Suppose $r \geq s$, $\sum_{j=1}^{r} \alpha_j = 1$, and $\alpha_j > 0$, $j = 1, 2, \ldots, r$. Then the multistep Runge-Kutta methods defined by (1.2), (1.3), and (1.5) are all optimally $B$-convergent of order at least $s$ and $B$-convergent of order $s + 1$.

**Remark 1.** Specializing Theorems 2.4 and 2.5 to the case of $r = 1$, we obtain immediately the well-known related results for Runge-Kutta methods presented by Frank et al. [6, 7] and Burrage and Hundsdorfer [2].

**Remark 2.** Specializing Theorem 2.6 to the case of $r = 1$, we obtain immediately the well-known result that the implicit midpoint rule is optimally $B$-convergent of order 2 (cf. [9, 10]).

**Remark 3.** For existence and uniqueness of the solution to the equation (1.2a), we refer to [12]; if the space $X$ is of finite dimension, see also [3, 4, 5, 7, 8].

### 3. Some examples

**Example 1.** Consider the $r$-step one-stage multistep Runge-Kutta method

$$
\begin{align*}
Y &= \sum_{j=1}^{r} a_j y_{n-1+j} + hbf(Y), \\
y_{n+r} &= \sum_{j=1}^{r} \alpha_j y_{n-1+j} + h\gamma f(Y),
\end{align*}
$$

or equivalently,

$$
\begin{align*}
y_{n+r} &= \sum_{j=1}^{r} \alpha_j y_{n-1+j} + h\gamma f \left( \beta y_{n+r} + \sum_{j=1}^{r} (a_j - \beta \alpha_j)y_{n-1+j} \right),
\end{align*}
$$

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
where

\[ r \geq 1, \quad \gamma = r - \sum_{j=1}^{r} \alpha_j (j - 1), \quad a_j = \frac{\alpha_j}{\gamma} (r + 1 - j), \quad j = 1, 2, \ldots, r, \]

\[ b = \frac{1}{2\gamma} \left[ r^2 - \sum_{j=1}^{r} \alpha_j (j - 1)(2r + 1 - j) \right], \quad \beta = \frac{b}{\gamma}, \]

the real parameters \( \alpha_1, \alpha_2, \ldots, \alpha_r \) satisfy \( \sum_{j=1}^{r} \alpha_j = 1 \) and \( \alpha_j > 0, \quad j = 1, 2, \ldots, r \). It is easily seen that the method satisfies the assumptions of Theorem 2.6 with \( w = 2 \) and \( \eta = \xi = 1 \), and the condition (2.5) with \( \tau = 1 \) is trivial since \( s = 1 \). Therefore, in view of Theorem 2.6, the method (3.1) or (3.2) is optimally \( B \)-convergent of order 2.

**Example 2.** For \( r = s = 2 \), the coefficients of a series of methods which satisfy the assumptions of Theorem 2.7 have been computed; some of them are as follows:

(i) \[
\alpha = \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix}, \quad \gamma = \begin{bmatrix} 0.8570633514 \\ 0.3929366486 \end{bmatrix}, \quad A = \begin{bmatrix} 0.2352842040 & 0.7647157960 \\ 0.7592738744 & 0.2407261256 \end{bmatrix}, \]

\[ B = \begin{bmatrix} 0.4290266119 \\ -0.1374873664 \end{bmatrix}, \quad C = \begin{bmatrix} 0.4682336621 \\ 0.5714724214 \end{bmatrix}; \]

(ii) \[
\alpha = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \quad \gamma = \begin{bmatrix} 0.9106438658 \\ 0.5893561342 \end{bmatrix}, \quad A = \begin{bmatrix} 0.5049603372 & 0.4950396628 \\ 0.9165272661 & 0.0834727339 \end{bmatrix}, \]

\[ B = \begin{bmatrix} 0.4553667456 \\ -0.1031014246 \end{bmatrix}, \quad C = \begin{bmatrix} 0.4991787090 \\ 0.7597573923 \end{bmatrix}; \]

(iii) \[
\alpha = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}, \quad \gamma = \begin{bmatrix} 0.9560446375 \\ 0.7939553625 \end{bmatrix}, \quad A = \begin{bmatrix} 0.7597573923 & 0.2402426077 \\ 0.9744099679 & 0.02559003213 \end{bmatrix}, \]

\[ B = \begin{bmatrix} 0.4782650437 \\ -0.08644371227 \end{bmatrix}, \quad C = \begin{bmatrix} 0.8750152175 \\ 0.5036002413 \end{bmatrix}; \]

However, for all these methods, condition (2.5) with \( \tau = 2 \) does not seem to be satisfied, so we can only conclude that these methods are optimally \( B \)-convergent of order 2 and \( B \)-convergent of order 3.

**Bibliography**


**Department of Mathematics, Xiangtan University, Hunan Province, People's Republic of China**