SIAM J. NUMER. ANAL.
Vol. 24, No. 3, June 1987

006

ON THE ALGEBRAIC EQUATIONS IN IMPLICIT RUNGE-KUTTA METHODS*

W. H. HUNDSDORFER† AND M. N. SPIJKER‡

Abstract. This paper is concerned with the system of (nonlinear) algebraic equations which arise in the application of implicit Runge-Kutta methods to stiff initial value problems. Without making the classical assumption that the stepsize $h>0$ is small, we derive transparent conditions on the method that guarantee existence and uniqueness of solutions to the equations. Besides, we discuss the sensitivity of the Runge-Kutta procedure with respect to perturbations in the algebraic equations.

Key words. numerical analysis, stiff initial value problems, implicit Runge-Kutta methods, nonlinear algebraic equations, stability

AMS(MOS) subject classifications. Primary 65L05; secondary 47H15, 65H10

1. Introduction. We shall deal with the numerical solution of the system of n ordinary differential equations

$$\frac{d}{dt}U(t)=f(t, U(t)) \qquad (t \geqq t_0), \tag{1.1}$$

under an initial condition $U(t_0)=u_0$. Here $t_0 \in \mathbb{R}$, $u_0 \in \mathbb{K}^n$ and $f:\mathbb{R}\times\mathbb{K}^n \to \mathbb{K}^n$ is a given continuous function. To cope simultaneously with real and with complex differential equations, the set $\mathbb{K}$ will stand consistently for either $\mathbb{R}$ or $\mathbb{C}$. Further, $\langle \cdot, \cdot \rangle$ is an arbitrary inner product on $\mathbb{K}^n$, and $|\xi| = \langle \xi, \xi \rangle^{1/2}$ (for $\xi \in \mathbb{K}^n$).

In order to introduce the problem treated in this article we assume

$$\mathrm{Re}\,\langle f(t,\tilde{\xi})-f(t,\xi), \tilde{\xi}-\xi\rangle \leqq 0 \qquad (\text{for all } t\in\mathbb{R} \text{ and } \tilde{\xi}, \xi \in \mathbb{K}^n). \tag{1.2}$$

This condition implies (cf. e.g. [9]) that for any two solutions U, $\tilde{U}$ to (1.1) the norm $|\tilde{U}(t)-U(t)|$ does not increase when t increases.

Let $h>0$ denote a stepsize and $t_k = t_{k-1}+h$ $(k=1, 2, 3, \cdots)$. Using an implicit Runge-Kutta method, approximations u_k to $U(t_k)$ are computed (for $k \geqq 1$) by

$$u_k = u_{k-1} + h\sum_{i=1}^{m} b_i f(t_{k-1}+c_i h, y_i), \tag{1.3a}$$

$$y_i = u_{k-1} + h\sum_{j=1}^{m} a_{ij} f(t_{k-1}+c_j h, y_j) \qquad (1 \leqq i \leqq m). \tag{1.3b}$$

Here $m \geqq 1$ and a_{ij}, b_j are real parameters, $c_i = a_{i1}+a_{i2}+\cdots+a_{im}$. We define the $m\times m$ matrices $A=(a_{ij})$, $B=\mathrm{diag}\,(b_1, b_2, \cdots, b_m)$ and the vector $b=(b_1, b_2, \cdots, b_m)^T \in \mathbb{R}^m$.

During these last years *algebraically stable* Runge-Kutta methods have gained much interest. These methods can be characterized by the property that B is positive

* Received by the editors December 3, 1984; accepted for publication (in revised form) October 14, 1985.
† Centre for Mathematic and Computer Science, Amsterdam, the Netherlands.
‡ Institute of Applied Mathematics and Computer Science, University of Leiden, Wassenaarseweg 80, Leiden, the Netherlands.

definite while $(BA+A^TB-bb^T)$ is positive semidefinite. In [1], [4] this property was shown to imply the important *contractivity* relation

$$|\tilde{u}_k-u_k|\leqq|\tilde{u}_{k-1}-u_{k-1}| \qquad (k\geqq 1),$$

for any two sequences $\{u_k\}$, $\{\tilde{u}_k\}$ computed from (1.3) with the same arbitrary stepsize $h>0$. However, algebraic stability does not guarantee that the system of algebraic equations (1.3b) has a solution for arbitrary $h>0$ (see [5]).

It was proved by Crouzeix (cf. [6], [5], [10]) that, whenever (1.2) is fulfilled and

(1.4) there is a positive definite diagonal matrix D such that $DA+A^TD$ is positive definite,

then the system (1.3b) does have a unique solution (for arbitrary $h>0$). Some well-known algebraically stable methods satisfy (1.4) (the Gauss methods, the Radau IA and IIA methods, the 2-stage Lobatto IIIC method—see [13]). But, e.g., the 3-stage Lobatto IIIC method is known to violate (1.4) (see [13], [10], [11], [12]).

The theory in the present paper provides a simple condition on A which is less restrictive than (1.4) and which still implies the existence of a unique solution to (1.3b) (for arbitrary $h>0$). The 3-stage Lobatto IIIC method fulfills this new condition.

In [2], [8], [3] contractivity (and stability) relations were derived under assumptions on f that are more general than assumption (1.2). Our main theorem on the existence of solutions to (1.3b) will also cope with f satisfying such generalized assumptions.

An important tool in obtaining our existence and unicity results consists in a study of the sensitivity of the solution of the algebraic equations with respect to (so-called internal) perturbations. As a by-product we thus shall obtain generalizations of results on this sensitivity already given in [13], [10], [12].

In § 2 we shall state and discuss our main result (Theorem 2.1) on the existence and uniqueness of solutions to (1.3b). In § 3 we derive the material that is basic for the proof of Theorem 2.1. We also apply this material in a study of the sensitivity of u_k (see (1.3)) with respect to internal perturbations. The final § 4 contains the proof of Theorem 2.1.

Remark 1.1. The Runge–Kutta step (1.3) is often written in the form

$$u_k=u_{k-1}+\sum_{i=1}^{m} b_ix_i, \tag{1.5a}$$

$$x_i=hf\left(t_{k-1}+c_ih,\ u_{k-1}+\sum_{j=1}^{m} a_{ij}x_j\right) \qquad (i\leqq i\leqq m). \tag{1.5b}$$

Our results on the existence of solutions to (1.3b) are also relevant to (1.5b), since (1.5b) has a unique solution iff (1.3b) has such a solution (see Lemma 4.1).

Remark 1.2. The results of this paper are also applicable to *general linear methods* (cf. [2]). The systems of algebraic equations arising in such methods are essentially of type (1.3b) (or (1.5b)).

2. Existence and uniqueness.

2.1. Formulation of the main theorem. Let α, β be given real constants. We consider the following three conditions on f, A and h.

(2.1) The function $f:\mathbb{R}\times\mathbb{K}^n\to\mathbb{K}^n$ is continuous, and
$\operatorname{Re}\langle f(t,\tilde{\xi})-f(t,\xi),\tilde{\xi}-\xi\rangle\leqq\alpha|f(t,\tilde{\xi})-f(t,\xi)|^2+\beta|\tilde{\xi}-\xi|^2$
(for all $t\in\mathbb{R}$ and $\xi,\tilde{\xi}\in\mathbb{K}^n$).

(2.2) There are real diagonal matrices $D=\operatorname{diag}(\delta_1,\delta_2,\cdots,\delta_m)$,

$S = \text{diag}(\sigma_1, \sigma_2, \cdots, \sigma_m)$ and $T = \text{diag}(\tau_1, \tau_2, \cdots, \tau_m)$ such that the matrix $DA + A^T D - S - A^T TA$ is positive semidefinite

(2.3) $\mathcal{M}_1$ and $\mathcal{M}_2$ are disjoint index sets with $\mathcal{M}_1 \cup \mathcal{M}_2 = \{1, 2, \cdots, m\}$;
$\delta_i \geqq 0$, $\sigma_i - 2h^{-1}\alpha\delta_i \geqq 0$, $\tau_i - 2h\beta\delta_i \geqq 0$ (if $1 \leqq i \leqq m$);
$\sigma_i - 2h^{-1}\alpha\delta_i > 0$ if either $i \in \mathcal{M}_1$ or ($i \in \mathcal{M}_2$ and $\alpha\delta_i \neq 0$);
$\tau_i - 2h\beta\delta_i > 0$ if either $i \in \mathcal{M}_2$ or ($i \in \mathcal{M}_1$ and $\beta\delta_i \neq 0$).

THEOREM 2.1. *Assume* (2.1), (2.2), (2.3). *Then the system* (1.3b) *has a unique solution* $y_1, y_2, \cdots, y_m \in \mathbb{K}^n$.

We note that the index sets occurring in condition (2.3) are allowed to be empty. Condition (2.1) on f is a generalization of the well-known one-sided Lipschitz condition (where $\alpha = 0$, see e.g. [1], [7], [13]) and of the circle condition in [9] (where $\beta = 0$). It was also used in [17], [8].

If $\alpha \geqq 0$, then there exist functions f satisfying (2.1) with arbitrarily large Lipschitz constants. It follows that initial value problems (1.1) are covered that can be arbitrarily stiff.

We conclude this section with a lemma which gives some more insight into condition (2.1) and which simplifies the application of the main Theorem 2.1. For given $\alpha, \beta \in \mathbb{R}$ we denote the class of functions f satisfying (2.1) by $\mathcal{F}(\alpha, \beta)$.

LEMMA 2.2. *Let* $\alpha, \beta \in \mathbb{R}$.

(a) *Suppose* $\beta_1 \in \mathbb{R}$, $\beta_1 > \beta$ *and* $\alpha \neq 0$. *Then there exists a number* $\alpha_1 < \alpha$ *such that* $\mathcal{F}(\alpha, \beta) \subset \mathcal{F}(\alpha_1, \beta_1)$.

(b) *Suppose* $\alpha_1 \in \mathbb{R}$, $\alpha_1 > \alpha$ *and* $\beta \neq 0$. *Then there exists a number* $\beta_1 < \beta$ *such that* $\mathcal{F}(\alpha, \beta) \subset \mathcal{F}(\alpha_1, \beta_1)$.

Proof. We shall only prove part (a) of this lemma. A proof of part (b) can be given along the same lines. Suppose first $\alpha < 0$ and $\beta_1 > \beta$. Let $f \in \mathcal{F}(\alpha, \beta)$, and let $t \in \mathbb{R}$, $\tilde{\xi}, \xi \in \mathbb{K}^n$ be arbitrary. Put $v = \tilde{\xi} - \xi$, $w = f(t, \tilde{\xi}) - f(t, \xi)$. We have

$$\text{Re}\langle v, w\rangle \leqq \alpha|w|^2 + \beta|v|^2.$$

Using the Schwarz inequality it follows that

$$\alpha|w|^2 + \beta|v|^2 + |w||v| \geqq 0.$$

Hence there is a $\gamma_0 > 0$ (only depending on α and β) such that

$$|w|^2 \leqq \gamma_0|v|^2.$$

Take $\alpha_1 < \alpha$ such that $(\beta_1 - \beta)/(\alpha - \alpha_1) \geqq \gamma_0$. We then have

$$\alpha|w|^2 + \beta|v|^2 \leqq \alpha_1|w|^2 + \beta_1|v|^2,$$

from which it is easily seen that $f \in \mathcal{F}(\alpha_1, \beta_1)$.

We now consider the case where $\alpha > 0$, $\beta_1 > \beta$. For any $\alpha_1 \in (\frac{1}{2}\alpha, \alpha)$ and $v, w \in \mathbb{K}^n$ satisfying

$$\text{Re}\langle v, w\rangle > \alpha_1|w|^2 + \beta_1|v|^2,$$

we have

$$|v||w| > \tfrac{1}{2}\alpha|w|^2 + \beta_1|v|^2.$$

It follows that there is a constant $\gamma_1 > 0$ (only depending on α and β_1) such that

$$|w|^2 \leqq \gamma_1|v|^2.$$

Take $\alpha_1 \in (\frac{1}{2}\alpha, \alpha)$ such that $(\beta_1 - \beta)/(\alpha - \alpha_1) \geqq \gamma_1$. Assume $f \in \mathcal{F}(\alpha, \beta)$ but $f \notin \mathcal{F}(\alpha_1, \beta_1)$. Then we know there are $t \in \mathbb{R}$ and $\tilde{\xi}, \xi \in \mathbb{K}^n$ such that

$$\alpha_1|w|^2 + \beta_1|v|^2 < \text{Re}\langle v, w\rangle \leqq \alpha|w|^2 + \beta|\nu|^2,$$

and

$$|w|^2 \leqq [(\beta_1 - \beta)/(\alpha - \alpha_1)]|v|^2$$

with $v = \tilde{\xi} - \xi$, $w = f(t, \tilde{\xi}) - f(t, \xi)$. This yields a contradiction. □

2.2. Application of the main theorem. From Theorem 2.1 one easily obtains

COROLLARY 2.3. *Assume* $f: \mathbb{R} \times \mathbb{K}^n \to \mathbb{K}^n$ *is continuous and satisfies* (1.2). *Suppose* (2.2) *holds with*

$$\delta_i \geqq 0, \quad \sigma_i \geqq 0, \quad \tau_i \geqq 0, \quad \sigma_i + \tau_i > 0 \quad (\textit{for } 1 \leqq i \leqq m).$$

Then (1.3b) *has a unique solution.*

When $\tau_i = 0$ $(1 \leqq i \leqq m)$, the corollary is proved by applying Theorem 2.1 with $\mathcal{M}_1 = \{1, 2, \cdots, m\}$, $\mathcal{M}_2 = \emptyset$, and when $\sigma_i = 0$ $(1 \leqq i \leqq m)$ it is proved with $\mathcal{M}_1 = \emptyset$, $\mathcal{M}_2 = \{1, 2, \cdots, m\}$. In the general case one can choose $\mathcal{M}_1 = \{i \mid \sigma_i > 0\}$, $\mathcal{M}_2 = \{i \mid \sigma_i = 0$ and $\tau_i > 0\}$.

The above corollary is a generalization of [6, Thm. 5.4], [5, Thm. 1] and [10, Lem. 4.2], where (1.4) was required. Condition (1.4) implies that the assumption on (2.2) in the corollary is fulfilled (with $\tau_i = 0$). On the other hand, (2.2) can be fulfilled with $\delta_i \geqq 0$, $\sigma_i \geqq 0$, $\tau_i \geqq 0$, $\sigma_i + \tau_i > 0$ while (1.4) is violated. An example of this situation is provided by the 3-stage Lobatto IIIC method referred to in the Introduction (see also § 2.3).

COROLLARY 2.4. *Let* $h > 0$ *and* $\alpha, \beta \in \mathbb{R}$ *be given. Suppose* $\kappa, \lambda \in \mathbb{R}$ *and* $D = \operatorname{diag}(\delta_1, \delta_2, \cdots, \delta_m)$ *are such that the matrix*

$$DA + A^T D - \kappa D - \lambda A^T D A$$

is positive semidefinite. Assume further $\delta_i > 0$ $(1 \leqq i \leqq m)$, $2\alpha h^{-1} \leqq \kappa$, $2\beta h \leqq \lambda$ *and* $2\alpha h^{-1} + 2\beta h < \kappa + \lambda$. *Then* (1.3b) *has a unique solution whenever f satisfies* (2.1).

Proof. For the cases $[2\alpha h^{-1} \leqq \kappa,\ 2\beta h < \lambda,\ \alpha \neq 0]$ and $[2\alpha h^{-1} < \kappa,\ 2\beta h \leqq \lambda,\ \beta \neq 0]$ the proof easily follows by combining Theorem 2.1 and Lemma 2.2. If $[2\alpha h^{-1} \leqq \kappa,\ 2\beta h < \lambda,\ \alpha = 0]$, Theorem 2.1 can be applied directly with $\mathcal{M}_1 = \emptyset$, and if $[2\alpha h^{-1} < \kappa,\ 2\beta h \leqq \lambda,\ \beta = 0]$, we take $\mathcal{M}_2 = \emptyset$ in Theorem 2.1. □

We note that if $\alpha = \kappa = 0$, the content of the above corollary reduces to a theorem formulated in [15, Thm. 4.3.1]. The latter theorem in its turn generalizes results on the system (1.3b) formulated in [12, Thms. 5.3.9, 5.3.12].

2.3. Examples.

Example 2.5. The algebraically stable, 3-stage Lobatto IIIC method is given by

$$A = \begin{pmatrix} 1/6 & -1/3 & 1/6 \\ 1/6 & 5/12 & -1/12 \\ 1/6 & 2/3 & 1/6 \end{pmatrix}, \qquad b = \begin{pmatrix} 1/6 \\ 2/3 \\ 1/6 \end{pmatrix}.$$

Condition (1.4) is not fulfilled (see e.g. [13]). However, with the choice $\delta_1 = 1$, $\delta_2 = 4$, $\delta_3 = 1$, $\tau_1 = 2$, $\sigma_2 = 2$, $\tau_3 = 2$ and the other τ_i, σ_i equal to zero, condition (2.2) is fulfilled. From Corollary 2.3 we thus see that (1.3b) always has a unique solution when f is continuous and satisfies (1.2).

We note that this Runge–Kutta method does not satisfy (2.2) with any $\delta_i \geqq 0$, $\sigma_i > 0$, $\tau_i = 0$ $(1 \leqq i \leqq m)$ or with $\delta_i \geqq 0$, $\sigma_i = 0$, $\tau_i > 0$ $(1 \leqq i \leqq m)$.

Example 2.6. Consider an arbitrary method that is algebraically stable. Applying Corollary 2.4 with $\kappa = \lambda = 0$, it follows that (1.3b) has a unique solution whenever f satisfies (2.1) with some $\alpha \leqq 0$, $\beta \leqq 0$, $\alpha + \beta < 0$ (which is a bit stronger than (1.2)). This result provides an extension of [6, Remark 5.7], [5, Cor. and Remark 3, p. 90].

Example 2.7. Consider a method satisfying (1.4). From Corollary 2.4 it can be seen that there exist κ_0, $\lambda_0 > 0$ such that (1.3b) has a unique solution for any $h > 0$ and f satisfying (2.1) with $\alpha h^{-1} \leqq \kappa_0$ and $\beta h \leqq \lambda_0$. This generalizes a related result on the system (1.3b) formulated in [12, Thms. 5.3.9, 5.3.12] where $\alpha = 0$ is assumed.

3. Stability with respect to internal perturbations.

3.1. Notation. For given column vectors $x_1, x_2, \cdots, x_m \in \mathbb{K}^n$ we denote the column vector $(x_1^T, x_2^T, \cdots, x_m^T)^T \in \mathbb{K}^{nm}$ by $[x_i]$. On the space $\mathbb{K}^{nm}$ we deal with the norm

$$\|x\| = (|x_1|^2 + |x_2|^2 + \cdots + |x_m|^2)^{1/2}$$

for $x = [x_i] \in \mathbb{K}^{nm}$, where $|\cdot|$ denotes the norm of § 1. For any linear mapping L from $\mathbb{K}^{nm}$ into $\mathbb{K}^{nm}$ we define $\|L\| = \sup\{\|Lx\|: x \in \mathbb{K}^{nm} \text{ with } \|x\| = 1\}$.

$\mathcal{M}_1$ and $\mathcal{M}_2$ are disjoint sets with $\mathcal{M}_1 \cup \mathcal{M}_2 = \{1, 2, \cdots, m\}$, and the projections $I_j: \mathbb{K}^{nm} \to \mathbb{K}^{nm}$ (for $j = 1, 2$) are defined by $I_j x = y$ for $x = [x_i]$ with $y = [y_i]$ given by

$$y_i = x_i \quad (\text{when } i \in \mathcal{M}_j), \qquad y_i = 0 \quad (\text{when } i \notin \mathcal{M}_j).$$

Let $u_{k-1} \in \mathbb{K}^n$, $h > 0$ and t_{k-1} be given. We define the functions $f_i: \mathbb{K}^n \to \mathbb{K}^n$ $(1 \leqq i \leqq m)$ and $F: \mathbb{K}^{nm} \to \mathbb{K}^{nm}$ by

$$f_i(\xi) = hf(t_{k-1} + c_i h, u_{k-1} + \xi) \quad (\text{for } \xi \in \mathbb{K}^n),$$

$$Fx = [f_i(x_i)] \quad (\text{for } x = [x_i] \in \mathbb{K}^{nm}).$$

Further we define $H: \mathbb{K}^{nm} \to \mathbb{K}^{nm}$ by $Hz = [h_i(z)]$ (for $z = [z_i] \in \mathbb{K}^{nm}$) with

$$h_i(z) = z_i - \sum_{j \in \mathcal{M}_1} a_{ij} f_j(z_j) - \sum_{j \in \mathcal{M}_2} a_{ij} z_j \quad (\text{if } i \in \mathcal{M}_1),$$

$$h_i(z) = z_i - f_i\left(\sum_{j \in \mathcal{M}_1} a_{ij} f_j(z_j) + \sum_{j \in \mathcal{M}_2} a_{ij} z_j\right) \quad (\text{if } i \in \mathcal{M}_2).$$

The $n \times n$ identity matrix is denoted by $I^{(n)}$ and the Kronecker product by $\otimes$. We define

$$\mathbf{b} = b \otimes I^{(n)}, \qquad \mathbf{A} = A \otimes I^{(n)}, \qquad \mathbf{a}_i = a_i \otimes I^{(n)}.$$

Here b, A are as in § 1, and a_i^T denotes the ith row of the matrix A (for $1 \leqq i \leqq m$).

We define the mappings (from $\mathbb{K}^{nm}$ to $\mathbb{K}^{nm}$)

$$F_j = I_j F, \qquad H_j = I_j H, \qquad \mathbf{A}_j = I_j \mathbf{A} \quad (\text{for } j = 1, 2).$$

Remark that, with $I = I_1 + I_2$ denoting the $nm \times nm$ identity mapping, we have

$$H = I - (I_1 + F_2)\mathbf{A}(F_1 + I_2). \tag{3.1}$$

3.2. Runge–Kutta methods with internal perturbations. The main purpose of this subsection is a discussion of the following four equalities and of their relations to the Runge–Kutta method (1.3).

$$y - \mathbf{A}Fy = p, \tag{3.2}$$

$$x - F\mathbf{A}x = q, \tag{3.3}$$

$$Hz = r, \tag{3.4}$$

$$y - \mathbf{A}x = s, \qquad x - Fy = t. \tag{3.5}$$

LEMMA 3.1.

(a) (3.2) *implies* (3.4) *with*

$$z = (I_1 + F_2)y, \qquad r = I_1 p + (F_2 y - F_2(y - p));$$

(3.4) *implies* (3.2) *with*

$$y=[I_1+\mathbf{A}_2(F_1+I_2)]z, \qquad p=(I_1+\mathbf{A}I_2)r.$$

(b) (3.3) *implies* (3.4) *with*

$$z=(\mathbf{A}_1+I_2)x, \qquad r=(\mathbf{A}_1I_1+I_2)q+(F_2\mathbf{A}x-F_2\mathbf{A}(x-I_1q));$$

(3.4) *implies* (3.3) *with*

$$x=(F_1+I_2)z, \qquad q=(F_1z-F_1(z-r))+I_2r.$$

(c) (3.5) *implies* (3.4) *with*

$$z=I_1y+I_2x, \qquad r=I_1s+(\mathbf{A}_1I_1+I_2)t+(F_2y-F_2(y-s-\mathbf{A}I_1t));$$

(3.4) *implies* (3.5) *with*

$$x=(F_1+I_2)z, \quad y=I_1z+\mathbf{A}_2x, \quad s=I_1r, \quad t=I_2r.$$

Using (3.1) the proof of this lemma is straightforward, and we omit it.

With the notation of § 3.1 we can rewrite the Runge–Kutta step (1.3) as

$$u_k=u_{k-1}+\mathbf{b}^TFy, \qquad y-\mathbf{A}Fy=0, \tag{3.6}$$

and (1.5) can be written in the form

$$u_k=u_{k-1}+\mathbf{b}^Tx, \qquad x-F\mathbf{A}x=0. \tag{3.7}$$

Applying Lemma 3.1 (with $p=q=r=0$), we see that both (3.6) and (3.7) are equivalent to the following formulation of the Runge–Kutta method,

$$u_k=u_{k-1}+\mathbf{b}^T(F_1+I_2)z, \qquad Hz=0. \tag{3.8}$$

If any numerical procedure is applied to solve the equation $Hz=0$, we obtain, in general, only an approximation, say $\tilde{z}$, to the true z. Denoting the corresponding numerical approximation to u_k by $\tilde{u}_k$ we thus have

$$\tilde{u}_k=u_{k-1}+\mathbf{b}^T(F_1+I_2)\tilde{z}, \tag{3.9a}$$

$$H\tilde{z}=r \tag{3.9b}$$

with a residual vector $r\in\mathbb{K}^{nm}$, $r\approx 0$. We note that the relations (3.9) with ($\mathcal{M}_1=\{1,2,\cdots,m\}$) and a different interpretation of the vector r also occur in the interesting investigations of B-consistency by Frank, Schneid and Ueberhuber (cf. [13], [14]). We call the components $r_i\in\mathbb{K}^n$ of $r=[r_i]\in\mathbb{K}^{nm}$ *internal perturbations* in the Runge–Kutta step (3.8).

A question of great practical and theoretical importance is whether $\|\tilde{z}-z\|$ and $|\tilde{u}_k-u_k|$ are small (uniformly for all f satisfying (2.1)) whenever $\|r\|$ is small (cf. (3.8), (3.9)). The results of § 3.3 are relevant to this question for $\|\tilde{z}-z\|$, and those of § 3.4 for $|\tilde{u}_k-u_k|$.

In practice one usually computes u_k from (3.6) or from (3.7). These cases are covered by our considerations since (3.8), (3.9) reduce to (3.6), (3.16) when $\mathcal{M}_1=\{1,2,\cdots,m\}$, while (3.8), (3.9) reduce to (3.7), (3.17) when $\mathcal{M}_2=\{1,2,\cdots,m\}$.

3.3. Internal stability. We shall investigate, for arbitrary z, $\tilde{z}\in\mathbb{K}^{nm}$, the sensitivity of $\tilde{z}-z$ with respect to $H\tilde{z}-Hz$, where the latter difference can be interpreted as the difference between two (different) internal perturbations (cf. (3.9b)). The results we obtain are basic for the proof in § 4 of Theorem 2.1.

Let $z, \tilde{z}$ be arbitrary vectors in $\mathbb{K}^{nm}$. In view of Lemma 3.1(c) we define

$$(3.10)\qquad \begin{aligned} x &= (F_1+I_2)z, & y &= I_1 z + \mathbf{A}_2 x, \\ \tilde{x} &= (F_1+I_2)\tilde{z}, & \tilde{y} &= I_1\tilde{z} + \mathbf{A}_2\tilde{x}. \end{aligned}$$

LEMMA 3.2. *Assume* (2.1), (2.2), (2.3). *Then there is a constant* γ_0 (*only depending on* $D, S, T, h^{-1}\alpha, h\beta$) *such that*

$$\|I_1(\tilde{x}-x)\| + \|I_2(\tilde{y}-y)\| \leqq \gamma_0 \|H\tilde{z} - Hz\|$$

whenever $z, \tilde{z} \in \mathbb{K}^{nm}$ *and* $x, \tilde{x}, y, \tilde{y}$ *are defined by* (3.10).

Proof. We define $u=[u_i]$, $v=[v_i]$, $w=[w_i]$, $p=[p_i]$, $q=[q_i] \in \mathbb{K}^{nm}$ by

$$u = \tilde{x} - x, \qquad v = \tilde{y} - y, \qquad w = F\tilde{y} - Fy,$$
$$p = I_1(H\tilde{z} - Hz), \qquad q = I_2(H\tilde{z} - Hz).$$

By the last part of Lemma 3.1 we thus have

$$(3.11)\qquad v - \mathbf{A}u = p, \qquad u - w = q.$$

From (2.1) it follows that

$$\operatorname{Re}\langle v_i, w_i\rangle \leqq \bar{\alpha}|w_i|^2 + \bar{\beta}|v_i|^2$$

where $\bar{\alpha} = h^{-1}\alpha$, $\bar{\beta} = h\beta$. Substituting $v_i = \mathbf{a}_i^T u + p_i$, $w_i = u_i - q_i$ (cf. (3.11)) in this inequality and using $\langle p_i, q_i\rangle = 0$, we obtain

$$\operatorname{Re}\langle \mathbf{a}_i^T u, u_i\rangle - \bar{\alpha}|u_i|^2 - \bar{\beta}|\mathbf{a}_i^T u|^2 \leqq \operatorname{Re}\langle u_i, -p_i - 2\bar{\alpha}q_i\rangle + \operatorname{Re}\langle \mathbf{a}_i^T u, q_i + 2\bar{\beta}p_i\rangle + \bar{\beta}|p_i|^2 + \bar{\alpha}|q_i|^2.$$

From (2.2) and Lemma 2.2 in [7] it can be seen that

$$\sum_{i=1}^m 2\delta_i \operatorname{Re}\langle \mathbf{a}_i^T u, u_i\rangle \geqq \sum_{i=1}^m \sigma_i |u_i|^2 + \sum_{i=1}^m \tau_i |\mathbf{a}_i^T u|^2.$$

A combination of the last two inequalities yields

$$(3.12)\qquad \begin{aligned} &\sum_{i=1}^m \left(\frac{1}{2}\sigma_i - \bar{\alpha}\delta_i\right)|u_i|^2 + \sum_{i=1}^m \left(\frac{1}{2}\tau_i - \bar{\beta}\delta_i\right)|\mathbf{a}_i^T u|^2 \\ &\qquad \leqq \sum_{i=1}^m \delta_i\{|u_i|\cdot|p_i + 2\bar{\alpha}q_i| + |\mathbf{a}_i^T u|\cdot|q_i + 2\bar{\beta}p_i| + \bar{\beta}|p_i|^2 + \bar{\alpha}|q_i|^2\}. \end{aligned}$$

Let $\xi, \eta, \lambda, \mu \in \mathbb{R}^m$ be column-vectors with components $\xi_i = (\frac{1}{2}\sigma_i - \bar{\alpha}\delta_i)^{1/2}|u_i|$, $\eta_i = (\frac{1}{2}\tau_i - \bar{\beta}\delta_i)^{1/2}|\mathbf{a}_i^T u|$, $\lambda_i = (\frac{1}{2}\sigma_i - \bar{\alpha}\delta_i)^{-1/2}\delta_i|p_i + 2\bar{\alpha}q_i|$, $\mu_i = (\frac{1}{2}\tau_i - \bar{\beta}\delta_i)^{-1/2}\delta_i|q_i + 2\bar{\beta}p_i|$ $(1 \leqq i \leqq m)$ (we use the convention $0^{-1/2} = 0$). Putting

$$\varepsilon = \sum_{i=1}^m \delta_i\{\bar{\beta}|p_i|^2 + \bar{\alpha}|q_i|^2\},$$

we see from (2.3) that (3.12) is equivalent to

$$\xi^T\xi + \eta^T\eta \leqq \xi^T\lambda + \eta^T\mu + \varepsilon.$$

After an application of Schwarz's inequality a little calculation shows that

$$(\xi^T\xi + \eta^T\eta)^{1/2} \leqq \tfrac{1}{2}(\lambda^T\lambda + \mu^T\mu)^{1/2} + \tfrac{1}{2}(\lambda^T\lambda + \mu^T\mu + 4\varepsilon)^{1/2}.$$

Hence

$$(3.13)\qquad \sum_{i=1}^m (\sigma_i - 2\bar{\alpha}\delta_i)|u_i|^2 + \sum_{i=1}^m (\tau_i - 2\bar{\beta}\delta_i)|\mathbf{a}_i^T u|^2 \leqq \gamma_1 \sum_{i=1}^m |h_i(\tilde{z}) - h_i(z)|^2$$

with a constant γ_1 only depending on the parameters $\delta_i, \sigma_i, \tau_i, \bar{\alpha}, \bar{\beta}$.

The proof is completed by applying (2.3) and substituting $\mathbf{a}_i^T u = v_i$ (for $i \in \mathcal{M}_2$; see (3.11)) into (3.13). □

Using the above lemma we shall prove the following theorem, which is the main result of this section.

THEOREM 3.3. *Assume* (2.1), (2.2), (2.3). *Then there exists a function* $\phi : \mathbb{K}^{nm} \times [0, \infty) \to [0, \infty)$ *with the properties*

(i) $\phi(z; \cdot)$ *is isotone on* $[0, \infty)$ (*for each* $z \in \mathbb{K}^{nm}$),

(ii) $\phi(z; \rho) \to \phi(z; 0) = 0$ (*as* $\rho \to 0+$; *for each* $z \in \mathbb{K}^{nm}$),

(iii) $\|\tilde{z} - z\| \leqq \phi(z; \|H\tilde{z} - Hz\|)$ (*for all* $z, \tilde{z} \in \mathbb{K}^{nm}$).

Moreover, if $\mathcal{M}_2 = \emptyset$, *then* (i), (ii) *and* (iii) *hold with* $\phi(z, \rho) \equiv \gamma\rho$ *where* γ *is a constant only depending on* A, $h^{-1}\alpha$, $h\beta$ (*and not on* z, f *or the dimension* n).

Proof. Let $z, \tilde{z} \in \mathbb{K}^{nm}$ be given. Defining u, v, w, p, q as in the proof of Lemma 3.2, we have the representation

$$\tilde{z} - z = I_1 v + I_2 u.$$

From (3.11) and Lemma 3.2 we obtain

$$\|I_2 u\| \leqq \|q\| + \|F_2\tilde{y} - F_2 y\| \leqq \|q\| + \psi(z; \gamma_0\|H\tilde{z} - Hz\|)$$

where

$$\psi(z; \rho) = \sup\{\|F_2(y+e) - F_2 y\| : e \in \mathbb{K}^{nm} \text{ with } \|I_2 e\| \leqq \rho\}, \tag{3.14}$$

$$y = I_1 z + \mathbf{A}_2(F_1 + I_2)z.$$

Using (3.11) and Lemma 3.2 once more, we thus obtain

$$\begin{aligned}\|I_1 v\| &\leqq \|p\| + \|\mathbf{A}_1 I_1\| \cdot \|I_1 u\| + \|\mathbf{A}_1 I_2\| \cdot \|I_2 u\| \\ &\leqq \|p\| + \|\mathbf{A}_1 I_1\| \cdot \gamma_0 \cdot \|H\tilde{z} - Hz\| + \|\mathbf{A}_1 I_2\|\{\|q\| + \psi(z; \gamma_0\|H\tilde{z} - Hz\|)\}.\end{aligned}$$

It follows that property (iii) holds with

$$\phi(z; \rho) = (2 + \|\mathbf{A}_1 I_2\| + \gamma_0\|\mathbf{A}_1 I_1\|)\rho + (1 + \|\mathbf{A}_1 I_2\|)\psi(z; \gamma_0\rho). \tag{3.15}$$

The remaining properties stated in the theorem follow from the continuity of f (see (2.1)) and from the fact that for any $m \times m$ matrix M the norm $\|M \otimes I^{(n)}\|$ is independent of n (which can be proved e.g. by using Lemma 2.2 in [7]). □

If $\mathcal{M}_2 \neq \emptyset$ the function ϕ defined by (3.15) depends through ψ on the (local) Lipschitz constant of f. If $\alpha \geqq 0$ this Lipschitz constant can be arbitrarily large. In this case the upper bound on $\|\tilde{z} - z\|$ provided by the theorem thus only holds for the particular function f under consideration, and not uniformly for all f satisfying (2.1).

We note that when $\mathcal{M}_2 = \emptyset$ and $\alpha = 0$, the content of Theorem 3.3 is similar to the (so-called BSI-stability) results formulated in [13, Thm. 4.1, Cor. 4.1], [12, Thm. 5.3.7].

3.4. External stability. We deal with the effect of the internal perturbation r on the difference $\tilde{u}_k - u_k$ where u_k, $\tilde{u}_k$ satisfy (3.8), (3.9). The following theorem provides a condition under which a bound for $|\tilde{u}_k - u_k|$ in terms of $\|r\|$ holds uniformly for all f satisfying (2.1). This condition can be fulfilled in cases where no analogous uniform bound holds for $\|\tilde{z} - z\|$.

THEOREM 3.4. *Assume* (2.1), (2.2), (2.3). *Suppose there exist real* d_j (*for* $j \in \mathcal{M}_2$) *such that*

$$b_i = \sum_{j \in \mathcal{M}_2} d_j a_{ji} \quad (\textit{for all } i \in \mathcal{M}_2).$$

Then there is a constant γ only depending on A, b, $h^{-1}\alpha$, $h\beta$ (and not on u_{k-1}, z, f or the dimension n) such that

$$|\tilde{u}_k - u_k| \leqq \gamma \|r\|$$

whenever u_k, $\tilde{u}_k$, r satisfy (3.8), (3.9).

Proof. We define

$$d_i = b_i - \sum_{j \in \mathcal{M}_2} d_j a_{ji} \quad (\text{for all } i \in \mathcal{M}_1),$$

and

$$d = (d_1, d_2, \cdots, d_m)^T, \qquad \mathbf{d} = d \otimes I^{(n)}.$$

One easily verifies that, with these definitions,

$$\mathbf{b}^T = \mathbf{d}^T I_1 + \mathbf{d}^T \mathbf{A}_2.$$

From (3.8), (3.9) it follows that

$$\tilde{u}_k - u_k = [\mathbf{d}^T I_1 + \mathbf{d}^T \mathbf{A}_2][(F_1 \tilde{z} - F_1 z) + I_2(\tilde{z} - z)].$$

Defining $x, \tilde{x}, y, \tilde{y}$ by (3.10) we have

$$F_1 \tilde{z} - F_1 z = I_1(\tilde{x} - x), \qquad \mathbf{A}_2[(F_1 \tilde{z} - F_1 z) + I_2(\tilde{z} - z)] = \mathbf{A}_2(\tilde{x} - x) = I_2(\tilde{y} - y).$$

Consequently

$$\tilde{u}_k - u_k = \mathbf{d}^T [I_1(\tilde{x} - x) + I_2(\tilde{y} - y)].$$

An application of Lemma 3.2 completes the proof. □

In order to formulate some interesting corollaries to the above theorem, we define for any index set $\mathcal{N} \subset \{1, 2, \cdots, m\}$ the $m \times m$ matrix $A(\mathcal{N})$ by

$$A(\mathcal{N}) = (c_{ij}), \quad c_{ij} = a_{ij} \ (\text{if } i \in \mathcal{N}, j \in \mathcal{N}), \quad c_{ij} = \delta_{ij} \ (\text{otherwise}),$$

where δ_{ij} denotes the Kronecker delta.

COROLLARY 3.5. *Suppose* (2.2) *holds with*

$$\delta_i \geqq 0, \quad \sigma_i \geqq 0, \quad \tau_i \geqq 0, \quad \sigma_i + \tau_i > 0 \quad (\text{for } 1 \leqq i \leqq m).$$

Let $\mathcal{M}_1$, $\mathcal{M}_2$ be disjoint, $\mathcal{M}_1 \cup \mathcal{M}_2 = \{1, 2, \cdots, m\}$, with

$$\{i \mid \sigma_i = 0\} \subset \mathcal{M}_2 \subset \{i \mid \tau_i > 0\},$$

and $\operatorname{Rank}[A(\mathcal{M}_2)^T, b] = \operatorname{Rank}[A(\mathcal{M}_2)^T]$. *Then there is a constant γ (only depending on A, b) such that*

$$|\tilde{u}_k - u_k| \leqq \gamma \|r\|,$$

whenever u_k, $\tilde{u}_k$, r satisfy (3.8), (3.9) *and the continuous $f: \mathbb{R} \times \mathbb{K}^n \to \mathbb{K}^n$ fulfills* (1.2).

This corollary completes some results on external stability for $\mathcal{M}_1 = \{1, 2, \cdots, m\}$ derived under assumptions (1.4), (1.2) in [10, Cor. 4.3].

COROLLARY 3.6. *Let $h > 0$ and $\alpha, \beta, \kappa, \lambda \in \mathbb{R}$ be given numbers, $D = \operatorname{diag}(\delta_1, \delta_2, \cdots, \delta_m)$, and let $\mathcal{M}_1$, $\mathcal{M}_2$ be disjoint index sets with $\mathcal{M}_1 \cup \mathcal{M}_2 = \{1, 2, \cdots, m\}$. Assume the following four conditions hold.*

(i) *$DA + A^T D - \kappa D - \lambda A^T D A$ is positive semidefinite;*

(ii) *$\delta_i > 0$ $(1 \leqq i \leqq m)$, $2\alpha h^{-1} \leqq \kappa$, $2\beta h \leqq \lambda$, $2\alpha h^{-1} + 2\beta h < \kappa + \lambda$;*

(iii) $\operatorname{Rank}[A(\mathcal{M}_2)^T, b] = \operatorname{Rank}[A(\mathcal{M}_2)^T]$;

(iv) *if $\alpha = \kappa = 0$ then either $\mathcal{M}_1 = \varnothing$ or A is regular.*

Then there is a constant γ (only depending on A, b, αh^{-1} and βh) such that

$$|\tilde{u}_k - u_k| \leqq \gamma \|r\|$$

whenever $\tilde{u}_k$, u_k, r satisfy (3.8), (3.9) *and f fulfills* (2.1).

Proof. By applying Lemma 2.2 to the function hf, the proof follows from Theorem 3.4 for the case $[2\alpha h^{-1} \leqq \kappa,\ 2\beta h < \lambda,\ \alpha \neq 0]$.

If $[\alpha = \kappa = 0,\ 2\beta h < \lambda,\ \mathcal{M}_1 = \emptyset]$, Theorem 3.4 may be applied directly.

In case $[\alpha = \kappa = 0, 2\beta h < \lambda,\ A$ regular$]$ we take $S = \kappa_1 D$, $T = \lambda_1 D$ in (2.2) with $\lambda_1 \in (2\beta h, \lambda)$, $\kappa_1 > \kappa$ and $\kappa_1 - \kappa$ sufficiently small. The assumptions of Theorem 3.4 are then fulfilled.

Similarly, if $[2\alpha h^{-1} < \kappa,\ 2\beta h \leqq \lambda]$ we choose $S = \kappa_1 D$, $T = \lambda_1 D$ with $\kappa_1 \in (2\alpha h^{-1}, \kappa)$, $\lambda_1 > \lambda$ and $\lambda_1 - \lambda$ sufficiently small. □

Let the Runge–Kutta method (1.3) be *algebraically stable.* Consider along with (3.6), (3.7), the perturbed relations

$$\tilde{u}_k = u_{k-1} + \mathbf{b}^T F\tilde{y}, \qquad \tilde{y} - \mathbf{A}F\tilde{y} = p, \tag{3.16}$$

$$\tilde{u}_k = u_{k-1} + \mathbf{b}^T \tilde{x}, \qquad \tilde{x} - F\mathbf{A}\tilde{x} = q, \tag{3.17}$$

respectively. For given $h > 0$, $\alpha \leqq 0$, $\beta \leqq 0$, $\alpha + \beta < 0$, Corollary 3.6 (with $\kappa = \lambda = 0$) proves the existence of a constant γ such that

$$(3.7), (3.17) \Rightarrow |\tilde{u}_k - u_k| \leqq \gamma \|q\|$$

uniformly for all f satisfying (2.1) (note that $\text{Rank}\,[A^T, b] = \text{Rank}\,[A^T]$ since $x^T(A^T B x) \geqq \frac{1}{2}(x^T b)^2$ (for all $x \in \mathbb{R}^m$)). Under the same assumptions the corollary also proves the existence of a γ such that

$$(3.6), (3.16) \Rightarrow |\tilde{u}_k - u_k| \leqq \gamma \|p\|$$

uniformly for all f satisfying (2.1), provided we assume additionally that

$$\alpha < 0, \quad \text{or } A \text{ is regular.}$$

We note that when $\alpha = 0$ this stability result for (3.16) also follows from [12, Thm. 5.3.7]. On the other hand, Corollary 3.6 implies the general bound for $|\tilde{u}_k - u_k|$ in terms of $\|p\|$ (cf. (3.6), (3.16)) that also follows from [12, Thm. 5.3.7].

3.5. Examples.

Example 3.7. Consider the 3-stage Labotto IIIC method (cf. Example 2.5) and let f satisfy (1.2). Choosing $\mathcal{M}_1 = \{2\}$, $\mathcal{M}_2 = \{1, 3\}$, it follows from Corollary 3.5 that

$$|\tilde{u}_k - u_k| \leqq \gamma \cdot \|r\|$$

whenever (3.8), (3.9) hold. Here γ is independent of $h > 0$ and f. The formulation (3.8) of the Runge–Kutta step for which this stability result is valid, reads in full

$$u_k = u_{k-1} + \tfrac{1}{6}(z_1 + 4f_2(z_2) + z_3), \tag{3.18a}$$

$$\begin{aligned} z_1 &= f_1(\tfrac{1}{6}(z_1 - 2f_2(z_2) + z_3)), \\ z_2 &= \tfrac{1}{12}(2z_1 + 5f_2(z_2) - z_3), \\ z_3 &= f_3(\tfrac{1}{6}(z_1 + 4f_2(z_2) + z_3)) \end{aligned} \tag{3.18b}$$

with $f_i(\xi) = hf(t_{k-1} + c_i h,\ u_{k-1} + \xi)$, $c_0 = 0$, $c_1 = \frac{1}{2}$, $c_2 = 1$.

For $\|\tilde{z} - z\|$ there is no analogous upper bound valid in terms of $\|r\|$.

If we define $\tilde{u}_k$, $\tilde{y}$ by (3.16), it can be proved that not only

$$\sup\{\|\tilde{y} - y\|: p \in \mathbb{K}^{3n}, \|p\| \leqq 1, f \text{ satisfies } (1.2)\} = \infty$$

(cf. [10, ex. 4.4], [12, ex. 5.9.2]), but also

$$\sup\{|\tilde{u}_k - u_k|: p \in \mathbb{K}^{3n}, \|p\| \leqq 1, f \text{ satisfies } (1.2)\} = \infty.$$

In practical applications the use of (3.18) thus seems to have an advantage over the use of (1.3). A small residual vector in the process (3.18) has generally a substantially smaller effect on the approximation to $U(t_k)$ than in the process (1.3).

Example 3.8. Consider an arbitrary method satisfying condition (1.4) (e.g. Gauss, Radau IA or IIA—see [13]).

Applying Corollary 3.6 it can be seen that, for any disjoint $\mathcal{M}_1, \mathcal{M}_2$ with $\mathcal{M}_1 \cup \mathcal{M}_2 = \{1, 2, \cdots, m\}$, there exist $\kappa_0 > 0$, $\lambda_0 > 0$, $\gamma > 0$ such that

$$(3.8), (3.9) \Rightarrow |\tilde{u}_k - u_k| \leqq \gamma \|r\|$$

uniformly for all $h > 0$ and f satisfying (2.1) with

$$\alpha h^{-1} \leqq \kappa_0, \qquad \beta h \leqq \lambda_0.$$

In particular we thus have

$$(3.6), (3.16) \Rightarrow |\tilde{u}_k - u_k| \leqq \gamma \|p\| \quad \text{and} \quad (3.7), (3.17) \Rightarrow |\tilde{u}_k - u_k| \leqq \gamma \|q\|$$

uniformly for $h > 0$ and f as above. This completes a so-called BS-stability result on (3.6), (3.16) with $\alpha = 0$ given in [13, Thm. 4.1, Cor. 4.1], [12, Thm. 7.4.1].

It thus follows that a small residual, e.g. in the numerical solution of either (1.3b) or (1.5b), only slightly disturbs the corresponding u_k computed via (1.3a) or (1.5a), respectively (uniformly for $\alpha h^{-1} \leqq \kappa_0$, $\beta h \leqq \lambda_0$).

Example 3.9. We finally give a counterexample showing that assumption (iv) in Corollary 3.6 cannot be omitted.

Consider Euler's method ($m = 1$, $A = 0$, $b = 1$). The conditions (i), (ii), (iii) of the corollary are fulfilled with

$$\delta_1 = 1, \quad \kappa = 0, \quad \lambda = 1, \quad \alpha = 0, \quad \beta = 0, \quad h = 1, \quad \mathcal{M}_2 = \varnothing.$$

Applying (3.6), (3.16) with $u_{k-1} = 0$, $f(t, \xi) \equiv \mu\xi$, $\mu < 0$, we have

$$\tilde{u}_k - u_k = \mu p.$$

Letting $\mu \to -\infty$ we see that the conclusion of Corollary 3.6 is not valid.

4. The proof of Theorem 2.1. Theorem 2.1 is easily proved by using Lemma 4.1 and by a combination of Theorem 3.3 with the subsequent Lemma 4.2.

LEMMA 4.1. *Each of the following systems* (4.1)–(4.4) *has a unique solution iff any of the other systems has a unique solution.*

$$y - \mathbf{A}Fy = 0, \tag{4.1}$$

$$x - F\mathbf{A}x = 0, \tag{4.2}$$

$$Hz = 0, \tag{4.3}$$

$$y - \mathbf{A}x = 0, \qquad x - Fy = 0. \tag{4.4}$$

Proof. Apply Lemma 3.1. □

LEMMA 4.2. *Let E be a finite dimensional vector space over $\mathbb{K}$ with norm $\|\cdot\|$, and let $G: E \to E$ be a given continuous function. Assume $\phi: E \times [0, \infty) \to [0, \infty)$ has the properties*

(a) $\phi(z; \cdot)$ *is isotone on* $[0, \infty)$ (*for all* $z \in E$),

(b) $\phi(z; 0) = 0$ (*for all* $z \in E$),

(c) $\|\tilde{z} - z\| \leqq \phi(z; \|G\tilde{z} - Gz\|)$ (*for all* $z, \tilde{z} \in E$).

Then there is a unique $z^* \in E$ with $Gz^* = 0$.

Proof. G is a continuous one-to-one mapping defined on E. The domain-invariance theorem (cf. [18]) thus implies that $G(E)$ is open.

Property (c) implies that $\|Gz\| \to \infty$ (when $\|z\| \to \infty$). Therefore a *bounded* sequence $z_1, z_2, z_3 \cdots$ exists with

$$\lim_{k\to\infty} \|Gz_k\| = r, \qquad r = \inf\{\|Gz\|: z \in E\}.$$

Consequently there is a subsequence $\{y_k\}$ of $\{z_k\}$ with

$$\lim_{k\to\infty} y_k = z^*, \quad \lim_{k\to\infty} Gy_k = Gz^*, \quad \|Gz^*\| = r$$

for some $z^* \in E$.

Since $G(E)$ is open, we have $r = 0$. □

We note that theorems with much resemblance to the above lemma can be found in the literature (see e.g. [16, Thm. 13.5], [19, Thm. 5.3.8]).

REFERENCES

[1] K. BURRAGE AND J. BUTCHER, *Stability criteria for implicit Runge–Kutta methods*, this Journal, 16 (1979), pp. 46–57.

[2] ———, *Nonlinear stability for a general class of differential equation methods*, BIT, 20 (1980), pp. 185–203.

[3] G. J. COOPER, *A generalization of algebraic stability for Runge–Kutta methods*, IMA J. Numer. Anal., 4 (1984), pp. 427–440.

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

[5] M. CROUZEIX, W. H. HUNDSDORFER AND M. N. SPIJKER, *On the existence of solutions to the algebraic equations in implicit Runge–Kutta methods*, BIT, 23 (1983), pp. 84–91.

[6] M. CROUZEIX AND P. A. RAVIART, Unpublished lecture notes, Université de Rennes, 1980.

[7] G. DAHLQUIST, *Error analysis for a class of methods for stiff nonlinear initial value problems*, Lecture Notes in Mathematics, 506 (1976), Springer-Verlag, Berlin, pp. 60–72.

[8] ———, *G-stability is equivalent to A-stability*, BIT, 18 (1978), pp. 384–401.

[9] G. DAHLQUIST AND R. JELTSCH, *Generalized disks of contractivity for explicit and implicit Runge-Kutta methods*, Report TRITA-NA-7906, Dept. Comp. Sci., Roy. Inst. Techn., Stockholm, 1979.

[10] K. DEKKER, *Error bounds for the solution of the algebraic equations in Runge–Kutta methods*, BIT, 24 (1984), pp. 347–356.

[11] K. DEKKER AND E. HAIRER, *A necessary condition for BSI-stability*, BIT, 25 (1985), pp. 285–288.

[12] K. DEKKER AND J. G. VERWER, *Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations*, North-Holland, Amsterdam, 1984.

[13] R. FRANK, J. SCHNEID AND C. W. UEBERHUBER, *Stability properties of implicit Runge–Kutta methods*, this Journal, 22 (1985), pp. 497–514.

[14] ———, *Order results for implicit Runge–Kutta methods applied to stiff systems*, this Journal, 22 (1985), pp. 515–534.

[15] W. H. HUNDSDORFER, *The numerical solution of nonlinear stiff initial value problems*, Ph.D. thesis, Univ. Leiden; CWI-tract, Amsterdam, 1984.

[16] TH. MEIS AND U. MARCOWITZ, *Numerical Solution of Partial Differential Equations*, Springer-Verlag, New York, 1981.

[17] O. NEVANLINNA, *On the numerical integration of nonlinear initial value problems by linear multistep methods*, BIT, 17 (1977), pp. 58–71.

[18] J. T. SCHWARTZ, *Nonlinear Functional Analysis*, Gordon and Breach, New York, 1969.

[19] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.