A FAMILY OF FIFTH-ORDER RUNGE-KUTTA PAIRS

S. N. PAPAKOSTAS AND G. PAPAGEORGIOU

ABSTRACT. The construction of a Runge-Kutta pair of order 5(4) with the minimal number of stages requires the solution of a nonlinear system of 25 order conditions in 27 unknowns. We define a new family of pairs which includes pairs using 6 function evaluations per integration step as well as pairs which additionally use the first function evaluation from the next step. This is achieved by making use of Kutta's simplifying assumption on the original system of the order conditions, i.e., that all the internal nodes of a method contributing to the estimation of the endpoint solution provide, at these nodes, cost-free second-order approximations to the true solution of any differential equation. In both cases the solution of the resulting system of nonlinear equations is completely classified and described in terms of five free parameters. Optimal Runge-Kutta pairs with respect to minimized truncation error coefficients, maximal phase-lag order and various stability characteristics are presented. These pairs were selected under the assumption that they are used in Local Extrapolation Mode (the propagated solution of a problem is the one provided by the fifth-order formula of the pair). Numerical results obtained by testing the new pairs over a standard set of test problems suggest a significant improvement in efficiency when using a specific pair of the new family with minimized truncation error coefficients, instead of some other existing pairs.

1. INTRODUCTION

One of the methods of solution of the system of first-order differential equations

(1) $y' = f(x, y), \quad y(0) = y_0, \quad x \ge 0, \quad f: R \times R^m \to R^m$

is an explicit Runge-Kutta (RK) pair characterized by the extended Butcher tableau

$$egin{array}{c} c & A \ \hline b \\ \hat{b} \end{array}$$

where b^T , \hat{b}^T , $c \in \mathbb{R}^s$ and $A \in \mathbb{R}^{s \times s}$ is strictly lower triangular. The vectors b, \hat{b} characterize two RK methods (usually of adjacent orders $p \ge q$) which share the same function evaluations (A, c are common to both methods). In the following we use letters with hats to denote quantities pertaining to the lower-order method of a pair. The methods studied in this article obey the simplifying assumption $A \cdot e = c$, $e = (1, 1, \ldots, 1)^T$.

$$-\underbrace{(1,1,\ldots,1)}_{s}$$

Received by the editor September 7, 1993 and, in revised form, September 5, 1994 and April 5, 1995.

¹⁹⁹¹ Mathematics Subject Classification. Primary 65L05.

Key words and phrases. Initial Value Problems, Runge-Kutta, pairs of embedded methods, phase-lag.

Let $x_0 = 0, x_n \ge 0, y_n \approx y(x_n)$ and $K = (k_1^T, k_2^T, \dots, k_s^T)^T, k_i \in \mathbb{R}^m$ for i = 1 (1) s. We define

$$F: R^m \times R^{ms} \to R^{ms}: F(\lambda, K) = \left(f(\lambda_1, k_1)^T, f(\lambda_2, k_2)^T, \dots, f(\lambda_s, k_s)^T\right)^T,$$

 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s)^T, \lambda_i \in R$. Let *I* denote the $s \times s$ identity matrix, $h_n = x_{n+1} - x_n$ and let the fixed point problem in R^{ms}

(2)
$$K_n = F(x_n e + h_n c, e \otimes y_n + h_n (A \otimes I) K_n)$$

be solved with respect to K_n (in the case of explicit methods) by recursive substitution. Provided that the solution at x_n is known, the solution of (1) at the next point x_{n+1} is evaluated from the equation

(3)
$$y(x_{n+1}) \approx y_{n+1} = y_n + h_n (b \otimes I) K_n$$

Equations (2) and (3) describe, for n = 1, 2, ..., the algorithm for the application of an explicit *s*-stage RK method to the problem (1). At the same time, an error estimation for the *q*th-order formula of the pair may be obtained from the relation

$$e(x_{n+1}) \approx y_{n+1} - \hat{y}_{n+1} = h_n\left(\left(b - \hat{b}\right) \otimes I\right) K_n.$$

A Runge-Kutta method is of order p if and only if

 $X(\tau) = 0 \ \forall \tau \in T_i, \text{ for } i = 1(1) p,$

where T_i is the set of rooted trees of order i and

(4)
$$X(\tau) = \frac{1}{\sigma(\tau)} \left(\Phi(\tau) - \frac{1}{\gamma(\tau)} \right).$$

Here, σ , γ are integer-valued functions of τ and Φ is a certain composition of A, b, c, the skeleton of which depends only on τ . Relation (4) defines a set of order conditions, which are linear in the components of b and nonlinear in the components of A, c (see, for example, Hairer, Nørsett and Wanner [9] or Butcher [2]). In the following the symbol $T^{(i)}$ denotes a vector whose elements are all the elements of the set $X(T_i)$ in some prescribed (but otherwise arbitrary) order.

It is known that the minimal number of stages required for the construction of a fifth-order RK method and a 5(4) pair is six (see Butcher [1] or Shanks [17] for the negative part of this statement and, for example, Fehlberg [8] for the affirmative part). A complete characterization of the solution of the 17 order conditions for a fifth-order method is given by Cassity [3]. However, the results of his study have not found any practical implementation. No analogous study exists for fifth-order pairs. All methods and pairs of orders five or higher are constructed according to certain types of simplifying assumptions, applied to the original system of order conditions.

The most popular RK pairs currently in use are those constructed by Fehlberg [8] (most notably the pair FE4(5)#2 used for example in Matlab) and by Dormand and Prince [5] (DP5(4)7M or simply DP5(4)). The latter is undoubtedly better when both pairs are applied in local extrapolation mode (see below). The pair of Fehlberg belongs to a two-parameter family of pairs. Fehlberg selected the values of these free parameters in order to minimize the truncation error coefficients of the fourth-order method of the pair. It can be shown that the restriction $c_5 = 1$ that Fehlberg imposed on the pairs of the family he proposed is not essential, and

it seems that the only reason for its use was the simplification of the otherwise very laborious, at that time, necessary calculations. After performing extensive numerical testing, Shampine in [16] suggested that, from the numerical point of view, it is advantageous to propagate the higher-order solution of a pair (*Local Extrapolation or Higher Order Mode*). Later on, Dormand and Prince extended a family of fifth-order methods defined by Butcher in [1]. The family they proposed uses the first function evaluation from the next step in order to embed a fourth-order method to the fifth-order one, at effectively no additional cost. An individual pair of their four-parameter family, with minimized truncation error coefficients of its fifth-order method, is until now widely regarded as the best fifth-order pair (used in local extrapolation mode).

In general, some norm (usually the Euclidean or the maximum) of the truncation error coefficients of that formula of the pair that propagates the numerical solution is regarded as a good indication of its numerical performance. Among families of pairs of the same order and of the same number of effective function evaluations, usually the best pair that can be constructed belongs to the family with the greater number of free parameters. The latter number is usually disproportionate to the number of simplifying assumptions that were used for the derivation of the respective family. This observation is justified among others in [15] and has motivated the study presented in this article.

We intend to construct a family of 5(4) RK pairs under the *single* simplifying assumption (of Kutta [12]) that the function evaluations corresponding to the internal stages of a method, which contribute to the estimation of the endpoint solution, provide cost-free second-order approximations at these nodes (for any differential system, as the one described by (1)). The solution of the resulting simplified system of equations with respect to *five* free parameters is provided and all cases, depending on the values of these parameters, are described and characterized completely. In our analysis we exclude, after a theoretical justification, those cases that do not lead to practically implementable pairs.

Sometimes it may be desirable to apply an explicit RK pair to the solution of a periodic initial value problem describing free oscillations or free oscillations of low frequency with forced oscillations of high frequency superimposed. For problems of this type, as is shown in [14], it seems advantageous to use specially constructed pairs with a high *phase-lag order*. For an explanation of the term and the resulting order conditions for the case of RK methods, the reader might wish to consult [10] or [14].

A search has been carried out among the pairs of the newly constructed family, in order to find those possessing a fifth-order method with minimized truncation error coefficients, maximized phase-lag order or enhanced stability characteristics. Our search resulted in the selection of some pairs which were derived employing the FSAL device (First Step As Last) and some others that were derived without it. Some numerical testing reported in the last section of this paper, employing a fifth-order pair with minimized truncation error coefficients, seems to point in favor of its reliability and improved efficiency.

2. The construction of the New Pairs

The order conditions that define an explicit RK method of order not exceeding four are linear in the elements of A. For higher orders this useful property is lost.

One reason for the employment of simplifying assumptions for the solution of these order conditions (and an inevitable one according to our opinion) is because in this way this *strong nonlinearity* can be avoided. Moreover, these assumptions are chosen in such a way as to inherit an algebraic compatibility with respect to the original system of order conditions. This also results in a significant reduction of their number. The nonlinearity of the resulting simplified system of order conditions, in terms of the elements of c, causes no problem as long as the latter remain free parameters of the solution.

In the following we denote by the integer power of a vector, v^i , the componentwise multiplication $\underbrace{v \cdot v \cdots v}_{i}$. We use the customary symbol $v_{(i,j)}$ to denote the vector whose elements are the elements of v from index i through j, and $v_{(i+)}$ to denote the vector resulting from v by dropping its first i-1 elements. For a set of vectors these symbols apply to every vector in the set (it is assumed that the same convention applies when applying these symbols to both parts of an equation). We also define C = diag(c).

When we are not interested in the estimation of the values of the truncation error coefficients of a method, but only to solve the order conditions for a specific method or pair, we may ignore the term $1/\sigma(\tau)$ in equations (4) and we may alternatively use one of two different formulations that may be used to describe these equations. The first formulation is by writing for a specific order, say p, all of them except one as a series of orthogonality conditions. In this case it is usually advantageous to retain the equation for the bushy tree (that is $bc^{p-1} = 1/p$) and express the rest as their difference from this equation. A second way to express all or some of the equations of them (as was proposed by Curtis [4]). This is essentially the use of a Gauss-type elimination, when the equations resulting from (4) are considered as a multilinear system in terms of the elements of A. Henceforth, both for theoretical and practical reasons, we shall most frequently use here both of these alternative approaches.

Hypothesis (A). We assume that A, b, \hat{b} , c characterize a 5(4) RK pair (with distinct b, \hat{b}) which effectively uses six stages (i.e., $b_7 = 0$), so that $c_7 = 1$ and $a_{7i} = b_i$ for i = 1 (1) 6 (FSAL device). Consequently, we may conveniently assume that if $\hat{b}_7 = 0$ (non-FSAL pairs), then $c_7 = 0$ and $a_{7i} = 0$.

Simplifying Assumption (B). All the function evaluations corresponding to the internal nodes of a RK pair, except of the second, provide at these nodes second-order approximations of the theoretical solution of any differential equation as in (1). This is equivalent to requiring

$$\left(Ac - \frac{c^2}{2}\right)_{(3,6)} = 0.$$

If the equations for the bushy trees (see (7b) below) are satisfied, the above equation holds for index seven as well, irrespective of whether or not the FSAL device is used.

Remark 1. For a pair satisfying Hypothesis (A), we have $b_6 \cdot c_2 \neq 0$. Otherwise, we would be led to an essentially five-stage, fifth-order method, which is impossible.

We first give a generalization and the formal statement of a result of major importance in the construction of RK methods and pairs, because of its motivation for the use of various types of simplifying assumptions during their construction (see Curtis [4] and Butcher [2] for an informal presentation in a more narrow context). In the following, I represents an identity matrix of suitable dimension.

Proposition 1. Suppose A, b, c characterize an s-stage RK method for which the quadrature order conditions $bc^i = \frac{1}{i+1}$ for i = 0 (1) p, $p \ge 3$ hold. For nonnegative integers k, l such that $k + l \le p - 3$, consider

$$\begin{split} \chi_i &= \left(b \left(C^k \left(A + \frac{C}{k+1} \right) - \frac{I}{k+1} \right) \right)_i, \\ \psi_i &= \left(A c^l - \frac{c^{l+1}}{l+1} \right)_i, \end{split}$$

If all χ_i , ψ_i equal zero except possibly one (say χ_{μ} or ψ_{μ}), then this one is also identically equal to zero provided that $b_{\mu}c_{\mu} \neq 0$.

Proof. The proof follows by establishing the relation

$$\begin{split} \chi_{\mu} &= \left(b \left(C^{k} \left(A + \frac{C}{k+1} \right) - \frac{I}{k+1} \right) \right)_{\mu} c_{\mu}^{l} \\ &= \left(b \left(C^{k} \left(A + \frac{C}{k+1} \right) - \frac{I}{k+1} \right) \right) c^{l} \\ &= b C^{k} A c^{l} + \frac{1}{k+1} b C^{k+1} c^{l} - \frac{1}{k+1} b c^{l} \\ &= b C^{k} \left(A c^{l} - \frac{c^{l+1}}{l+1} \right) + \frac{1}{l+1} b C^{k} c^{l+1} + \frac{1}{k+1} b C^{k+1} c^{l} - \frac{1}{k+1} b c^{l} \\ &= b C^{k} \left(A c^{l} - \frac{c^{l+1}}{l+1} \right) + \frac{1}{(k+l+2)(l+1)} \\ &+ \frac{1}{(k+l+2)(k+1)} - \frac{1}{(k+1)(l+1)} \\ &= \left(b C^{k} \right)_{\mu} \left(A c^{l} - \frac{c^{l+1}}{l+1} \right)_{\mu} \\ &= b_{\mu} c_{\mu}^{k} \psi_{\mu}. \quad \Box \end{split}$$

Obviously, Proposition 1, in case of a RK pair, is applicable to both of its methods.

Lemma 1 (Simplified System of Equations). Under Hypothesis (A) the original system of order conditions for a 5(4) RK pair, satisfying Simplifying Assumption (B), is equivalent to the following simplified system:

$$Ae = c$$
 Hypothesis,

(5)
$$\left(Ac - \frac{c^2}{2}\right)_{(3,6)} = 0$$
 Simplifying Assumption (B),

$$\begin{array}{c} b_2 = 0\\ (bA)_2 = 0\\ (bcA)_2 = 0\\ (bA^2)_2 = 0\\ \hat{b}_2 = 0\\ (\hat{b}A)_2 = 0 \end{array} \end{array} \right\} \qquad Subsidiary \ Simplifying \ Assumptions, \\ \end{array}$$

(7a-e)

$$be = 1$$
, $bc = \frac{1}{2}$, $bc^2 = \frac{1}{3}$, $bc^3 = \frac{1}{4}$, $bc^4 = \frac{1}{5}$ Quadrature Conditions in terms of the weights b_i ,

(7f-i) $\hat{b}c = 1$, $\hat{b}c = \frac{1}{2}$, $\hat{b}c^2 = \frac{1}{3}$, $\hat{b}c^3 = \frac{1}{4}$ Quadrature Conditions in terms of the weights \hat{b}_i ,

(8a-e)
$$bAc^2 = \frac{1}{3 \cdot 4}, \quad bAc^3 = \frac{1}{4 \cdot 5}, \quad bcAc^2 = \frac{1}{3 \cdot 5}, \quad bA^2c^2 = \frac{1}{3 \cdot 4 \cdot 5}, \\ \hat{b}Ac^2 = \frac{1}{3 \cdot 4}.$$

Equations (7a-e) and (8a-e) is all that remains from the original system of order conditions.

Proof. First we prove the subsidiary simplifying assumption (6a-f). The others follow in a similar way. From Simplifying Assumption (B) we have

$$bA^2\left(Ac-\frac{c^2}{2}\right)=0 \quad \Leftrightarrow \quad \left(bA^2\right)_2\left(-\frac{c_2^2}{2}\right)=0,$$

and since $c_2 \neq 0$ (see Remark 1), it follows that $(bA^2)_2 = 0$. Next we may prove in a similar way that the original equations based on the values of $\Phi(\tau)$: $b\Phi_1\Phi_2Ac$, $b\Phi_1Ac$ become equivalent to those of $b\Phi_1\Phi_2c^2$, $b\Phi_1c^2$ whenever the symbols Φ_1 , Φ_2 assume one of the values in $\{A, C, I\}$. Finally, we need to show that the original order condition $b(Ac)^2 = \frac{1}{2\cdot2\cdot5}$ is equivalent to $bC^2Ac = \frac{1}{2\cdot5}$, which follows from the identity

$$b(Ac)^{2} - bC^{2}Ac + \frac{1}{4}bc^{4} = b\left(Ac - \frac{c^{2}}{2}\right)^{2} = 0.$$

Lemma 2. Suppose that Hypothesis (A) and Simplifying Assumption (B) hold. Let

$$v = \left(A\left(Ac^2 - \frac{c^2}{2}\right)\right)_{(3,6)}$$

and for i = 3, 4, ..., 7,

$$S^{i} = \left\{c, c^{2}, c^{3}
ight\}_{(3,i)} \text{ and } T^{i} = S^{i} \cup \left\{A\left(Ac - \frac{c^{2}}{2}
ight)
ight\}_{(3,i)}$$

Then

- (i) $v \neq 0;$
- (ii) The vectors in S^7 are linearly independent;
- (iii) The vectors in T^7 are linearly independent;
- (iv) If $\hat{b}_7 \neq 0$, then b(A + C I) = 0.

Proof. (i) If the assertion were not true, from Simplifying Assumption (B) and Remark 1 it would follow that all the nodes after c_2 vanish (which, in view of the Vandermonde system of order conditions, leads to a contradiction).

(ii) In case the assertion were not true, for suitable real numbers κ , λ , μ not all zero, we would have

(9)
$$\left(\kappa c + \lambda c^2 + \mu c^3\right)_{(3,7)} = 0$$

Multiplying this relation from the left by $b_{(3,7)}$, $(bC)_{(3,7)}$ and using (7a-e), we obtain

(10)
$$\frac{\frac{\kappa}{2} + \frac{\lambda}{3} + \frac{\mu}{4} = 0,}{\frac{\kappa}{3} + \frac{\lambda}{4} + \frac{\mu}{5} = 0.}$$

Equations (10), together with the last row of (9), $\kappa + \lambda + \mu = 0$, lead us to the contradiction $\kappa = \lambda = \mu = 0$.

(iii) The proof is based on case (ii). Moreover, it is similar to the latter case if one takes into account the conditions $(bA)_2 = 0$, $(bCA)_2 = 0$ of Lemma 1.

(iv) The second component of the vector b(A + C - I) is, by Lemma 1, zero and the first, by Proposition 1, is also zero, provided that the rest are zero as well. A direct estimation using some of the original order conditions shows that the vectors in

$$\left\{b-\hat{b},b\left(A+C-I
ight)
ight\}_{(3,7)}$$

are orthogonal to the four linearly independent vectors (from case (iii)) of T^7 , and thus they are linearly dependent. Since $(b(A + C - I))_7 = 0$ and $(b - \hat{b})_7 \neq 0$, the assertion is proved.

One of the necessary conditions that a RK pair must satisfy, if it is to be used in practical implementations, is the nonvanishing of the (q + 1)st-order truncation error coefficients of its lower-order formula. Methods that do not satisfy this condition are those that fail on quadrature problems (*Quadrature Defective Methods*), as for example, the pairs of Fehlberg of orders six or higher (in Local Extrapolation Mode) [7]. In general, we call *defective* those p(p-1) pairs for which some of the *p*th-order truncation error coefficients of their lower-order formula vanish. Usually, known pairs that exhibit this kind of behavior fall into singular cases of more general families of pairs. We do not intend, however, to fully justify this claim in the present article.

Condition (C). It is assumed that the nodes c_i of Hypothesis (A) satisfy the restrictions

$$c_i \neq c_j, \quad c_i \neq 0 \quad \text{for} \quad i, j = 3, 4, 5, 6.$$

When Condition (C) is violated, the subsequent analysis must be modified, and it becomes somewhat lengthy because of the rather large number of singular cases that must be considered. The following lemma justifies our decision to disregard, in the rest of this article, cases that violate Condition (C).

Lemma 3. If a RK pair satisfies Hypothesis (A), Simplifying Assumption (B), but fails to satisfy Condition (C), then it is Quadrature Defective.

Proof. First we note that, irrespective of the value of \hat{b}_7 , because of Lemma 2 (iv) and the remark made in Hypothesis (A), $((c - c_6 e) c)_i$ is equal to zero for i = 1, 6, 7. Furthermore, according to the hypothesis, there are suitable values j, k such that $((c - c_j e) (c - c_k e) (c - c_6 e) c)_{(3,7)} = 0$. Next from the relation

$$\left(b-\hat{b}\right)\left(c-c_{j}e\right)\left(c-c_{k}e\right)\left(c-c_{6}e\right)c=0$$

it follows that $\hat{b}c^4 = \frac{1}{5}$, which is a contradiction, and the lemma is proved.

Lemma 4. Under Hypothesis (A), Simplifying Assumption (B) and Condition (C) the following holds:

- (i) The vectors in S^5 are linearly independent;
- (ii) Let $u = (A(C c_3I)c)_{(3,6)}$ and let v be defined as in Lemma 2. If $\hat{b}_7 = 0$, then the vectors in $U_k^6 = S^6 \cup \{u + kv\}$ are linearly dependent for every value of k;

(iii) If
$$\hat{b}_7 = 0$$
, then $c_4 = \frac{c_3}{2(5c_3^2 - 4c_3 + 1)}$;

(iv) If $c_6 = 1$, then b(A + C - I) = 0.

Proof. (i) The resulting Vandermonde matrix has distinct elements.

(ii) This case of the lemma follows from the fact that the nonzero vector $(b-\hat{b})_{(3,6)}$ is orthogonal to the four vectors of U_k^6 .

(iii) Let $\eta = \frac{2(c_2-c_3)}{c_2}$. From case (ii) of the present lemma we see that there exist real numbers κ, λ, μ (not all of them zero) such that

(11)
$$u + \eta v = \left(\kappa c + \lambda c^2 + \mu c^3\right)_{(3,6)}$$

A direct calculation, taking into account Simplifying Assumption (B), shows that the first two components of (11) are

(12)
$$\begin{array}{l} 0 = \kappa c_3 + \lambda c_3^2 + \mu c_3^3, \\ 0 = \kappa c_4 + \lambda c_4^2 + \mu c_4^3. \end{array}$$

Multiplying successively equation (11) from the left by $b_{(3,6)}$, $(bC)_{(3,6)}$, we find

(13)
$$\int_{0}^{1} \int_{0}^{x} (x - c_3) x \, dx \, dx = \frac{\kappa}{2} + \frac{\lambda}{3} + \frac{\mu}{4}, \\ \int_{0}^{1} x \int_{0}^{x} (x - c_3) x \, dx \, dx = \frac{\kappa}{3} + \frac{\lambda}{4} + \frac{\mu}{5}.$$

Equations (12), (13) must be satisfied simultaneously. A trivial calculation gives the required relation expressing c_4 in terms of c_3 (the singular cases $c_3 = \frac{1}{2}$, $c_4 = 1$ and $c_3 = \frac{19 \pm \sqrt{41}}{40}$, $c_4 = \frac{2}{3}$ still satisfy the required relation). (iv) Using some of the original order conditions, we find that $(b (A+C-I))_{(3,5)}$

(iv) Using some of the original order conditions, we find that $(b(A+C-I))_{(3,5)}$ is orthogonal to the three linearly independent vectors (from case (i) of the present lemma) of S^5 and consequently is equal to zero.

The following result is based on Lemmas 3 and 4.

Corollary 1. An effectively six-stage, nondefective, 5(4) RK pair must satisfy the relations

$$\hat{b}_7 \neq 0 \quad and \quad b \left(A + C - I
ight) = 0,$$

or $\hat{b}_7 = 0 \quad and \quad c_4 = rac{c_3}{2 \left(5 c_3^2 - 4 c_3 + 1
ight)}.$

Lemma 5. Assume that $c_4 = \frac{c_3}{2(5c_3^2 - 4c_3 + 1)}$. The equation $bA^2c^2 = \frac{1}{3\cdot 4\cdot 5}$ is a consequence of all the other equations in Lemma 1, not including $\hat{b}Ac^2 = \frac{1}{3\cdot 4}$. Moreover, the vectors in U_n^6 are linearly dependent.

Proof. We shall first prove the second assertion of the lemma. Let u, v, U_k^6 be defined as in Lemma 4 (ii), η as in the proof of Lemma 4 (iii), and let

$$egin{array}{rcl} ilde{u} &=& (A\left(C-c_{3}I
ight)c
ight)_{(3,5)}\,, & ilde{v} = \left(A\left(Ac-rac{c^{2}}{2}
ight)
ight)_{(3,5)}\,, \ & ilde{b} &=& b_{(3,5)} & ext{and} & ilde{c} = c_{(3,5)}\,. \end{array}$$

The four vectors of R^3 in the set $S^5 \cup \{\tilde{u} + \eta \tilde{v}\}$ are linearly dependent, and because of Lemma 4 (i) we find that for suitable numbers κ , λ , μ

(14)
$$\tilde{u} + \eta \tilde{v} = \kappa \tilde{c} + \lambda \tilde{c}^2 + \mu \tilde{c}^3$$

We notice that equations (12) are valid in this case as well. Multiplying (14) from the left by $(b(C - c_6 I))_{(3,5)}$, we find that

(15)
$$\int_{0}^{1} x \int_{0}^{x} (x - c_{3}) x \, dx \, dx - c_{6} \int_{0}^{1} \int_{0}^{x} (x - c_{3}) x \, dx \, dx$$
$$= \kappa \left(\frac{1}{3} - \frac{c_{6}}{2}\right) + \lambda \left(\frac{1}{4} - \frac{c_{6}}{3}\right) + \mu \left(\frac{1}{5} - \frac{c_{6}}{4}\right).$$

Using the assumed expression concerning c_4 , we may show that the above values of κ , λ , μ satisfy both equations (13) as well. From the sequence of relations

$$\begin{split} b_6 \left(u_6 + \eta v_6 \right) &= b \left(u + \eta v \right) - b \left(\tilde{u} + \eta \tilde{v} \right) \\ &= \int_{0^+}^1 \int_{0^+}^x \left(x - c_3 \right) x \, dx \, dx - \left(k \tilde{b} \tilde{c} + \lambda \tilde{b} \tilde{c}^2 + \mu \tilde{b} \tilde{c}^3 \right) \\ &= \frac{\kappa}{2} + \frac{\lambda}{3} + \frac{\mu}{4} - \left(\kappa \left(bc - b_6 c_6 \right) + \lambda \left(bc^2 - b_6 c_6^2 \right) + \mu \left(bc^3 - b_6 c_6^3 \right) \right) \\ &= b_6 \left(\kappa c_6 + \lambda c_6^2 + \mu c_6^3 \right) \end{split}$$

it follows that (11) is also satisfied for the index value of 6, and consequently the vectors in U_n^6 are linearly dependent.

The first assertion of the lemma concerning equation $bA^2c^2 = \frac{1}{3\cdot 4\cdot 5}$ may be obtained by multiplying (11) from the left by $(b(A+C-I))_{(3,6)}$.

Lemma 6. Let $c_6 = 1$. The equations $bA^2c^2 = \frac{1}{3\cdot 4\cdot 5}$, $(bA^2)_2 = 0$ are a consequence of all the other equations of Lemma 1.

Proof. The proof is based on Lemma 4 (iv).

Lemma 7. Let $c_4 = \frac{c_3}{2(5c_3^2 - 4c_3 + 1)}$. The equations $\hat{b}Ac^2 = \frac{1}{3\cdot 4}$, $(bA^2)_2 = 0$ are a consequence of all the other equations in Lemma 1.

Proof. Using Lemma 5 and multiplying (11) from the left by $(b - \hat{b})_{(3,6)}$, we may easily prove the result concerning the equation $\hat{b}Ac^2 = \frac{1}{3\cdot 4}$.

For the proof of the second equation we distinguish two cases. If $c_6 = 1$, we use Lemma 6. Otherwise, it may be shown that the vectors in $\left\{b\left(A+C-I\right), b-\hat{b}\right\}_{(3,6)}$ are orthogonal to the vectors in S^6 and consequently they are linearly dependent. Moreover, for some real k we have $\left(b\left(A+C-I\right)=\kappa\left(b-\hat{b}\right)\right)_{(3,6)}$. Multiplying this relation from the right by $\left(A^2\right)_{(3,6)}$, we easily see that indeed in this case $\left(bA^2\right)_2 = 0$ holds.

We use the results of the analysis in this section (particularly Lemmas 2, 6, 7) to present a general algorithm for the construction of effectively six-stage, nondefective, RK pairs of order 5(4), a full description of which may be found in the Appendix.

Theorem 1. The set of coefficient values in A, b, \hat{b} , c provided by the algorithm in the Appendix determines an effectively six-stage, nondefective 5(4) RK pair.

The major advantage of the algorithm of the Appendix, over some other similar existing ones, is the rather unifying approach it offers in the treatment of the problem of solving the relevant system of the order conditions.

3. The selection of new pairs and numerical results

The new family of 5(4) pairs we studied in the previous section is mainly characterized by one more degree of freedom in the selection of its nodes than the other known families, which in view of Theorem 1, can be seen to form a proper subfamily of ours. Since for nonstiff problems, the dependence of the numerical performance of a RK pair on the values of its nodes is crucial, we expect this property of the new family to have some beneficial effects on the selection of pairs with enhanced accuracy and stability characteristics.

First a modified Newton minimization algorithm was used in conjunction with the algorithm described in the Appendix of this article for the purposes of selecting new pairs with minimized sixth-order truncation error coefficients with respect to their higher-order methods (i.e., $||T^{(6)}||_2$). It may be shown that this quantity does not depend on \hat{b}_6 , and this happens both for the six-stage and the seven-stage FSAL pairs of the new family. Consequently, for all the new pairs presented here it is assumed that, depending on the occasion (i.e., FSAL or non-FSAL pairs), \hat{b}_6 or \hat{b}_7 is equal to 1/20. The absolute minimum of $||T^{(6)}||_2$ for the six-stage pairs is attained for the free parameters taking the following values, in the form of rational approximations accurate to three significant digits,

$$(c_2, c_3, c_5, c_6) = \left(\frac{33}{133}, \frac{23}{81}, \frac{77}{94}, \frac{15}{17}\right),$$

and the resulting pair, designated as NEW5(4), has the defining values included in Table 1. In this table are also included the respective values of all the other new pairs mentioned in this article, as well as those of some older ones.

TABLE 1. The main characteristics of the KK pairs discussed in this article along with
the values of the free parameters determining them. For the pairs proposed here, \hat{b}_6 or \hat{b}_7 ,
depending on the occasion (i.e., FSAL or non-FSAL pairs), are equal to $1/20$. Some other
fifth-order pairs may be found in [8] and [14].
I_{IM} : Imaginary Stability Interval. I_R : Real Stability Interval.
$\begin{array}{ccc} & & & \\ & & & & \\ & & & \\ & & & & & \\ & & & & \\ & & & & & \\ & & & & & \\ & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\ & & & & & \\$
$B_2 = \frac{\ \hat{T}_{(5)}\ _{s}^{2}}{\ \hat{T}_{(5)}\ _{s}^{2}}, C_2 = \frac{\ \hat{T}_{(5)}\ _{s}^{2}}{\ \hat{T}_{(5)}\ _{s}^{2}},$

I_{IM}	(0,0.99)	(0,0)	(0, 0)	(0, 0)	(0, 0)	(0, 0.26)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	(0, 0)	
I_R	(-3.30, 0)	(-3.73, 0)	(-5.70, 0)	(-4.43, 0)	(-3.67, 0)	(-3.52, 0)	(-3.68, 0)	(-3.83, 0)	(-3.83, 0)	(-5.70, 0)	(-5.70, 0)	(-6.20, 0)	(-6.20, 0)	
D_{∞}	11.59	3.43	3.85	3.60	80	13.73	1	7.81	3.71	8.97	33.56	11.63	30.65	
3°	1.66	1.34	1.74	1.72	1.36	1.82	1.36	2.00	1.51	1.66	1.32	1.63	1.60	
B_2	1.54	1.28	4.97	2.75	1.63	1.82	1.37	2.01	1.52	2.15	1.06	2.15	2.01	
$T^{(7)}$	$3.99 \cdot 10^{-4}$	$1.22 \cdot 10^{-3}$	$1.81 \cdot 10^{-3}$	$1.48 \cdot 10^{-3}$	$3.35 \cdot 10^{-3}$	$6.54 \cdot 10^{-5}$	$8.69 \cdot 10^{-4}$	$3.1\cdot 10^{-4}$	$8.99 \cdot 10^{-4}$	$1.68\cdot10^{-3}$	$2.17\cdot 10^{-3}$	$1.73 \cdot 10^{-3}$	$2.11 \cdot 10^{-3}$	
c5 or c6	8/9	. 1	2/3	2/3	1/2	10/00	15/17	16/15	1001/1000	57/100	7/8	26/47	19/20	
c4 of c5	4/5	2/3	5/9	6/13	1	17/18	77/94	23/21	128/125	140/297	11/12	1104/2329	1003/1000	
3	3/10	3/10	1/3	3/10	3/8	21/64	23/81	8/23	$\left(15+\sqrt{5}\right)/55$	9/28	$\left(47 + \sqrt{359}\right)/185$	15/46	$(233 + \sqrt{9889}) / 925$	
c2	1/5	1/5	2/9	1/5	1/4	9/40	33/133	21/83	40/59	6/29	7/24	4/19	149/500	
Pair	DP5(4)7M	DP5(4)6M	DP5(4)7S	DP5(4)7C	FE5(4)#2	NEW5(4)F	NEW5(4)	NEW5(4)FP	NEW5(4)P	NEW5(4)FS	NEW5(4)S	NEW5(4)FSR	NEW5(4)SR	
	L													



TABLE 2. Coefficients of NEW5(4)F (exact rationals)

The seven-stage FSAL pairs, however, can attain arbitrarily small values of $||T^{(6)}||_2$, but with the undesirable side effect of some elements of A, b becoming very large. Specifically, in this minimization process both the values of c_4 , c_5 tend to become equal to 1. So, we assigned to both of them some two-digit number rationals close to 1, and we found the values of the other free parameters that gave us the pair which minimized the value of $||T^{(6)}||_2$ and an adequate value of D_{∞} (see the legend to Table 1). The coefficient parameters of the so constructed pair NEW5(4)F are given in Table 2. It is remarkable that the value of $||T^{(6)}||_2$ of this pair is almost an order of magnitude less than the respective value of the pair DP5(4).

The behavior of explicit RK pairs on mildly stiff problems and problems that are mildly stiff oscillatory is mainly dependent on the form of the stability polynomial P, which in turn, for the pairs studied here, is exclusively dependent on the value of $t_6 = bA^4c$. The latter value is the one that determines the phase-lag order of these pairs as well (see [14]).

Using the algorithm of the Appendix and with the assistance of a symbolic algebra manipulation package, we can easily determine

(16)
$$t_6 = \begin{cases} -\frac{c_3 (5c_3 - 2)}{480 (5c_3^2 - 4c_3 + 1)}, & \text{six-stage pairs,} \\ \frac{(5c_3 - 2) c_4}{240}, & \text{seven-stage FSAL pairs.} \end{cases}$$

Equation (16) shows that since for the pairs of the new family there is no dependence of the value of t_6 on that of c_2 , the maximal phase-lag order that the pairs of the new family can attain, is exactly the same as that of those belonging to their respective subfamilies studied in [14]. Specifically, a six-stage pair of phase-lag order 8(4), with globally minimized truncation error coefficients with respect to its fifth-algebraic-order method (NEW5(4)P) may be constructed, using the algorithm of the Appendix and the values of the free parameters contained in Table 1. A sevenstage FSAL pair of phase-lag order 8(4) with similar properties (NEW5(4)FP) is also included in this table. In selecting the values of the parameters of NEW5(4)FP we chose not to increase the phase-lag order of the lower-algebraic-order formula of the pair, but instead tried to utilize the extra degree of freedom in choosing c_4 so as to decrease further the value of $||T^{(6)}||_2$. The numerical experiments in [14] seem to support our decision.

Figure 1 is a three-dimensional plot of the radius R of the boundary of the stability region of the higher-order formula of any six-stage, fifth-order pair as a function of the angle ϑ and the parameter t_6 , when $(\vartheta, t_6) \in \left[\frac{\pi}{2}, \pi\right] \times \left[-0.0125, 0.0083\right]$ (the latter interval is the range of values that assumes t_6 when $(c_3, c_4) \in [0, 1] \times [0, 1]$).



FIGURE 1. Graphical representation of $R(t_6, \vartheta)$

This plot seems to confirm the observations of Lawson [13]. Specifically, two choices of t_6 will attract our interest here. The first is when $t_6 = 1/1296$ (see [5]), the other when $t_6 = 1/1370$. The latter value is the one yielding pairs with almost maximal Real Stability Interval and one that much restricts the undesirable characteristic (already pointed in [13]) of the rough folding of the boundary of the stability region near the real axis. Pairs of this type might be suitable for mildly stiff problems with their large-modulus eigenvalues along the negative real axis. Figure 2 shows the stability region of the pairs based on both of these selections of t_6 . Moreover, as $|P(i\vartheta)| - 1 = \frac{1-720t_6}{360} \vartheta^6$, the first of (16) shows that all the non-FSAL pairs of the new family have a vanishing Imaginary Stability Interval.

Various choices of the free parameters which result in pairs with minimized values of $||T^{(6)}||_2$ (for the selections $t_6 = 1/1296$ and $t_6 = 1/1370$) are also presented in Table 1. We note that, in the case of the pair NEW5(4)SR, further minimization of $||T^{(6)}||_2$ might be possible, but this results in rather big values of D_{∞} . In short, all the pairs of Table 1 (except of NEW5(4)F and NEW5(4)SR) may be considered as corresponding to their respective global minima of $||T^{(6)}||_2$.

We tested the general-purpose FSAL pairs NEW5(4)F and DP5(4) on the whole set of problems of DETEST (see Hull et al. [11]), in Error Per Step Mode, for tolerances ranging from 10^{-3} to 10^{-9} . Figure 3 represents the graphical interpretation of the results obtained, in the form of a semilogarithmic plot of the geometric mean of the maximum global errors achieved over all DETEST problems, against the respective geometric mean of function evaluations, for the range of the above tolerances. Under this interpretation of the numerical results, each problem of DETEST contributes exactly the same as any other in the overall picture.

In this figure we also included the pair FE5(4). In this way we find that NEW5(4) is better than DP5(4), on this set of test problems, by as much as DP5(4) is better



FIGURE 2. Stability regions of six-stage, fifth-order RK pairs when $t_6 = \frac{1}{1296}$ — and $t_6 = \frac{1}{1370}$ - - -



FIGURE 3. Efficiency curves for NEW5(4)F —, DP5(4) - - - and FE5(4) · - ·

than FE5(4). This outcome remains virtually unchanged even if we consider endpoint errors. We may further verify this global picture by inspecting the respective figures for each problem separately, each one of which proved to be of a form similar to that provided in Figure 3. Under a different type of interpretation of the numerical results (as suggested by Enright and Pryce [6], in the tabular format as standardized in [15]) NEW5(4) is in total 15.8% more efficient than DP5(4) (which in turn is more efficient than FE5(4) by about the same amount—17.7%).

APPENDIX. Algorithm for the construction of Runge-Kutta pairs of orders 5(4)

Distinguish two cases:

- (a) $\hat{b}_7 \neq 0$ (FSAL). Choose $c_2 \neq 0$ and c_3 , c_4 , c_5 arbitrary but satisfying Condition (C). Set $c_6 = 1$.
- (b) $\hat{b}_7 = 0$. Choose $c_2 \neq 0$ and c_3 , c_5 , c_6 arbitrary but satisfying Condition (C) (if $c_6 = 1$ choose $c_3 \neq 1/2$). Set $c_4 = \frac{c_3}{2(5c_3^2 4c_3 + 1)}$.

$$\begin{split} a_{64} &= -\frac{60a_{54}b_5c_4\left(c_4-c_5\right)\left(c_3-c_4\right)+5c_3\left(2c_5-1\right)-5c_5+3}{60b_6c_4\left(c_3-c_4\right)\left(c_4-c_5\right)},\\ a_{54} &= -\frac{5c_3\left(4c_6-3\right)-2\left(5c_6-4\right)}{120b_5c_4\left(c_3-c_4\right)\left(c_5-c_6\right)},\\ a_{43} &= -\frac{2a_{42}c_2-c_4^2}{2c_3},\\ a_{53} &= -\frac{2a_{52}c_2+2a_{54}c_4-c_5^2}{2c_3},\\ a_{63} &= -\frac{2a_{62}c_2+2a_{64}c_4+2a_{65}c_5-c_6^2}{2c_3},\\ b_1 &= 1-b_3-b_4-b_5-b_6, \quad \hat{b}_1 &= 1-\hat{b}_3-\hat{b}_4-\hat{b}_5-\hat{b}_6-\hat{b}_7,\\ a_{21} &= c_2, \quad a_{31} &= c_3-a_{32}, \quad a_{41} &= c_4-a_{42}-a_{43},\\ a_{51} &= c_5-a_{52}-a_{53}-a_{54}, \quad a_{61} &= c_6-a_{62}-a_{63}-a_{64}-a_{65}.\\ \text{In case (a) set} \\ a_{71} &= b_1, \ a_{72} &= 0, \ a_{73} &= b_3, \ a_{74} &= b_4, \ a_{75} &= b_5, \ a_{76} &= b_6. \end{split}$$

Acknowledgments

The authors thank Dr. Ch. Tsitouras for his contribution in compiling the interpretation of the numerical tests.

References

- J. C. Butcher, On Runge-Kutta processes of high order, J. Austral. Math. Soc. 4 (1964), 179–194. MR 29:2972
- _____, The numerical analysis of ordinary differential equations, John Wiley and Sons, Chichester, 1987. MR 88d:65002
- C. R. Cassity, The complete solution of the fifth order Runge-Kutta equations, SIAM J. Numer. Anal. 6 (1969), 432–436. MR 40:8268
- A. R. Curtis, High order explicit Runge-Kutta formulae, their uses and their limitations, J. Inst. Math. Appl. 16 (1975), 35-55. MR 52:4630
- J. R. Dormand and P. J. Prince, A family of embedded Runge-Kutta formulae, J. Comput. Appl. Math. 6 (1980), 19-26. MR 81g:65098
- W. H. Enright and J. D. Pryce, Two FORTRAN packages for assessing initial value methods, ACM Trans. Math. Software 13 (1987), 1–27.
- E. Fehlberg, Classical fifth, sixth, seventh, and eighth order Runge-Kutta formulas with stepsize control, TR R-287, NASA, 1968.
- Low order classical Runge-Kutta formulas with stepsize control and their application to some heat-transfer problems, TR R-315, NASA, 1969.
- 9. E. Hairer, S. P. Nørsett, and G. Wanner, Solving ordinary differential equations I, first ed., Springer, Berlin, 1987. MR 87m:65005
- P. J. van der Houwen and B. P. Sommeijer, Explicit Runge-Kutta(-Nyström) methods with reduced phase errors for computing oscillating solutions, SIAM J. Numer. Anal. 24 (1987), 595-617. MR 88e:65088
- T. E. Hull, W. H. Enright, B. M. Fellen, and A. E. Sedgwick, Comparing numerical methods for ordinary differential equations, SIAM J. Numer. Anal. 9 (1972), 603–637. MR 50:3577
- J. D. Lawson, An order five Runge-Kutta process with extended region of stability, SIAM J. Numer. Anal. 3 (1966), 593-597. MR 35:7589
- G. Papageorgiou, Ch. Tsitouras, and S. N. Papakostas, Runge-Kutta pairs for periodic initial value problems, Rep. NA 93-1, Nat. Tech. Univ. Athens, Dept. Math., 1993.

- S. N. Papakostas, Ch. Tsitouras, and G. Papageorgiou, A general family of explicit Runge-Kutta pairs of orders 6(5), SIAM J. Numer. Anal. 33 (1996).
- L. F. Shampine, Local extrapolation in the solution of ordinary differential equations, Math. Comp. 27 (1973), 91–97. MR 48:10135
- B. E. Shanks, Solutions of differential equations by evaluations of functions, Math. Comp. 20 (1966), 21–38. MR 32:4858

Department of Mathematics, Zografou Campus, National Technical University of Athens, Athens 157 $80,\,\mathrm{Greece}$

E-mail address: spapakos@theseas.ntua.gr E-mail address: papag@nisyros.ntua.gr