A family of embedded Runge-Kutta formulae

J. R. Dormand and P. J. Prince (*)

ABSTRACT

A family of embedded Runge-Kutta formulae RK5(4) are derived. From these are presented formulae which have (a) 'small' principal truncation terms in the fifth order and (b) extended regions of absolute stability.

1. INTRODUCTION

Consider the problem of solving numerically the first order system of ordinary differential equations

$$\underline{y}'(x) = \underline{f}[x, \underline{y}(x)], \tag{1.1}$$

with $\underline{y}(x_0)$ known.

Without loss of generality [1] the first order autonomous system

$$\underline{y}'(x) = \underline{f}[\underline{y}(x)], \tag{1.2}$$

with $\underline{y}(x_0)$ known,

can be considered. Under suitable continuity and differentiability conditions approximations $\underline{y}_n$ to the true solution $\underline{y}(x_n)$ at the points x_n, where

$x_{n+1} = x_n + h_n$, $h_n = \theta(x_n)h$ and $0 < \theta(x_n) \leqslant 1$,

$n = 0, 1, 2, \ldots$, can be obtained using an explicit Runge-Kutta (RK) formula given by

$$\underline{y}_{n+1} = \underline{y}_n + h_n \underline{\Phi}(\underline{y}_n, h_n) = \underline{y}_n + \sum_{i=1}^{s} b_i \underline{k}_i \tag{1.3}$$

where

$$\underline{k}_1 = h_n \underline{f}(\underline{y}_n),$$

$$\underline{k}_i = h_n \underline{f}(\underline{y}_n + \sum_{j=1}^{i-1} a_{ij} \underline{k}_j), \quad i = 2, 3, \ldots, s,$$

and usually $\underline{y}_0 = \underline{y}(x_0)$.

The local truncation error $\underline{t}_{n+1}$, of this method at x_{n+1} is given by

$$\underline{t}_{n+1} = \underline{y}(x_n) + h_n \underline{\Phi}[\underline{y}(x_n), h_n] - \underline{y}(x_{n+1})$$

which using the Taylor expansion about x_n may be written

$$\underline{t}_{n+1} = h_n \{\underline{\Phi}[\underline{y}(x_n), h_n] - \underline{\Delta}[\underline{y}(x_n), h_n]\},$$

where

$$\underline{\Delta}[\underline{y}(x), h] = \sum_{r=1}^{\infty} \frac{h^{r-1}}{r!} \underline{y}^{(r)}(x).$$

If $\underline{\phi}$ and $\underline{\Delta}$ agree to $0(h^p)$ then the process is said to be a pth order RK formula (RKp) and $\underline{t}_{n+1}$ can then be written

$$\underline{t}_{n+1} = \sum_{j=1}^{\infty} h_n^{p+j} \underline{\phi}_{p+j-1}[\underline{y}(x_n)], \tag{1.4}$$

where

$$\underline{\phi}_r[\underline{y}(x)] = \sum_{i=1}^{n_{r+1}} a_i^{(r+1)} \underline{F}_i^{(r+1)}[\underline{y}(x)], \quad r = 1, 2, \ldots,$$

are termed error functions, $\underline{F}_i^{(r+1)}$, $i = 1, 2, \ldots, n_{r+1}$, being the elementary differentials [1] of order $r+1$ of $\underline{f}$. Note that if the formula is of order p, then $\underline{\phi}_r \equiv 0$, $r = 1, 2, \ldots, p-1$. This implies

$$a_i^{(r+1)} = 0, \quad i = 1, 2, \ldots, n_{r+1}, \quad r = 1, 2, \ldots, p-1. \tag{1.5}$$

For consistency (Lambert [2]) the following equation must be satisfied :

$$a_1^{(1)} = \sum_{i=1}^{s} b_i - 1 = 0, \quad (n_1 = 1).$$

This equation together with (1.5) are termed equations of condition for the RKp formula. Butcher [1] has listed the expressions for the elementary differentials for up to order 8 and Harris [3] has considered the computer derivation of the equations of condition. Table 1 contains the required equations for orders up to 6.

It is now widely accepted that the Runge-Kutta embedding technique is an efficient method for the numerical solution of non-stiff initial value problems. In this technique two RK formulae of orders p and q ($q > p$), usually $q = p + 1$, are obtained which share the same function evaluations, i.e. they have the same a_{ij}. It is common practice that the higher order formula uses more stages than the lower order formula but in this work we shall also allow the converse to be true. From

(*) J. R. Dormand, P. J. Prince, Department of Mathematics and Statistics, Teesside Polytechnic, Middlesbrough, Cleveland (UK).

Table 1. Equations of condition for orders 1 to 6

Order	r	1	2	3	4	5	6	7
No of elementary differentials	n_r	1	1	2	4	9	20	48

Notes

(i) $c_i = \sum_{j=1}^{s} a_{ij}$, $i = 1, 2, \ldots, s$ where s is number of stages.

(ii) $a_{ij} = 0$, $j \geqslant i$ for an explicit RK.

(iii) All subscripts run from 1 to s.

---

1 $a_1^{(1)} = \sum_i b_i - 1$

---

2 $a_1^{(2)} = \sum_i b_i c_i - \frac{1}{2}$

---

3 $a_1^{(3)} = \frac{1}{2} \sum_i b_i c_i^2 - \frac{1}{6}$

4 $a_2^{(3)} = \sum_{ij} b_i a_{ij} c_j - \frac{1}{6}$

---

5 $a_1^{(4)} = \frac{1}{6} \sum_i b_i c_i^3 - \frac{1}{24}$

6 $a_2^{(4)} = \sum_{ij} b_i c_i a_{ij} c_j - \frac{1}{8}$

7 $a_3^{(4)} = \frac{1}{2} \sum_{ij} b_i a_{ij} c_j^2 - \frac{1}{24}$

8 $a_4^{(4)} = \sum_{ijk} b_i a_{ij} a_{jk} c_k - \frac{1}{24}$

---

9 $a_1^{(5)} = \frac{1}{24} \sum_i b_i c_i^4 - \frac{1}{120}$

10 $a_2^{(5)} = \frac{1}{2} \sum_{ij} b_i c_i^2 a_{ij} c_j - \frac{1}{20}$

11 $a_3^{(5)} = \frac{1}{2} \sum_{ijk} b_i a_{ij} c_j a_{ik} c_k - \frac{1}{40}$

12 $a_4^{(5)} = \frac{1}{2} \sum_{ij} b_i c_i a_{ij} c_j^2 - \frac{1}{30}$

13 $a_5^{(5)} = \frac{1}{6} \sum_{ij} b_i a_{ij} c_j^3 - \frac{1}{120}$

14 $a_6^{(5)} = \sum_{ijk} b_i c_i a_{ij} a_{jk} c_k - \frac{1}{30}$

15 $a_7^{(5)} = \sum_{ijk} b_i a_{ij} c_j a_{jk} c_k - \frac{1}{40}$

16 $a_8^{(5)} = \frac{1}{2} \sum_{ijk} b_i a_{ij} a_{jk} c_k^2 - \frac{1}{120}$

17 $a_9^{(5)} = \sum_{ijkm} b_i a_{ij} a_{jk} a_{km} c_m - \frac{1}{120}$

---

18 $a_1^{(6)} = \frac{1}{120} \sum_i b_i c_i^5 - \frac{1}{720}$

19 $a_2^{(6)} = \frac{1}{6} \sum_{ij} b_i c_i^3 a_{ij} c_j - \frac{1}{72}$

20 $a_3^{(6)} = \frac{1}{2} \sum_{ijk} b_i c_i a_{ij} c_j a_{ik} c_k - \frac{1}{48}$

21 $a_4^{(6)} = \frac{1}{4} \sum_{ij} b_i c_i^2 a_{ij} c_j^2 - \frac{1}{72}$

22 $a_5^{(6)} = \frac{1}{2} \sum_{ijk} b_i a_{ij} c_j^2 a_{ik} c_k - \frac{1}{72}$

23 $a_6^{(6)} = \frac{1}{6} \sum_{ij} b_i c_i a_{ij} c_j^3 - \frac{1}{144}$

24 $a_7^{(6)} = \frac{1}{24} \sum_{ij} b_i a_{ij} c_j^4 - \frac{1}{720}$

25 $a_8^{(6)} = \frac{1}{2} \sum_{ijk} b_i c_i^2 a_{ij} a_{jk} c_k - \frac{1}{72}$

26 $a_9^{(6)} = \sum_{ijkm} b_i a_{ij} a_{ik} c_k a_{jm} c_m - \frac{1}{72}$

27 $a_{10}^{(6)} = \sum_{ijk} b_i c_i a_{ij} c_j a_{jk} c_k - \frac{1}{48}$

28 $a_{11}^{(6)} = \frac{1}{2} \sum_{ijk} b_i a_{ij} c_j^2 a_{jk} c_k - \frac{1}{120}$

29 $a_{12}^{(6)} = \frac{1}{2} \sum_{ijkm} b_i a_{ij} a_{jk} c_k a_{jm} c_m - \frac{1}{240}$

30 $a_{13}^{(6)} = \frac{1}{2} \sum_{ijk} b_i c_i a_{ij} a_{jk} c_k^2 - \frac{1}{144}$

31 $a_{14}^{(6)} = \frac{1}{2} \sum_{ijk} b_i a_{ij} c_j a_{jk} c_k^2 - \frac{1}{180}$

32 $a_{15}^{(6)} = \frac{1}{6} \sum_{ijk} b_i a_{ij} a_{jk} c_k^3 - \frac{1}{720}$

33 $a_{16}^{(6)} = \sum_{ijkm} b_i c_i a_{ij} a_{jk} a_{km} c_m - \frac{1}{144}$

34 $a_{17}^{(6)} = \sum_{ijkm} b_i a_{ij} c_j a_{jk} a_{km} c_m - \frac{1}{180}$

35 $a_{18}^{(6)} = \sum_{ijkm} b_i a_{ij} a_{jk} c_k a_{km} c_m - \frac{1}{240}$

36 $a_{19}^{(6)} = \frac{1}{2} \sum_{ijkm} b_i a_{ij} a_{jk} a_{km} c_m^2 - \frac{1}{720}$

37 $a_{20}^{(6)} = \sum_{ijkmn} b_i a_{ij} a_{jk} a_{km} a_{mn} c_n - \frac{1}{720}$

---

the embedding an estimate $\underline{E}_{n+1} = \underline{y}_{n+1} - \hat{\underline{y}}_{n+1}$ (see [5]), of $\underline{t}_{n+1}$, the local truncation error in the pth order formula can be obtained. This can be used to monitor the local error and hence control the step size. For example the formulae

$$h_{n+1} = 0.9\, h_n \left[\frac{\delta}{\| \underline{E}_{n+1} \|_\infty} \right]^{1/p+1}, \quad \text{(error per step control)}$$

and

$$h_{n+1} = 0.9 h_n \left[\frac{\delta}{\left\| \frac{\underline{E}_{n+1}}{h_n} \right\|_\infty} \right]^{1/p}, \quad \text{(error per unit step control)}$$

[4] are widely used, δ being the maximum allowable local error. Applied in local extrapolation (higher order) mode {RKq(p) instead of RKp(q) mode}, which practical results indicate is preferable, ([5], [6]) the embedded algorithm takes the form

$$\hat{\underline{y}}_{n+1} = \underline{y}_n + h_n \hat{\underline{\Phi}}(\underline{y}_n, h_n) = \underline{y}_n + \sum_{i=1}^{s} \hat{b}_i \underline{k}_i$$

and

$$\underline{y}_{n+1} = \hat{\underline{y}}_n + h_n \underline{\Phi}(\hat{\underline{y}}_n, h_n) = \hat{\underline{y}}_n + \sum_{i=1}^{s} b_i \underline{k}_i, \qquad (1.6)$$

where

$$\underline{k}_1 = h_n \underline{f}(\hat{\underline{y}}_n)$$

and

$$\underline{k}_i = h_n \underline{f}\left(\hat{\underline{y}}_n + \sum_{j=1}^{i-1} a_{ij} \underline{k}_j\right), \qquad i = 2, 3, \ldots, s.$$

In this case s is the number of stages (vector function evaluations per step) used by the combined process. To distinguish between the two formulae caps are used to indicate the qth order formula. Fehlberg ([7], [8]) has developed embedded RK formulae which have a 'small' principal truncation term in the lower order formula. Since this term gives the local error estimate an artificially small term could lead to poor step size control (see [11]). The aim of this paper is to develop RK5(4) formulae which (a) have a 'small' principal truncation term in the fifth order (this will be elaborated further in section 3 below) and (b) have an extended region of absolute stability.

2. A CLASS OF RK5(4) EMBEDDED FORMULAE

Reference to table 1 shows that there are 25 equations to be satisfied for an RK5(4). Here the possibility of using seven stages is considered, where the first evaluation at the nth step is the same as the last evaluation at the previous step ([9], [11]). This implies

$$\hat{b}_7 = 0$$
$$c_7 = 1 \qquad (2.1)$$
and
$$a_{7j} = \hat{b}_j, \qquad j = 1, 2, \ldots, 6.$$

After the first step the effective number of function evaluations per step is still six although one extra function evaluation is lost after any step rejection. In the following, the 17 governing equations for the 5th order formula will be referred to as equations $\hat{①}, \hat{②}, \ldots, \hat{⑰}$ where all subscripts run from 1 to 6, and the eight governing equations for the 4th order will be referred to as equations ①, ②, ..., ⑧, where all subscripts run from 1 to 7.

A solution of the 25 equations is considered by imposition of the following equations :

$$\sum_{i=1}^{6} \hat{b}_i a_{ij} = \hat{b}_j (1 - c_j), \quad j = 1, 2, \ldots, 6, \qquad (2.2)$$

$$\hat{b}_2 = b_2 = 0,$$

$$\sum_{j=1}^{6} a_{ij} c_j = \frac{1}{2} c_i^2, \quad \sum_{j=1}^{6} a_{ij} c_i^2 = \frac{1}{3} c_i^3,$$

$$i = 3, 4, 5, 6, \qquad (2.3)$$

three of which are dependent [10]. This leaves the following independent equations from the original 25 : $\hat{①}, \hat{②}, \hat{③}, \hat{⑤}, \hat{⑤}, \hat{⑭}$, ①, ②, ③, ⑤ and ⑧. Equations $\hat{⑭}$ and ⑧ can further be simplified to give

$$\sum_{i=1}^{6} \hat{b}_i c_i a_{i2} = 0, \qquad (2.4)$$

and

$$\sum_{i=1}^{6} b_i a_{i2} = 0 \qquad (2.5)$$

respectively. $\hat{b}_1$ and b_1 will be determined from $\hat{①}$ and ① respectively since these are the only equations in which they occur. From (2.2) $c_6 = 1$ and therefore $\hat{②}$, $\hat{③}$, $\hat{⑤}$ and $\hat{⑨}$ yield the following expressions for $\hat{b}_3, \hat{b}_4, \hat{b}_5$ and $\hat{b}_6$:

$$\hat{b}_3 = \frac{3 - 5(c_4 + c_5) + 10 c_4 c_5}{60 c_3 (1 - c_3)(c_3 - c_4)(c_3 - c_5)},$$

$$\hat{b}_4 = \frac{3 - 5(c_3 + c_5) + 10 c_3 c_5}{60 c_4 (1 - c_4)(c_4 - c_3)(c_4 - c_5)},$$

$$\hat{b}_5 = \frac{3 - 5(c_3 + c_4) + 10 c_3 c_4}{60 c_5 (1 - c_5)(c_5 - c_3)(c_5 - c_4)},$$

$$\hat{b}_6 = \frac{12 - 15(c_3 + c_4 + c_5) + 20(c_3 c_4 + c_4 c_5 + c_5 c_3) - 30 c_3 c_4 c_5}{60 (1 - c_3)(1 - c_4)(1 - c_5)} \qquad (2.6)$$

With i = 3 and 4 (2.3) gives

$$c_2 = \frac{2}{3} c_3, \quad a_{32} = \frac{3}{4} c_3,$$

$$a_{42} = \frac{3 c_4^2 (3 c_3 - 2 c_4)}{4 c_3^2}, \quad a_{43} = \frac{c_4^2 (c_4 - c_3)}{c_3^2},$$

and (2.2) with $j = 2$ and (2.4) gives

$$\sum_{i=3}^{5} \hat{b}_i a_{i2} (1 - c_i) = 0,$$

which yields a_{52}. Hence a_{62} follows from (2.4). Equations (2.3) with $i = 5$ may be solved for a_{53} and a_{54}, equation (2.2) with $j = 5$ gives a_{65}, allowing a_{63} and a_{64} to be determined from (2.3) with $i = 6$. The remaining three equations from (2.2) and (2.3) are dependent and therefore satisfied. The equations ②, ③, ⑤ and (2.5) are linear in the b's and can be used to determine b_3, b_4, b_5 and b_6 given a value for b_7 which is arbitrary.

3. CHOICE OF THE DEGREES OF FREEDOM

In this model of an RK5(4) there are four degrees of freedom : c_3, c_4, c_5 and b_7. The value of b_7 will be chosen to tune the embedded process [11] so that a reasonable error prediction, and hence step size control, may be attained. Since the local extrapolation mode is being adopted one way of choosing the three c's would be to make the principal truncation term $(h_n^{p+1} \hat{\underline{\phi}}_p)$ 'small' for the 5th order formula (see [12]). For general systems of equations this is of course a prohibitive task. Practical tests [8], however, indicate that, in spite of our comments in section 1, good results for the lower order mode are obtained by attempting to make the a_i as small as possible.
In this paper the parameters are chosen to give a small $\| \hat{\underline{a}}^{(6)} \|_2$. For the present model with the assumptions made it is found that (see table 1)

$$\hat{a}_1^{(6)} = \frac{1}{10} \hat{a}_2^{(6)} = \frac{1}{15} \hat{a}_3^{(6)} = \frac{1}{10} \hat{a}_4^{(6)} = \frac{1}{10} \hat{a}_5^{(6)}$$

$$= -\frac{1}{5} \hat{a}_7^{(6)} = -\frac{1}{30} \hat{a}_{11}^{(6)} = -\frac{1}{15} \hat{a}_{12}^{(6)} = -\frac{1}{20} \hat{a}_{14}^{(6)},$$

$$\hat{a}_6^{(6)} = \frac{1}{3} \hat{a}_{10}^{(6)} = \hat{a}_{13}^{(6)} = -\hat{a}_{15}^{(6)} = -\frac{1}{3} \hat{a}_{18}^{(6)} = -\hat{a}_{19}^{(6)},$$

$$\hat{a}_8^{(6)} = \hat{a}_9^{(6)} = -\frac{1}{2} \hat{a}_{17}^{(6)}, \text{ and } \hat{a}_{20}^{(6)} = -\hat{a}_{16}^{(6)}, \text{ where}$$

$$\hat{a}_1^{(6)} = \frac{\{2 - 3(c_3 + c_4) + 5c_3c_4\} - c_5 \{3 - 5(c_3 + c_4) + 10c_3c_4\}}{7200}$$

$$\hat{a}_6^{(6)} = \frac{(1 - 2c_3) - c_4 (2 - 5c_3)}{720},$$

$$\hat{a}_8^{(6)} = \frac{\{2 - 3c_4 (3 - 5c_3)\} - 3c_5 \{1 - 5c_4 (1 - 2c_3)\}}{720}$$

and

$$\hat{a}_{20}^{(6)} = \frac{3c_4 (2 - 5c_3) - 1}{720}. \qquad (3.1)$$

In addition we impose the constraints that $|c_i| \leq 1$ and that the c_i are reasonably 'distinct', thus avoiding large values of b_i (see 2.6) and a_{ij} which could cause considerable rounding errors in practical applications. Following Lawson [13] another choice of the parameters aims to extend the region of absolute stability. In this case where there are two RK formulae the region of absolute stability should be enlarged for both formulae. For the 5th order formula applied to $y' = \lambda y$ (λ complex) the region of absolute stability is the region in the left hand part of the complex plane where

$$\left| \sum_{r=0}^{5} \frac{z^r}{r!} + dz^6 \right| < 1, \quad (z = x + iy)$$

where

$$d = \hat{a}_{20}^{(6)} + \frac{1}{6!} = \frac{c_4 (2 - 5c_3)}{240} \qquad (3.2)$$

from (3.1).
The corresponding region for the 4th order formula is defined by

$$\left| \sum_{r=0}^{4} \frac{z^r}{r!} + ez^5 + fz^6 + gz^7 \right| < 1, \qquad (3.3)$$

where

$$e = a_9^{(5)} + \frac{1}{5!}, \quad f = a_{20}^{(6)} + \frac{1}{6!} \quad \text{and} \quad g = a_{48}^{(7)} + \frac{1}{7!}.$$

Other choices for the degrees of freedom are possible such as those which make $\hat{a}_1^{(6)}$ or $\hat{a}_{20}^{(6)}$ zero. In the former case the higher order formula is then 6th order for equations of the type $\underline{y}' = \underline{f}(x)$, (i.e. quadrature problems) and in the latter case the higher order formula is 6th order for equations of the form $\underline{y}' = A\underline{y} + \underline{b}x$ (A a constant matrix and $\underline{b}$ a constant vector). Concerning the quadrature problem, it has been noted previously [14] that many embedded formulae of high order, such as those of Fehlberg [7] with $p \geq 5$, and Dormand & Prince [11] fail because, in this case, the two formulae yield the same numerical approximation for $\underline{y}(x_{n+1})$.
Consequently the local error estimate is zero, preventing step-size control. Shampine [15] has developed a modification to the RKF7 formula of Fehlberg [7] which overcomes this difficulty but it is preferable to develop formulae which are free from this deficiency. The formulae presented in this paper are effective in the quadrature (or near quadrature) case and it is intended to present higher order formulae with the same property in a future paper.
The choice of the parameter b_7 leads to two cases :
(a) $b_7 \neq 0$.
As mentioned previously b_7 is now the 'tuning' parameter. In this case the b_i will be distinct from the $\hat{b}_i$ so that there are differing formulae. The choices $c_3 = 3/10$, $c_4 = 4/5$, $c_5 = 8/9$ and $b_7 = 1/40$ lead to the formula RK5(4)7M presented in table 2.

Table 2. Coefficients for RK5(4)7M

c_i	a_{ij}						$\hat{b}_i$	b_i
0							$\frac{35}{384}$	$\frac{5179}{57600}$
$\frac{1}{5}$	$\frac{1}{5}$						0	0
$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$					$\frac{500}{1113}$	$\frac{7571}{16695}$
$\frac{4}{5}$	$\frac{44}{45}$	$-\frac{56}{15}$	$\frac{32}{9}$				$\frac{125}{192}$	$\frac{393}{640}$
$\frac{8}{9}$	$\frac{19372}{6561}$	$-\frac{25360}{2187}$	$\frac{64448}{6561}$	$-\frac{212}{729}$			$-\frac{2187}{6784}$	$-\frac{92097}{339200}$
1	$\frac{9017}{3168}$	$-\frac{355}{33}$	$\frac{46732}{5247}$	$\frac{49}{176}$	$-\frac{5103}{18656}$		$\frac{11}{84}$	$\frac{187}{2100}$
1	$\frac{35}{384}$	0	$\frac{500}{1113}$	$\frac{125}{192}$	$-\frac{2187}{6784}$	$\frac{11}{84}$	0	$\frac{1}{40}$

For this formula $\| \hat{\underline{a}}^{(6)} \|_2 = 3.99 \times 10^{-4}$, $d = \frac{1}{600}$, $e = \frac{1097}{120000}$, $f = \frac{161}{120000}$ and $g = \frac{1}{24000}$ which give the regions of absolute stability sketched in figure 1. For RKF4 [8], which practical results ([5], [6]) indicate is preferable in local extrapolation mode, $\| \hat{\underline{a}}^{(6)} \|_2 = 3.36 \times 10^{-3}$ which is about 8 times larger than that for the RK5(4)7M. The stability regions for RKF4 are also sketched in figure 1.

Lawson [12] has shown that an enlarged stability region is obtained if $d = \frac{1}{1280}$ giving $c_4 = \frac{3}{16(2-5c_3)}$ (3.2). The remaining parameters, c_3, c_5 and b_7, have then been selected to give small $\| \hat{\underline{a}}^{(6)} \|_2$ and extended stability region for the 4th order. It can be argued [16] that high stability for the lower order formula is unnecessary when local extrapolation is used. However, an absolute stability failure on the lower order for-

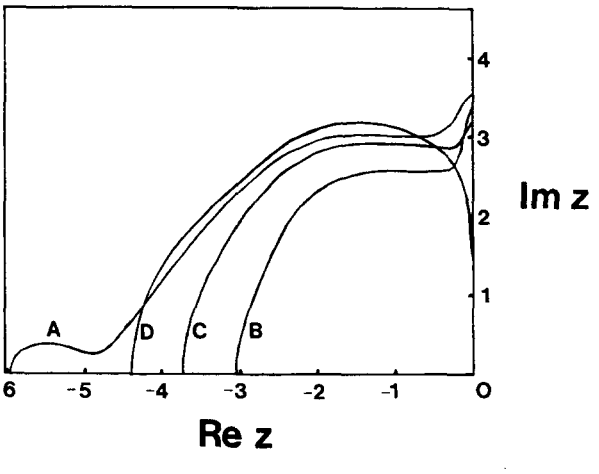

Fig. 1(b). Stability regions for fourth order formulae.

A: RK5(4)7S, B: RKF4, C: RK5(4)6M, D: RK5(4)7M.

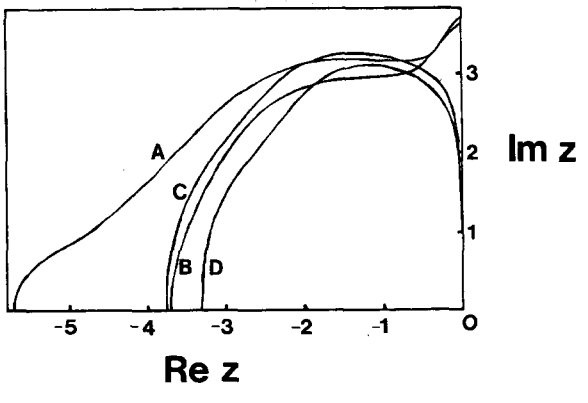

Fig. 1(a). Stability regions for fifth order formulae.

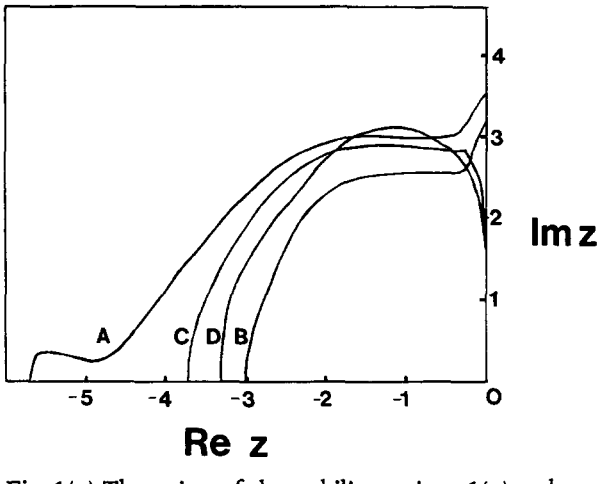

Fig. 1(c) The union of the stability regions 1(a) and 1(b) for each formulae pair. This may be interpreted as the stability regions for composite pairs.

mula could cause severe problems in step size control, resulting in a loss of computational efficiency. Consequently it is preferable to make the stability region of the lower order formula comparable to that of the higher order formula. With $c_3 = 1/3$ and $c_5 = 2/3$ $\|\underline{\hat{a}}^{(6)}\|_2 = 1.81 \times 10^{-3}$ and the best choice of b_7 to give comparable stability regions is 1/150. However this results in a poorly 'tuned' formulae pair and a compromised choice is $b_7 = -1/50$. The formula in this case is termed the RK5(4) 7S and is given in table 3. Figure 1 gives the stability regions.

Table 3. Coefficients for RK5(4)7S

c_i	a_{ij}						$\hat{b}_i$	b_i
0							$\frac{19}{200}$	$\frac{431}{5000}$
$\frac{2}{9}$	$\frac{2}{9}$						0	0
$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{4}$					$\frac{3}{5}$	$\frac{333}{500}$
$\frac{5}{9}$	$\frac{55}{324}$	$-\frac{25}{108}$	$\frac{50}{81}$				$-\frac{243}{400}$	$-\frac{7857}{10000}$
$\frac{2}{3}$	$\frac{83}{330}$	$-\frac{13}{22}$	$\frac{61}{66}$	$\frac{9}{110}$			$\frac{33}{40}$	$\frac{957}{1000}$
1	$-\frac{19}{28}$	$\frac{9}{4}$	$\frac{1}{7}$	$-\frac{27}{7}$	$\frac{22}{7}$		$\frac{7}{80}$	$\frac{193}{2000}$
1	$\frac{19}{200}$	0	$\frac{3}{5}$	$-\frac{243}{400}$	$\frac{33}{40}$	$\frac{7}{80}$	0	$-\frac{1}{50}$

(b) $b_7 = 0$.

In this case the seventh evaluation is not required, and it is found that unless the matrix

$$\begin{bmatrix} c_3 & c_4 & c_5 & 1 \\ c_3^2 & c_4^2 & c_5^2 & 1 \\ c_3^3 & c_4^3 & c_5^3 & 1 \\ a_{32} & a_{42} & a_{52} & a_{62} \end{bmatrix}$$

is singular then $b_i = \hat{b}_i$, $i = 1, 2, \ldots, 6$, i.e. the two RK formulae are identical. The previous matrix is singular [17] if

$$c_4 = \frac{c_3}{2(5c_3^2 - 4c_3 + 1)}, \qquad (3.3)$$

thus b_3, b_4 and b_5 are determined from ②, ③ and ⑤ given a value of the 'tuning' parameter b_6, which must not be chosen equal to $\hat{b}_6$.

Following (a) c_3 and c_5 are chosen to give 'small' $\|\underline{\hat{a}}^{(6)}\|_2$, and $c_3 = 3/10$, $c_5 = 2/3$ and $b_6 = 1/20$ give $\|\underline{\hat{a}}^{(6)}\|_2 = 1.23 \times 10^{-3}$ leading to the formula RK5(4) 6M (table 4, stability regions figure 1). This formula is sixth order for quadrature problems.

Table 4. Coefficients for RK5(4)6M

c_i	a_{ij}						$\hat{b}_i$	b_i
0							$\frac{19}{216}$	$\frac{31}{540}$
$\frac{1}{5}$	$\frac{1}{5}$						0	0
$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$					$\frac{1000}{2079}$	$\frac{190}{297}$
$\frac{3}{5}$	$\frac{3}{10}$	$-\frac{9}{10}$	$\frac{6}{5}$				$-\frac{125}{216}$	$-\frac{145}{108}$
$\frac{2}{3}$	$\frac{226}{729}$	$-\frac{25}{27}$	$\frac{880}{729}$	$\frac{55}{729}$			$\frac{81}{88}$	$\frac{351}{220}$
1	$-\frac{181}{270}$	$\frac{5}{2}$	$-\frac{266}{297}$	$-\frac{91}{27}$	$\frac{189}{55}$		$\frac{5}{56}$	$\frac{1}{20}$

The second choice of extending the region of absolute stability leads to unacceptable formulae. Choosing $d = \frac{1}{1275}$, (3.2) and (3.3) yield the two possible pairs of values (i) $(c_3, c_4) = (2/13, 13/85)$ and (ii) $(c_3, c_4) = (16/45, 72/85)$. Both of these are unsuitable because they result in 'large' a_{ij}, $\hat{b}_i$ and b_i which are unsatisfactory with regard to rounding error. Comparing cases (a) and (b) it can be seen that use of the seventh evaluation yields an extra degree of freedom which can be used to give smaller $\|\underline{\hat{a}}^{(6)}\|_2$ or an extended region of absolute stability. These advantages are offset by the loss of an extra evaluation following a rejected step.

4. NUMERICAL RESULTS

The algorithms described above have been tested on a wide range of problems including those given by Hull et al. [3] in the DETEST implementation (see table 5). According to the criteria laid down by Hull et al. the two "minimum truncation error" formulae RK5(4)7M, RK5(4)6M are more efficient than the RKF4 (Fehlberg [8]). A feature which is not considered by Hull et al., but which the authors believe to be of some importance, is the measurement of global error, and a modified version of DETEST used in connection with this work computes the maximum global error (over all steps and variables). Rather than a complete presentation of results over the 25 problems in DETEST the efficiency curves for two problems, viz A3 and D5, are shown here.

Table 5. DETEST results taken over seven tolerances 10^{-3-r}, r = 0, 1,..., 6

Formula	FCN calls	No of steps	Max Error	Fraction deceived	Fraction bad deceived
RKF4 (in 5(4) mode)	226208	36363	12.4	0.005	0.000
RK5(4)7M	209305	33126	9.7	0.008	0.000
RK5(4)6M	267324	43394	1.2	0.000	0.000
RK5(4)7S	186229	29158	15.2	0.027	0.001

(a) Problem A3 :

$y' = y \cos x$, $y(0) = 1$, $x \in [0, 20]$.

Four efficiency curves for this problem are shown in figure 2. It is clear that the RK5(4)7M is much superior to the other formulae in this case, 800 function evaluations being sufficient for a maximum global error (absolute) of 10^{-6}, compared with 1450 for the RKF4 (in local extrapolation mode).

(b) Problem D5 : (Two-body gravitational problem)

$$y_1' = y_3, \qquad y_1(0) = 1 - e,$$

$$y_2' = y_4, \qquad y_2(0) = 0,$$

$$y_3' = -y_1/(y_1^2 + y_2^2)^{3/2}, \qquad y_3(0) = 0,$$

$$y_4' = -y_2/(y_1^2 + y_2^2)^{3/2}, \qquad y_4(0) = \sqrt{\frac{1+e}{1-e}},$$

e = 0.9 (eccentricity of orbit), $x \in [0, 20]$.

This problem represents a severe test for the step-size control procedure since the step length must vary by about two orders of magnitude. Figure 3 shows the efficiency curve for this problem and again it is apparent that the RK5(4)7M is most efficient.

It should be emphasized that these tests on the four formulae have been conducted under identical circumstances. The number of function evaluations is inclusive of rejected steps and so represents a machine independent measure of the relative efficiencies of the methods under test.

Any of the embedded formulae may be applied successfully to a moderately stiff system of differential equations. However, the step-size will depend on the ratio of the stability limit to the modulus of the largest eigenvalue of the Jacobian matrix rather than a realistic error estimate unless very small tolerances are used. Thus formulae with only moderate stability ranges will require small steps and the RK5(4)7S will permit a substantial reduction in computing time *provided* only a modest global error is required. If high accuracy is needed the RK5(4)7S offers no advantage.

5. CONCLUSIONS

The above experiments justify our attempt to find 'minimum' truncation error formulae since the formula with the higher asymptotic applicability is most efficient. The extra degree of freedom obtained by allowing seven function evaluations for the lower order

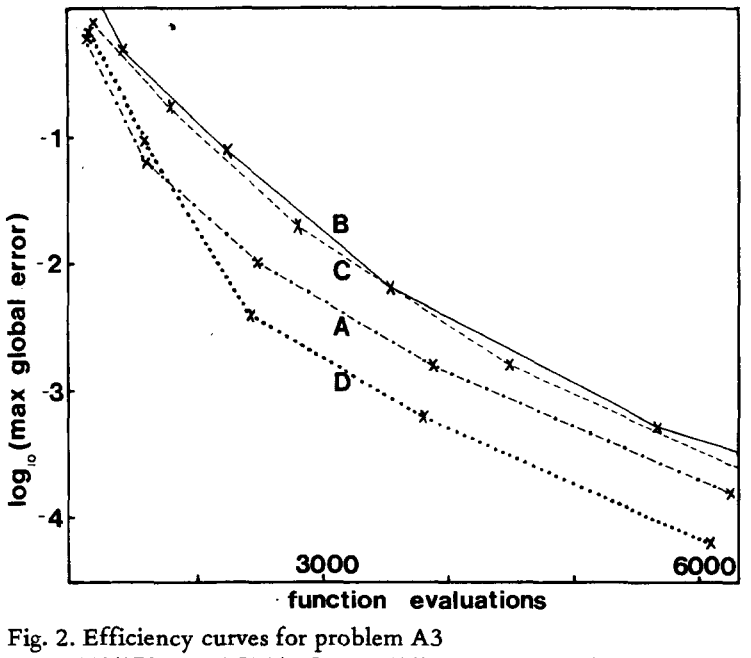

Fig. 2. Efficiency curves for problem A3
A : RK5(4)7S, B : RKF4, C : RK5(4)6M, D : RK5(4)7M.

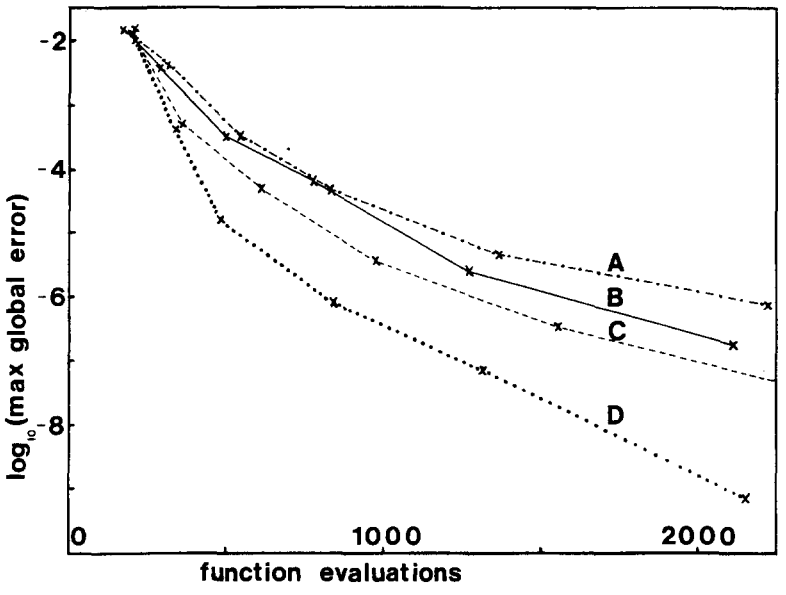

Fig. 3. Efficiency curves for problem D5
A : RK5(4)7S, B : RKF4, C : RK5(4)6M, D : RK5(4)7M.

formula permits lower truncation terms at a penalty of the loss of one extra evaluation when a step rejection occurs but this seems to be worthwhile. It also allows the derivation of a practical high stability formula RK5(4)7S.

REFERENCES

1. BUTCHER J. C. : "Coefficients for the study of Runge-Kutta integration processes", J. Australian Math. Soc., Vol. 3, 1963, pp. 185-201.
2. LAMBERT J. D. : *Computational methods in ordinary differential equations,* John Wiley, London, 1973.
3. HARRIS R. P. : *Runge-Kutta processes,* Proceedings of Fourth Computer Conference, Adelaide, South Australia, 1969.
4. HULL T. E., ENRIGHT W. H., FELLEN B. M. and SEDGWICK A. E. : "Comparing numerical methods for ordinary differential equations", SIAM J. Num. Anal., Vol. 9, No. 4, Dec. 1972.
5. SHAMPINE L. F. and WATTS H. A. : "Global error estimation for ordinary differential equations", ACM TOMS, Vol. 2, No. 2, June 1976, pp. 172-186.
6. JACKSON K. R., ENRIGHT W. H. and HULL T. E. : "A theoretical criterion for comparing Runge-Kutta formulas", Technical Report, No. 101, Dept. of Computer Science, University of Toronto, Jan. 1977.
7. FEHLBERG E. : "Classical fifth, sixth, seventh and eighth order Runge-Kutta formulas with stepsize control", NASA TR R 287, Oct. 1968.
8. FEHLBERG E. : "Low order classical Runge-Kutta formulas with step-size control and their application to some heat transfer problems", NASA TR R-315, 1969.
9. FEHLBERG E. : "Classical eighth and lower order Runge-Kutta-Nystrom formulas with stepsize control for special second order differential equations", NASA TR R-381, March 1972.
10. BUTCHER J. C. : "On Runge-Kutta processes of high order", J. Australian Math. Soc., Vol. 4, 1964, pp. 179-194.
11. DORMAND J. R. and PRINCE P. J. : "New Runge-Kutta algorithms for numerical simulation in dynamical astronomy", Celestial Mechanics, Vol. 18, 1978, pp. 223-232.
12. SHAMPINE L. F. and WATTS H. A. : *Mathematical software III,* ed. Rice J. R. (1977), Academic Press, pp. 257-275.
13. LAWSON J. D. : "An order five Runge-Kutta process with extended region of stability", SIAM J. Num. Anal., Vol. 3, 1966, pp. 593-597.
14. ENRIGHT W. H., BEDET R., FARKAS I. and HULL T. E. : "Test results on initial value methods for non-stiff ordinary differential equations", Technical Report, No. 68, Dept. of Computer Science, University of Toronto, May 1974.
15. SHAMPINE L. F. : "Quadrature and Runge-Kutta formulas", App. Math. and Computation, Vol. 2, 1976, pp. 161-171.
16. ALT R. : "A stable one-step method with step-size control for stiff systems of ODE's", J. of Computational and Applied Maths., Vol. 4, No. 1, March 1978, pp. 29-36.
17. ENGLAND R. : "Error estimates for Runge-Kutta type solutions to systems of ordinary differential equations", Computer J., Vol. 12, 1969, pp. 166-170.