Computing 36, 35–41 (1986)

Computing

The Order of B-Convergence of the Gaussian Runge-Kutta Method

K. Dekker, J. F. B. M. Kraaijevanger, and **M. N. Spijker,** Leiden

Received April 5, 1985; revised July 16, 1985

Abstract — Zusammenfassung

The Order of B-Convergence of the Gaussian Runge-Kutta Method. In this note the exact order of B-convergence is determined for the m-stage Gaussian Runge-Kutta method. For $m=1$ this order is 2, whereas for $m \geq 2$ this order turns out to be only m.

AMS Subject Classifications: 65L05, 65L20.

Key words: Numerical analysis, stiff initial value problems, Runge-Kutta methods, B-convergence.

Die B-Konvergenzordnung des Gauss'schen Runge-Kutta Verfahrens. In dieser Note bestimmen wir die exakte B-Konvergenzordnung des m-stufigen Gauss'schen Runge-Kutta Verfahrens. Für $m=1$ ist diese Ordnung 2, während sich für $m \geq 2$ nur m als Ordnung ergibt.

1. Introduction

In this paper we deal with the numerical solution of the initial value problem

$$U'(t)=f(t, U(t)) \quad (0 \leq t \leq T),\ U(0)=u_0. \tag{1.1}$$

Here u_0 denotes a given vector in the n-dimensional complex vectorspace $\mathbb{C}^n$, and $f: \mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$ denotes a given function. It is assumed throughout that

$$\operatorname{Re}\langle f(t, \tilde{\xi})-f(t, \xi), \tilde{\xi}-\xi\rangle \leq 0 \quad (\text{for all } t \in \mathbb{R} \text{ and } \xi, \tilde{\xi} \in \mathbb{C}^n) \tag{1.2}$$

where $\langle \cdot, \cdot \rangle$ denotes an arbitrary, but fixed, inner product on $\mathbb{C}^n$.

If (1.1) has a solution $U(t)$ with a continuous j-th order derivative $U^{(j)}(t)$ (for $0 \leq t \leq T$) we use the notation

$$\| U \|_j = \max \left\{ | U^{(i)}(t)| : 0 \leq t \leq T \text{ and } i=0,1,\ldots,j \right\}.$$

Here $|\xi|$ stands for $\sqrt{\langle \xi, \xi \rangle}$ when $\xi \in \mathbb{C}^n$.

The class of problems (1.1), (1.2) comprises (non-linear) initial value problems that can be *arbitrarily stiff* while $\| U \|_j$ is of moderate size (see e.g. [6], [7]).

The general Runge-Kutta method applied to (1.1) reads

$$y_i = u_k + h \sum_{j=1}^{m} a_{ij} f(t_k + c_j h, y_j) \quad (i=1,2,\ldots,m) \tag{1.3 a}$$

$$u_{k+1} = u_k + h \sum_{j=1}^{m} b_j f(t_k + c_j h, y_j) \tag{1.3b}$$

where a_{ij}, b_j are given real coefficients and $c_j = a_{j1} + a_{j2} + \ldots + a_{jm}$ $(j = 1, 2, \ldots, m)$. With $h > 0$ we denote the stepsize, $t_k = k\,h$ and the vectors $u_k \cong U(t_k)$ are computed by applying (1.3) successively for $k = 0, 1, 2, \ldots$.

The stability properties of method (1.3) with respect to problems (1.1) satisfying (1.2) were investigated in [2], [4], where the concept of *B-stability* was introduced. A convergence property related to B-stability has been given in the important paper [7]. We formulate a slightly adapted definition:

Definition 1.1: Method (1.3) is *B-convergent of exact order p* if p is the largest number for which there are constants $\gamma > 0$, $h_0 > 0$, $j > 0$ such that

$$|U(t_k) - u_k| \le \gamma h^p \cdot \|U\|_j \quad (\text{for } k \ge 0,\ 0 < t_k \le T,\ 0 < h \le h_0) \tag{1.4}$$

whenever the u_k satisfy (1.3) and $U(t)$ is a solution to (1.1) with a continuous j-th order derivative.

In this definition γ, h_0, j may only depend on a_{ij}, b_j, T (and not on f, u_0, n or the inner product $\langle \cdot, \cdot \rangle$). The right-hand member of (1.4) is thus independent of the stiffness of (1.1).

The purpose of this paper is to study the exact order of B-convergence for the well-known Gaussian Runge-Kutta method (which is characterized at the beginning of section 4).

In section 2 we review some known results on orders of convergence of the Gaussian method, and we formulate our main result in theorem 2.3. Our proof of this theorem is based on lemma 2.2 which is also formulated in section 2.

In section 3 we prove a simple lemma giving upper bounds for orders of convergence for the case of arbitrary Runge-Kutta methods of type (1.3).

In section 4 we turn again to the Gaussian method. We prove our key lemma 2.2 by using the material from section 3 and a few special properties of the Gaussian method.

2. Known Results and New Results on Orders of Convergence of the Gaussian Runge-Kutta Methods

We first list the results on B-convergence known to us from the literature.

In [7] it was proved that the m-stage Gaussian method satisfies (1.4) with $p = m$.

In [8], [10] this order result was improved for the case where $m = 1$. It was shown that the Gaussian method with $m = 1$ (the implicit midpoint rule) satisfies (1.4) with $p = 2$.

In order to continue our review of the relevant literature we have to consider first the simple scalar test problem

$$U'(t)=\lambda\big(U(t)-g(t)\big)+g'(t) \quad (0\leq t\leq T),\ U(0)=u_0. \tag{2.1}$$

Here $u_0=g(0)$, $\lambda\in\mathbb{C}$, $\operatorname{Re}\lambda\leq 0$, and $g:\mathbb{R}\to\mathbb{C}$ is a given differentiable function.

Definition 2.1: Method (1.3) is *A-convergent of order p* if there are constants $\gamma>0$, $h_0>0, j>0$ such that (1.4) holds whenever the u_k are generated by applying method (1.3) to problem (2.1) and $U(t)$ is a solution to (2.1) with a continuous j-th order derivative.

In this definition γ, h_0, j may again only depend on a_{ij}, b_j, T (and not on $\lambda\in\mathbb{C}$ with $\operatorname{Re}\lambda\leq 0$ or on g).

In [1] arguments were presented which strongly suggest that the Gaussian methods are A-convergent of order $(m+1)$ (for all odd $m\geq 1$), and not A-convergent of order $>m$ (for all even $m\geq 2$).

In the present paper we arrive at the following lemma contrasting with the above suggestion.

Lemma 2.2: *For $m=1$ the Gaussian Runge-Kutta method is not A-convergent of an order >2, whereas for $m\geq 2$ this method is not A-convergent of an order $>m$.*

A combination of this lemma with the results on B-convergence reviewed above easily leads to the following final theorem.

Theorem 2.3: *For $m=1$ the Gaussian Runge-Kutta method is B-convergent of exact order 2. For $m\geq 2$ this method is B-convergent of exact order m.*

Note: Suppose definition 1.1 would be modified so as to deal only with initial value problems (1.1), (1.2) where $u_0\in\mathbb{R}^n, f:\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$. Then the corresponding modified version of theorem 2.3 remains true. This follows from lemma 2.2 by identifying $\mathbb{C}$ in the usual way with $\mathbb{R}^2$.

3. Upper Bounds for Orders of A-Convergence for Arbitrary Runge-Kutta Methods

We assume that $q\geq 0$ is an integer such that

$$(c_i)^k-k\sum_{j=1}^{m}a_{ij}(c_j)^{k-1}=\begin{cases}0 & (\text{for } 1\leq i\leq m, 1\leq k\leq q)\\ s_i & (\text{for } 1\leq i\leq m, k=q+1),\end{cases} \tag{3.1 a}$$

$$1-k\sum_{j=1}^{m}b_j(c_j)^{k-1}=\begin{cases}0 & (\text{for } 1\leq k\leq q)\\ s_0 & (\text{for } k=q+1).\end{cases} \tag{3.1 b}$$

Here $s_0, s_1, \ldots, s_m\in\mathbb{R}$ and we write $s=(s_1, s_2, \ldots, s_m)^T\in\mathbb{R}^m$.

In the following we deal with the test problem (2.1) where $g(t)=t^{q+1}$. Thus we have

$$U(t)=t^{q+1},\ f(t,\xi)=\lambda\xi+\big(g'(t)-\lambda g(t)\big),$$

and one easily sees that

$$U(t_k+c_i h)=U(t_k)+h\sum_{j=1}^{m} a_{ij} f(t_k+c_j h, U(t_k+c_j h))+h^{q+1} s_i \;\; (1\le i\le m), \tag{3.2a}$$

$$U(t_{k+1})=U(t_k)+h\sum_{j=1}^{m} b_j f(t_k+c_j h, U(t_k+c_j h))+h^{q+1} s_0. \tag{3.2b}$$

Subtracting the relations (1.3) from (3.2) we obtain a recurrence relation for the errors

$$d_k=U(t_k)-u_k.$$

This recurrence relation can be written in the form

$$d_{k+1}=\phi(\zeta)\,d_k+h^{q+1}\{s_0+\psi(\zeta)\,s\} \tag{3.3}$$

where we use the notations

$$\phi(\zeta)=1+\zeta\, b^T(I-\zeta A)^{-1} e,\ \psi(\zeta)=\zeta\, b^T(I-\zeta A)^{-1},\ \zeta=h\lambda,$$
$$b=(b_1,b_2,\ldots,b_m)^T\in\mathbb{R}^m,\ e=(1,1,\ldots,1)^T\in\mathbb{R}^m,\ A=(a_{ij}).$$

Lemma 3.1: *Let* $q\ge 0$ *be such that* (3.1) *holds.*

a) Assume there is a $\zeta\in\mathbb{C}$ *such that*

$$s_0+\psi(\zeta)\,s\ne 0 \text{ and } \operatorname{Re}\zeta\le 0. \tag{3.4}$$

Then there is no A-convergence of order $>q+1$.

b) Assume there is a $\zeta\in\mathbb{C}\cup\{\infty\}$ *such that*

$$\phi(\zeta)=1,\ s_0+\psi(\zeta)\,s\ne 0 \text{ and } \operatorname{Re}\zeta\le 0 \ (\text{if } \zeta\ne\infty). \tag{3.5}$$

Then there is no A-convergence of order $>q$.

Proof:

a) We write $\sigma=s_0+\psi(\zeta)\,s$. Let h be given with $0<h\le T$. Applying (3.3) with $k=0$, $\lambda=h^{-1}\zeta$ we obtain $|d_1|=h^{q+1}|\sigma|$. Assuming A-convergence of order p there follows $\gamma h^p\|g\|_j\ge h^{q+1}|\sigma|$ (for $0<h\le h_0$). Letting $h\to 0$ we arrive at $p\le q+1$.

b) First assume (3.5) with $\zeta\ne\infty$. Let k be an integer ≥ 1, $h=Tk^{-1}$. Applying (3.3) with $\lambda=h^{-1}\zeta$ we have $d_k=k\,h^{q+1}\sigma$. Consequently, A-convergence of order p implies $\gamma h^p\|g\|_j\ge Th^q|\sigma|$ (for $k\ge T(h_0)^{-1}$). Letting $k\to\infty$ we arrive at $p\le q$. Next assume $\zeta=\infty$. Choose $\tilde\zeta\in\mathbb{C}$, $\operatorname{Re}\tilde\zeta\le 0$, $\tilde\zeta\cong\infty$. Applying (3.3) with $\lambda=h^{-1}\tilde\zeta$ we have $d_{k+1}=\phi(\tilde\zeta)\,d_k+h^{q+1}\{s_0+\psi(\tilde\zeta)\,s\}$. When $\tilde\zeta\to\infty$ we have $\phi(\tilde\zeta)\to 1$, $s_0+\psi(\tilde\zeta)\,s\to\sigma\ne 0$ and consequently $d_k\to k\,h^{q+1}\sigma$. Assuming A-convergence of order p we again arrive at $p\le q$. □

Remark 3.2: The function $s_0+\psi(\zeta)\,s$ is also considered in [3], [5] and [9]. In [9] it is used to obtain upper bounds for the order of B-consistency of Runge-Kutta methods; in [3] the function occurs in conditions determining lower bounds for the order of B-convergence on a class of semi-linear problems. [5] gives error estimates in which $s_0+\psi(\zeta)\,s$ plays a role, for a special scalar differential equation.

4. An Analysis of the Gaussian Runge-Kutta Method

4.1. The Gaussian Method

Let $m \geq 1$ be a given integer. Then the corresponding Gaussian method is the unique method of type (1.3) satisfying (see e.g. [6], [7])

$$(c_i)^k - k \sum_{j=1}^{m} a_{ij} (c_j)^{k-1} = 0 \quad (1 \leq k \leq m, 1 \leq i \leq m) \tag{4.1 a}$$

$$1 - k \sum_{j=1}^{m} b_j (c_j)^{k-1} = 0 \quad (1 \leq k \leq 2m). \tag{4.1 b}$$

Consequently, (3.1) holds with $q = m$ and $s_0 = 0$ for the Gaussian method.

It is well known that the c_i are the zeros of the shifted Legendre polynomial of degree m, so that

$$0 < c_i < 1, \; c_i \neq c_j \; \text{(for } 1 \leq i \leq m, 1 \leq j \leq m, i \neq j). \tag{4.2}$$

Finally we note that the method has so-called classical order $p = 2m$ (see e.g. [6], [7]). In view of (4.1), (3.3) this implies

$$\psi(\zeta) s = O(\zeta^m) \quad (\text{for } \zeta \to 0). \tag{4.3}$$

4.2. The Function $s_0 + \psi(\zeta) s$

In view of condition (3.4) occurring in lemma 3.1 the following lemma (cf. also [9]) is useful.

Lemma 4.1: *For the Gaussian method we have $\psi(\infty) s \neq 0$*

Proof: Writing $c^j = ((c_1)^j, (c_2)^j, \ldots, (c_m)^j)^T \in \mathbb{R}^m$ we have from (3.1 a) (with $q = m$) the expression $s = c^{m+1} - (m+1) A c^m$. Consequently

$$\psi(\infty) s = -b^T A^{-1} (c^{m+1} - (m+1) A c^m) = (m+1) b^T c^m - b^T A^{-1} c^{m+1}.$$

In view of (3.1 b) (with $q = m$, $s_0 = 0$) there follows

$$\psi(\infty) s = 1 - b^T A^{-1} c^{m+1}. \tag{4.4}$$

We introduce the diagonal matrices $C = \mathrm{diag}(c_1, c_2, \ldots, c_m)$, $M = \mathrm{diag}(1, 2, \ldots, m)$ and the $m \times m$ Vandermonde matrix V with columns $c^0, c^1, \ldots, c^{m-1}$. From (3.1 a) we have $A V M = C V$, and therefore $b^T A^{-1} c^{m+1} = b^T V M V^{-1} C^{-1} c^{m+1}$. By (3.1 b) we have $b^T V M = e^T$, which implies

$$b^T A^{-1} c^{m+1} = e^T V^{-1} c^m. \tag{4.5}$$

Writing $r = (r_1, r_2, \ldots, r_m)^T = V^{-1} c^m$ we see that the polynomial $P(t) = t^m - (r_1 + r_2 t + \ldots + r_m t^{m-1})$ has the m distinct zeros $t = c_1, c_2, \ldots, c_m$. Since $P(t)$ has at most m different zeros we conclude that $P(1) \neq 0$. An application of (4.4), (4.5) completes the proof. □

4.3. *The Function* $\phi(\zeta)$

In view of requirement $\phi(\zeta)=1$ occurring in part b of lemma 3.1 the following lemma is useful.

Lemma 4.2: *For the Gaussian method there are precisely m different numbers* $\zeta_1, \zeta_2, \ldots, \zeta_m$ *with* $\zeta_j \in \mathbb{C} \cup \{\infty\}$, $\operatorname{Re} \zeta_j = 0$ *(if* $\zeta_j \neq \infty$*)*, $\phi(\zeta_j)=1$ *(for* $j=1,2,\ldots,m$*).*

Proof: It is well known (see e.g. [6]) that ϕ satisfies

$$\phi(\zeta)\,\phi(-\zeta) \equiv 1$$

and that ϕ has m different poles ζ, which all have $\operatorname{Re}\zeta > 0$. It follows that ϕ has m zeros $\alpha_1, \alpha_2, \ldots, \alpha_m$ with $\operatorname{Re}\alpha_j < 0$ $(1 \leq j \leq m)$. Consequently, the function f defined by

$$w = f(z) = \phi\,[(z-1)/(z+1)]$$

has m zeros $a_j = (1+\alpha_j)/(1-\alpha_j)$ with $|a_j| < 1$, while $|w| = 1$ when $|z| = 1$.

Let k denote the winding number with respect to $w=0$ of the curve $\Gamma: w = f(z)$, $z = \exp(i\,t)$ $(0 \leq t \leq 2\pi)$. By the principle of the argument we have $k = m$. Therefore there are precisely m different values $z_1, z_2, \ldots, z_m$ with $|z_j| = 1, f(z_j) = 1$. This proves the lemma with $\zeta_j = (z_j - 1)/(z_j + 1)$. □

4.4. *The Proof of Lemma* 2.2

First assume $m=1$. In view of (4.1) and lemma 4.1 we can apply lemma 3.1 a with $q=1$, $s_0=0$. It follows that the method is not A-convergent of an order >2.

Next assume $m \geq 2$. Since $\phi(0)=1$ we see from lemma 4.2 that there is some ξ with

$$\xi \in \mathbb{C} \cup \{\infty\},\ \xi \neq 0,\ \operatorname{Re}\xi = 0 \text{ (if } \xi \neq \infty),\ \phi(\xi) = 1.$$

From the definition of $\psi(\zeta)$ we see that, if $\psi(\zeta)\,s$ is no constant, the function $\psi(\zeta)\,s$ has at most m zeros (counted with their multiplicities). From (4.3) and lemma 4.1 we thus conclude that

$$\psi(\xi)\,s \neq 0.$$

An application of lemma 3.1 b completes the proof. □

References

[1] Ascher, V., Bader, G.: Stability of collocation at Gaussian points. Tech. Rep. 84-13, Departm. Comp. Sc. Univ. British Columbia, Vancouver (1984).

[2] Burrage, K., Butcher, J. C.: Stability criteria for implicit Runge-Kutta methods. SIAM J. Numer. Anal. *16*, 46–57 (1979).

[3] Burrage, K., Hundsdorfer, W. H., Verwer, J. G.: A study of *B*-convergence of Runge-Kutta methods. Computing (this issue).

[4] Crouzeix, M.: Sur la *B*-stabilité des méthodes de Runge-Kutta. Numer. Math. *32*, 75–82 (1979).

[5] Crouzeix, M., Raviart, P. D.: Méthodes de Runge-Kutta. Lecture notes, Université de Rennes (1980).

[6] Dekker, K., Verwer, J. G.: Stability of Runge-Kutta methods for stiff nonlinear differential equations. Amsterdam, New York, Oxford: North Holland Publ. Comp. (1984).

[7] Frank, R., Schneid, J., Ueberhuber, C. W.: Order results for implicit Runge-Kutta methods applied to stiff systems. SIAM J. Numer. Anal. *22,* 515–534 (1985).

[8] Kraaijevanger, J. F. B. M.: *B*-convergence of the implicit midpoint rule and the trapezoidal rule. Rep. 01-85, Inst. Appl. Math. Comp. Sc. Univ. Leiden (1985). To appear in BIT.

[9] Montijano, J. I.: Estudio de los metodos SIRK para la resolucion numérica de ecuaciones diferenciales de tipo stiff. Thesis, University of Zaragoza (1983).

[10] Stetter, H. J.: Zur *B*-Konvergenz der impliziten Trapez- und Mittelpunktregel. Unpublished note (1982).

K. Dekker, J. F. B. M. Kraaijevanger,
and M. N. Spijker
Institute of Applied Mathematics and
Computer Science
University of Leiden
Wassenaarseweg 80
2333 AL Leiden
The Netherlands