



---

Order Barriers and Characterizations for Continuous Mono-Implicit Runge- Kutta Schemes

Author(s): Paul Muir and Brynjulf Owren

Source: *Mathematics of Computation*, Vol. 61, No. 204 (Oct., 1993), pp. 675-699

Published by: [American Mathematical Society](#)

Stable URL: <http://www.jstor.org/stable/2153247>

Accessed: 10/07/2013 10:51

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Mathematical Society is collaborating with JSTOR to digitize, preserve and extend access to *Mathematics of Computation*.

<http://www.jstor.org>

## ORDER BARRIERS AND CHARACTERIZATIONS FOR CONTINUOUS MONO-IMPLICIT RUNGE-KUTTA SCHEMES

PAUL MUIR AND BRYNJULF OWREN

**ABSTRACT.** The mono-implicit Runge-Kutta (MIRK) schemes, a subset of the family of implicit Runge-Kutta (IRK) schemes, were originally proposed for the numerical solution of initial value ODEs more than fifteen years ago. During the last decade, a considerable amount of attention has been given to the use of these schemes in the numerical solution of boundary value ODE problems, where their efficient implementation suggests that they may provide a worthwhile alternative to the widely used collocation schemes. Recent work in this area has seen the development of some software packages for boundary value ODEs based on these schemes. Unfortunately, these schemes lead to algorithms which provide only a discrete solution approximation at a set of mesh points over the problem interval, while the collocation schemes provide a natural continuous solution approximation. The availability of a continuous solution is important not only to the user of the software but also within the code itself, for example, in estimation of errors, defect control, mesh selection, and the provision of initial solution estimates for new meshes. An approach for the construction of a continuous solution approximation based on the MIRK schemes is suggested by recent work in the area of continuous extensions for explicit Runge-Kutta schemes for initial value ODEs. In this paper, we describe our work in the investigation of continuous versions of the MIRK schemes: (i) we give some lower bounds relating the stage order to the minimal number of stages for general continuous IRK schemes, (ii) we establish lower bounds on the number of stages needed to derive continuous MIRK schemes of orders 1 through 6, and (iii) we provide characterizations of these schemes having a minimal number of stages for each of these orders.

### INTRODUCTION

An implicit Runge-Kutta (IRK) scheme (see, for example, Butcher [9]) can be used to compute an approximation to the solution of the initial value problem (IVP)

$$(1.1) \quad y'(x) = f(x, y(x)), \quad y(x_0) = y_0,$$

where  $y(x) \in R^m$ ,  $f: R \times R^m \rightarrow R^m$ , and  $y_0 \in R^m$  is a given initial vector. The IRK scheme is used in a stepwise fashion to proceed from the initial conditions through a sequence of points,  $x_i$ , and solution approximations,  $y_i$ , to the

---

Received by the editor November 26, 1991 and, in revised form, June 15, 1992.

1991 *Mathematics Subject Classification.* Primary 65L05, 65L10.

*Key words and phrases.* Runge-Kutta methods, boundary value ODEs, continuous extensions.

This work was supported by the Natural Sciences and Engineering Research Council of Canada and the Information Technology Research Center of Ontario.

©1993 American Mathematical Society  
0025-5718/93 \$1.00 + \$.25 per page

desired termination point. During the  $i$ th step, we obtain an approximation,  $y_{i+1}$ , to the true solution,  $y(x)$ , evaluated at the point  $x_{i+1} = x_i + h_i$ , of the form

$$(1.2) \quad y_{i+1} = y_i + h_i \sum_{r=1}^s b_r K_r,$$

where

$$(1.3) \quad K_r = f \left( x_i + c_r h_i, y_i + h_i \sum_{j=1}^s a_{rj} K_j \right), \quad r = 1, \dots, s,$$

and  $c_r = \sum_{j=1}^s a_{rj}$ . These schemes are sometimes given in the form of a tableau containing their coefficients, which, for the above scheme, would have the following form:

$c_1$	$a_{11}$	$\dots$	$a_{1s}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$c_s$	$a_{s1}$	$\dots$	$a_{ss}$
	$b_1$	$\dots$	$b_s$

The above tableau is sometimes written in the more compact form:

$c$	$A$
	$b^T$

where  $c = Ae$ ,  $c = (c_1, c_2, \dots, c_s)^T$ ,  $b = (b_1, b_2, \dots, b_s)^T$ ,  $A$  is the  $s$  by  $s$  matrix whose  $(i, j)$ th component is  $a_{ij}$ , and  $e$  is the vector of 1's of length  $s$ . Note that in (1.3) above, each stage is defined implicitly in terms of itself and the other stages. Hence, in order to obtain approximate values for the stages, it is necessary to solve a system of, in general, nonlinear equations. This must usually be done using some form of modified Newton iteration, which makes the calculation of the stages a somewhat computationally expensive process.

A number of interesting subclasses of the IRK schemes have been identified and investigated in the literature. These schemes attempt a trade-off between the higher accuracy of the IRK schemes and schemes which can be implemented more efficiently. Examples of such schemes are singly-implicit Runge-Kutta schemes (Burrage [5]), diagonally implicit Runge-Kutta schemes (Nørsett [24]), and the mono-implicit Runge-Kutta (MIRK) schemes of Cash and Singhal [13] (also known as implicit endpoint quadrature formulas (van Bokhoven [3])).

In this paper, we will primarily focus on the class of MIRK schemes. These schemes are defined by means of a slightly alternative representation to that of the IRK schemes, which involves the introduction of  $s$  new parameters  $v_r$ ,  $r = 1, \dots, s$ , that allow for explicit dependence on  $y_{i+1}$  in the definition

of the stages of the scheme. The standard form of this class is

$$(1.4) \quad y_{i+1} = y_i + h_i \sum_{r=1}^s b_r K_r,$$

where

$$(1.5) \quad K_r = f \left( x_i + c_r h_i, (1 - v_r) y_i + v_r y_{i+1} + h_i \sum_{j=1}^{r-1} x_{rj} K_j \right),$$

$$r = 1, \dots, s,$$

and  $c_r = v_r + \sum_{j=1}^{r-1} x_{rj}$ . They are usually represented by a modified tableau; the above method would have the tableau

$c_1$	$v_1$	0	0	...	0	0
$c_2$	$v_2$	$x_{21}$	0	...	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$c_{s-1}$	$v_{s-1}$	$x_{s-1,1}$	$x_{s-1,2}$	...	0	0
$c_s$	$v_s$	$x_{s,1}$	$x_{s,2}$	...	$x_{s,s-1}$	0
		$b_1$	$b_2$	...	$b_{s-1}$	$b_s$

which is sometimes written in the more compact form

$c$	$v$	$X$
		$b^T$

where  $v = (v_1, v_2, \dots, v_s)^T$ ,  $c = v + X e$ , and  $X$  is strictly lower triangular. The MIRK scheme (1.4), (1.5) is equivalent to the general IRK scheme (1.2), (1.3) with  $A = X + v b^T$  (cf. [18]). These schemes are interesting because they are implicit only in  $y_{i+1}$ ; this leads to a more efficient implementation than is possible for the fully implicit IRK schemes (van Bokhoven [3]).

The use of IRK schemes for the solution of boundary value ODE problems (BVP) has been well known for some time (Weiss [30]). In this paper, we assume that the boundary value ODEs are expressed in the general form

$$(1.6) \quad y'(x) = f(x, y(x)), \quad x \in [a, b],$$

where  $y \in R^m$  and  $f : R \times R^m \rightarrow R^m$ , with boundary conditions

$$(1.7) \quad g(y(a), y(b)) = \underline{0},$$

where  $g : R^m \times R^m \rightarrow R^m$ . The main step in a Runge-Kutta-based algorithm for the numerical solution of a boundary value ODE is the solution of a discrete

system,  $\Phi(Y) = 0$ , where  $Y$  is a vector whose  $i$ th component,  $y_i$ , is an approximation to the exact solution evaluated at the  $i$ th point of an  $(N + 1)$ -point mesh which subdivides the problem interval. The function  $\Phi(Y)$  can be expressed in a component-wise form with one component of size  $m$  for the boundary conditions  $g$ , and  $N$  components  $\phi_i$ ,  $i = 0, \dots, N - 1$ , each of size  $m$ , where  $\phi_i$  is defined by an IRK scheme. We have,

$$(1.8) \quad \phi_i = y_{i+1} - y_i - h_i \sum_{r=1}^s b_r K_r,$$

with the stages  $K_r$  defined as in (1.3). As in the IVP case, when all the  $a_{rj}$  coefficients are nonzero, the stages are defined implicitly, and in order to compute them, we will, in general, have to solve a system of nonlinear equations.

The implicit Runge-Kutta formulas include the collocation formulas which have been a popular approach for the numerical solution of boundary value ODEs. The collocation formulas based on Gauss-point rules have been implemented in the widely used code COLSYS (Ascher, Christiansen, and Russell [1]), and its successor COLNEW (Bader and Ascher [2]). However, compared to the situation for IVPs, there has been relatively little work done on efficient subclasses of the IRK schemes for boundary value ODEs. Most of the subclasses presented to date have been based on the class of MIRK schemes. A single formula from this subclass of schemes was first suggested for use in the numerical solution of boundary value ODEs by Cash and Moore [12]. This result was generalized by Cash and Singhal [14], where a symmetric subclass of the MIRK schemes was presented. The use of these schemes within an iterated deferred correction scheme for the numerical solution of boundary value problems has been described by Cash [10, 11], and Cash and Wright [15]. This work has led to the development of a code called HAGRON, which has been shown to compare favorably with COLSYS and D02GAF, a code from the NAG library based on PASVA3 (Lentini and Pereyra [23]). The entire class of MIRK schemes has been considered for use in the numerical solution of boundary value ODEs by Gupta [21] and Enright and Muir [18].

When a MIRK scheme is used within the algorithm outlined above for the solution of boundary value ODEs, it defines equations having the same form as in (1.8), but the corresponding stages are of the mono-implicit type as in (1.5). However, in the BVP context, since approximations to both  $y_i$  and  $y_{i+1}$  are available from the current Newton iterate, the stages of the MIRK scheme can be computed explicitly, and thus very efficiently. These improvements in efficiency are not gained, however, by giving up good stability as in the case of explicit Runge-Kutta schemes. There are many instances of MIRK schemes which are symmetric and  $A$ -stable, two important properties for discretization schemes used in the numerical solution of boundary value ODEs.

An important advantage of the collocation formulas is that a continuous approximation to the solution over the entire problem interval is naturally obtained. This can be very useful when the user requires solution information at off-mesh points, but the continuous solution approximation can also be useful to the code itself, for example, for error estimation, defect control, provision of initial estimates for Newton iterates, or mesh refinement and redistribution. Unlike the collocation formulas, many other schemes do not have an associated natural continuous approximation to the solution. Examples of codes based

on such schemes are D02GAF, mentioned earlier, some codes based on special one-sided Runge-Kutta schemes (Kreiss, Nichols, and Brown [22], Brown and Lorenz [4]), and the HAGRON code, mentioned above.

In the area of initial value ODE problems, the idea of extending the discrete solution approximation to get a continuous solution has received considerable attention over the last few years. A number of authors have demonstrated the possibility of generating inexpensive interpolants for explicit Runge-Kutta formulas which are not of the collocation type, in the context of the numerical solution of initial value problems (see, for example, Enright et al. [17], Gladwell et al. [20], and references therein for some of the earlier work). These interpolants are obtained for Runge-Kutta formula pairs, by constructing extra stages within the current step, thus preserving the one-step nature of the formula. The most recent work on these continuous explicit Runge-Kutta schemes is by Owren and Zennaro [26, 27] and Verner [29].

Nørsett and Wanner [25] considered perturbed collocation schemes and proved that any nonconfluent Runge-Kutta scheme is equivalent to a generalized type of collocation with a polynomial of degree  $s$ . The usefulness of the perturbed collocation polynomial as a continuous extension of the Runge-Kutta scheme has been limited by some undesirable properties of these approximations. For instance, as the authors point out, the higher derivatives of the approximation may become unbounded as the step size tends to zero. A slightly different approach is taken by Zennaro [31], where it is proved that any Runge-Kutta scheme of order  $p$  possesses *natural continuous extensions* (NCE) of some degree  $d \leq p$ . These continuous approximations have uniform order  $d$  and satisfy certain orthogonality conditions, which make them particularly useful for solving ODEs containing forcing terms. Collocation schemes and, in fact, also a subclass of the perturbed collocation schemes, are themselves NCEs. In [32] the related class of projection schemes is considered. Introducing a linear projector  $Q_d : C^0[x_i, x_i + h_i] \rightarrow \Pi_{d-1}$ , where  $\Pi_{d-1}$  is the set of polynomials of degree  $\leq d - 1$ , he considers the polynomial  $u(x)$  defined by  $u'(x) = Q_d f(x, u(x))$ . Again, it is interesting to observe that the collocation schemes turn out to be a subclass.

The basic form of a continuous Runge-Kutta (CRK) scheme on the  $i$ th step is a polynomial in  $\theta$  of the form

$$(1.9) \quad u(x_i + \theta h_i) = y_i + h_i \sum_{r=1}^s b_r(\theta) K_r,$$

where

$$(1.10) \quad K_r = f \left( x_i + c_r h_i, y_i + h_i \sum_{j=1}^s a_{rj} K_j \right), \quad r = 1, \dots, s.$$

Observe that the stage values  $K_r$ ,  $r = 1, \dots, s$ , are the same as for the discrete Runge-Kutta scheme, (1.3). We assume that  $b_r(\theta)$ ,  $r = 1, \dots, s$ , are weight polynomials of some maximum degree  $d$ . It is easy to see that in order to obtain continuity of the global piecewise continuous approximation furnished by a CRK scheme, one must require that  $b_r(0) = 0$  for all the weight polynomials (see, e.g., Verner [29]). We will tacitly assume this property for the schemes we consider. Notice also that any CRK scheme has an underlying

discrete scheme defined by the coefficient matrix  $A$  and with discrete weights  $b_r = b_r(1)$ ,  $r = 1, \dots, s$ .

For boundary value ODEs, the extension of discrete solution approximations to continuous ones has received substantially less attention. There is only the work of Pruess [28], who is concerned with the development of interpolants for collocation schemes for boundary value problems, where the information required to construct a continuous approximation is not available and one has only solution approximations at the mesh points.

In this paper, we investigate continuous versions of the MIRK schemes (which we will refer to as CMIRK schemes) with the idea of providing continuous solutions for boundary value ODEs. The CMIRK schemes have the form as above in (1.9) except that the corresponding stages will be of mono-implicit type, (1.5). The paper is organized as follows. In §2, we present some theoretical results concerning continuous extensions for general IRK schemes. In §3, these results are applied to derive order barriers for CMIRK schemes of orders 1 through 6. Characterizations of families of CMIRK schemes of orders 1 through 6 are given in §4. We conclude in §5 with a summary of our results and a discussion of future work.

## 2. SOME GENERAL THEORY FOR CONTINUOUS RUNGE-KUTTA SCHEMES

In this section we will develop some general tools for the derivation and order barrier analysis of CRK schemes, (1.9), (1.10), following the ideas of Owren and Zennaro [26, 27]. Without loss of generality we shall henceforth consider continuous approximations to  $y(x)$  on the first step interval (from  $x_0$  to  $x_1$ ), and we shall omit step indices where it does not cause ambiguity. We define the *uniform order* of a CRK scheme (1.9), (1.10) as the greatest integer  $p$  for which

$$(2.1) \quad \max_{0 \leq \theta \leq 1} |y(x_0 + \theta h) - u(x_0 + \theta h)| = O(h^{p+1}),$$

where  $u(x_0) = y(x_0)$  and  $|\cdot|$  is any norm on  $R^m$ . One of the goals of this paper is to establish, for the CMIRK schemes, the minimum number of stages  $s$  required for uniform order  $p$ . It is well known that the lower bound of  $s = p$  can be achieved for the class of fully implicit CRK schemes.

In the following we will use the theory of trees and order conditions introduced by Butcher [7, 8] without further citations. We recommend the excellent text [9] for an account of this material. The adaptations of this theory to continuous schemes involve generalizations of the ideas of Owren and Zennaro [26]. In order to clarify our notation, we will briefly review some of the main results from that paper. It was remarked by Zennaro [31] that the maximum degree  $d$  of the weight polynomials must be at least  $p$  for (2.1) to be satisfied, and that  $d > p$  can lead to unbounded derivatives of  $u(x)$  as  $h$  tends to zero; hence, we will always assume that  $d = p$ . This requirement and (2.1) cause the CRK scheme (1.9), (1.10) to be a natural continuous extension of the underlying discrete scheme, as defined in [31]. Let  $N_p$  be the total number of rooted trees of order at most  $p$ , and let  $t_1, \dots, t_{N_p}$  be any ordering of these trees such that  $\rho(t_i) > \rho(t_j)$  only if  $i > j$ , where  $\rho(t_k)$  is the order of  $t_k$ . Following Butcher [9, p.163], we introduce the elementary weight  $\phi_{ij} := \phi_j(t_i)$  for each tree  $t_i$  and stage  $j$ , and we let  $\gamma(t_i)$  be the *density* of  $t_i$ . Then (2.1) is satisfied if and

only if

$$\sum_{j=1}^s \phi_{ij} b_j(\theta) = \frac{\theta^{\rho(t_i)}}{\gamma(t_i)}, \quad i = 1, \dots, N_p.$$

Since  $\rho(t_i) \geq 1$  for all  $t_i$ , and since  $b_i(0) = 0$  for all the weight polynomials, we may write

$$(2.2) \quad b_i(\theta) = \sum_{j=1}^p b_{ij} \theta^j \quad \text{and} \quad \frac{\theta^{\rho(t_i)}}{\gamma(t_i)} = \sum_{l=1}^p q_{il} \theta^l.$$

Thus, by defining the  $N_p \times s$  matrix  $\Phi := ((\phi_{ij}))$ , the  $s \times p$  matrix  $B := ((b_{ij}))$ , and the  $N_p \times p$  matrix  $Q := ((q_{il}))$ , we arrive at the matrix equation

$$(2.3) \quad \Phi B = Q.$$

From classical order theory for Runge-Kutta schemes we know that  $\Phi$  only depends on the  $s \times s$  matrix  $A$ , and from (2.2) we see that  $Q$  is a constant matrix, in fact it has the simple form  $q_{il} = \gamma(t_i)^{-1} \delta_{\rho(t_i), l}$ . (Recall that  $\delta_{i,j} = 1$  if  $i = j$  and  $\delta_{i,j} = 0$  if  $i \neq j$ .) Clearly, the columns of  $Q$  must be linearly independent, so  $\text{rank}(Q) = p$ . It will be convenient to view the matrix  $\Phi$  as the image of a nonlinear operator whose domain is the union over all positive  $s$  of the sets of all  $s \times s$  matrices  $A$ , and we denote this operator by  $F_p(A)$ . We will also need another operator with the same domain, namely  $G_p(A) := (F_p(A)|Q)$ , i.e., the matrix  $\Phi$  augmented by the columns of  $Q$ . We summarize some properties of these matrices.

**Lemma 2.1.** *For all  $s \times s$  matrices  $A$ ,*

$$(2.4) \quad \text{rank}(G_p(A)) \geq p,$$

$$(2.5) \quad \text{rank}(G_{p+1}(A)) \geq \text{rank}(G_p(A)) + 1,$$

$$(2.6) \quad \text{rank}(F_p(A)) \leq s,$$

$$(2.7) \quad \text{rank}(F_p(A)) \geq p \quad \text{if at least } p \text{ of the } c_i \text{'s are distinct.}$$

*Proof.* (2.6) is obvious. (2.4) and (2.5) follow easily by considering the last  $p$  columns of  $G_p(A)$ . We see that (2.7) holds when at least  $p$  of the  $c_i$ 's are distinct because the rows of  $F_p(A)$  corresponding to the trees  $\tau$  and  $[\tau^k]$ ,  $1 \leq k \leq p - 1$ , define a Vandermonde system (see, e.g., Verner [29]).  $\square$

It is well known that the rows of  $F_p(A)$  can be generated recursively by forming componentwise products between previously generated rows, or by multiplying a previously generated row by the matrix  $A$ . We will sometimes find it convenient to write  $zw = (z_1 w_1, \dots, z_s w_s)^T$ , where  $z = (z_1, \dots, z_s)^T$  and  $w = (w_1, \dots, w_s)^T$ . Similarly, we write  $z^k = (z_1^k, \dots, z_s^k)^T$ . Finally, we define  $C$  to be the diagonal matrix with  $c_1, \dots, c_s$  on the diagonal. The following result is an obvious generalization of a result by Owren and Zennaro [26].

**Proposition 2.2.** *An  $s \times s$  matrix  $A$  defines the coupling coefficients of an  $s$ -stage CRK scheme of order  $p$  if and only if  $\text{rank}(F_p(A)) = \text{rank}(G_p(A))$ .*

Another important result regarding explicit Runge-Kutta schemes is Theorem 2.4 of [26]. If  $A$  is a strictly lower triangular  $s \times s$  matrix such that

$\text{rank}(F_p(A)) = \text{rank}(G_p(A)) = \rho < s$ , then there exists a strictly lower triangular  $\rho \times \rho$  matrix  $A^*$  such that  $\text{rank}(F_p(A^*)) = \text{rank}(G_p(A^*)) = \rho$ . This means that when one is seeking continuous explicit Runge-Kutta schemes with the minimal number of stages, it is sufficient to consider  $s \times s$  matrices  $A$  which satisfy

$$(2.8) \quad \text{rank}(F_p(A)) = \text{rank}(G_p(A)) = s.$$

By carefully modifying the proof of this theorem, it can be verified that a similar result holds for many other classes of CRK schemes. Let  $M$  be such a class, and let  $\hat{M}$  be the corresponding set of matrices  $A$  associated with schemes in  $M$ . Then, when seeking schemes in  $M$  with the minimal number of stages, it is sufficient to consider matrices  $A \in \hat{M}$  which satisfy (2.8). This holds when  $M$  is the class of general CRK schemes, as well as continuous (singly-) diagonally implicit Runge-Kutta schemes (C(S)DIRK schemes) and for the class of CMIRK schemes, to which we will pay particular attention in §§3 and 4.

One may observe that the requirement  $\text{rank}(F_p(A)) = \text{rank}(G_p(A)) = s$  is equivalent to the associated set of weight polynomials  $b_1(\theta), \dots, b_s(\theta)$  being unique. A trivial case for which  $\text{rank}(F_p(A)) < s$  is when  $A$  has two identical rows. The associated scheme is then reducible. We state without proof the following proposition.

**Proposition 2.3.** *Assume that the  $s \times s$  matrix  $A$  defines a CRK scheme with the minimal number of stages. Then it cannot have two identical rows.*

For our further discussion, we need a definition introduced by Butcher [8].

**Definition 2.4.** Consider the conditions

$$(2.9) \quad k \sum_{j=1}^s a_{ij} c_j^{k-1} = c_i^k, \quad 1 \leq k \leq \xi.$$

If an  $s \times s$  matrix  $A$  is such that (2.9) holds for  $1 \leq i \leq s$ , we say that  $A$  satisfies  $C(\xi)$ . (If the matrix  $A$  is the matrix of coupling coefficients of an IRK scheme, it is common to also say that the IRK scheme itself satisfies  $C(\xi)$ .) If (2.9) holds for  $i = l$ , we say that stage  $l$  of the corresponding scheme satisfies  $C(\xi)$ .

It is well known from the classical theory for Runge-Kutta schemes that when  $C(\xi)$ ,  $\xi > 1$ , is imposed on  $A$ , the complexity of the system of order conditions is significantly reduced. This is indeed the case also for CRK schemes. To quantify this, we will first present an enumeration result for a subset of the rooted trees. We will say that a rooted tree contains a  $k$ -leaf if it has a nonterminal node, different from the root, which has at most  $k - 1$  children, all of which must be leaves.

**Theorem 2.5.** *Let  $\alpha_j$ ,  $j \geq 1$ , be the number of rooted trees of order  $j$  which do not contain any  $k$ -leaves. Then  $\alpha_1, \alpha_2, \dots$  satisfy the relation*

$$\alpha_1 + \alpha_2 x + \alpha_3 x^2 + \dots = (1 - x)^{-1} \prod_{n=k+1}^{\infty} (1 - x^n)^{-\alpha_n}.$$

*Proof.* Similar to that of Butcher [9, Theorem 145A, p. 89].  $\square$

**Corollary 2.6.** *For  $1 \leq j \leq 2k + 2$ ,  $\alpha_j = 2^q$ , where  $q = \max\{0, j - k - 1\}$ .*

These enumeration results may be used to provide upper bounds for the number of stages needed to construct a CRK scheme of a given order. We have

**Theorem 2.7.** Assume that an  $s \times s$  matrix  $A$  satisfies  $C(\xi)$  for some  $\xi \geq 1$ . Let  $\alpha_j$  be the number of rooted trees of order  $j$  without  $\xi$ -leaves, and for each  $p$  define  $r_p = \sum_{j=1}^p \alpha_j$ . Then

$$\text{rank}(G_p(A)) \leq r_p.$$

*Proof.* It is clear (see, e.g., [9, pp. 194, 214]), that when  $A$  satisfies  $C(\xi)$ , any row of  $G_p(A)$  that originates from a tree  $t_i$  containing a  $\xi$ -leaf will be proportional to another row of  $G_p(A)$  which originates from a tree  $t_j$  that does not contain a  $\xi$ -leaf, provided  $\rho(t_i) = \rho(t_j)$ .  $\square$

**Theorem 2.8.** The  $s \times s$  matrix  $A$  defines the stages of an  $s$ -stage CRK scheme of order  $s$  if and only if it satisfies  $C(s-1)$  and the numbers  $c_i = \sum_{j=1}^s a_{ij}$  are all distinct.

*Proof.* Assume  $C(s-1)$  holds for distinct  $c_i$ 's. Then according to Corollary 2.6, Theorem 2.7, and (2.4) we get  $\text{rank}(G_s(A)) = s$ . Since the  $c_i$ 's are distinct, we have  $\text{rank}(F_s(A)) = s$  by (2.6) and (2.7); thus, by Proposition 2.2,  $A$  defines the desired scheme. Conversely, assume that  $A$  represents a scheme of order  $s$ . Then, we must similarly have  $\text{rank}(G_s(A)) = s$ . Thus, (2.4) and (2.5) imply that rows  $i$  and  $j$  must be proportional if  $\rho(t_i) = \rho(t_j)$ . In particular, this means that the pairs of rows originating from  $[\tau^l]$  and  $[[\tau^{l-1}]]$  must be proportional for  $2 \leq l \leq s-1$ . But this means that  $c^l = lAc^{l-1}$  for  $1 \leq l \leq s-1$ , which is exactly the condition  $C(s-1)$ . Furthermore, it follows that the row space of  $F_s(A)$  is spanned by the vectors  $c^k$ ,  $0 \leq k \leq s-1$ , and hence  $\text{rank}(F_s(A)) = s$  if and only if the  $c_i$ 's are distinct.  $\square$

The theorem above shows that the CRK schemes of order  $p$  that have the minimal number of stages are the ones which satisfy  $C(p-1)$ . In fact, these schemes can be recognized as perturbed collocation (PECO) schemes where the collocation polynomial is identical to the continuous approximation  $u(x)$ . From Nørsett and Wanner [25] we have the perturbation operator  $P : \Pi_s \rightarrow \Pi_s$ , defined by

$$Pu(x) := u(x) + \sum_{j=1}^s N_j(\theta) u^{(j)}(x_0) h^j, \quad 0 \leq \theta = \frac{x - x_0}{h} \leq 1,$$

where the  $N_j(\theta)$ 's are polynomials of degree  $\leq s$ . Then the PECO scheme is

$$\begin{aligned} u(x_0) &= y_0, \\ u'(x_0 + c_r h) &= f(x_0 + c_r h, Pu(x_0 + c_r h)), \quad r = 1, \dots, s, \\ y_1 &= u(x_0 + h), \end{aligned}$$

where the  $c_r$ 's are distinct. We have the following relationship between the schemes of Theorem 2.8 and PECO schemes.

**Theorem 2.9.** A CRK scheme of order  $p$  with  $p$  stages is equivalent to a PECO scheme with  $N_1(\theta) = \dots = N_{p-1}(\theta) \equiv 0$  and  $N_p(\theta) \in \Pi_{p-1}$ , a polynomial of degree  $p-1$ , such that

$$N_p(c_r) = -\frac{1}{p!} \left( c_r^p - p \sum_{j=1}^p a_{rj} c_j^{p-1} \right), \quad r = 1, \dots, p.$$

Prior to proving this theorem, we consider the following lemma.

**Lemma 2.10.** *Let  $A$  define a CRK scheme of order  $p$  with weight polynomials  $b_j(\theta)$ ,  $j = 1, \dots, s$ . Assume that exactly  $p$  of these polynomials,  $b_{i_1}(\theta), \dots, b_{i_p}(\theta)$ ,  $i_r \in \{1, \dots, s\}$ ,  $r = 1, \dots, p$ , are nonzero. Then*

$$K_{i_r} = u'(x_0 + c_{i_r}h), \quad r = 1, \dots, p,$$

$$b'_{i_r}(\theta) = L_{i_r}(\theta), \quad r = 1, \dots, p,$$

where  $L_{i_r}(\theta)$  is the  $i_r$ th elementary Lagrange interpolation polynomial with respect to the points  $\{c_{i_1}, \dots, c_{i_p}\}$ .

*Proof.* Such CRK schemes must in particular satisfy the quadrature conditions

$$\sum_{r=1}^p b_{i_r}(\theta)c_{i_r}^{k-1} = \frac{\theta^k}{k}, \quad k = 1, \dots, p.$$

Differentiating, we get

$$\sum_{r=1}^p b'_{i_r}(\theta)c_{i_r}^{k-1} = \theta^{k-1}, \quad k = 1, \dots, p.$$

Thus the polynomials  $b'_{i_r}(\theta)$ ,  $r = 1, \dots, p$ , form a basis for  $\Pi_{p-1}$ . In fact, for any  $p(\theta) \in \Pi_{p-1}$  we have

$$p(\theta) = \sum_{r=1}^p p(c_{i_r})b'_{i_r}(\theta),$$

from which it follows that  $b'_{i_r}(\theta) = L_{i_r}(\theta)$ ,  $r = 1, \dots, p$ . Since  $u'(x_0 + \theta h) = \sum b'_{i_r}(\theta)K_{i_r}$ , we have  $u'(x_0 + c_{i_r}h) = K_{i_r}$ ,  $r = 1, \dots, p$ .  $\square$

We now employ this lemma in the proof of Theorem 2.9.

*Proof of Theorem 2.9.* Let  $w := (N_p(c_1), \dots, N_p(c_p))^T$ , and let  $b_p$  be the  $p$ th column of the weight matrix  $B$ . Let  $M$  be the  $p \times p$  matrix whose  $(i, j)$ th element is  $b_j(c_i)$ . We argue that for the CRK schemes of Theorem 2.8 we have  $A = M + (p!)wb_p^T$ . To show this, we multiply  $M + (p!)wb_p^T$  with the  $p \times p$  matrix  $V$  whose columns are  $c^0, c^1, \dots, c^{p-1}$ . The quadrature conditions imply  $Mc^{k-1} = c^k/k$ ,  $k = 1, \dots, p$ . By (2.3) we get  $b_p^T c^{k-1} = \delta_{pk}/k$ . Hence, for  $k \leq p - 1$  we get  $(M + (p!)wb_p^T)c^{k-1} = c^k/k = Ac^{k-1}$ , since  $C(p - 1)$  holds for  $A$  and  $(M + (p!)wb_p^T)c^{p-1} = c^p/p + (p!)w/p = Ac^{p-1}$ . Thus,  $AV = (M + (p!)wb_p^T)V$ , which implies  $A = M + (p!)wb_p^T$ , since  $V$  is nonsingular (because the  $c_i$ 's are distinct). Now we apply Lemma 2.10 with  $s = p$  to obtain

$$u'(x_0 + c_r h) = K_r = f \left( x_0 + c_r h, y_0 + h \sum_{j=1}^p a_{rj} K_j \right)$$

$$= f \left( x_0 + c_r h, y_0 + h \sum_{j=1}^p b_j(c_r) K_j + (p!)h \sum_{j=1}^p N_p(c_r) b_{jp} K_j \right),$$

$r = 1, \dots, p.$

Since  $u^{(p)}(x_0)h^p = (p!)h \sum_{j=1}^p b_{jp}K_j$ , we get

$$u'(x_0 + c_r h) = f(x_0 + c_r h, u(x_0 + c_r h) + N_p(c_r)h^p u^{(p)}(x_0)),$$

and we have proved that our CRK scheme is indeed the PECO scheme mentioned above.  $\square$

For an  $s \times s$  matrix  $A$  we introduce the following equivalence relation on the set of indices  $\{1, \dots, s\}$ :

$$i \equiv j \text{ if and only if } c_i = c_j.$$

Thus, there will be exactly one equivalence class for each distinct abscissa value. Let  $s^*$  be the number of equivalence classes, which we label  $S_1, \dots, S_{s^*}$ .

In the following, we shall be concerned with schemes of order, say,  $q$ , which fail to satisfy  $C(q - 1)$ . Under such circumstances the following theorem provides a lower bound for the minimum number of stages needed to obtain a given order.

**Theorem 2.11.** *For an  $s \times s$  matrix  $A$ , let  $k \geq 2$  be the largest integer such that  $C(k - 1)$  holds for  $A$ , and let  $u = (u_1, \dots, u_s)^T$  be the residual of the  $C(k)$  condition. That is,  $u_i = c_i^k - k \sum_{j=1}^s a_{ij}c_j^{k-1}$  for  $1 \leq i \leq s$ , and  $u_i$  gives the amount by which the  $i$ th stage of the method fails to satisfy the  $C(k)$  condition.*

(a) *Then for any  $q \geq k + 1$ ,*

$$(2.10) \quad \text{rank}(G_q(A)) \geq q + 1.$$

(b) *In particular, if  $q > k + 1$  and  $\text{rank}(G_q(A)) = \text{rank}(F_q(A)) = q + 1$ , then there exists  $L \in \{1, \dots, s^*\}$  such that*

- (i)  $|S_L| > 1$ ;
- (ii) *for all stages  $i \notin S_L$ ,  $C(q - 1)$  holds, and for all stages  $i \in S_L$ , there exist real numbers  $\beta_l$  such that*

$$(2.11) \quad c_i^l - l \sum_{j=1}^s a_{ij}c_j^{l-1} = \beta_l u_i, \quad k + 1 \leq l \leq q - 1;$$

(iii) *there exists  $\lambda \in \mathbf{R}$  such that*

$$(2.12) \quad \sum_{i \in S_L} u_i (A - \lambda I) e_i = 0.$$

*Remark 2.12.* In part (b) of this theorem, we are considering the case where  $A$  defines a method of order  $q$  satisfying condition  $C(k - 1)$  for  $k < q - 1$ . The point of this part of the theorem is to indicate that if the lower bound,  $\text{rank}(G_q(A)) = q + 1$ , is to be achieved when the order  $q$  of the method is at least two orders higher than the stage order  $k - 1$ , then several fairly restrictive conditions, (i), (ii), and (iii), must hold.

Condition (i) simply says that at least two of the abscissa must be the same. If, in addition, we have  $s = q + 1$ , then a slightly stronger result is possible: since  $s^* \geq q$  (see, e.g., Zennaro [31]), if at least two of the abscissa are in  $S_L$ , then all the others must be distinct and thus define their own equivalence classes in order that  $s^* = q$ . However, since there are exactly  $q + 1$  abscissa, we must have  $|S_L| = 2$  and  $|S_i| = 1, i = 1, \dots, q + 1, i \neq L$ .

Condition (ii) says that all the stages corresponding to distinct abscissa must satisfy the  $C(q - 1)$  condition, while for the stages corresponding to the equal abscissa, the residuals of the  $C(l)$  conditions,  $c_i^l - l \sum_{j=1}^s a_{ij} c_j^{l-1}$ , for  $l = k + 1, \dots, q - 1, i \in S_L$ , must be proportional to the corresponding residuals from the  $C(k)$  condition,  $u_i$ . The only components of  $u$  that can be nonzero are those corresponding to these equal abscissa.

*Proof of Theorem 2.11.* For the proof of part (a), suppose that  $\text{rank}(G_q(A)) \leq q$ . Since, by Lemma 2.1,  $\text{rank}(G_q(A)) \geq q$ , we need only consider the case where  $\text{rank}(G_q(A)) = q$ . Using a proof similar to that of Theorem 2.8, it is easily shown that  $\text{rank}(G_q(A)) = q$  implies that  $C(q - 1)$  must hold. This gives a contradiction since we have only  $C(k - 1)$ , where  $k \leq q - 1$ , and part (a) is proved.

For the proof of part (b)(ii), we begin by noting that we have assumed that  $\text{rank}(F_q(A)) = \text{rank}(G_q(A)) = q + 1$  for  $q > k + 1$ , and then observe, in view of (2.5), that  $\text{rank}(G_{k+2}(A)) = k + 3$ . Then, in view of Theorem 2.7, we construct the matrix  $G'_{k+2}(A)$  by deleting all rows of  $G_{k+2}(A)$  that correspond to trees containing  $(k - 1)$ -leaves. Clearly,  $\text{rank}(G'_{k+2}(A)) = \text{rank}(G_{k+2}(A))$ . The matrix  $G'_{k+2}(A)$  has the form:

$$\left( \begin{array}{ccccccccc} (c^0)^T & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ (c^1)^T & 0 & \frac{1}{2} & 0 & \dots & 0 & 0 & 0 & 0 \\ (c^2)^T & 0 & 0 & \frac{1}{3} & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ (c^{k-1})^T & 0 & 0 & 0 & \dots & \frac{1}{k} & 0 & 0 & 0 \\ (c^k)^T & 0 & 0 & 0 & \dots & 0 & \frac{1}{k+1} & 0 & 0 \\ (Ac^{k-1})^T & 0 & 0 & 0 & \dots & 0 & \frac{1}{k(k+1)} & 0 & 0 \\ (c^{k+1})^T & 0 & 0 & 0 & \dots & 0 & 0 & \frac{1}{k+2} & 0 \\ (Ac^k)^T & 0 & 0 & 0 & \dots & 0 & 0 & \frac{1}{(k+1)(k+2)} & 0 \\ (cAc^{k-1})^T & 0 & 0 & 0 & \dots & 0 & 0 & \frac{1}{k(k+2)} & 0 \\ (A^2c^{k-1})^T & 0 & 0 & 0 & \dots & 0 & 0 & \frac{1}{k(k+1)(k+2)} & 0 \end{array} \right).$$

⏟  
s
⏟  
k

Simply by examining the last  $k + 2$  columns of  $G'_{k+2}(A)$  we observe that the first  $k$  rows are linearly independent of the last six rows. Furthermore, since  $\text{rank}(G'_{k+2}(A)) = \text{rank}(G_{k+2}(A)) = k + 3$ , at most three of the last six rows can be linearly independent. Let us next use the independent rows whose first  $s$  components are  $c^k$  and  $c^{k+1}$  to eliminate the nonzero entries in the last two columns of the other four rows (whose first  $s$  components are  $Ac^{k-1}$ ,  $Ac^k$ ,  $cAc^{k-1}$ , and  $A^2c^{k-1}$ ). The resulting four reduced rows, whose first  $s$  components are now  $c^k - kAc^{k-1}$ ,  $c^{k+1} - (k + 1)Ac^k$ ,  $c^{k+1} - kcAc^{k-1}$ , and  $c^{k+1} - k(k + 1)A^2c^{k-1}$ , are clearly linearly independent of the two we used to perform the elimination. Therefore, if we are to have  $\text{rank}(G_{k+2}(A)) = k + 3$ , these four reduced rows must all be proportional. Hence, considering the 1st, 3rd, and 4th of these rows, there must exist constants  $\mu_1, \mu_2 \in \mathbf{R}$  such that

$$(2.13) \quad \mu_1(c^k - kAc^{k-1}) = c^{k+1} - kcAc^{k-1},$$

$$(2.14) \quad \mu_2(c^k - kAc^{k-1}) = c^{k+1} - k(k+1)A^2c^{k-1},$$

and, considering the 1st and 2nd of these reduced rows, there must exist  $\beta_{k+1} \in \mathbf{R}$  such that

$$(2.15) \quad \beta_{k+1}(c^k - kAc^{k-1}) = c^{k+1} - (k+1)Ac^k.$$

When  $q > k + 2$ , we must consider  $G_q(A)$ , which is obtained by extending  $G'_{k+2}(A)$  by adding on appropriate rows and columns. For each  $l = k + 3, \dots, q$ , we now consider the pairs of rows of  $G_q(A)$  which have  $c^{l-1}$  and  $Ac^{l-2}$  as their first  $s$  components. As before, we use the first row of each pair to eliminate the only nonzero element among the last  $q$  components of the second member of the pair. In each case the second row is reduced to the form  $c^{l-1} - lAc^{l-2}$ ,  $l = k + 3, \dots, q$ . Using the same reasoning as above, we then observe that each row whose first  $s$  components have the form  $c^{l-1}$ ,  $l = k + 3, \dots, q$ , will contribute 1 to the rank, and thus, if  $\text{rank}(G_q(A))$  is to be  $q + 1$ , all the reduced rows must be proportional to each other. With (2.15) included, this means that there must exist constants  $\beta_{k+1}, \dots, \beta_{q-1} \in \mathbf{R}$  such that

$$(2.16) \quad \beta_l(c^k - kAc^{k-1}) = c^l - lAc^{l-1}, \quad k + 1 \leq l \leq q - 1.$$

It then follows from (2.13) that  $c_i u_i = \mu_1 u_i$ , or  $(c_i - \mu_1)u_i = 0$ ,  $i = 1, \dots, s$ . Thus, for all stages  $i$  with  $c_i \neq \mu_1$ , we must have  $u_i = 0$ , and then from (2.16) it follows that these stages must satisfy the  $C(q - 1)$  condition. The remaining stages  $i$  must all have  $c_i = \mu_1$ , and these are all members of  $S_L$ . Thus,  $c^k - kAc^{k-1}$  can be expressed solely in terms of its nonzero components, i.e.,  $c^k - kAc^{k-1} = \sum_{i \in S_L} u_i e_i$ . From (2.16) we get

$$\beta_l \left( \sum_{i \in S_L} u_i e_i \right) = c^l - lAc^{l-1}, \quad k + 1 \leq l \leq q - 1,$$

and for the  $i$ th component

$$\beta_l u_i = c_i^l - l \sum_{j=1}^s a_{ij} c_j^{l-1}, \quad k + 1 \leq l \leq q - 1, i \in S_L,$$

which proves part (b)(ii).

To prove part (b)(iii), we use  $c^k - kAc^{k-1} = \sum_{i \in S_L} u_i e_i$  in (2.14); we get

$$\begin{aligned} \mu_2 \sum_{i \in S_L} u_i e_i &= c^{k+1} - (k+1)A \left( c^k - \sum_{i \in S_L} u_i e_i \right) \\ &= c^{k+1} - (k+1)Ac^k + (k+1)A \sum_{i \in S_L} u_i e_i. \end{aligned}$$

Then using (2.15), we get

$$\mu_2 \sum_{i \in S_L} u_i e_i = \sum_{i \in S_L} \beta_{k+1} u_i e_i + (k+1) \sum_{i \in S_L} u_i A e_i,$$

and hence,

$$(2.17) \quad \sum_{i \in S_L} u_i ((k+1)A - (\mu_2 - \beta_{k+1})I) e_i = 0 = \sum_{i \in S_L} u_i (A - \lambda I) e_i,$$

where  $\lambda = (\mu_2 - \beta_{k+1})/(k + 1)$ , which proves part (b)(iii).

To prove part (b)(i), we argue that  $S_L$  must contain more than one element, since otherwise, for the single stage  $i \in S_L$ , an examination of the order condition  $\sum_{j=1}^s b_j(\theta)u_j = 0$  would give  $b_i(\theta)u_i = 0$ , since we have  $u_j = 0, j \neq i$ , and thus  $b_i(\theta) \equiv 0$ . Furthermore, (2.17) and  $u_j = 0, j \neq i$ , imply that  $a_{ij} = 0, i \neq j$ . These two observations imply that the scheme would be reducible in the sense of Dahlquist and Jeltsch [16]. Hence,  $S_L$  must have more than one element, and (b)(i) is proved.  $\square$

An interesting question is whether there exist CRK schemes with coefficient matrix  $A$  satisfying conditions (i)–(iii) in Theorem 2.11. Indeed, it can be shown that, for instance, with  $k = 2$  (only  $C(1)$  satisfied) one can construct CRK schemes of order 4 having five stages. This can be done even with the CMIRK schemes, where the form of the  $A$ -matrix is somewhat restricted. The above theorem will be used in the next section to show that certain methods cannot satisfy  $C(k - 1), k \leq q - 1$ , and also achieve the lower bound,  $\text{rank}(G_q(A)) = q + 1$ , with  $s = q + 1$ , because they violate one or more of the conditions (i), (ii), or (iii).

### 3. ORDER BARRIERS FOR CMIRK SCHEMES

In this section we will derive the minimum number of stages  $\text{CMN}(p)$  needed to construct CMIRK schemes of orders  $p = 1, \dots, 6$ . It is known from Burrage, Chipman, and Muir [6] that the largest integer  $k$  for which  $C(k)$  can be satisfied for a CMIRK scheme is 3. We may therefore immediately conclude from Theorem 2.8 and (2.4) that  $\text{CMN}(p) = p$  for  $1 \leq p \leq 4$ . The same theorem in conjunction with Theorem 2.7 proves that  $\text{CMN}(5) = 6$ . From these theorems we see that a 5th-order CMIRK scheme with six stages can be constructed by imposing  $C(3)$  on  $A$ . We shall see that this condition is also necessary. We have

**Theorem 3.1.** *Let the  $s \times s$ -matrix  $A$  define a CMIRK scheme which does not satisfy  $C(3)$ . If  $s = \text{rank}(G_q(A)) = \text{rank}(F_q(A))$  with  $q \geq 5$ , then  $s \geq q + 2$ .*

*Proof.* We prove that if  $C(3)$  does not hold then the requirement  $\text{rank}(G_q(A)) = q + 1$  implies that either two rows of  $A$  are identical, which is impossible in view of Proposition 2.3 and (2.10), or that  $s^* < q$ , which is impossible (see, e.g., Zennaro [31]).

With  $C(1)$  or  $C(2)$  we can use the second part of Theorem 2.11, with  $k = 2$  or  $k = 3$ , and Remark 2.12. Thus, there is a set  $S_L$  which contains exactly two indices. Furthermore, stages  $i \notin S_L$  satisfy  $C(q - 1)$ . We assume without loss of generality that  $u_i \neq 0$  for  $i \in S_L$ . Notice that by combining the two order conditions of order  $k + 1$  in Theorem 2.11 one gets  $\sum_{i \in S_L} b_i(\theta)u_i \equiv 0$ . Using this for CMIRK schemes at  $\theta = 1$ , and recalling that  $A = X + vb^T$ , we can write (2.12) as

$$\sum_{i \in S_L} u_i(A - \lambda I)e_i = \sum_{i \in S_L} u_i(X - \lambda I)e_i = 0.$$

Let  $j = \min\{i : i \in S_L\}$ . Since  $X$  is strictly lower triangular, and since  $u_j \neq 0$ , it follows that  $\lambda = 0$  and  $x_{j+1,j} = 0$ . Moreover, for CMIRK schemes, the first stage satisfies  $C(k)$  for arbitrary  $k$ , so  $1 \notin S_L$ . Also, since the first three

stages of a CMIRK scheme cannot all satisfy  $C(4)$  (see Burrage, Chipman, and Muir [6]), we only have to investigate the cases where stage 2, stage 3, or both are contained in  $S_L$ . Before we proceed with this, consider the  $C(1)$  case. Since it is necessary that  $\text{rank}(G_5(A)) = 6$ , we get, by using the condition corresponding to the tree  $[[\tau], [\tau]]$ , that

$$c^4 - 4(Ac)^2 = \xi(c^2 - 2Ac) = \xi \sum_{i \in S_L} u_i e_i$$

for some  $\xi \in \mathbf{R}$ . Hence,

$$\xi \sum_{i \in S_L} u_i e_i = 2c^2(c^2 - 2Ac) - (c^2 - 2Ac)^2 = 2c_L^2 \sum_{i \in S_L} u_i e_i - \sum_{i \in S_L} u_i^2 e_i.$$

Since  $u_i \neq 0$  for  $i \in S_L$ , we get  $u_i = u_j$  for  $i, j \in S_L$  when only  $C(1)$  is satisfied.

(i)  $\{2, 3\} \subset S_L$ . Here  $x_{32} = 0$  with  $c_2 = c_3$ , and either of the conditions  $C(2)$  or  $u_2 = u_3$  implies that rows 2 and 3 in the matrix  $A$  are identical.

(ii)  $2 \in S_L, 3 \notin S_L$ . Since  $x_{32} = 0$ , we can permute the second and third stages to obtain one of the cases below.

(iii)  $\{3, 4\} \subset S_L$ . Here  $x_{43} = 0$ . In the  $C(1)$  case the conditions  $c_3 = c_4, u_3 = u_4$  immediately imply that stages 3 and 4 are identical. In the  $C(2)$  case we compute  $u_j$  and  $w_j = c_j^4 - \sum_{l=1}^s a_{jl} c_l^3$  to obtain

$$u_j = c_3^2(c_3 - 1) - x_{j2}, \quad w_j = c_3^2(c_3^2 - 1) - 2x_{j2}, \quad j = 3, 4.$$

By Theorem 2.11, (2.11), it is necessary that  $w_j = \beta_4 u_j$  for some  $\beta_4 \in \mathbf{R}$ . Imposing this with the above formulas, we get

$$c_3^2(c_3 - 1)^2(x_{42} - x_{32}) = 0.$$

If  $x_{42} = x_{32}$ , then it is easy to see that stages 3 and 4 are identical. Since  $C(4)$  holds for stages 1 and 2, we have  $c_1 = 0$  and  $c_2 = 1$  (or vice versa) so that  $c_3 \in \{0, 1\}$  would imply  $s^* < 5$ , which is impossible.

(iv)  $\{3, j\} \subset S_L, j > 4$ . Here  $x_{43} = 0$ , so we may permute stage 3 and stage 4. The first three stages of an irreducible CMIRK scheme cannot satisfy  $C(4)$  (Burrage, Chipman, and Muir [6]).  $\square$

In particular, this theorem shows that all 5th-order CMIRK schemes with six stages satisfy  $C(3)$ . Turning to the order 6 case, we have

**Theorem 3.2.** *There are no 6th-order CMIRK schemes with seven stages or less.*

*Proof.* By (2.10), a 6th-order scheme must have at least seven stages. Assume that such a scheme exists with seven stages. Then, according to Theorem 3.1 it must satisfy  $C(3)$  and we may apply the second part of Theorem 2.11 with  $k = 4$ , and Remark 2.12. Thus, there is a set  $S_L$  that consists of exactly two indices. At least one of these must be in  $\{1, 2, 3\}$ , since stage  $i$  satisfies  $C(5)$  for  $i \notin S_L$ . But neither stage 1 nor stage 2 is of interest, since both these stages satisfy  $C(k)$  for arbitrary  $k$  when  $C(3)$  is imposed. Thus, we must have  $S_L = \{3, j\}$ , where  $j > 3$ . Hence,  $x_{43} = 0$ . If  $j = 4$ , then  $c_3 = c_4$ , and stages 3 and 4 are identical. If  $j > 4$ , we permute the 3rd and 4th stages, and hence, the first three stages must satisfy  $C(5)$ , which is impossible.  $\square$

Hence,  $\text{CMN}(6) \geq 8$ , and in §4 we shall see that CMIRK schemes of order 6 with eight stages exist, thus  $\text{CMN}(6) = 8$ .

Summarizing, we have the following table, where  $p$  is the order,  $\text{CMN}(p)$  is the minimal number of stages a CMIRK scheme of order  $p$  can have, and  $\text{MN}(p)$  is the minimal number of stages a (discrete) MIRK scheme of order  $p$  can have (Burrage, Chipman, and Muir [6]):

$p$	$\text{CMN}(p)$	$\text{MN}(p)$
1	1	1
2	2	1
3	3	2
4	4	3
5	6	4
6	8	5

#### 4. CHARACTERIZATIONS OF CONTINUOUS MONO-IMPLICIT RUNGE-KUTTA SCHEMES

In this section we present characterizations for a number of families of low-order CMIRK schemes having uniform orders  $p = 1, 2, 3, 4, 5,$  and  $6,$  respectively. We consider only schemes having a minimal number of stages, and within each such class we will consider all possible  $C(\xi)$  restrictions which preserve the minimal stage requirement. All the CMIRK schemes presented here having order 4, 5, or 6, and all schemes of order 3 satisfying  $C(3)$ , are  $C^1$  continuous. In fact, when the order is 3 or greater, it can be easily shown that for all CMIRK schemes which use only the minimal number of stages and satisfy  $C(3)$ , we must have  $C^1$  continuity.

**4.1. CMIRK schemes of order 1.** A CMIRK scheme of uniform order 1 can be obtained using one stage. For the family of 1-stage CMIRK schemes we have the tableau:

$c$	$c$	$0$
		$b(\theta)$

where  $b(\theta) = \theta$  and  $c$  is a free parameter. These schemes satisfy  $C(1)$  and include the continuous versions of the explicit and implicit Euler schemes.

**4.2. CMIRK schemes of order 2.** A CMIRK of order 2 can be obtained using two stages. If we assume that only the  $C(1)$  condition holds, we get the following 3-parameter family of 2-stage CMIRK schemes, of uniform order 2:

$c_1$	$c_1$	$0$	$0$
$c_2$	$v_2$	$c_2 - v_2$	$0$
		$b_1(\theta)$	$b_2(\theta)$

where

$$c_1 \neq c_2, \quad b_1(\theta) = \frac{\theta}{c_2 - c_1} \left( c_2 - \frac{\theta}{2} \right), \quad \text{and} \quad b_2(\theta) = \frac{\theta}{c_1 - c_2} \left( c_1 - \frac{\theta}{2} \right).$$

If we further require the schemes to satisfy  $C(2)$ , we get the following 1-parameter family of CMIRK schemes having uniform order 2:

0	0	0	0
$c_2$	$c_2^2$	$c_2(1 - c_2)$	0
		$b_1(\theta)$	$b_2(\theta)$

where

$$c_2 \neq 0, \quad b_1(\theta) = \frac{\theta}{c_2} \left( c_2 - \frac{\theta}{2} \right), \quad \text{and} \quad b_2(\theta) = \frac{\theta^2}{2c_2}.$$

This family includes the continuous version of the trapezoidal rule.

**4.3. CMIRK schemes of order 3.** A third-order CMIRK scheme requires at least three stages. Furthermore, in order to have no more than three stages, it must satisfy  $C(2)$ . Therefore, we will consider CMIRK schemes satisfying  $C(2)$  and  $C(3)$ . For the  $C(2)$  case we get the following 3-parameter family of schemes:

0	0	0	0	0
$c_2$	$c_2^2$	$c_2(1 - c_2)$	0	0
$c_3$	$v_3$	$c_3 - v_3 - \frac{c_3^2 - v_3}{2c_2}$	$\frac{c_3^2 - v_3}{2c_2}$	0
		$b_1(\theta)$	$b_2(\theta)$	$b_3(\theta)$

where

$$b_1(\theta) = \frac{\theta(6c_2c_3 - 3\theta(c_2 + c_3) + 2\theta^2)}{6c_2c_3},$$

$$b_2(\theta) = \frac{\theta^2(3c_3 - 2\theta)}{6c_2(c_3 - c_2)}, \quad \text{and} \quad b_3(\theta) = \frac{\theta^2(3c_2 - 2\theta)}{6c_3(c_2 - c_3)},$$

and  $c_2$ ,  $c_3$ , and  $v_3$  are the parameters, with the restrictions that  $c_2 \neq c_3$ ,  $c_2 \neq 0$ , and  $c_3 \neq 0$ .

Let us now consider the case where we impose the maximum  $C(3)$  condition. The CMIRK scheme tableau is further restricted, and we get the following 1-parameter family:

0	0	0	0	0
1	1	0	0	0
$c_3$	$c_3^2(3 - 2c_3)$	$c_3(c_3 - 1)^2$	$c_3^2(c_3 - 1)$	0
		$b_1(\theta)$	$b_2(\theta)$	$b_3(\theta)$

where  $c_3 \neq 0, 1$  and

$$b_1(\theta) = \frac{\theta(6c_3 - 3\theta(1 + c_3) + 2\theta^2)}{6c_3},$$

$$b_2(\theta) = \frac{\theta^2(3c_3 - 2\theta)}{6(c_3 - 1)}, \quad \text{and} \quad b_3(\theta) = \frac{\theta^2(2\theta - 3)}{6c_3(c_3 - 1)}.$$

This family includes the continuous version of the 3-point Lobatto rule.

**4.4. CMIRK schemes of order 4.** A CMIRK scheme of uniform order 4 can have the minimum number of stages, 4, if and only if it satisfies  $C(3)$ . Below we give a 3-parameter family of 4-stage, continuous 4th-order CMIRK schemes satisfying  $C(3)$ :

0	0	0	0	0	0
1	1	0	0	0	0
$c_3$	$c_3^2(3 - 2c_3)$	$c_3(c_3 - 1)^2$	$c_3^2(c_3 - 1)$	0	0
$c_4$	$v_4$	$\frac{\alpha}{6c_3}$	$\frac{\gamma}{6(c_3 - 1)}$	$\frac{-3(c_4^2 - v_4) + 2(c_4^3 - v_4)}{6c_3(c_3 - 1)}$	0
		$b_1(\theta)$	$b_2(\theta)$	$b_3(\theta)$	$b_4(\theta)$

where

$$b_1(\theta) = \frac{-\theta(3\theta^3 - 4\theta^2(1 + c_3 + c_4) + 6\theta(c_3 + c_4 + c_3c_4) - 12c_3c_4)}{12c_3c_4},$$

$$b_2(\theta) = \frac{\theta^2(3\theta^2 - 4\theta(c_3 + c_4) + 6c_3c_4)}{12(c_3 - 1)(c_4 - 1)},$$

$$b_3(\theta) = \frac{\theta^2(3\theta^2 - 4\theta(c_4 + 1) + 6c_4)}{12c_3(c_3 - 1)(c_3 - c_4)},$$

$$b_4(\theta) = \frac{\theta^2(3\theta^2 - 4\theta(c_3 + 1) + 6c_3)}{12c_4(c_4 - 1)(c_4 - c_3)},$$

with  $\alpha = 6c_3(c_4 - v_4) - 3c_3(c_4^2 - v_4) - 3(c_4^2 - v_4) + 2(c_4^3 - v_4)$ ,  $\gamma = 3c_3(c_4^2 - v_4) - 2(c_4^3 - v_4)$ , and the restrictions that  $c_3 \neq c_4$ ,  $c_3 \neq 0, 1$ , and  $c_4 \neq 0, 1$ . In this case,  $F_4(A)$  is a Vandermonde system over the abscissae  $\{0, 1, c_3, c_4\}$  and the resulting weights, as indicated in Lemma 2.10, satisfy

$$b_r(\theta) = \int_0^\theta L_r(t) dt, \quad r = 1, 2, 3, 4,$$

where  $L_r(t)$  is the  $r$ th elementary Lagrange interpolating polynomial for the abscissa set  $\{0, 1, c_3, c_4\}$ .

**4.5. CMIRK schemes of order 5.** As we saw in §3, it is when we consider 5th-order schemes that we first encounter the situation where it is no longer possible to find CMIRK schemes having the same number of stages as the uniform order. In fact, a CMIRK scheme of uniform order 5 must have at least six stages, and, by Theorem 3.1, must satisfy  $C(3)$  in order to attain this minimum.

If we simply apply the  $C(3)$  condition plus the order conditions up to and including order 5, we will get a 10-parameter family of CMIRK schemes, but the expressions for the coefficients and weights of the family are long, complicated expressions. However, for a given choice of the ten free parameters it is trivial to obtain the corresponding CMIRK scheme.

We can simplify things slightly by imposing a restriction on this family of CMIRK schemes. We require all stages, except the third, to satisfy  $C(4)$ . This reduces the number of degrees of freedom in the family and leaves us with a 7-parameter family of 6-stage, 5th-order CMIRK schemes, which we present below:

0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0
$c_3$	$c_3^2(3 - 2c_3)$	$c_3(c_3 - 1)^2$	$c_3^2(c_3 - 1)$	0	0	0	0
$c_4$	$\frac{c_4^2(3c_4^2 - 4c_3c_4 + 6c_3 - c_4)}{2c_3 - 1}$	$x_{41}$	$x_{42}$	$x_{43}$	0	0	0
$c_5$	$v_5$	$x_{51}$	$x_{52}$	$x_{53}$	$x_{54}$	0	0
$c_6$	$v_6$	$x_{61}$	$x_{62}$	$x_{63}$	$x_{64}$	$x_{65}$	0
		$b_1(\theta)$	$b_2(\theta)$	$b_3(\theta)$	$b_4(\theta)$	$b_5(\theta)$	$b_6(\theta)$

Here,

$$\begin{aligned}
 x_{41} &= \frac{c_4(c_4 - 1)^2(4c_3^2 - 3c_3c_4 - 2c_3 + c_4)}{2c_3(2c_3 - 1)}, \\
 x_{42} &= \frac{c_4^2(c_4 - 1)(4c_3^2 - 3c_3c_4 + 2c_4 - 3c_3)}{2(c_3 - 1)(2c_3 - 1)}, \\
 x_{43} &= \frac{c_4^2(c_4 - 1)^2}{2c_3(c_3 - 1)(2c_3 - 1)}, \\
 x_{51} &= \frac{12c_3c_4(c_5 - v_5) - 6(c_3 + c_4 + c_3c_4)(c_5^2 - v_5) + 4(c_3 + c_4 + 1)(c_5^3 - v_5) - 3(c_5^4 - v_5)}{12c_3c_4}, \\
 x_{52} &= -\frac{6c_3c_4(c_5^2 - v_5) - 4(c_3 + c_4)(c_5^3 - v_5) + 3(c_5^4 - v_5)}{12(c_3 - 1)(c_4 - 1)}, \\
 x_{53} &= \frac{6c_4(c_5^2 - v_5) - 4(c_4 + 1)(c_5^3 - v_5) + 3(c_5^4 - v_5)}{12c_3(c_3 - 1)(c_3 - c_4)}, \\
 x_{54} &= \frac{6c_3(c_5^2 - v_5) - 4(c_3 + 1)(c_5^3 - v_5) + 3(c_5^4 - v_5)}{12c_4(c_4 - 1)(c_4 - c_3)}, \\
 x_{62} &= \frac{12c_3c_4c_5(c_6 - v_6 - x_{61}) - 6(c_3c_4 + c_3c_5 + c_4c_5)(c_6^2 - v_6) + 4(c_3 + c_4 + c_5)(c_6^3 - v_6) - 3(c_6^4 - v_6)}{12(c_3 - 1)(c_4 - 1)(c_5 - 1)}, \\
 x_{63} &= \frac{12c_4c_5(c_6 - v_6 - x_{61}) - 6(c_4 + c_5 + c_4c_5)(c_6^2 - v_6) + 4(1 + c_4 + c_5)(c_6^3 - v_6) - 3(c_6^4 - v_6)}{12(1 - c_3)(c_4 - c_3)(c_5 - c_3)}, \\
 x_{64} &= \frac{12c_3c_5(c_6 - v_6 - x_{61}) - 6(c_3 + c_5 + c_3c_5)(c_6^2 - v_6) + 4(1 + c_3 + c_5)(c_6^3 - v_6) - 3(c_6^4 - v_6)}{12(1 - c_4)(c_3 - c_4)(c_5 - c_4)},
 \end{aligned}$$

and

$$x_{65} = \frac{12c_3c_4(c_6 - v_6 - x_{61}) - 6(c_3 + c_4 + c_3c_4)(c_6^2 - v_6) + 4(1 + c_3 + c_4)(c_6^3 - v_6) - 3(c_6^4 - v_6)}{12(1 - c_5)(c_3 - c_5)(c_4 - c_5)}.$$

The free parameters are  $c_3, c_4, c_5, c_6, v_5, v_6$ , and  $x_{61}$ . The restrictions on the parameters  $c_3, c_4, c_5$ , and  $c_6$  are that they must be distinct and not equal to 0 or 1.

The weights are determined by the rows of  $F_5(A)$ . We perform an elimination step on the row of  $F_5(A)$  corresponding to the single nonquadrature condition, as follows. Replace

$$(Xc^3 + v/4) \quad \text{with} \quad c^4 - 4(Xc^3 + v/4).$$

This makes the right-hand sides corresponding to this row equal to 0 and allows us to observe that the  $C(4)$  condition applied to all but the 3rd stage implies that the new row reduces to the form

$$[ \ 0 \ 0 \ c_3^4 - 4([Xc^3]_3 + v_3/4) \ 0 \ 0 \ 0 \ ],$$

where  $[Xc^p]_j$  is the  $j$ th component of the vector  $Xc^p$ . Since the right-hand sides for this row are all 0, this implies that all the coefficients for the third weight are zero, i.e.,  $b_3(\theta) = 0$ . Furthermore, if we then remove the third column and third row from  $F_5(A)$ , the result is a Vandermonde system over the abscissae  $\{0, 1, c_4, c_5, c_6\}$ . The resulting weight expressions are given, as indicated in Lemma 2.10, by

$$b_r(\theta) = \int_0^\theta L_r(t) dt, \quad r = 1, 2, \quad b_3(\theta) = 0,$$

$$b_r(\theta) = \int_0^\theta L_{r-1}(t) dt, \quad r = 4, 5, 6,$$

where  $L_r(t)$  is the  $r$ th elementary Lagrange interpolating polynomial for the abscissa set  $\{0, 1, c_4, c_5, c_6\}$ .

**4.6. CMIRK schemes of order 6.** From Theorem 3.2 of the previous section, we saw that a CMIRK scheme of order 6 must have at least eight stages. In this subsection we will construct a family of 8-stage CMIRK schemes of order 6, satisfying  $C(3)$ . The situation here is somewhat different than in the previous cases, because when we apply the  $C(3)$  condition to the continuous order conditions up to and including order 6, we are left with ten rather than eight conditions. If we are to obtain a family of CMIRK schemes of order 6 with only eight stages, we must choose some of our free parameters to ensure that the ten order conditions lead to only eight independent rows in  $G_6(A)$  and not less than eight independent rows in  $F_6(A)$ . In the following discussion we outline one strategy for the determination of the coefficients of the CMIRK schemes which allows us to achieve this goal. Of the ten order conditions present, six are quadrature conditions, and there are four nonquadrature conditions,

$$b(\theta)^T(Xc^3 + v/4) = \theta^5/20, \quad b(\theta)^T c(Xc^3 + v/4) = \theta^6/24,$$

$$b(\theta)^T(Xc^4 + v/5) = \theta^6/30, \quad b(\theta)^T(X(Xc^3 + v/4) + v/20) = \theta^6/120.$$

Our strategy involves forcing the four nonquadrature conditions to reduce to only two independent rows. In order to simplify the discussion, we first perform an elimination step on the four rows of  $G_6(A)$  corresponding to the nonquadrature conditions. From each nonquadrature row we subtract a multiple of the corresponding quadrature row, and then scale the resulting row. The resulting rows in  $F_6(A)$  are

$$\begin{aligned} \text{(i)} \quad & c^5 - 4(Xc^3 + v/4), & \text{(ii)} \quad & c^6 - 4c(Xc^3 + v/4), \\ \text{(iii)} \quad & c^6 - 5(Xc^4 + v/5), & \text{(iv)} \quad & c^6 - 20(X(Xc^3 + v/4) + v/20), \end{aligned}$$

and the corresponding right-hand sides are 0. This simplification has also been used by Verner [29]. We now give a description of the strategy for the determination of the 8-stage, 6th-order CMIRK family.

The first step follows from an idea of the previous subsection. Here we begin by imposing the restriction that all stages, except the third and fourth, satisfy the  $C(4)$  condition. It can then be shown, from an inspection of the resulting forms of the order conditions, that it follows that the  $C(5)$  condition must also be imposed for all but the third and fourth stages. The rows of  $F_6(A)$  derived from the above nonquadrature conditions become

$$\begin{aligned} \text{(i)} \quad & [ 0 \ 0 \ u_3 \ u_4 \ 0 \ 0 \ 0 \ 0 ], \\ \text{(ii)} \quad & [ 0 \ 0 \ c_3u_3 \ c_4u_4 \ 0 \ 0 \ 0 \ 0 ], \\ \text{(iii)} \quad & [ 0 \ 0 \ w_3 \ w_4 \ 0 \ 0 \ 0 \ 0 ], \\ \text{(iv)} \quad & [ r_1 \ r_2 \ r_3 \ r_4 \ r_5 \ r_6 \ r_7 \ r_8 ], \end{aligned}$$

where  $u_j = c_j^4 - 4([Xc^3]_j + v_j/4)$ ,  $j = 3, 4$ ,  $w_j = c_j^5 - 5([Xc^4]_j + v_j/5)$ ,  $j = 3, 4$ ,  $r_j = 5(u_3x_{j3} + u_4x_{j4})$ ,  $j = 1, 2, 5, 6, 7, 8$ , and  $r_j = w_j + 5(u_3x_{j3} + u_4x_{j4})$ ,  $j = 3, 4$ . If we can force the last of the above rows to also have nonzeros in only positions 3 and 4, then we will have four rows with that same structure, of which exactly two must be independent. These two rows together with the six rows corresponding to the quadrature conditions will allow us to determine a family of 8-stage, 6th-order schemes. Thus, our strategy reduces to determining coefficients of the family so that

$$r_1 = r_2 = r_5 = r_6 = r_7 = r_8 = 0.$$

It follows immediately from the strictly lower triangular structure of  $X$  that  $r_1 = r_2 = 0$ . It turns out that it is also possible to satisfy the remaining conditions and we get some restrictions on some of the coefficients of the family. We are left with a 9-parameter family of 8-stage, 6th-order CMIRK schemes. This defines the  $c$ ,  $v$ , and  $X$  coefficients in terms of  $c_3, c_4, c_5, c_6, c_7, v_7, c_8, v_8$ , and  $x_{81}$ . Before presenting these coefficients, we consider the determination of the weights  $b_j(\theta)$ ,  $j = 1, \dots, 8$ . These depend on the rows of  $F_6(A)$ .

Within  $F_6(A)$ , there are six rows corresponding to the quadrature conditions, plus the two independent rows, with nonzeros only in positions 3 and 4, corresponding to the reduced nonquadrature conditions. There is no need to work out the actual values for  $w_3, w_4$  or  $r_3, r_4$ , since the values of the nonzeros in the two independent rows are not relevant. Since the right-hand sides of these rows are zero for all orders, it can be observed that the two independent rows

having nonzeros in only the third and fourth positions imply that all the coefficients corresponding to  $b_3(\theta)$  and  $b_4(\theta)$  must be zero, i.e.,  $b_3(\theta) = b_4(\theta) = 0$ . With the 3rd and 4th columns and rows of  $F_6(A)$  removed, the resulting system is Vandermonde, and it then follows from Lemma 2.10 that

$$b_r(\theta) = \int_0^\theta L_r(t) dt, \quad r = 1, 2, \quad b_3(\theta) = b_4(\theta) = 0,$$

$$b_r(\theta) = \int_0^\theta L_{r-2}(t) dt, \quad r = 5, 6, 7, 8,$$

where  $L_r(\theta)$  is the  $r$ th elementary Lagrange interpolating polynomial for the abscissa set  $\{0, 1, c_5, c_6, c_7, c_8\}$ .

In summary, a family of 6th-order, 8-stage CMIRK schemes can be obtained by expressing some of the coefficients for each stage in terms of others of that stage, using the  $C(3)$  condition for stages 3 and 4, and the  $C(5)$  condition for the remaining stages. Further relationships among the coefficients of the stages are obtained by requiring  $r_5 = r_6 = r_7 = r_8 = 0$ , as given earlier in this subsection. The completion of this process allows us to write all the  $c$ ,  $v$ , and  $x$  coefficients in terms of the nine free parameters,  $c_3, c_4, c_5, c_6, c_7, v_7, c_8, v_8$ , and  $x_{81}$ . The definition of the CMIRK family is completed by determining the weight polynomials in terms of the integrals of the elementary Lagrange interpolating polynomials, as indicated above.

We have used this approach, and have set up and solved the above conditions, to explicitly obtain expressions for all the coefficients and weight polynomials of this family. As might be expected, the coefficients of this CMIRK family are very long and complicated expressions, so we do not present them here. Rather, we have arbitrarily selected values for the free parameters, inserted these values into the expressions for the coefficients and weight polynomials, and below display the resulting 8-stage, 6th-order CMIRK scheme. The existence of this scheme establishes the order barrier for 6th-order CMIRK schemes at 8.

0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	0	0	0	0
$\frac{1}{8}$	$\frac{11}{256}$	$\frac{49}{512}$	$-\frac{7}{512}$	0	0	0	0	0	0
$\frac{7}{8}$	$\frac{1813}{1813}$	$\frac{203}{203}$	$-\frac{609}{203}$	$-\frac{7}{203}$	0	0	0	0	0
$\frac{8}{8}$	$\frac{1280}{37}$	$\frac{512}{387}$	$\frac{2560}{39}$	$\frac{10}{159}$	$-\frac{15}{159}$	0	0	0	0
$\frac{4}{3}$	$\frac{928}{441}$	$\frac{12992}{1791}$	$\frac{12992}{-877}$	$\frac{182}{-265}$	$\frac{812}{25}$	0	0	0	0
$\frac{4}{2}$	$\frac{928}{2}$	$\frac{12992}{694}$	$\frac{12992}{-8114}$	$\frac{812}{265424}$	$\frac{812}{-5008}$	$\frac{1}{2}$	0	0	0
$\frac{5}{3}$	$\frac{5}{3}$	$\frac{109375}{328125}$	$\frac{328125}{-37292}$	$\frac{1640625}{204368}$	$\frac{328125}{-3856}$	$\frac{566}{-88514}$	$\frac{-10954}{-550682}$	0	0
$\frac{5}{5}$	$\frac{5}{5}$	0	$\frac{815625}{815625}$	$\frac{1359375}{1359375}$	$\frac{271875}{271875}$	$\frac{815625}{815625}$	$\frac{3171875}{3171875}$	$\frac{8756}{45675}$	0
		$b_1(\theta)$	$b_2(\theta)$	$b_3(\theta)$	$b_4(\theta)$	$b_5(\theta)$	$b_6(\theta)$	$b_7(\theta)$	$b_8(\theta)$

where

$$\begin{aligned}
 b_1(\theta) &= \frac{-\theta}{216}(-216 + 1134\theta - 2968\theta^2 + 4113\theta^3 - 2880\theta^4 + 800\theta^5), \\
 b_2(\theta) &= \frac{\theta^2}{216}(108 - 684\theta + 1713\theta^2 - 1920\theta^3 + 800\theta^4), \quad b_3(\theta) = 0, \\
 b_4(\theta) &= 0, \quad b_5(\theta) = \frac{64\theta^2}{189}(54 - 234\theta + 411\theta^2 - 330\theta^3 + 100\theta^4), \\
 b_6(\theta) &= \frac{-64\theta^2}{189}(18 - 110\theta + 261\theta^2 - 270\theta^3 + 100\theta^4), \\
 b_7(\theta) &= \frac{-125\theta^2}{1512}(270 - 1440\theta + 2865\theta^2 - 2496\theta^3 + 800\theta^4), \\
 b_8(\theta) &= \frac{-125\theta^2}{1512}(180 - 1060\theta + 2385\theta^2 - 2304\theta^3 + 800\theta^4).
 \end{aligned}$$

## 5. CONCLUSIONS

In this paper we have derived families of CMIRK schemes which use a minimal number of stages, for orders 1 through 6. The underlying theoretical investigation has explored the relationship between the stage order and minimal number of stages required to obtain schemes of a desired order of accuracy, in the more general setting of CRK schemes. Applications of these general results to the particular family of CMIRK schemes lead to the order barriers upon which the characterizations of the families were based.

For orders 1 through 4 complete characterizations of the CMIRK families are presented. For order 5 a complete characterization is available but in the interest of shortening the presentation, a somewhat restricted subfamily is presented. For the 6th-order case, a particular strategy is used to obtain a family of schemes. The particular families presented in this paper for the 4th-, 5th-, and 6th-order cases are interesting because they can be viewed as instances of continuous schemes derived from the more traditional interpolation-based approach for the construction of continuous extensions to Runge-Kutta schemes (Enright et al. [17]). In each of these cases, the stages which do not satisfy  $C(p-1)$ , for the  $p$ th-order family, have zero weight polynomials.

The schemes developed in this paper are complementary to those discussed by Pruess [28], since they require neither higher derivative evaluations nor function evaluations outside the current subinterval. As is the case for all one-step schemes, the MIRK schemes have the advantage of allowing the calculations for each subinterval to be done independently of those on the other subintervals. This is important when we consider applying these schemes in a parallel computing environment where it is preferable to be able to perform the work on each subinterval independently. An application of the characterizations of these continuous families is in the determination of particular CMIRK schemes for use in a defect-controlled boundary value ODE solver (Enright and Muir [18]). After establishing suitable criteria for the selection of optimal schemes for this application, the next step will be to explore the parameter spaces of these families in order to compare various schemes, and ultimately select good schemes for implementation.

## BIBLIOGRAPHY

1. U. Ascher, J. Christiansen, and R. D. Russell, *Collocation software for boundary value ODE's*, ACM Trans. Math. Software **7** (1981), 209–222.
2. G. Bader and U. Ascher, *A new basis implementation for a mixed order boundary value ODE solver*, SIAM J. Sci. Statist. Comput. **8** (1987), 483–500.
3. W. M. G. van Bokhoven, *Efficient higher order implicit one-step methods for integration of stiff differential equations*, BIT **20** (1980), 34–43.
4. D. L. Brown and J. Lorenz, *A high-order method for stiff boundary value problems with turning points*, SIAM J. Sci. Statist. Comput. **8** (1987), 790–808.
5. K. Burrage, *A special family of Runge-Kutta methods for solving stiff differential equations*, BIT **18** (1978), 22–41.
6. K. Burrage, F. Chipman, and P. Muir, *Order results for mono-implicit Runge-Kutta methods*, SIAM J. Numer. Anal. (submitted).
7. J. C. Butcher, *Coefficients for the study of Runge-Kutta integration processes*, J. Austral. Math. Soc. **3** (1963), 185–201.
8. ———, *Implicit Runge-Kutta processes*, Math. Comp. **18** (1964), 50–64.
9. ———, *The numerical analysis of ordinary differential equations*, Wiley, Toronto, 1987.
10. J. R. Cash, *On the numerical integration of nonlinear two-point boundary value problems using iterated deferred corrections, Part 1: A survey and comparison of some one-step formulae*, Comput. Math. Appl. **12a** (1986), 1029–1048.
11. ———, *On the numerical integration of nonlinear two-point boundary value problems using iterated deferred corrections, Part 2: The development and analysis of highly stable deferred correction formulae*, SIAM J. Numer. Anal. **25** (1988), 862–882.
12. J. R. Cash and D. R. Moore, *A high order method for the numerical solution of two-point boundary value problems*, BIT **20** (1980), 44–52.
13. J. R. Cash and A. Singhal, *Mono-implicit Runge-Kutta formulae for the numerical integration of stiff differential systems*, IMA J. Numer. Anal. **2** (1982), 211–227.
14. ———, *High order methods for the numerical solution of two-point boundary value problems*, BIT **22** (1982), 184–199.
15. J. R. Cash and M. H. Wright, *A deferred correction method for nonlinear two-point boundary value problems: implementation and numerical evaluation*, SIAM J. Sci. Statist. Comput. **12** (1991), 971–989.
16. G. Dahlquist and R. Jeltsch, *Generalized disks of contractivity for explicit and implicit Runge-Kutta methods*, Technical Report TRITA-NA-7906, The Royal Institute of Technology, Stockholm, 1979.
17. W. H. Enright, K. R. Jackson, S. P. Nørsett, and P. G. Thomsen, *Interpolants for Runge-Kutta formulas*, ACM Trans. Math. Software **12** (1986), 193–218.
18. W. H. Enright and P. H. Muir, *Efficient classes of Runge-Kutta methods for two-point boundary value problems*, Computing **37** (1986), 315–334.
19. ———, *A mono-implicit Runge-Kutta type BVODE code with defect control*, SIAM J. Sci. Statist. Comput. (submitted).
20. I. Gladwell, L. F. Shampine, L. S. Baca, and R. W. Brankin, *Practical aspects of interpolation in Runge-Kutta codes*, SIAM J. Sci. Statist. Comput. **8** (1987), 322–341.
21. S. Gupta, *An adaptive boundary value Runge-Kutta solver for first order boundary value problems*, SIAM J. Numer. Anal. **22** (1985), 114–126.
22. H.-O. Kreiss, N. K. Nichols, and D. L. Brown, *Numerical methods for stiff two-point boundary value problems*, SIAM J. Numer. Anal. **23** (1986), 325–368.
23. M. Lentini and V. Pereyra, *An adaptive finite difference solver for nonlinear two-point boundary value problems with mild boundary layers*, SIAM J. Numer. Anal. **14** (1977), 91–111.
24. S. P. Nørsett, *Semi-explicit Runge-Kutta methods*, Mathematics and Computation No.6/74, University of Trondheim, 1974.

25. S. P. Nørsett and G. Wanner, *Perturbed collocation and Runge-Kutta methods*, Numer. Math. **38** (1981), 193–208.
26. B. Owren and M. Zennaro, *Order barriers for continuous explicit Runge-Kutta methods*, Math. Comp. **56** (1991), 645–661.
27. ———, *Derivation of optimal continuous explicit Runge-Kutta methods*, SIAM J. Sci. Statist. Comput. **13** (1992), 1488–1501.
28. S. Pruess, *Interpolation schemes for collocation solutions of two-point boundary value problems*, SIAM J. Sci. Statist. Comput. **7** (1986), 322–333.
29. J. Verner, *Differentiable interpolants for high-order Runge-Kutta methods*, Technical Report 1990-9, Queen's University, 1990.
30. R. Weiss, *The application of implicit Runge-Kutta and collocation methods to boundary value problems*, Math. Comp. **28** (1974), 449–464.
31. M. Zennaro, *Natural continuous extensions of Runge-Kutta methods*, Math. Comp. **46** (1986), 119–133.
32. ———, *Natural Runge-Kutta and projection methods*, Numer. Math. **53** (1988), 423–438.

DEPARTMENT OF MATHEMATICS AND COMPUTING SCIENCE, SAINT MARY'S UNIVERSITY, HALIFAX, NOVA SCOTIA, CANADA B3H 3C3  
*E-mail address*: muir@husky1.stmarys.ca

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF TRONDHEIM, N-7055 DRAGVOLL, NORWAY  
*E-mail address*: bryn@sima.sintef.no