Analysis of the Convergence of More General Linear Iteration Scheme on the Implementation of Implicit Runge-Kutta Methods to Stiff Differential Equations

R.Vigneswaran* and S.Kajanthan

Abstract—**A modified Newton scheme is typically used to solve large sets of non-linear equations arising in the implementation of implicit Runge-Kutta methods. As an alternative to this scheme, iteration schemes, which sacrifice superlinear convergence for reduced linear algebra costs, have been proposed. A more general linear iterative scheme of this type proposed by Cooper and Butcher in 1983 for implicit Runge-Kutta methods, and he has applied the successive over relaxation technique to improve the convergence rate. In this paper, we establish the convergence result of this scheme by proving some theoretical results suitable for stiff problems. Also these convergence results are verified by two and three stage Gauss method and Radue IIA method.**

Index Terms—**Implementation, Implicit Runge-Kutta methods, Rate of convergence, Stiff systems, Convergence results**

I. Introduction

The numerical integration of a stiff system of n ordinary differential equations is commonly carried out using an implicit numerical method. This may be an s-stage implicit Runge-kutta method or, more generally, an implicit multi value method with s internal stages. A modified Newton iteration is often used to solve the algebraic equations that arise from the s internal stages. Each step of the iteration requires the solution of a set of sn linear equations and schemes may be developed to solve these linear equations efficiently. This approach has been discussed by Chipman [6]. Butcher [2] developed a scheme of this type, using a similarity transformation of the coefficient matrix associated with the internal stages of the method, which is particularly effective when the matrix has a single point spectrum. Enright [16] further developed this approach for methods whose coefficient matrix has a more general (real) spectrum and Varah [23] suggested the use of complex arithmetic to cope with methods whose coefficient matrix has a complex spectrum. These schemes use particular similarity transformations to implement the same modified Newton iteration but, for the purpose of a convergence analysis, any convenient canonical transformation may be used.

In another approach, schemes based directly on iterative procedures may be developed. Frank and Ueberhuber [17] consider the use of iterated defect corrections and Butcher [3] describes a variety of iteration schemes. Cooper and Butcher [8] examine a generalisation of one of these schemes which explicitly uses the Jacobian of the differential system. Extension have been described by Cooper and Vigneswaran[11] [12] . In this type of scheme, each step of the iteration consists of a number of sub-steps, with each sub-step requiring the solution set of n linear equations so that these schemes compete with the modified Newton iteration despite the lack of superlinear convergence. These schemes are similar in form to a canonical representation of the modified Newton iteration and the same type of convergence analysis is applicable, modified to cope with the lack of superlinear convergence. For these schemes attention may be restricted to the real domain.

There is current interest in the numerical solution of ordinary differential equations on parallel computers. Jackson and Norsett [19] give an extensive survey of recent work and examine possible approaches. In one approach, new methods may be designed which are particularly suitable for implementation on parallel processors, using a modified Newton iteration. Karakashian and Rust [20] point out that it is advantageous to use a method with a coefficient matrix similar to a (real) diagonal matrix. Other methods have been considered by Iserles and Norsett [18] who examine Runge-Kutta methods and by Butcher who examine special type of multivalue methods. Another approach is to design alternative iteration schemes to implement existing methods efficiently on parallel processors. A scheme of this type has been described by van der Houwen and Sommeijer [24] and Cooper [9] designed a scheme specially for singly implicit methods. Cooper and Vigneswaran [13] obtain a more general scheme, suitable whenever the coefficient matrix of the method has real eigenvalues, which converges in a finite number of iterations when applied to linear problems. They analyse convergence under assumptions suitable for stiff problems.

Here, the approach of Cooper and Vigneswaran [13] is adopted, where it is assumed that the differential system satisfies one sided Lipschitz condition and a corresponding condition is imposed on the Jacobian. The coefficient matrix is assumed to satisfy an algebraic condition which is known to guarantee the existence and uniqueness of a solution to the algebraic equations. Under these assumptions, convergence results are derived for the more general scheme proposed by Cooper and Butcher [8] and it is verified by two and three stage Gauss method and Radue IIA method. The convergence rate appears to depend on the condition number of a transformation matrix associated with the numerical

Manuscript received September 18, 2019; revised May 12, 2020.
R.Vigneswaran is corresponding author, with Faculty of Science, University of Jaffna, Sri Lanka; e-mail:vicky@jfn.ac.lk, rvicky58@gmail.com
S.Kajanthan is with Faculty of Technology, University of Jaffna, Sri Lanka;e-mail: tskajanthan@gmail.com, kajanthans@univ.jfn.ac.lk

method. When this is large, a small step length must be chosen.

II. Iteration Scheme

Consider an initial value problem for stiff system of $n(\geq 1)$ ordinary differential equations

$$x' = f(x(t)), \quad x(t_0) = x_0, \quad f : \mathbb{R}^n \to \mathbb{R}^n, \qquad (1)$$

where f is assumed to be as smooth as necessary. An s-stage implicit Runge-Kutta method computes an approximation x_{r+1} to the solution $x(t_{r+1})$ at grid point $t_{r+1} = t_r + h$ by

$$x_{r+1} = x_r + h\sum_{i=1}^{s} b_i f(y_i) \qquad (h > 0)$$

where the internal approximations $y_1, y_2, \cdots, y_s$ satisfy the sn equations

$$y_i = x_r + h\sum_{j=1}^{s} a_{ij} f(y_j), \quad i = 1, 2, \cdots, s \qquad (2)$$

$A = [a_{ij}]$ is the real coefficient matrix and $b = (b_1, b_2, \cdots, b_s)^T$ is the column vector of the Runge-Kutta method. Let $X_r = x_r \oplus x_r \oplus \cdots \oplus x_r$ and $Y = y_1 \oplus y_2 \oplus \cdots \oplus y_s$ be sn element of column vectors and let $F(Y) = f(y_1) \oplus f(y_2) \oplus \cdots \oplus f(y_s)$. Then equation (2) may be represented by the compact form

$$Y = X_r + h(A \otimes I_n)F(Y) \qquad (3)$$

where $A \otimes I_n$ is the Kronecker product of the matrix A with $n \times n$ identity matrix I_n and, in general $A \otimes B = [a_{ij}B]$. This article deals with methods suitable for stiff systems so that the matrix A is not strictly lower triangular.

Equation (3) may be solved by a modified Newton iteration. Let J be the Jacobian of f evaluated at some recent point x_r, updated infrequently. The modified Newton scheme evaluates $Y^1, Y^2, Y^3, \cdots$, to satisfy

$$(I_{sn} - hA \otimes J)(Y^m - Y^{m-1}) = D(Y^{m-1}), \qquad (4)$$
$$m = 1, 2, \cdots,$$

where D is the approximation defect, $D(Z) = X_r - Z + h(A \otimes I_n)F(Z)$. In each step of this iteration, a set of sn linear equations has to be solved. Schemes have been developed, to solve equation (4), which use the fact that J is constant [1], [6], [7]. In other schemes advantage is taken of the special forms of some implicit methods [2], [4], [5], [16].

In another approach, schemes based directly on iterative procedure have been developed [3], [8], [11], [12],[17],[25],[26],[27],[28],[29]. For a singly implicit method, there is a non-singular matrix S so that $S^{-1}AS = \lambda(I_s - L)^{-1}$, where L is zero except for some ones on the sub-diagonal. On applying this transformation, the scheme (4) becomes

$$[I_s \otimes (I_n - h\lambda J)]E^m = [(I_s - L)S^{-1} \otimes I_n]D(Y^{m-1}) + (L \otimes I_n)E^m, \qquad (5)$$
$$Y^m = Y^{m-1} + (S \otimes I_n)E^m, \quad m = 1, 2, 3 \cdots \cdot$$

Cooper and Butcher [8] proposed an iterative scheme, sacrificing superlinear convergence for reduced linear algebra cost, which may be regarded as a generalization of the scheme (5) for singly implicit methods. They considered the scheme

$$[I_s \otimes (I_n - h\lambda J)]E^m = (BS^{-1} \otimes I_n)D(Y^{m-1}) + (L \otimes I_n)E^m, \qquad (6)$$
$$Y^m = Y^{m-1} + (S \otimes I_n)E^m, \quad m = 1, 2, \cdots,$$

where B and S are real $s \times s$ non-singular matrices and L is strictly lower triangular matrix of order s, and $\Lambda = \lambda I$ be an $s \times s$ diagonal matrix with real diagonal elements. Cooper[9] proposed a scheme, with $L = 0$, suitable for implementing singly implicit methods on parallel processors. Cooper and Vigneswaran [13] extended this to allow Λ to have differing diagonal elements and gave a convergence analysis applicable to (real) schemes with $L = 0$ and a special choice of B and S. In this article, the analysis of the convergence of the scheme given by Cooper and Vigneswaran [13] is extended to more general linear iterative scheme given by Cooper and Butcher [8] of the form (6) for the real case. Sharper bounds are obtained by corresponding conditions are imposed for the general case (6).

III. Some Inequalities

The aim is to establish convergence of the iterative scheme (6) under assumptions on the differential system suitable for stiff problems. These conditions are described in this section and some preliminary results are also derived in this section to establish convergence result of the iterative scheme (6). Various authors have studied the existence and uniqueness of a solution of the Runge-Kutta equation (3) for a differential system (1). Let f be a continuous mapping of $\mathbb{R}^n$ to $\mathbb{R}^n$ satisfying a one-sided Lipschitz condition of the form

$$\langle f(v) - f(w), v - w\rangle \leq \nu|v - w|^2 \qquad \forall\ v, w \in \mathbb{R}^n \quad (7)$$

for a given inner product on $\mathbb{R}^n$ and corresponding norm $|v| = \langle v, v\rangle^{1/2}$. Associated with this inner product is a real symmetric positive definite matrix Q such that $\langle v, w\rangle = w^T Q v$ for all $v, w \in \mathbb{R}^n$. In this article $M \geq 0$ denotes that the real matrix M is non negative definite and $M > 0$ denotes that M is positive definite.

Crouzeix, Hundsdorfer and Spijker [14] considered the case f is monotone, where $\nu = 0$. They showed that there exist a unique solution of (3) if the Runge-Kutta method has a coefficient matrix A for which there is a diagonal matrix $R > 0$ such that $RA + A^T R > 0$. This implies, in particular, that the eigenvalues of A have positive real parts. Dekker [15] and Cooper [9] obtained similar results for the more general case $\nu > 0$ and recently Kraaijevanger and Schneid [21] have given necessary and sufficient conditions for the existence of a unique solution. For the case, $\nu > 0$, suppose there is an $\alpha > 0$ and diagonal matrix $R = [r_i] > 0$ such that

$$RA + A^T R - 2\alpha A^T R A > 0. \qquad (8)$$

Note that A may be singular but, if A has a zero eigenvalue of multiplicity r, condition (8) can hold only if there are r corresponding linearly independent eigenvectors. Note also that (8) gives

$$\alpha\lambda_i \leq 1, \quad i = 1, 2, \ldots, s, \qquad (9)$$

where $\lambda_i, \quad i = 1, 2, \ldots, s$ are the eigenvalues of A. It is assumed that this condition (9) applies to the diagonal elements of Λ as well.

With respect to any $s \times s$ positive definite diagonal matrix $D = [d_i]$, the inner product on $\mathbb{R}^n$ induces an inner product on $\mathbb{R}^{sn}$,

$$\langle\langle V, W\rangle\rangle_D = \sum_{i=1}^{s} d_i \langle v_i, w_i\rangle = W^T(D \otimes Q)V, \qquad (10)$$

with corresponding norm $\|V\|_D$. The argument given by Cooper [10] may be adapted to show that, for given positive step length h such that $\alpha - h\nu > 0$, there exist a unique solution Y of (3) and

$$\|Y - X_r\|_R \leq \frac{h}{\alpha - h\nu}\|F(X_r)\|_R,$$

where $X_r = x_r \oplus x_r \oplus \cdots \oplus x_r$, $F(X_r) = f(x_r) \oplus f(x_r) \oplus \cdots \oplus f(x_r)$, $x_r = x(t_r)$,
t_r is a grid point and $f(x_r)$ is defined in (1). Let D be a given positive definite diagonal matrix with spectral radius $\rho[D] = 1$ and define

$$|f| = \max_{1\leq i\leq s} |f(X_r)|.$$

Let $r_i's$ be the elements of the diagonal matrix R. The bound for $\|Y - Y_r\|_R$ gives

$$\|Y - X_r\|_D \leq \frac{h\beta|f|}{\alpha - h\nu}, \qquad \beta = \max_{1\leq i\leq s}\sqrt{\frac{\text{tr } R}{r_i}}. \qquad (11)$$

The constants α and β depend on the numerical method only, and ν is independent of the stiffness of the problem. It is assumed that $|f|$ is independent of stiffness also. If, however, transient solution components are significant the step length h must be chosen small.

The following conditions correspond to those given by Cooper and Vigneswaran [13]. It is assumed that f is continuously differentiable with derivative $f' : \mathbb{R}^n \to \mathbb{L}(\mathbb{R}^n, \mathbb{R}^n)$, where $\mathbb{L}(\mathbb{R}^n, \mathbb{R}^n)$ is the space of continuous linear maps of $\mathbb{R}^n$ into $\mathbb{R}^n$ with norm induced by the given norm on $\mathbb{R}^n$. let P be some non singular matrix such that $P^TP = Q$. The norm $|C|$ of $C \in \mathbb{L}(\mathbb{R}^n, \mathbb{R}^n)$ is given by

$$|C|^2 = \sup_{v\neq 0} \frac{v^TC^TQCv}{v^TQv} = \rho[P^{-T}C^TQCP^{-1}]. \qquad (12)$$

Assume that f' satisfies a Lipschitz condition

$$|f'(v) - f'(w)| \leq L|v - w|, \qquad \forall v, w \in \mathbb{R}^n$$

and observe that $L = 0$ for linear differential systems and may be small for nonlinear stiff systems. The norm on $\mathbb{R}^n$ defined by the inner product (10) induces a norm on $\mathbb{L}(\mathbb{R}^n, \mathbb{R}^n)$. Since $F'(V) = f'(v_1) \oplus f'(v_2) \oplus \cdots \oplus f'(v_s)$ is block diagonal it follows that

$$\|F'(V)\|_D = \max_{1\leq i\leq s} |f'(v_i)|,$$

with a corresponding result for any block diagonal element of $\mathbb{L}(\mathbb{R}^{sn}, \mathbb{R}^{sn})$. It follows that, with respect to these norms, F' satisfies a Lipschitz condition on $\mathbb{R}^{sn}$ with the same Lipschitz constant L. There is also a special result for the norm of any element of $\mathbb{L}(\mathbb{R}^{sn}, \mathbb{R}^{sn})$ of the form $B \otimes C$ where B is any $s \times s$ matrix and C any $n \times n$ matrix. In this case it can be shown that $\|B \otimes C\|_D = |B|_D|C|$. Here, a positive definite diagonal matrix D defines an inner product $\langle x, y\rangle_D = y^TDx$ on $\mathbb{R}^s$,with corresponding norm $|x|_D$. With respect to an $s \times s$ matrix B, the induced norm on $\mathbb{L}(\mathbb{R}^s, \mathbb{R}^s)$ is given by

$$|B|_D^2 = \sup_{x\neq 0} \frac{x^TB^TDBx}{x^TDx} = \rho[D^{-1/2}B^TDBD^{-1/2}]. \qquad (13)$$

In the following it is assumes that J is the Jacobian of f evaluated at x_p, or some difference approximation to $f'(x_p)$. In either case, suppose that

$$\|I \otimes J - F'(X_r)\|_D \leq \gamma L \qquad (14)$$

for given γ assumed to be independent of stiffness. If the solution is varying rapidly it may be necessary to choose h small. For Runge-Kutta methods an alternative is to evaluate the Jacobian at x_r to force $\gamma = 0$. It is also assumed that

$$\langle Jv, v\rangle \leq \nu|v|^2 \qquad \forall v \in \mathbb{R}^n \qquad (15)$$

which may be regarded as a local version of (7) since

$$f(v) - f(w) = \int_0^1 f'(v - (1 - t)(v - w))dt(v - w). \qquad (16)$$

In order to obtain the convergence result, It has to obtain the bound for some expressions using the above conditions and norms.

Lemma 1. *Let $h \leq H$ where $\alpha - H\nu > 0$. Then $I - h\Lambda \otimes J$ is non singular and*

$$\| (I - h\Lambda \otimes J)^{-1} \|_D \leq \frac{\alpha}{\alpha - H\nu}.$$

Proof: Since $I - h\Lambda \otimes J$ is block diagonal it suffices to consider a single block $I - h\lambda J$ with $\lambda \neq 0$ $(\in \mathbb{R})$. The Schwarz inequality gives $|(I - h\lambda J)v||\lambda v| \geq |\langle(I - h\lambda J)v, \lambda v\rangle|$ so that

$$|(I-h\lambda J)v||\lambda v| \geq |\langle(I-h\lambda J)v, \lambda v\rangle| = |\lambda|v|^2 - h\lambda^2\langle Jv, v\rangle|$$

Inequality (9) gives $\alpha\lambda \leq 1$. Apply these results with inequality (15), to give

$$|(I - h\lambda J)v| \geq \frac{\alpha - H\nu}{\alpha}|v| \qquad \forall v \neq 0.$$

This gives $\det(I-h\lambda J) \neq 0$ and establish the first inequality. ■

Lemma 2. *Let $h \leq H$ where $\alpha - H\nu > 0$.*

$$\| (I - h\Lambda \otimes J)^{-1}(I + h\Lambda \otimes J) \|_D \leq \frac{\alpha + H\nu}{\alpha - H\nu}.$$

Proof: Consider a typical diagonal block $(I - h\lambda J)^{-1}(I + h\lambda J)$ and the inequality (9) gives $0 \leq \alpha\lambda \leq 1$. Now consider
$(1 + h\lambda\nu)^2|(I - h\lambda J)v|^2 - (1 - h\lambda\nu)^2|(I + h\lambda J)v|^2$

$$\begin{aligned} &= 4h\lambda[\nu|v|^2 - (1 + h^2\lambda^2\nu^2)\langle Jv, v\rangle + h^2\lambda^2\nu^2|Jv|^2] \\ &= 4h\lambda(1 - h^2\lambda^2\nu^2)[\nu|v|^2 - \langle Jv, v\rangle] + 4h^3\lambda^3\nu|\nu v - Jv|^2 \end{aligned}$$

which is non negative by virtue of (15). It follows that $(1+h\lambda\nu)|(I-h\lambda J)v| \geq (1-h\lambda\nu)|(I+h\lambda J)v|$ for all v and hence that

$$|(I-h\lambda J)^{-1}(I+h\lambda J)w| \leq \frac{1+h\lambda\nu}{1-h\lambda\nu} \leq \frac{\alpha+H\nu}{\alpha-H\nu} \qquad \forall w.$$

This establishes the required inequality. ∎

Lemma 3. *Let $h \leq H$ where $\alpha - H\nu > 0$. Let L be a strictly lower triangular $s \times s$ matrix, with $L^p = 0$, which commutes with Λ. Then*

$$\| \, [(I-L)\otimes I - h\Lambda \otimes J]^{-1} \, \|_D \leq \frac{\alpha\sigma}{\alpha - H\nu},$$

where $\sigma = 1 + \dfrac{\alpha \mid L \mid_D}{\alpha - Hv} + \cdots + \left(\dfrac{\alpha \mid L \mid_D}{\alpha - Hv}\right)^{p-1}.$

Proof: Let $M = (I - h\Lambda \otimes J)^{-1}(L \otimes I)$ and apply Lemma 1 to give

$$\|M\|_D \leq \|(I-h\Lambda\otimes J)^{-1}\|_D \|L\otimes I\|_D \leq \frac{\alpha|L|_D}{\alpha - H\nu}.$$

Note that L^p and $L\Lambda = \Lambda L$ gives $M^p = 0$ and hence $(I-M)^{-1} = I + M + \cdots + M^{p-1}$. The result follows from the identity
$[(I-L)\otimes I - h\Lambda\otimes J]^{-1} = (I-M)^{-1}(I-h\Lambda\otimes J)^{-1}$. ∎

IV. Convergence Result

To examine the convergence of the scheme (6) to the solution Y of (3),
define $G : \mathbb{R}^n \to \mathbb{R}^{sn}$ by

$$G(Z) = F(Y) - F(Y-Z) - (I\otimes J)Z. \qquad (17)$$

The approximation defect may be now expressed as

$$D(Y-Z) = (I - hA\otimes J)Z - h(A\otimes I)G(Z). \qquad (18)$$

The following lemma gives the bound for the norm of $G(Z)$ which is independent of stiffness.

Lemma 4. *Let $h \leq H$ where $\alpha - H\nu > 0$. Then for a given diagonal matrix $D > 0$,*

$$\| \, G(Z) \, \|_D \leq L[\gamma + \frac{H\beta|f|}{\alpha - H\nu}] \, \| \, Z \, \|_D + \frac{L}{2} \, \| \, Z \, \|_D^2.$$

Proof: The integral expression (16) applied to (17) gives

$$G(Z) = \int_0^1 [F'(Y-(1-t)Z) - F'(X_r)]dt \; Z + [F'(X_r) - I\otimes J]Z.$$

Now apply the Lipschitz condition for F' and inequality (14), namely $\|F'(X_r) - I\otimes J\|_D \leq \gamma L$, to give

$$\|G(Z)\|_D \leq L[\gamma + \|Y - X_r\|_D]\|Z\|_D + \frac{L}{2}\|Z\|_D^2.$$

The result follows from (11). ∎

Consider the convergence of the sequence of iterates $\{Y^m\}$ to the solution Y. Define $V^m = (S^{-1}\otimes I)(Y - Y^m), \quad m = 0,1,2,\ldots,$ and let $\bar{A} = S^{-1}AS$. It follows from (6) and (18) that

$$V^m = M_1 V^{m-1} + hM_2(S^{-1}\otimes I)G((S\otimes I)V^{m-1}), \quad (19)$$

where

$$\begin{aligned} M_1 = [(I-L)\otimes I - h\Lambda\otimes J]^{-1}[((I-L-B)\otimes I) \\ -h(\Lambda - B\bar{A})\otimes J] \quad (20) \\ M_2 = [(I-L)\otimes I - h\Lambda\otimes J]^{-1}[B\bar{A}\otimes I]. \end{aligned}$$

The convergence result is derived from (19) by showing that, for any $k > 0$, there is a diagonal matrix $D > 0$ such that $\|M_1\|_D \leq k$, and by obtaining a bound for $\|M_2\|_D$ and $\|(S^{-1}\otimes I)G((S\otimes I)V^{m-1})\|_D$. The Lemmas $1, 2, 3$ and 4 gives the bound for those.

A bound for the norm of M_2, given by Lemma 3, is independent of stiffness and sufficient for an analysis of convergence. However, the quality of the bound for the norm of M_1 is vital and it becomes necessary to treat M_1 separately for each type of iteration scheme examined.

Theorem 5. *Suppose that there exists a diagonal matrix $D > 0$ such that $\|M_1\|_D \leq k < 1$. Let ϵ be given with $k < \epsilon < 1$. There exist positive constants H and δ such that, for any step length $h \leq H$ and any starting value Y^0 with $\| \, V^0 \, \|_D \leq \delta$, the sequence Y^m converges to Y and*

$$\| \, V^m \, \|_D \leq \epsilon \, \| \, V^{m-1} \, \|_D, \quad m = 1,2,3,\ldots,$$

where $V^m = (S^{-1}\otimes I)(Y - Y^m)$.

Proof: Choose H so that $\alpha - H\nu > 0$. Apply Lemma 3 and Lemma4 to (19) to obtain the inequality

$$\| \, V^m \, \|_D \leq (k + hK) \, \| \, V^{m-1} \, \|_D + hC \, \| \, V^{m-1} \, \|_D^2, \quad (21)$$

where C and K are given, in terms of the condition number $c(S) = |S|_D|S^{-1}|_D$, by

$$\begin{aligned} C &= \frac{\alpha\sigma}{\alpha - H\nu}|B\bar{A}|_D C(S)|S|_D\frac{L}{2}, \\ K &= \frac{\alpha\sigma}{\alpha - H\nu}|B\bar{A}|_D C(S)L[\gamma + \frac{H\beta|f|}{\alpha - Hv}]. \end{aligned}$$

Now choose H and δ so that $k + H(K + C\delta) \leq \epsilon$. Then $\| \, V^{m-1} \, \|_D \leq \delta$ implies $\| \, V^m \, \|_D \leq \epsilon \, \| \, V^{m-1} \, \|_D \leq \delta$ and the result follows. ∎

It has been pointed out that the quality of the bound $\|M_1\|_D \leq k$ is crucial in this result. In this case the convergence rate may be improved, in the sense that ϵ may be made arbitrary small, by decreasing the step length h. It is important that D is chosen so that k is close to $\rho[M_1]$. In the inequality (21) the term in K is controlled by a factor h, and even by h^2 when $\gamma = 0$. For Runge-Kutta methods this can be achieved by choosing $J = f'(x_r)$. The term in C is also controlled by a factor h and is second order in the iterate so that the magnitude of C is not critical if Y^0 is a good approximation to Y. The expressions for K and C both involve the product $\sigma|B\bar{A}|_D c(S)$ and each factor depends on the choice of diagonal matrix D. Some schemes are discussed using the established basic convergence result. For each s-stage method there is a real matrix S such that

$$S^{-1}AS = \bar{A} = A_1 \oplus A_2 \oplus \cdots \oplus A_r \qquad (22)$$

a real block diagonal matrix. The sub matrices are chosen to have the form

$$A_i = \begin{bmatrix} a_i & a_i - b_i \\ a_i + b_i & a_i \end{bmatrix}, \quad i = 1,2,\ldots,r, \; (23)$$

with $b_i > a_i, \quad i = 1, 2, \ldots, r$ and except that, when s is odd, $A_r = [a_r]$. Many iterative methods have coefficient matrices which may be transformed to real block diagonal matrices of the same form as (22).

A. Gauss Method

For the $s-$stage Gauss methods of order $2s$ there is a real matrix S such that $S^{-1}AS$ has the block diagonal form as given by (22). When $s = 2$, S is the identity matrix and when $s > 2$, S may be computed by noting that the columns are linearly independent eigenvectors of $[a_iI - A]^2, \quad i = 1, 2, \ldots, r$. For the Gauss method the scheme is applied with the real transformation (22) and with $L = L_1 \oplus L_2 \oplus \cdots \oplus L_r$ and $B = B_1 \oplus B_2 \oplus \cdots \oplus B_r$ corresponding to the block diagonal form of $S^{-1}AS$. The cases $s = 2$ and $s = 3$ are considered separately. For $s = 2$, where $S = I$, the basic scheme is given by

$$\lambda = b_1, \quad L_1 = \begin{bmatrix} 0 & 0 \\ & \\ 2 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & \frac{b_1-a_1}{a_1+b_1} \\ & \\ -2 & \frac{2a_1}{a_1+b_1} \end{bmatrix}, \tag{24}$$

with $a_1 = \frac{1}{4}$ and $b_1 = \frac{\sqrt{3}}{6}$. Expression (20) gives

$$M_1 = \left(\frac{b_1 - a_1}{b_1 + a_1}\right) \begin{bmatrix} M_{11} & 0 \\ 0 & [M_{11}]^2 \end{bmatrix} \left[\begin{pmatrix} 0 & -1 \\ 0 & -1 \end{pmatrix} \otimes I\right] \tag{25}$$

which is an upper triangular matrix so that it is appropriate to choose $D = [d^2, 1], \ 0 < d \leq 1$, where $M_{11} = (I - hb_1J)^{-1}(I + hb_1J)$. With this choice Lemma 2 gives

$$\|M_1\|_D \leq \frac{b_1 - a_1}{b_1 + a_1}\sqrt{1 + d^2}\left(\frac{\alpha + H\nu}{\alpha - H\nu}\right)^2 \tag{26}$$

For this case, $s = 2$, $\dfrac{b_1 - a_1}{b_1 + a_1} = 7 - 4\sqrt{3} \simeq 0.0718$, and the bound for $\|M_1\|_D$ is close to this minimum, even when d is close to 1, provided that $H\nu$ is small relative to α. Note that $B\bar{A}$ is also upper triangular so that $|B\bar{A}|_D$ decreases with d but that $|L|_D = 2d^{-1}$. It may be shown that

$$|B\bar{A}|_D = \sqrt{\frac{\beta + \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}}, \qquad \sigma = 1 + \frac{2\alpha}{d(\alpha - H\nu)} \tag{27}$$

where $\alpha = (a_1 + b_1)^2, \ \ \beta = a_1^2b_1^2 + 2a_1b_1^3 + 5b_1^4 + a_1^2b_1^2d^2 - 2a_1b_1^3d^2 + b_1^4d^2$ and
$\gamma = 4b_1^6$, this gives, for $s = 2$, $|B\bar{A}|_D < 0.314$ for $d \leq 1$. Note that d needs to be chosen close to 1 to control σ. Since $c(S) = 1$ the convergence theorem holds, with ϵ close to $7 - 4\sqrt{3}$, without undue restriction on h.
When $s = 3$ the basic scheme proposed by Cooper and Butcher has a similar form with $\lambda = b_1$ and L_1 and B_1 given by (24). In addition

$$L_2 = [0], \qquad B_2 \ = \ \left[\frac{2b_1}{b_1 + a_2}\right]. \tag{28}$$

Approximate values of the coefficients are $a_1 \simeq 0.1423$, $b_1 \simeq 0.1967$ and $a_2 \simeq 0.2153$. Again expression (20) gives

$$M_1 = \left(\frac{b_1 - a_1}{b_1 + a_1}\right) \begin{bmatrix} M_{11} & 0 & 0 \\ 0 & [M_{11}]^2 & 0 \\ 0 & 0 & M_{11} \end{bmatrix} \left(\begin{bmatrix} 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & l \end{bmatrix} \otimes I\right).$$

where $l = \dfrac{(b_1 + a_1)(a_2 - b_1)}{(b_1a_1)(b_1 + a_2)}$ and in this case also M_1 is an upper triangular and the suitable choice of diagonal matrix is $D = [d^2, 1, 1]$. This gives the same bound (26) for $\|M_1\|_D$ and, provided $H\nu$ is small relative to α, this bound is close to the value $\dfrac{b_1 - a_1}{b_1 + a_1} \simeq 0.160$. Again (27) holds giving $|B\bar{A}|_D < 0.235$ for $d \leq 1$ and d needs to be chosen near 1 to control σ. Values for the condition number $c(S)$ can be calculated and the value of $c(S)$ is increases with stage s, suggesting a need for some restriction on h to get the convergence theorem holds.

B. Radue IIA Method

For the $s-$stage Radue IIA methods of order $2s-1$ there is a real matrix S such that $S^{-1}AS$ has the block diagonal form as given by (22) with $L = L_1 \oplus L_2 \oplus \cdots \oplus L_r$ and $B = B_1 \oplus B_2 \oplus \cdots \oplus B_r$ corresponding to the block diagonal form of $S^{-1}AS$. The cases $s = 2$ and $s = 3$ are considered separately. For $s = 2$, the basic scheme is given by (24) with $a_1 = \frac{1}{3}$ and $b_1 = \frac{1}{\sqrt{6}}$. With these values, expression M_1 gives an upper triangular matrix same as (25). With this choice and the appropriate choice of $D = [d^2, 1]$, $0 < d \leq 1$, Lemma 2 gives (26) with the corresponding a_1 and b_1. For this case $\dfrac{b_1 - a_1}{b_1 + a_1} \simeq 0.101$, and the bound for $\|M_1\|_D$ is close to this minimum, even when d is close to 1, provided that $H\nu$ is small relative to α. The matrix $B\bar{A}$ is also upper triangular so that $|B\bar{A}|_D$ decreases with d but that $|L|_D = 2d^{-1}$. The bound $|B\bar{A}|_D$ is given by (27) and $|B\bar{A}|_D < 0.459$ for $d \leq 1$. The minimum of $c(S) = 3.98$. Similar way as for the 2-stage Gauss method we can show that the convergence Theorem holds, with ϵ close to 0.101.
When $s = 3$ the basic scheme proposed by Cooper and Butcher [8] has a similar form with $\lambda = b_1$ and L_1 and B_1 given by (24). In addition L_2 and B_2 are given by (28) with the approximate values of the coefficients are $a_1 \simeq 0.1626$, $b_1 \simeq 0.2462$ and $a_2 \simeq 0.2749$. With these values $\|M_1\|_D$ give the same bound (26) and this bound is close to the value $\dfrac{b_1 - a_1}{b_1 + a_1} \simeq 0.205$, with the suitable choice of diagonal matrix is $D = [d^2, 1, 1]$. Again (27) holds giving $|B\bar{A}|_D < 0.308$ for $d \leq 1$ and d needs to be chosen near 1 to control σ. In this case also, we can show that the convergence Theorem holds, with ϵ close to 0.205 with minimum $c(S) \simeq 5.62$.

V. Conclusion

We established the convergence result for the more general linear iteration scheme proposed by Cooper and Butcher [8] and this convergence result was verified by two and three stage Gauss method and Radue IIA method.

References

[1] T.A.Bickart, “An efficient solution process for implicit Runge-Kutta methods”, *SIAM J. Numer. Anal.*, vol.14, pp.1022-1027, 1977. https://doi.org/10.1137/0714069.

[2] J.C.Butcher, “On the implementation of implicit Runge-Kutta methods”, *BIT*, vol.16, pp.237-240, 1976. https://doi.org/10.1007/BF01932265.

[3] J.C.Butcher, “Some implementation schemes for implicit Runge-Kutta methods”, *Proceeding of the Dundee Conference on Numerical Analysis, Lecture Notes in Mathematics*, vol.773, pp.12-24, 1980. https://doi.org/10.1007/BFb0094160.

[4] J.C.Butcher and J.R.Cash, “Towards efficient Runge-Kutta methods for stiff systems”, *SIAM J.Numer. Anal.*, vol.27, pp.753-761, 1990. https://doi.org/10.1137/0727044.

[5] J.R.Cash, “On a class of implicit Runge-Kutta procedures”, *IMA J. Appl. Math.*, vol.19, pp.455-470, 1977. https://doi.org/10.1093/imamat/19.4.455.

[6] F.H.Chipman, “The implementation of Runge-Kutta implicit processes”, *BIT*, vol.13, vol.391-393, 1973. https://doi.org/10.1007/BF01933400.

[7] A.G.Collings and G.J.Tee, “An analysis of Euler and implicit Runge-Kutta numerical integration schemes for structural dynamic problems”, *Proceeding of the Sixth Australasian Conference on the Mechanics of Structures and Materials*, vol.1, pp.147-154, 1977.

[8] G.J.Cooper and J.C.Butcher, “An iteration scheme for implicit Runge-Kutta methods”, *IMA J. Numer. Anal.*, vol.3, pp.127-140, 1983. https://doi.org/10.1093/imanum/3.2.127.

[9] G.J.Cooper, “On the implementation of singly implicit Runge-Kutta methods”, *Math. Comp.*, vol.57, pp.663-672, 1991. https://doi.org/10.2307/2938710.

[10] G.J.Cooper, “On the existence of solutions for algebraically stable Runge-Kutta methods”, *IMA J. Numer. Anal.*, vol.6, pp.325-330, 1986. https://doi.org/10.1093/imanum/6.3.325.

[11] G.J.Cooper and R.Vignesvaran, “A scheme for the implementation of implicit Runge-Kutta methods”, *Computing*, vol.45, pp.321-332, 1990. https://doi.org/10.1007/BF02238800.

[12] G.J.Cooper and R.Vignesvaran, “Some schemes for the implementation of implicit Runge-Kutta methods”, *J. Comp. App. Math.*, vol.45, 213-225, 1993. https://doi.org/10.1016/0377-0427(93)90276-H.

[13] G.J.Cooper and R.Vignesvaran, “On the Use of Parallel Processors for Implicit Runge-Kutta Methods”, *Computing*, vol.51, pp.135-150, 1993. https://doi.org/10.1007/BF02243848.

[14] M.Crouzeix, W.H.Hundsdorfer and M.N.Spijker, “On the existence of solutions to the algebraic equations in implicit Runge-Kutta methods”, *BIT*, vol.23, pp.84-91, 1983. https://doi.org/10.1016/S0377-0427(99)00129-6.

[15] K.Dekker, “Error bounds for the solution to the algebraic equations in Runge-Kutta methods”, *BIT*, vol.24, pp.347-356, 1984. https://doi.org/10.1007/BF02136033.

[16] W.H. Enright, “Improving the efficiency of matrix operations in the numerical solution of ODEs”, *Technical Report no.98 (Computer Science Dept. Univ. of Toronto)*, 1976.

[17] R. Frank and C.W. Ueberhuber, “Iterated defect correction for the efficient solution of stiff systems of ordinary differential equations”, *BIT*, vol.17, pp.146-159, 1977. https://doi.org/10.1007/BF01932286.

[18] A.Iserles and S.P.Norsett, “On the theory of parallel Runge-Kutta methods”, *IMA J. Numer. Anal.*, vol.10, pp.463-488, 1990. https://doi.org/10.1093/imanum/10.4.463.

[19] K. R.Jackson and S. P.Norsett, “The potential for parallelism in Runge-Kutta methods”, *Part 1: RK formulas in standard form, Tech. Report no.239/90 (Computer Science Dept. Univ. of Toronto)*, 1990.

[20] O.A.Karakashian and W.Rust, “On the parallel implementation of implicit Runge-Kutta methods”, *SIAM J. Sci. Stat. Comput.*, vol.9, pp.1085-1090, 1988. https://doi.org/10.1137/0909074.

[21] J.F.B.M.Kraaijevanger and J.Schneid, “On the unique solvability of the Runge-Kutta equations”, *Numer. Anal, Report no.90-6 (Mathematics Dept, Massachusetts Institute of Technology)*, 1990.

[22] H.O.Kreiss, “Difference methods for stiff ordinary differential equations”, *SIAM J. Numer. Anal.*, vol.15, pp.21-58, 1978. https://doi.org/10.1137/0715003.

[23] J.M.Varah, “On the efficient implementation of implicit Runge-Kutta methods”, *Math. Comp.*, vol.33, pp.557-561, 1979. https://doi.org/10.2307/2006296.

[24] P.J.Van der Houwen and B.P.Sommeijer, “Iterated Runge-Kutta methods on parallel computers”, *SIAM J. Sci. Comput.*, vol.12, pp.1000-1028, 1991. https://doi.org/10.1137/0912054.

[25] R.Vigneswaran, “A Non-Linear Scheme based on projection method for Two stage Gauss Method”, *International Journal of Pure and Applied Mathematics(IJPAM)*, vol.101, no.4, pp.463-476, 2015.

[26] R.Vigneswaran and S.Kajanthan, “Some Efficient Implementation Schemes for Implicit Runge- Kutta Methods”, *International Journal of Pure and Applied Mathematics(IJPAM)*, vol.93, no.4, pp.525–540, 2014. https://doi.org/10.12732/ijpam.v93i4.4.

[27] R.Vigneswaran and S.Kajanthan, “Improving the Rate of Convergence for a scheme with Extra sub-step for Two Stage Gauss Method”, *International Journal of Pure and Applied Mathematics(IJPAM)*, vol.116, no.1, pp.243–261, 2017. https://doi.org/10.12732/ijpam.v116i1.24.

[28] R.Vigneswaran and S.Kajanthan, “Accelerating the Rate of Convergence of a More General Linear Iteration Scheme by using the Projection Method for Implicit Runge-Kutta Methods”, *Proceedings of the 4^{th} Jaffna University International Research Conference (JUICe2018)*, vol.4, pp.573–577, 2018.

[29] R. Vigneswaran and S.Kajanthan, “A Class of s-step Non-linear Iteration Scheme Based On Projection Method For Gauss Method”, *Advances in Mathematical Sciences*, vol.2, pp.26–30, 2019. https://doi.org/10.37516/adv.math.sci.2019.0061