Texas A&M University Campus Libraries
Courier
ILLiad TN: 2740290

Journal Title: International Journal of Computer Mathematics

Volume: 21
Issue: 1
Month/Year: 1987
Pages: 65-84

Article Author: Ala H. Al-Rabeh

Article Title: Embedded dirk methods for the numerical integration of stiff systems of odes

8/7/2014 2:27:51 PM
(Please update within 24 hours)

Call #: QA76 .I59

Location: evans

Not Wanted Date: 02/03/2015

Status: Graduate/Professional Student
Phone: 618-780-8290
E-mail: pmaginot@neo.tamu.edu

Name: Maginot, Peter

Pickup at Evans

Address:
621 Navarro Dr
College Station, TX 77845

Intern. J. Computer Math., 1987, Vol. 21, pp. 65–84
Photocopying permitted by license only
© 1987 Gordon and Breach, Science Publishers, Inc.
Printed in Great Britain

Embedded DIRK Methods for the Numerical Integration of Stiff Systems of ODEs

ALA H. AL-RABEH

Department of Systems Engineering, University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

(*Received January, 1986*)

In this paper, the optimal order of embedded pairs of Diagonally Implicit Runge–Kutta (DIRK) methods is examined. It is shown that a q-stage DIRK method of order p embedded in a $q+1$ stage DIRK method of order $p+1$ cannot have $p=q+1$. Thus adopting embedding techniques to estimate the local truncation error results in giving up an order of accuracy for $q<6$. Embedded pairs of orders two and three for the basic method are derived with the additional stage being either explicit or implicit. Numerical results indicate that significant savings are realized when the extra stage is explicit.

An analysis of A and L-stability properties of q-stage order q DIRK methods with unequal diagonal elements is presented. Necessary and sufficient conditions for A and L-stability are derived.

To assess the potential of such methods, a number of embedded DIRK formulas are implemented. Numerical results for selected test problems are presented.

KEY WORDS: Stiff systems of ordinary differential equations, Diagonally Implicit Runge–Kutta methods, A-stability.

C.R. CATEGORIES: G.1.7. Ordinary differential equations, stiff equations.

1. INTRODUCTION

In this paper we are concerned with the approximate numerical integration of the m-dimensional stiff initial value problem,

$$y' = f(x, y), \qquad y(x_0) = y_0 \tag{1}$$

using Diagonally Implicit Runge–Kutta (DIRK) formula:

$$K_i^{(n)}=f\left(x_n+c_ih, y_n+h\sum_{j=1}^{i} a_{ij}K_j^{(n)}\right), \quad i=1(1)q$$

and,

$$y_{n+1}=y_n+h\sum_{i=1}^{q} b_iK_i^{(n)}. \tag{2}$$

If we further have, $a_{ii}=\gamma$, $i=1(1)q$, then the class of methods (2) is called Singly Diagonally Implicit Runge–Kutta methods (SDIRK). Such formulae have been studied by Norsett [17], Crouzeix [9], Alexander [1], Cash [7], Al-Rabeh [2, 3], and Cooper and Safy [10]. Cash [8] suggested a variable order/variable stepsize SDIRK algorithm.

Using the DIRK method (2) requires a way of estimating the local truncation error. Alexander [1] used extrapolation techniques to approximate the local truncation error. In Alexander's implementation, see also Gaffney [16], two matrices of the form,

$$I-\gamma h\frac{\partial f}{\partial y}, \quad \text{and} \quad I-2\gamma h\frac{\partial f}{\partial y}$$

have to be LU-decomposed with each Jacobian update and a stepsize change. Thus, obtaining such an estimate may require a considerable increase in computational cost. An alternative class of methods for estimating the local truncation error is the embedding techniques. Norsett [17] and Cash [8] implemented embedded pairs of DIRK methods, Burrage *et al.* [5] used embedding techniques to estimate the local truncation error for the singly implicit Runga–Kutta methods.

Define,

$$K_{q+1}^{(n)}=f\left(x_n+c_{q+1}h,\ y_n+h\sum_{j=1}^{q+1} a_{ij}K_j^{(n)}\right)$$

and,

$$\hat{y}_{n+1}=y_n+h\sum_{i=1}^{q+1}\hat{b}_iK_i^{(n)}. \tag{3}$$

If the basic DIRK method (2) and the embedding method defined by (2) and (3) are of orders p and $p+1$ respectively, then the principal local truncation error in the basic DIRK method could be estimated by,

$$\begin{aligned}T_{n+1}&=\hat{y}_{n+1}-y_{n+1}\\&=h\left[\sum_{i=1}^{q}(\hat{b}_i-b_i)K_i^{(n)}+\hat{b}_{q+1}K_{q+1}^{(n)}\right].\end{aligned} \tag{4}$$

Norsett [17] and Cash [7] considered the case in which the $q+1$ stage is implicit, thus

$$a_{ii}=\gamma, \qquad i=1(1)q+1. \tag{5}$$

Al-Rabeh [2], and Norsett *et al.* [19], suggested that the $q+1$ stage could be explicit, and in particular,

$$a_{ii}=\begin{cases}\gamma, & \text{for} \quad i=1(1)q\\ 0, & \text{for} \quad i=\mathrm{q}+1\end{cases}. \tag{6}$$

Hereafter, we will refer to the DIRK pair defined by (2) and (3) as a matched pair of DIRK methods $(q/q+1,\ p/p+1)$. In Section 2, it is shown that a matched pair of DIRK methods $(q/q+1, p/p+1)$ cannot have $p=q+1$, whereas generally the optimum order of a q-stage DIRK method is $q+1$ for $q<6$, Al-Rabeh [2], Cooper and Sayfy [10]. In Section 3, optimal order embedded pairs of DIRK methods of orders two and three for the basic DIRK method are derived. In Section 4, the A and L-stability properties of the (q, q) DIRK method where the diagonal elements are not necessary equal are analyzed. Necessary and sufficient conditions for A and L-stability are obtained. In Section 5, the implementation details of three pairs of DIRK methods are presented.

2. OPTIMUM ORDER EMBEDDED DIRK METHODS

In this section we show that a matched pair of DIRK methods

$(q/q+1;\ p/p+1)$ cannot have order $p=q+1$. Following Dalquist *et al.* [11] we state,

DEFINITION A DIRK method is nonconfluent if all c_i are distinct and confluent otherwise.

As in Butcher [6], we define the following relationship on the coefficients a_{ij}, nodes c_i and weights b_i.

$$B(p): g_k = 0, \quad \text{for} \quad k = 1(1)p,$$

where,

$$(1+g_k)/k = \sum_{i=1}^{q} c_i^{k-1} b_i. \tag{7}$$

THEOREM 1 *Assume that the basic q-stage DIRK method is of order q then an embedding nonconfluent DIRK method of order $q+1$ with weights $(\hat{b}_i)$ and nodes (c_i) must have at least one additional quadrature node corresponding to a non-zero weight.*

Proof Follows directly from (7).

This theorem implies that for the embedding method to be of one order higher, it must have at least one extra stage, thus we justify the definition of the embedded DIRK methods given by (2) and (3).

THEOREM 2 *A matched pair of nonconfluent DIRK methods $(q/q-1,\ p/p+1)$ cannot have $p=q+1$.*

Proof Assume that the basic DIRK method is of order $q+1$ then $B(q+1)$ i.e.

$$\sum_{i=1}^{q} c_i^{k-1} b_i = 1/k, \quad k = 1(1)q+1. \tag{8}$$

Since all the c_i, $i=1(1)q$ are distinct, then the first q equations of (8) determine the weights b_i, $i=1(1)q$ uniquely. In fact,

$$b = V_q^{-1} \delta_q \tag{9}$$

where, $b=[b_1,\ldots,b_q]^t$, $\delta_q=[1,\ 1/2,\ldots,1/q]^t$, and

$$V_q=\begin{bmatrix} 1 & \cdots & 1 \\ c_1 & \cdots & c_q \\ & & \\ c_1^{q-1} & \cdots & c_q^{q-1} \end{bmatrix}$$

is the $q\times q$ Vandermond matrix.

The last component of Eq. (8) gives

$$cb=1/(q+1),$$

where, $c=[c_1^q\ldots c_q^q]$, which using (9) reduces to,

$$cV_q^{-1}\delta_q=1/(q+1). \tag{10}$$

If the embedding DIRK method is of order $q+2$, then $B(q+2)$ i.e.

$$\sum_{i=1}^{q+1} c_i^{k-1}\hat{b}_i=1/k, \quad k=1(1)q+2. \tag{11}$$

The first $q+1$ equations give,

$$\hat{b}=V_{q+1}^{-1}\delta_{q+1}. \tag{12}$$

Using (10), the last component of (12) is,

$$\hat{b}_{q+1}=(cV_q^{-1}\delta_q-1/(q+1))\Big/\prod_{i=1}^{q}(c_i-c_{q+1})=0.$$

Hence for (8) and (11) to be consistent we must have,

$$\hat{b}_i=b_i, \quad i=1(1)q$$

i.e., the embedding and the basic DIRK methods are identical, a contradiction.

The implication of the previous theorem is that adopting an embedding technique to estimate the local truncation error requires giving up an order of accuracy. As the optimum order of a q-stage DIRK method is $q+1$, for $q<6$, Al-Rabeh [2], Cooper and Safy [10]. Note that the use of extrapolation to estimate the local truncation error does not impose such a penalty.

3. THE CONSTRUCTION OF EMBEDDED DIRK METHODS

In this section we construct matched pairs of DIRK methods of orders two and three for the basic DIRK method.

3.1 (2/3, 2/3)

Recall that for the basic DIRK method to be of order two, we require, Butcher [6],

$$\phi=1,\ [\phi]=1/2,$$

whereas for the embedding DIRK method to be of order three, we require,

$$\hat{\phi}=1,\ [\hat{\phi}]=1/2,\ [\hat{\phi}^2]=1/3,\quad \text{and}[{}_2\hat{\phi}]_2=1/6,$$

where ϕ, $[\phi], [\phi^2], \ldots$ etc. are the elementary weights of Butcher [6].

Furthermore we let,

$$a_{11}=a_{22}=\gamma,\ a_{33}=\begin{cases}0, \text{ for expl. err. estimat.}\\ \gamma, \text{ for impl. err. estimat.}\end{cases}$$

In all we have eleven parameters and nine constraints. However, two interesting choices of parameters, that may lead to reductions in computational cost, are presented.

Case 1 Take, $c_3=\gamma+1$, $a_{31}=b_1$, $a_{32}=b_2$, and $a_{33}=\gamma$, then (2) and (3) give,

$$K_3^{(n)}=K_1^{(n+1)}$$

provided that the stepsize is unaltered for at least two successive steps. Hence over two successive and accepted steps with no change in h five iteration cycles are required only. The matched pair of DIRK methods are characterized by the following table.

$$
\begin{array}{c|cc}
\gamma & \gamma & 0 \\
\beta & \beta-\gamma & \gamma \\
\hline
 & \dfrac{2\beta-1}{2(\beta-\gamma)} & \dfrac{2\gamma-1}{2(\gamma-\beta)}
\end{array}
\tag{13}
$$

$$
\begin{array}{c|ccc}
\gamma & \gamma & 0 & 0 \\
\beta & \beta-\gamma & \gamma & 0 \\
1+\gamma & \dfrac{2\beta-1}{2(\beta-\gamma)} & \dfrac{2\gamma-1}{2(\gamma-\beta)} & \gamma \\
\hline
 & \dfrac{A-1/6}{\beta-\gamma} & \dfrac{6\gamma^2-1}{6(\gamma-\beta)(\gamma-\beta+1)} & \dfrac{A+1/3}{(\gamma-\beta+1)}
\end{array}
\tag{14}
$$

where, $A=\gamma\beta-(\gamma+\beta)/2$, and $\beta=3\gamma(-4\gamma^2+2\gamma+1)$, given that $\gamma\neq\beta-1$, $\gamma\neq\beta$, $\gamma\neq(1\pm 1/\sqrt{3})/2$.

Case 2 Let, $a_{31}=b_1$, $a_{32}=b_2$, and $a_{33}=0$. Then the matched pair of DIRK methods could be characterized by the following table,

$$
\begin{array}{c|cc}
\gamma & \gamma & 0 \\
\beta & \beta-\gamma & \gamma \\
\hline
 & \dfrac{2\beta-1}{2(\beta-\gamma)} & \dfrac{2\gamma-1}{2(\gamma-\beta)}
\end{array}
\tag{15}
$$

$$
\begin{array}{c|ccc}
\gamma & \gamma & 0 & 0 \\
\beta & \beta-\gamma & \gamma & 0 \\
1 & \dfrac{2\beta-1}{2(\beta-\gamma)} & \dfrac{2\gamma-1}{2(\gamma-\beta)} & 0 \\
\hline
 & \dfrac{A}{(\beta-\gamma)(1-\gamma)} & \dfrac{B}{(\gamma-\beta)(1-\beta)} & \dfrac{C}{(1-\gamma)(1-\beta)}
\end{array}
\tag{16}
$$

where, $A=\beta/2-1/6$, $B=\gamma/2-1/6$, $C=\gamma\beta-(\gamma+\beta)/2+1/3$, and, $\beta=\gamma(6\gamma^2-8\gamma+3)/(6\gamma^2-4\gamma+1)$, given that $\gamma\neq\beta\neq 1$, $\gamma\neq(1+1/\sqrt{3})/2$.

Note that for this choice,

$$K_3^{(n)}=f_{n+1}.$$

Case 3 Take $\gamma=1-1/\sqrt{2}$, $c_2=1$, then the matched pair of DIRK methods is given by,

$$\begin{array}{c|cc} \gamma & \gamma & 0 \\ 1 & 1-\gamma & \gamma \\ \hline & 1-\gamma & \gamma \end{array} \tag{17}$$

$$\begin{array}{c|ccc} \gamma & \gamma & 0 & 0 \\ 1 & 1-\gamma & \gamma & 0 \\ \beta & \delta & \beta-\delta & 0 \\ \hline & \dfrac{3\beta-1}{6(1-\gamma)(\beta-\gamma)} & \dfrac{2\gamma(3\beta-1)-1}{6(\beta-1)} & \dfrac{3\gamma-1}{6(\beta-\gamma)(\beta-1)} \end{array} \tag{18}$$

where $\delta=(\gamma-1)(4\beta^2-6\beta+1)$, $\beta\neq\gamma\neq 1$.

Note that the basic (2, 2) DIRK method is strongly S-stable, Alexander [1]. Moreover, Alexander implemented (17) using an extrapolation method to estimate the error.

3.2 (3/4, 3/4)

Consider the construction of the basic (3, 3) DIRK method, we require,

$$\phi=1,\ [\phi]=1/2,\ [\phi^2]=1/3, \quad \text{and} \quad [_2\phi]_2=1/6.$$

For the embedding (4, 4) DIRK method we require,

$$\hat{\phi}=1,\ [\hat{\phi}]=1/2,\ [\hat{\phi}^2]=1/3,\ [_2\hat{\phi}]_2=1/6,$$

$$[_3\hat{\phi}]_3=1/24,\ [_2\hat{\phi}^2]_2=1/12,\ [\hat{\phi}^3]=1/4,$$

$$[\hat{\phi}[\hat{\phi}]]=1/8.$$

Case I: Implicit error estimates Choose, $a_{11}=a_{22}=a_{33}=a_{44}=\gamma$, then lengthy algebraic manipulation yields the following two-parameter family of DIRK methods.

$$c_1=a_{11}=\gamma,$$

$$c_2=\alpha=\gamma(2\gamma-1)^2/(4\gamma^2-2\gamma+1/3),$$

$$c_3=\beta,$$

$$c_4=\delta=1-\gamma,$$

$$b_1=\frac{(\alpha-1/2)\beta-(\alpha/2-1/3)}{(\alpha-\gamma)(\beta-\gamma)},$$

$$b_2=\frac{(\gamma-1/2)\beta-(\gamma/2-1/3)}{(\gamma-\alpha)(\beta-\alpha)},$$

$$b_3=\frac{(\gamma-1/2)\alpha-(\gamma/2-1/3)}{(\gamma-\beta)(\alpha-\beta)},$$

$$a_{21}=\gamma-\alpha,$$

$$a_{32}=\frac{\gamma^2-\gamma+1/6}{b_3(\alpha-\gamma)},$$

$$a_{31}=\beta-\gamma-a_{32},$$

$$\hat{b}_1=\frac{A\alpha-B}{(\delta-\gamma)(\beta-\gamma)(\alpha-\gamma)},$$

$$\hat{b}_2 = \frac{A\gamma - B}{(\gamma-\alpha)(\beta-\alpha)(\delta-\alpha)},$$

where,

$$A = (1/2-\gamma)\beta - (1/3-\gamma)/2,$$

and,

$$B = (1/3-\gamma)\beta/2 - \gamma,$$

$$\hat{b}_3 = \frac{E(1-\gamma) - F}{(\gamma-\beta)(\alpha-\beta)(\delta-\beta)},$$

$$\hat{b}_4 = \frac{E\beta - F}{(\gamma-\delta)(\alpha-\delta)(\beta-\delta)},$$

where,

$$E = (\gamma - 1/2)\alpha - (\gamma/2 - 1/3),$$

and,

$$F = (\gamma/2 - 1/3)\alpha - (\gamma/3 - 1/4),$$

$$a_{43} = \frac{(\gamma^2 - \gamma + 1/6)(1/2-\alpha)}{\hat{b}_4(\beta-\gamma)(\beta-\alpha)},$$

$$a_{42} = \frac{(\gamma^2 - \gamma + 1/6)(1/2-\alpha) - \hat{b}_3(\alpha-\gamma)(\alpha-\beta)a_{32}}{\hat{b}_4(\alpha-\gamma)(\alpha-\beta)}.$$

Example Set, $\gamma = 0.4358665$, $\beta = 1-\gamma$, then the basic SDIRK method is L-stable, see Appendix I for a listing of the coefficients of this method.

Case II: Explicit error estimation Choose, $a_{11} = a_{22} = a_{33} = \gamma$ and $a_{44} = 0$, then lengthy algebraic manipulation yields the following three-parameter family of DIRK methods.

$$c_1 = a_{11} = \gamma,\ c_2 = \alpha,\ c_3 = \beta,\ c_4 = \delta,$$

$a_{31}, a_{32}, \hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4, c_4$ are the same as Case I.

$$a_{43} = \frac{g_1(1/2-\alpha) + \gamma(\gamma-\delta)(\alpha-\delta)\hat{b}_4}{(\beta-\alpha)(\beta-\delta)\hat{b}_4},$$

$$a_{42} = \frac{g_1(1/2-\beta) + \gamma(\gamma-\delta)(\beta-\delta)\hat{b}_4 + a_{32}(\alpha-\gamma)(\beta-\alpha)\hat{b}_3}{(\alpha-\beta)(\alpha-\gamma)\hat{b}_4}$$

$$a_{41} = \delta - a_{43} - a_{42} - \gamma,$$

where $g_1 = (\gamma^2 - \gamma + 1/6)/2$.

Example Take $\delta = 1$, $\beta = 1 - c_2$, and $\gamma = 0.4358665$, then the basic SDIRK method is L-stable, for a listing of the coefficients of this method see Appendix I.

4. STABILITY

Consider the application of the q-stage DIRK method given by (2) to the scalar equation,

$$y' = \lambda y, \quad \lambda \in \mathbb{C}, \quad \mathrm{Re}(\lambda) \leqq 0.$$

The resulting equation is,

$$y_{n+1} = E_q(z) y_n$$

where, $E_q(z) = 1 + zb^t(I - zA)^{-1}e$, $z = \lambda h$, e is the q-dimensional vector, whose elements are unity, $A = (a_{ij})$, and b^t is the vector of weights transposed.

Rewriting $E_q(z)$ as a rational function gives,

$$E_q(z) = \det \frac{(I - zA) + zb^t(I - zA)}{\det(I - zA)}.$$

Since A is lower triangular for DIRK methods then,

$$E_q(z)=\det\left|\begin{array}{c|c} I-zA & e \\ \hline -zb^t & 1 \end{array}\right| \Big/ \prod_{i=1}^{q}(1-za_{ii}).$$

Set,

$$a_{ii}=\gamma_i, \quad i=1(1)q$$

and,

$$t_1=\sum_i \gamma_i, \quad t_2=\sum_{i_1, i_2}\gamma_{i_1}\gamma_{i_2},$$

$$t_{q-1}=\sum_{i_1, i_2, \ldots, i_{q-1}}\gamma_{i_1}\gamma_{i_2}\cdots\gamma_{i_{q-1}}$$

$$t_q=\gamma_{i_1}\gamma_{i_2}\cdots\gamma_{i_q}.$$

with $t_0=1$ and indices i_j ranging from 1 to q, and $i_1<i_2\ldots<i_q$.

THEOREM 3 *For any q-stage DIRK method of order q or greater the growth function is given by,*

$$E_q(z,t_0,\ldots,t_q)=1+\frac{\sum_{i=0}^{q-1}\left[\sum_{j=0}^{i}\frac{(-1)^j t_j}{(i-j+1)!}\right]}{\sum_{i=0}^{q}(-1)^i t_i z^i}z.$$

Proof

$$E_q(z)=\det\left|\begin{array}{c|c} I-zA & e \\ \hline -zb^t & 1 \end{array}\right| \Big/ \prod_{i=1}^{q}(1-a_{ii}z)$$

$$=1+zP(z)\Big/\prod_{i=1}^{q}(1-a_{ii}z)$$

where,

$$P(z)=\sum_{i=0}^{q-1}\left(\sum_{j=0}^{i}(-1)^j[_{i-j}\phi]_{i-j}t_j\right)z^i.$$

Since the DIRK method is of order q then, Butcher [6],

$$[_j\phi]_j = 1/(j+1)!, \quad j = 0, 1, \ldots, k-1$$

giving the result.

We now attempt to derive conditions on the parameters of the basic DIRK method that would ensure A-stability or L-stability.

Define,

$$u_i = \begin{cases} \sum_{j=0}^{i} (-1)^j t_i/(i-j)!, & 1 \leqq i \leqq q. \\ 1 \qquad , & i = 0. \\ 0 \qquad , & \text{otherwise.} \end{cases}$$

THEOREM 4 *A (q,q) DIRK method is A-stable if and only if,*

i) $\gamma_i \geqq 0, \qquad i = 1(1)q.$

ii) $(u_{q-j}^2 - 2u_{q-j-1}u_{q-j+1} + 2u_{q-j-2}u_{q-j+2})$

$$\leqq (t_{q-j}^2 - 2t_{q-j-1}t_{q-j+1} + 2t_{q-j-2}t_{q-j+2}).$$

Furthermore, an A-stable (q,q) DIRK method is L-stable iff $u_q = 0$.

Proof Follows directly from the maximum modulus principle and Theorem 3.

Note that Cooper and Safy [10] derived conditions for A-stability for DIRK methods of any order p, see also Norsett [18].

Remark Norsett [18] derived expressions for $E_q(z)$ for the case of equal diagonal elements, and computed the range of γ for A and L-stability, see Burrage [4].

5. NUMERICAL IMPLEMENTATION

In this section we discuss the implementation of three DIRK formulae, the first is the strongly S-stable DIRK method defined by

A and *L*-stable SDIRK methods

q	Range of γ for A-stability	L-stability
1	$[1/2, \infty)$	1
2	$[1/4, \infty)$	$(2+/2)/2$
3	$[1/3, 1.06858]$	0.43587
4	$[0.394, 1.28058]$	0.57282
5	$[0.24651, 0.36180] U [0.42078, 0.47327]$	0.27805
6	$[0.28407, 0.54091]$	0.33414

(17, 18) with $\delta=0$ due to Alexander [1]. The second is the A-stable DIRK method defined by (15, 16) with $\gamma=0.26$. The third DIRK method was implemented by Norsett [17] and is defined by (13, 14) with $\gamma=1-1/\sqrt{2}$. The packages are called DIRKS, DIRKA, and DIRKL respectively. In all packages the Jacobian matrix (J) was evaluated numerically. An LU-decomposition of the iteration matrix $(I-\gamma hJ)$ is retained and may have to be updated if:

a) h changes,

b) the elements of J currently in use, become outdated causing difficulties in convergence.

Note that in case (a) the current J is used to recompute the LU-decomposition of the iteration matrix, and only in case (b) is J itself recomputed. To assess the performance of the DIRK methods tested, two types of measurements were taken, the first reflects cost, the second reflects reliability. The cost measurements include CPU time, number of function evaluations (FEV), number of Jacobian updates (JEV) and number of matrix inversions (MIV). The number of successful and unsuccessful integration steps were recorded. Two measures of reliability were included, the first, the maximum global error defined by,

$$\mathrm{MGE} = \max\,[e_i,\ 1 \leqq i \leqq M]$$

the second is the average global error, defined by

$$\mathrm{AGE} = \left(\sum_{i=1}^{M} e_i\right) \Big/ m\,M,$$

where,

$$e_i = \left\|(y(x_i) - y_i)/(1 + Abs(y(x_i))\right\|.$$

$\|\cdot\|$ is the root mean square norm, M is the total number of successful integration steps, and m is the dimension of (1). All packages were required to compute the solution at a predetermined set of points, using quadratic interpolation. In all packages y_{n+1} is used to advance the solution and $\hat{y}_{n+1}$ is used for error estimation purposes only. For a listing of test problems and numerical results see Appendix II.

6. CONCLUSIONS

Overall the explicit error estimator associated with DIRKA and DIRKS performed reasonably well compared with the implicit error estimator associated with DIRKL. The figures for the maximum and average global errors indicate that the explicit error estimate is as reliable as the implicit one. Moreover, as there is no iteration associated with the explicit stage significant savings in overheads, and function evaluations are achieved.

References

[1] R. Alexander, Diagonally Implicit Runge–Kutta methods for stiff ODE's, *SIAM J. of N.A.* **14** (1977), 1006–1021.

[2] A. Al-Rabeh, Semi-implicit Runge–Kutta methods for the numerical solution of stiff ODE's, Ph.D. Thesis, University of Manchester (1979).

[3] A. Al-Rabeh, A note on the Enright updating technique, *Comp. & Maths with Appls.* **10** (1984), 405–411.

[4] K. Burrage, A special family of Runge–Kutta methods for solving stiff ODE's, *BIT* **18** (1978), 22–41.

[5] K. Burrage, J. C. Butcher and F. H. Chipman, An implementation of singly implicit Runge–Kutta methods, *BIT* **20** (1980), 326–340.

[6] J. C. Butcher, Implicit Runge–Kutta processes, *Math. Comp.* **18** (1964), 50–64.

[7] J. R. Cash, Diagonally implicit Runge–Kutta formulae with error estimate, *Journ. Inst. Math. & Applics.* **24** (1979), 293–301.

[8] J. R. Cash and C. B. Leim, On the design of a variable order, variable step diagonally implicit Runge–Kutta algorithm, *J. Inst. Maths. & Applics.* **26** (1980), 87–98.

[9] M. Crouzeix, On the approximation of ordinary differential equations using Runge–Kutta methods. Thesis, Paris (1975).

[10] G. J. Cooper and A. Safy, Semi-explicit A-Stable Runge–Kutta methods, *Maths. of Comp.* **33** (1979), 541–556.
[11] G. Dahlquist and R. Jeltsch, *Generalized Disks of Contractivity for Explicit Runge–Kutta Methods*, Report TRITA-NA-7906, Dept. of Comp. Sci., Roy. Inst. of Techn., Stockholm (1979).
[12] B. Ehle, *A Comparison of Numerical Methods for Solving Certain Stiff ODE's*, Report No. 70, Dept. of Math., University of Victoria, Canada (1972).
[13] W. H. Enright, T. E. Hull and B. Lindberg, Comparing numerical methods for stiff systems of ODE's, *BIT* **15** (1975), 10–48.
[14] W. H. Enright, Improving the efficiency of matrix operations in numerical solution of stiff ODE's, *TOMS* **4** (1978), 127–136.
[15] I. Gladwell, J. A. L. Craigie and C. R. Crowther, *Testing Initial Value Problem Subroutines as Black Boxes*, N.A. Report 34, Dept. of Math., University of Manchester, England.
[16] P. Gaffney, A performance evaluation of some Fortran subroutines for the solution of stiff oscillatory ODE's, *TOMS* **10** (1984), 58–72.
[17] S. P. Norsett, *Semi Explicit Runge–Kutta Methods*, Math. Dept. University of Trondheim, Reprint No. 6/74.
[18] S. P. Norsett, C-polynomials for rational approximation to the exponentional function, *Numer. Math.* **25** (1975), 39–56.
[19] S. P. Norsett and P. G. Thomsen, Embedded SDIRK methods of basic order three, *BIT* **24** (1984), 634–646.
[20] L. F. Shampine, Evaluation of a test set for stiff ODE solvers, *TOMS* **7** (1981), 409–420.

Appendix I

Matched pairs of DIRK methods

$$
\begin{array}{c|ccc}
0.4358665 & 0.4358665 & 0 & 0 \\
0.0323722 & -0.4034943 & 0.4358665 & 0 \\
0.9676278 & -0.3298752 & 0.8616364 & 0.4358665 \\
\hline
 & -0.6307827 & 0.1413538 & 0.2278634
\end{array}
\qquad \text{(i)}
$$

$$
\begin{array}{c|cccc}
0.4358665 & 0.4358665 & 0 & 0 & 0 \\
0.0323722 & -0.4034943 & 0.4358665 & 0 & 0 \\
0.9676278 & -0.3298752 & 0.8616364 & 0.4358665 & 0 \\
1.0 & -0.7967301 & 1.1666770 & 0.6300530 & 0 \\
\hline
 & 0.559072 & 0.1651247 & 0.7670000 & -0.491197
\end{array}
\qquad \text{(ii)}
$$

0.4358665	0.4358665	0	0	0	
0.0323722	−0.4034943	0.4358665	0	0	(iii)
0.9676278	−0.3298751	0.8616364	0.4358665	0	
0.5641335	0.5575315	−0.1930865	−0.2361781	0.4358665	
	0.3153914	0.1846086	0.1846086	0.3153914	

Remark (i) and (ii) constitute a (3/4, 3/4) pair of DIRK methods with explicit error estimate. Whereas, (i) and (iii) constitute a (3/4, 3/4) pair of DIRK methods with implicit error estimate.

Appendix II

Test problems and numerical results

A total of 15 problems were selected for testing purposes. These were provided by Ehle [12], Gladwell *et al.* [15], Enright [14], and Enright *et al.* [13] incorporating the corresponding suggestions of Shampine [20]. A representative sample of these test problems is presented.

1. Test problem A: Linear with real eigenvalues.
   Problem A3 of Enright *et al.* [13].
2. Test problem B: Linear with complex eigenvalues.
   Problem 13 of Ehle [12].
3. Test problem C: Non-linear coupling.
   Problem C1 of Enright *et al.* [13].
4. Test problem D: Non-linear with real eigenvalues.
   Problem D2 of Enright *et al.* [13].
5. Test problem E: Non-linear with non-real eigenvalues.
   Problem E4 of Enright *et al.* [13].
6. Test problem F: Large linear problem.
   Problem 3 of Enright [14].

Test Problem A

TOL	Method	Time	FEV/JEV/MIV	NIS/NRS	AGE	MGE
	DIRKL	1.81	59/1/10	16/0	0.011	0.226
10^{-2}	DIRKS	1.83	55/1/11	18/0	0.009	0.265
	DIRKA	1.44	52/1/14	17/0	0.006	0.006
	DIRKL	6.01	180/1/14	55/0	0.026	0.540
10^{-4}	DIRKS	5.92	169/1/19	56/0	0.025	0.510
	DIRKA	5.28	151/1/17	50/0	0.020	0.415
	DIRKL	27.79	734/1/19	238/0	0.038	0.805
10^{-6}	DIRKS	27.85	729/1/19	242/0	0.036	0.835
	DIRKA	24.45	565/1/21	188/0	0.033	0.731

Test Problem B

TOL	Method	Time	FEV/JEV/MIV	NIS/NRS	AGE	MGE
	DIRKL	0.23	177/1/10	32/2	0.13	1.76
10^{-3}	DIRKS	0.22	170/1/8	36/0	0.12	1.66
	DIRKA	0.19	136/1/14	26/3	0.01	1.66
	DIRKL	0.67	652/1/14	127/3	0.87	11.71
10^{-5}	DIRKS	0.71	696/1/14	144/2	0.61	7.77
	DIRKA	0.51	490/1/19	97/6	0.71	9.26
	DIRKL	2.67	2778/1/20	548/7	4.29	55.88
10^{-7}	DIRKS	2.90	3093/1/17	621/3	3.48	39.40
	DIRKA	2.01	2137/1/21	429/6	3.63	45.68

Test Problem C

TOL	Method	Time	FEV/JEV/MIV	NIS/NRS	AGE	MGE
	DIRKL	1.45	79/1/10	18/0	0.03	0.41
10^{-2}	DIRKS	1.50	72/1/12	19/0	0.03	0.25
	DIRKA	1.37	66/1/16	17/0	0.01	0.17
	DIRKL	5.38	314/1/13	68/0	0.07	0.73
10^{-4}	DIRKS	5.38	290/1/16	67/0	0.07	0.71
	DIRKA	4.59	241/1/15	56/0	0.06	0.66
	DIRKL	24.76	1385/1/16	307/0	0.08	0.83
10^{-6}	DIRKS	24.05	1304/1/18	295/0	0.09	0.82
	DIRKA	20.85	1057/1/17	239/0	0.08	0.80

Test Problem D

TOL	Method	Time	FEV/JEV/MIV	NIS/NRS	AGE	MGE
	DIRKL	0.29	36/4/9	7/0	0.016	0.2
10^{-2}	DIRKS	0.31	32/3/9	8/0	0.011	0.211
	DIRKA	0.31	36/2/9	10/0	0.005	0.113
	DIRKL	0.80	104/2/9	18/0	0.052	0.471
10^{-4}	DIRKS	0.83	83/3/10	19/0	0.051	0.585
	DIRKA	0.93	99/3/14	22/0	0.036	0.364
	DIRKL	3.85	372/3/12	59/1	0.148	4.852
10^{-6}	DIRKS	4.18	404/3/14	68/1	0.086	1.149
	DIRKA	4.29	359/3/14	65/1	0.102	2.360

Test Problem E

TOL	Method	Time	FEV/JEV/MIV	NIS/NRS	AGE	MGE
	DIRKL	0.36	252/12/35	44/3	0.191	6.620
10^{-2}	DIRKS	0.36	271/9/34	58/4	0.119	2.629
	DIRKA	0.35	247/9/30	54/4	0.116	3.376
	DIRKL	0.98	935/13/40	160/7	0.895	27.480
10^{-4}	DIRKS	0.86	821/15/41	154/8	0.798	7.972
	DIRKA	0.77	696/13/38	130/6	0.693	19.398
	DIRKL	3.60	3817/13/42	686/8	4.15	96.339
10^{-6}	DIRKS	3.45	3708/16/40	692/8	3.654	53.256
	DIRKA	2.85	2928/17/48	553/11	3.291	76.325

Test Problem F

TOL	Method	Time	FEV/JEV/MIV	NIS/NRS	Global error at $x=5$
	DIRKL	5.31	34/1/3	6/0	0.025
10^{-2}	DIRKS	5.71	18/1/4	4/0	0.031
	DIRKA	6.36	35/1/4	7/1	0.023
	DIRKL	40.52	411/1/19	64/11	0.426
10^{-4}	DIRKS	34.96	303/1/19	50/11	0.629
	DIRKA	37.84	322/1/21	54/11	0.498
	DIRKL	184.75	2033/1/83	333/41	4.673
10^{-6}	DIRKS	164.96	1589/1/85	276/42	5.214
	DIRKA	176.23	1671/1/93	289/46	4.862