

## Limiting Formulas of Nine-stage Explicit Runge-Kutta Methods of Order Eight

HARUMI ONO<sup>†,☆</sup>

The attainable order of a nine-stage explicit Runge-Kutta formula is at most seven. However, by taking the limit as the distance between the first two and the last two abscissas approaches zero, the formula can achieve eighth order. In this paper, a family of nine-stage eighth-order limiting formulas with four free parameters is derived. Every other parameter is represented as a fractional expression using one or more of these free parameters. Two examples of the method are presented. One of them has a considerably large stability region, but its parameters require a large number of digits. The other has parameters requiring a comparatively small number of digits, but its stability region is not so large.

### 1. Introduction

The attainable order  $p$  of  $s$ -stage explicit Runge-Kutta methods is  $s - 1$  for  $s = 5, s = 6$ , and  $s = 7$ , and is  $s - 2$  for  $s = 8$  and  $s = 9$ . However, they can achieve order  $p + 1$  in the limiting case where the distance between some pairs of abscissas approaches zero. Such formulas are called limiting formulas. They require evaluation of the derivatives in addition to evaluation of the function. Five-stage fifth-order<sup>14)</sup>, six-stage sixth-order<sup>7)</sup> and eight-stage seventh-order<sup>8)</sup> limiting formulas have been derived.

In this paper, a family of nine-stage eighth-order limiting formulas is presented. It is derived in a similar way to the six-stage case. The nine-stage eighth-order limiting formula has four free parameters. Every other parameter is represented as a fractional expression using one or more of them. The polynomial that determines the stability region is a function of one of the free parameters. Thus, we can obtain various formulas by choosing the values of these free parameters to take account of the properties of the formula. The values of the product of the Jacobian matrix and the vector involved in this formula can be easily calculated by using automatic differentiation<sup>5),10)</sup>.

Here, two recommendable sets of free parameters and the methods corresponding to these sets are given. One of them has a fairly large stability region, but its parameters require a large number of digits. The other method's

stability region is not so large, but the numerators and denominators of the coefficients of this method are numbers with a comparatively small number of digits. The latter method is one of the most efficient methods for non-stiff systems, because high-order explicit Runge-Kutta methods are not used for stiff systems.

### 2. Limiting Formulas

We will consider the system of the initial value problem

$$\frac{dy}{dt} = f(t, y), \quad y(t_0) = y_0,$$

where  $f$  and  $y$  are vectors and  $f$  is assumed to be differentiable sufficiently often for the definition to be meaningful. The parameters of an  $s$ -stage explicit Runge-Kutta method are represented in the following Butcher array<sup>3)</sup>:

$$\begin{array}{c|cccc} c_2 & a_{21} & & & \\ c_3 & a_{31} & a_{32} & & \\ \vdots & \vdots & & \ddots & \\ c_s & a_{s1} & a_{s2} & \cdots & a_{s,s-1} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

The method is written in the form

$$f_1 = f(t_n, y_n), \quad y_i = y_n + h \sum_{j=1}^{i-1} a_{ij} f_j,$$

$$f_i = f(t_n + c_i h, y_i) \quad (i = 2, 3, \dots, s)$$

$$y_{n+1} = y_n + h \sum_{i=1}^s b_i f_i.$$

The properties of nine-stage seventh-order formulas are precisely reported and some effective formulas are proposed by Tanaka, et al.<sup>12),13)</sup>. A minute observation of his formulas shows that the truncation error of the formula

<sup>†</sup> Faculty of Engineering, Chiba University

<sup>☆</sup> Current address: 3-22-11, Hachimanyama, Setagaya-ku, Tokyo 156

**Table 1** Limiting formulas.

Stage	Limiting formula	Abscissa
5	5th-order	$c_2 \rightarrow 0$
	5th-order	$c_4 \rightarrow c_5 = 1$
6	6th-order	$c_2 \rightarrow 0$ and $c_5 \rightarrow c_6 = 1$
7	7th-order	$c_2, c_3 \rightarrow 0$ , and $c_6 \rightarrow c_7 = 1$
8	7th-order	$c_2 \rightarrow 0$

become small as the abscissa  $c_2$  approaches zero and  $c_8$  approaches  $c_9 (= 1)$  simultaneously. Thus, it is expected that a nine-stage formula can achieve eighth-order in the limiting case where  $c_2$  tends to 0 and  $c_8$  tends to  $c_9 (= 1)$ . Limiting formulas that use the values of derivatives were derived for five stages by Toda<sup>14)</sup>, for six stages and for eight stages by Ono<sup>7),8)</sup>, and for seven stages by Ono and Toda<sup>9)</sup>. They are listed in **Table 1**.

A nine-stage limiting formula will have a form similar to the six-stage limiting formula:

$$\begin{aligned}
 f_1 &= f(t_n, y_n), \quad F_2 = D(f(t_n, y_n)) \cdot v(f_1), \\
 y_3 &= y_n + h(a_{31}f_1 + h\alpha_3 F_2), \quad f_3 = f(t_n + c_3h, y_3), \\
 y_i &= y_n + h \left( \sum_{j=1, \neq 2}^{i-1} a_{ij}f_j + h\alpha_i F_2 \right), \\
 f_i &= f(t_n + c_ih, y_i) \quad (i = 4, 5, \dots, 8), \\
 \tilde{f}_9 &= \sum_{j=1, \neq 2}^8 A_{9j}f_j + h\alpha_9 F_2, \\
 F_9 &= D(f(t_n + h, y_8)) \cdot v(\tilde{f}_9), \\
 y_{n+1} &= y_n + h \left( \sum_{i=1, \neq 2}^8 b_i f_i + h\beta_2 F_2 + h\beta_9 F_9 \right),
 \end{aligned} \tag{1}$$

where  $D(f(t_p, y_p))$  denotes the Jacobian matrix of  $f$  at the point  $(t_p, y_p)$ ,  $v(f_1)$  denotes the vector  $(1, f_1^1, f_1^2, \dots, f_1^n)^T$ , and  $v(\tilde{f}_9)$  denotes the vector  $(1, \tilde{f}_9^1, \tilde{f}_9^2, \dots, \tilde{f}_9^n)^T$  (the superscripts denote the component numbers). The parameters of this limiting formula can be written in the following array analogous to the Butcher array:

$c_3$	$a_{31}$				$\alpha_3$	
$c_4$	$a_{41}$	$a_{43}$			$\alpha_4$	
$c_5$	$a_{51}$	$a_{53}$	$a_{54}$		$\alpha_5$	
$\vdots$	$\vdots$		$\ddots$		$\vdots$	
$c_8$	$a_{81}$	$a_{83}$	$a_{84}$	$a_{87}$	$\alpha_8$	
	$A_{91}$	$A_{93}$	$A_{94}$	$A_{97}$	$A_{98}$	$\alpha_9$
	$b_1$	$b_3$	$b_4$	$\dots$	$b_8$	
	$\beta_2$	$\beta_9$				

The relations between the parameters of limiting formulas and those of usual nine-stage for-

mulas are as follows:

Limiting formula	Usual nine-stage formula
$a_{i1}$	$\lim_{c_2 \rightarrow 0} (a_{i1} + a_{i2}) \quad (i = 3, \dots, 8)$
$\alpha_i$	$\lim_{c_2 \rightarrow 0} a_{i2}c_2 \quad (i = 3, \dots, 7)$
$A_{91}$	$\lim_{c_8 \rightarrow 1} (\lim_{c_2 \rightarrow 0} (a_{91} + a_{92} - (a_{81} + a_{82})) / (1 - c_8))$
$A_{9j}$	$\lim_{c_8 \rightarrow 1} (a_{9j} - a_{8j}) / (1 - c_8) \quad (j = 3, \dots, 7)$
$A_{98}$	$\lim_{c_8 \rightarrow 1} a_{98} / (1 - c_8)$
$\alpha_9$	$\lim_{c_8 \rightarrow 1} (\lim_{c_2 \rightarrow 0} (a_{92}c_2 - a_{82}c_2) / (1 - c_8))$
$b_1$	$\lim_{c_2 \rightarrow 0} (b_1 + b_2)$
$b_8$	$\lim_{c_8 \rightarrow 1} (b_8 + b_9)$
$\beta_2$	$\lim_{c_2 \rightarrow 0} b_2c_2$
$\beta_9$	$\lim_{c_8 \rightarrow 1} (-b_8(1 - c_8))$

**2.1 Order Conditions**

We shall restrict ourselves to the case in which  $b_3 = 0$  and the following simplifying assumptions hold:

$$\begin{aligned}
 c_8 = 1, \quad \alpha_3 &= \frac{c_3^2}{2}, \quad \sum_{j=3}^{i-1} a_{ij}c_j + \alpha_i = \frac{c_i^2}{2}, \\
 \sum_{j=3}^{i-1} a_{ij}c_j^2 &= \frac{c_i^3}{3} \quad (i = 4, 5, \dots, 8).
 \end{aligned} \tag{2}$$

Comparing the Taylor series expansion of Eq.(1) with that of the true value  $y(t_n + h)$  and matching the coefficients of each elementary differential, after tedious computation, we obtain the following equations of the order conditions:

$$a_{31} = c_3, \quad a_{i1} + \sum_{j=3}^{i-1} a_{ij} = c_i \quad (i = 4, 5, \dots, 8),$$

$$A_{91} + \sum_{j=3}^8 A_{9,j} = 1, \quad \alpha_9 + \sum_{j=3}^8 A_{9,j}c_j = 1, \tag{3}$$

$$\begin{aligned}
 \sum_{i=j+1}^8 b_i a_{ij} + \beta_9 A_{9j} &= b_j(1 - c_j) \quad (j = 4, 5, 6, 7), \\
 \beta_9 A_{98} &= -\beta_9,
 \end{aligned} \tag{4}$$

$$\sum_{i=4}^8 b_i a_{i3} + \beta_9 A_{93} = 0, \tag{5}$$

$$\sum_{i=5}^8 b_i \sum_{j=4}^{i-1} a_{ij}a_{j3} + \beta_9 \sum_{j=4}^8 A_{9j}a_{j3} = 0, \tag{6}$$

$$\begin{aligned}
 \sum_{i=6}^8 b_i \sum_{j=5}^{i-1} a_{ij} \sum_{k=4}^{j-1} a_{jk}a_{k3} \\
 + \beta_9 \sum_{j=5}^8 A_{9j} \sum_{k=4}^{j-1} a_{jk}a_{k3} = 0,
 \end{aligned} \tag{7}$$

$$\sum_{i=7}^8 b_i \sum_{j=6}^{i-1} a_{ij} \sum_{k=5}^{j-1} a_{jk} \sum_{l=4}^{k-1} a_{kl} a_{l3} + \beta_9 \sum_{j=6}^8 A_{9j} \sum_{k=5}^{j-1} a_{jk} \sum_{l=4}^{k-1} a_{kl} a_{l3} = 0, \quad (8)$$

$$\sum_{i=6}^8 b_i \sum_{j=5}^{i-1} a_{ij} \sum_{k=4}^{j-1} a_{jk} c_k a_{k3} + \beta_9 \sum_{j=5}^8 A_{9j} \sum_{k=4}^{j-1} a_{jk} c_k a_{k3} = 0, \quad (9)$$

$$b_1 + \sum_{i=4}^8 b_i = 1, \quad \sum_{i=4}^8 b_i c_i + \beta_2 + \beta_9 = \frac{1}{2}, \quad (10)$$

$$\sum_{i=4}^8 b_i c_i^2 + 2\beta_9 = \frac{1}{3}, \quad (11)$$

$$\sum_{i=5}^8 b_i \sum_{j=4}^{i-1} a_{ij} c_j^2 + \beta_9 \sum_{j=4}^8 A_{9j} c_j^2 = \frac{1}{12}, \quad (12)$$

$$\sum_{i=5}^8 b_i \sum_{j=4}^{i-1} a_{ij} c_j^3 + \beta_9 \sum_{j=4}^8 A_{9j} c_j^3 = \frac{1}{20}, \quad (13)$$

$$\sum_{i=5}^8 b_i \sum_{j=4}^{i-1} a_{ij} c_j^4 + \beta_9 \sum_{j=4}^8 A_{9j} c_j^4 = \frac{1}{30}, \quad (14)$$

$$\sum_{i=6}^8 b_i \sum_{j=5}^{i-1} a_{ij} \sum_{k=4}^{j-1} a_{jk} c_k^3 + \beta_9 \sum_{j=5}^8 A_{9j} \sum_{k=4}^{j-1} a_{jk} c_k^3 = \frac{1}{120}, \quad (15)$$

$$\sum_{i=5}^8 b_i \sum_{j=4}^{i-1} a_{ij} c_j^5 + \beta_9 \sum_{j=4}^8 A_{9j} c_j^5 = \frac{1}{42}, \quad (16)$$

$$\sum_{i=6}^8 b_i \sum_{j=5}^{i-1} a_{ij} \sum_{k=4}^{j-1} a_{jk} c_k^4 + \beta_9 \sum_{j=5}^8 A_{9j} \sum_{k=4}^{j-1} a_{jk} c_k^4 = \frac{1}{210}, \quad (17)$$

$$\sum_{i=7}^8 b_i \sum_{j=6}^{i-1} a_{ij} \sum_{k=5}^{j-1} a_{jk} \sum_{l=4}^{k-1} a_{kl} c_l^3 + \beta_9 \sum_{j=6}^8 A_{9j} \sum_{k=5}^{j-1} a_{jk} \sum_{l=4}^{k-1} a_{kl} c_l^3 = \frac{1}{840}, \quad (18)$$

$$\sum_{i=6}^8 b_i \sum_{j=5}^{i-1} a_{ij} c_j \sum_{k=4}^{j-1} a_{jk} c_k^3 + \beta_9 \sum_{j=5}^8 A_{9j} c_j \sum_{k=4}^{j-1} a_{jk} c_k^3 = \frac{1}{168}, \quad (19)$$

$$\sum_{i=5}^8 b_i \sum_{j=4}^{i-1} a_{ij} c_j^6 + \beta_9 \sum_{j=4}^8 A_{9j} c_j^6 = \frac{1}{56}, \quad (20)$$

$$\sum_{i=6}^8 b_i \sum_{j=5}^{i-1} a_{ij} \sum_{k=4}^{j-1} a_{jk} c_k^5 + \beta_9 \sum_{j=5}^8 A_{9j} \sum_{k=4}^{j-1} a_{jk} c_k^5 = \frac{1}{336}, \quad (21)$$

$$\sum_{i=7}^8 b_i \sum_{j=6}^{i-1} a_{ij} \sum_{k=5}^{j-1} a_{jk} \sum_{l=4}^{k-1} a_{kl} c_l^4 + \beta_9 \sum_{j=6}^8 A_{9j} \sum_{k=5}^{j-1} a_{jk} \sum_{l=4}^{k-1} a_{kl} c_l^4 = \frac{1}{1680}, \quad (22)$$

$$b_8 a_{87} a_{76} a_{65} a_{54} c_4^3 + \beta_9 \sum_{j=7}^8 A_{9j} \sum_{k=6}^{j-1} a_{jk} \sum_{l=5}^{k-1} a_{lm} c_m^3 = \frac{1}{6720}, \quad (23)$$

$$\sum_{i=6}^8 b_i \sum_{j=5}^{i-1} a_{ij} c_j \sum_{k=4}^{j-1} a_{jk} c_k^4 + \beta_9 \sum_{j=5}^8 A_{9j} c_j \sum_{k=4}^{j-1} a_{jk} c_k^4 = \frac{1}{280}, \quad (24)$$

$$\sum_{i=6}^8 b_i \sum_{j=5}^{i-1} a_{ij} c_j^2 \sum_{k=4}^{j-1} a_{jk} c_k^3 + \beta_9 \sum_{j=5}^8 A_{9j} c_j^2 \sum_{k=4}^{j-1} a_{jk} c_k^3 = \frac{1}{224}, \quad (25)$$

$$\sum_{i=7}^8 b_i \sum_{j=6}^{i-1} a_{ij} c_j \sum_{k=5}^{j-1} a_{jk} \sum_{l=4}^{k-1} a_{kl} c_l^3 + \beta_9 \sum_{j=6}^8 A_{9j} c_j \sum_{k=5}^{j-1} a_{jk} \sum_{l=4}^{k-1} a_{kl} c_l^3 = \frac{1}{1120}, \quad (26)$$

$$\sum_{i=7}^8 b_i \sum_{j=6}^{i-1} a_{ij} \sum_{k=5}^{j-1} a_{jk} c_k \sum_{l=4}^{k-1} a_{kl} c_l^3 + \beta_9 \sum_{j=6}^8 A_{9j} \sum_{k=5}^{j-1} a_{jk} c_k \sum_{l=4}^{k-1} a_{kl} c_l^3 = \frac{1}{1344}. \quad (27)$$

## 2.2 Solutions

In this section, we assume that abscissas  $c_i$  ( $i = 4, 5, 6, 7$ ) are distinct and are not equal to 0 and 1. We define, for later convenience, the following auxiliary parameters:

$$\sum_{i=j+1}^8 b_i a_{ij} + \beta_9 A_{9j} = b_j (1 - c_j) \stackrel{\text{def}}{=} \rho_j \quad (j = 4, 5, 6, 7),$$

$$\beta_9 A_{98} = \rho_8, \quad \sum_{j=k+1}^8 \rho_j a_{jk} \stackrel{\text{def}}{=} \sigma_k \quad (k = 4, 5, 6, 7),$$

$$\sum_{k=l+1}^7 \sigma_k a_{kl} \stackrel{\text{def}}{=} \tau_l \quad (l=4,5,6), \quad \sum_{k=l+1}^6 \tau_k a_{kl} \stackrel{\text{def}}{=} \phi_l \quad (l=4,5). \quad (28)$$

The outline of the derivation is shown in the appendix. Here we show only the results. Rewriting the equations of the order conditions by using Eq. (28), we obtain

$$\rho_i = \frac{28c_j c_k c_l - 14(c_k c_l + c_j c_l + c_j c_k) + 8(c_j + c_k + c_l) - 5}{840c_i^2(c_i - c_j)(c_i - c_k)(c_i - c_l)(c_i - 1)}$$

$$\rho_8 = (70c_i c_j c_k c_l - 42 \sum_{i,j,k} c_i c_j c_k + 28 \sum_{i,j} c_i c_j - 20 \sum_i c_i + 15) / (840(1 - c_4)(1 - c_5)(1 - c_6)(1 - c_7)),$$

$$\sigma_i = \frac{\rho_i(1 - c_i)}{2} \quad (i, j, k, l = 4, 5, 6, 7),$$

$$\tau_i = \frac{14c_j c_k - 6(c_j + c_k) + 3}{5040c_i^2(c_i - c_j)(c_i - c_k)} \quad (i, j, k = 4, 5, 6),$$

$$\phi_i = \frac{-8c_j + 3}{20160c_i^2(c_i - c_j)} \quad (i, j = 4, 5), \quad (29)$$

and the relation between  $c_4$  and  $c_5$ , which must be satisfied:

$$(56c_4^2 - 42c_4 + 9)c_5 - 3c_4 = 0. \quad (30)$$

If we assume that the values of  $c_i$  are in the interval  $(0, 1)$ , then from (30) we get

$$0 < c_4 < \frac{3}{8} \quad \text{or} \quad \frac{3}{7} < c_4 < 1. \quad (31)$$

All parameters of the method can be obtained successively in terms of  $c_i$  using  $\rho_i$ ,  $\sigma_i$ ,  $\tau_i$ , and  $\phi_i$ , provided that all denominators do not vanish. The parameters  $b_i$  and  $\beta_i$  are

$$b_i = \frac{\rho_i}{(1 - c_i)} \quad (i = 4, 5, 6, 7), \quad \beta_9 = -\rho_8,$$

$$b_8 = \frac{1}{3} - \left(\sum_{i=4}^7 b_i c_i^2 + 2\beta_9\right), \quad b_1 = 1 - \sum_{i=4}^8 b_i,$$

$$\beta_2 = \frac{1}{2} - \left(\sum_{i=4}^8 b_i c_i + \beta_9\right), \quad (32)$$

and  $a_{ij}$  and  $A_{9j}$  ( $j = 8, 7, 6, 5, 4$ ) are

$$A_{98} = -1, \quad a_{87} = \frac{\sigma_7}{\rho_8}, \quad A_{97} = \frac{\rho_7 - b_8 a_{87}}{\beta_9}, \quad a_{76} = \frac{\tau_6}{\sigma_7},$$

$$a_{86} = \frac{\sigma_6 - \rho_7 a_{76}}{\rho_8}, \quad A_{96} = \frac{\rho_6 - \sum_{i=7}^8 b_i a_{i6}}{\beta_9},$$

$$a_{65} = \frac{\phi_5}{\tau_6}, \quad a_{75} = \frac{\tau_5 - \sigma_6 a_{65}}{\sigma_7},$$

$$a_{85} = \frac{\sigma_5 - \sum_{i=6}^7 \rho_i a_{i5}}{\rho_8}, \quad A_{95} = \frac{\rho_5 - \sum_{i=6}^8 b_i a_{i5}}{\beta_9},$$

$$a_{54} = \frac{c_5^3(c_5 - c_4)}{c_4^3}, \quad a_{64} = \frac{\phi_4 - \tau_5 a_{54}}{\tau_6},$$

$$a_{74} = \frac{\tau_4 - \sum_{i=5}^6 \sigma_i a_{i4}}{\sigma_7}, \quad a_{84} = \frac{\sigma_4 - \sum_{i=5}^7 \rho_i a_{i4}}{\rho_8},$$

$$A_{94} = \frac{\rho_4 - \sum_{i=5}^8 b_i a_{i4}}{\beta_9}. \quad (33)$$

Finally,  $a_{i3}$ ,  $A_{9j}$  ( $j = 3, 1$ ),  $\alpha_i$ , and  $a_{i1}$  are

$$a_{43} = \frac{c_4^3}{3c_3^2}, \quad a_{i3} = \frac{c_i^3 - 3 \sum_{j=4}^{i-1} a_{ij} c_j^2}{3c_3^2} \quad (i = 5, 6, 7, 8),$$

$$A_{93} = -\frac{\sum_{i=4}^8 b_i a_{i3}}{\beta_9}, \quad A_{91} = 1 - \sum_{j=3}^8 A_{9j}, \quad \alpha_3 = \frac{c_3^2}{2},$$

$$\alpha_i = \frac{c_i^2}{2} - \sum_{j=3}^{i-1} a_{ij} c_j \quad (i = 4, 5, \dots, 8), \quad \alpha_9 = 1 - \sum_{j=3}^8 A_{9j} c_j,$$

$$a_{31} = c_3, \quad a_{i1} = c_i - \sum_{j=3}^{i-1} a_{ij} \quad (i = 4, 5, \dots, 8). \quad (34)$$

Now, we have obtained a set of parameters of the nine-stage eighth-order limiting formula with four free parameters,  $c_3$ ,  $c_4$ ,  $c_6$ , and  $c_7$ .

### 3. Determination of the Free Parameters

In this section we will consider how to determine the four free parameters. The stability region depends on only one free parameter,  $c_4$ . It is desirable to determine  $c_4$  so as to maximize the stability region. At the same time, it is preferable that every parameter be a number that requires a small number of digits and has a small magnitude. Here, we will present two sets of free parameters. Substituting the values of these sets into the solutions obtained in the previous section, we obtain two formulas. One of them has a comparatively large stability region, and the other has relatively simple numbers as parameters.

#### 3.1 Stability

The polynomial  $r$  that determines the stability region of the nine-stage eighth-order limiting formula (1) is given by

$$r(z) = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots + \frac{z^7}{7!} + \frac{z^8}{8!} + \gamma z^9,$$

where  $z$  is a complex number and

$$\gamma = \phi_5 a_{54} a_{43} \alpha_3 = \frac{c_4(3 - 8c_4)}{40320(56c_4^2 - 42c_4 + 9)}.$$

Let the simply connected interval  $(-d, 0)$  be the intersection of the stability region and the negative part of the real axis. The stability boundaries for several values of  $\gamma$  are shown in **Fig. 1**. The graph indicates that

$$\gamma \in \left( \frac{1}{620000}, \frac{1}{580000} \right) \quad (35)$$

gives the maximum stability region. The graph of  $\gamma(c_4)$  for the interval (31) is given in **Fig. 2**. The intervals for  $c_4$ , for which  $\gamma$  is in the interval

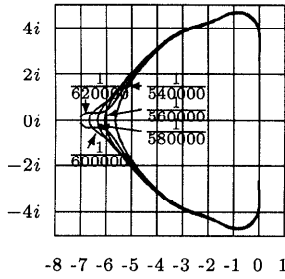


Fig. 1 Stability boundaries for several values of  $\gamma$ .

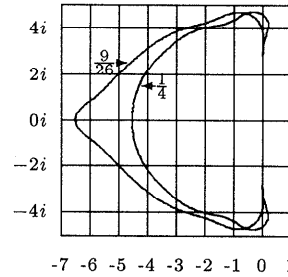


Fig. 3 Stability boundaries for  $c_4 = 9/26$  and  $1/4$ .

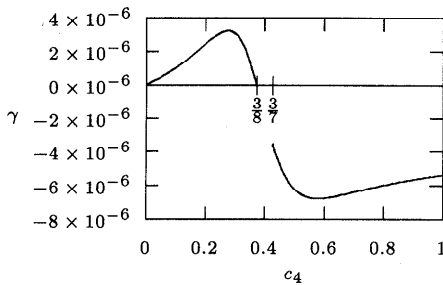


Fig. 2 Graph of  $\gamma(c_4)$  for  $0 < c_4 < 3/8$  and  $3/7 < c_4 < 1$ .

(35), are approximately

$$.1446 < c_4 < .1523 \text{ and } .3455 < c_4 < .3477. \quad (36)$$

### 3.2 Two Sets of Free Parameters

To find a method that has a comparatively large stability region, we will proceed as follows: (i) Let the rational number be a value of  $c_4$  that lies within the intervals given in Eq. (36) and whose numerator and denominator are numbers with at most two digits. (ii) For all such values of  $c_4$ , we look for a value of  $c_5$  whose numerator and denominator are also numbers with at most two digits (the value of  $c_5$  is determined by relation (30)). (iii) For every such pair of  $(c_4, c_5)$  and the values of  $c_6$  and  $c_7$  ( $0 < c_6, c_7 < 1$ ), we look for the minimum  $s$ , the sum of the magnitude of the parameters in Eqs. (32) and (33), because these parameters are independent of the value of  $c_3$ . (iv) As the values of  $c_6$  and  $c_7$ , we choose the simplest pair of numbers near the pair of  $c_6$  and  $c_7$  that gives the minimum  $s$ . (v) Finally, we look for the value of  $c_3$  in a similar way; that is, we look for the minimum sum of the magnitude of the parameters (34). We determine the simplest rational number as the value of  $c_3$  that gives nearly the minimum sum. Thus we obtain a set of free parameters:

$$c_3 = \frac{1}{3}, \quad c_4 = \frac{9}{26}, \quad c_6 = \frac{3}{4}, \quad c_7 = \frac{1}{4}. \quad (37)$$

For this set, the values of  $c_5$ ,  $\gamma$ , and  $d$  are

$$c_5 = \frac{39}{44}, \quad \gamma = \frac{1}{591360} \approx 1.691 \times 10^{-6}, \quad d \approx 6.5.$$

We will give another value of  $c_4$  for which all the parameters have a relatively small number of digits. First, we look for the pair of  $c_4$  and  $c_5$  such that all numerators and denominators are numbers with one digit. We determine  $1/4$  as the value of  $c_4$ , because it gives the largest stability region among such pairs. Then, we proceed as for  $c_4 = 9/26$ , and get

$$c_3 = \frac{1}{4}, \quad c_4 = \frac{1}{4}, \quad c_6 = \frac{7}{8}, \quad c_7 = \frac{3}{4}. \quad (38)$$

For this set, the values of  $c_5$ ,  $\gamma$ , and  $d$  are

$$c_5 = \frac{3}{8}, \quad \gamma = \frac{1}{322560} \approx 3.100 \times 10^{-6}, \quad d \approx 4.5.$$

This value of  $c_4$  is outside of the intervals in Eq. (36). Thus, the stability region is not so large, but the parameters of the formula require a far smaller number of digits than that for  $c_4 = 9/26$ . The stability boundaries for these two values of  $c_4$ ,  $9/26$  and  $1/4$ , are shown in Fig. 3.

### 3.3 Nine-stage Eighth-order Limiting Formulas

Substituting the set of  $c_i$  (38) into Eqs. (29), (32), (33), and (34), we obtain Formula 1. The stability region of this formula is not so large, but its parameters are numbers that require a comparatively small number of digits. The parameters of this formula are shown in Table 2.

For the set of  $c_i$  (37), we obtain Formula 2. This formula has a comparatively large stability region, but its parameters are more complicated than those of Formula 1. The parameters of Formula 2 are shown in Table 3.

## 4. Numerical Example and Conclusions

We will show that our derivations are correct. First, our formula is definitely eight-order. Though this can be verified by substituting the values from Table 2 and Table 3 into the error coefficients up to  $O(h^8)$  terms, here we will

Table 2 Formula 1.

$c_i$	$a_{ij} \quad (i = 3, 4, \dots, 8; j = 1, 3, \dots, 8)$								$\alpha_i$
1	1								1
4	4								32
1	1	12							1
4	6								96
3	3	9							0
8	32	64	27						
7	12607	2303	2695	490					539
8	2592	576	192	81					864
3	2297	3	207	38	54				199
4	2058	4	70	21	1715				1568
1	32183	832	600	320	1728	280			1345
	8967	183	61	183	2989	183			2562
$A_{9j}$	16106722	150016	470864	1243520	7922304	770224	-1		65822
	1640961	3721	18605	33489	911645	33489			26047
$b_j$	12289	0	704	2048	2048	64	10537		
	92610		4725	7875	8575	135	47250		
$\beta_2, \beta_9$	47	61							
	8820	6300							

Table 3 Formula 2.

$c_i$	$a_{ij} \quad (j = 1, 3, \dots, 8)$								$\alpha_i$
1	1								1
3	3								18
9	3897	2187							81
26	17576	17576							4394
39	8292271	14414517	38243179						342563
44	16866432	1874048	4216608						1874048
3	349085	3159	1184183						1597
4	3699072	2432	563616	27951					31616
1	63001339	351	7986095	661466	1164625				38219
4	299624832	2432	45652896	53578746	162				2560896
1	3578509	702	328398772	363416240	48640	912			21163
	8993673	73	38369457	720493137	41391	511			153738
$A_{9j}$	16288620394	7275528	3275107674488	2097338476640	281776384	114146528	-1		19731878
	3720382731	90593	79360826895	298043994339	17122077	3170755			31798143
$b_j$	1202603	0	501988136	2494357888	9728	2432	212561		
	8624070		1563686775	8636047875	19845	33075	803250		
$\beta_2, \beta_9$	857	73							
	147420	6300							

give a simpler demonstration using a numerical example. We give the errors in the numerical solution of a system of equations<sup>2)</sup>:

Example 1: Integrate

$$\begin{aligned} \frac{dy_1}{dt} &= y_2 y_3, & y_1(0) &= 0, \\ \frac{dy_2}{dt} &= -y_1 y_3, & y_2(0) &= 1, \\ \frac{dy_3}{dt} &= -k^2 y_1 y_2, & y_3(0) &= 1, \quad k^2 = 0.51 \end{aligned}$$

over the range [0, 60] by using Formula 1 and Formula 2. The largest errors in the last step for various values of  $h$  are shown in Fig. 4 and Table 4. The computations were performed in quadruple-precision arithmetic. For comparison, the results obtained by the eleven-stage eighth-order method given by Cooper and Verner<sup>4)</sup> are also shown. Figure 4 and Table 4 indicate that all formulas used here are exactly of order eight, because the accumulated truncation errors are proportional to  $h^8$ .

Next, to check the stability interval of our formulas, we present the results of an equation<sup>11)</sup>:

Example 2: Integrate

$$\frac{dy}{dt} = 100(\sin x - y), \quad y(0) = 0$$

up to 100 steps with various step sizes  $h$ . The

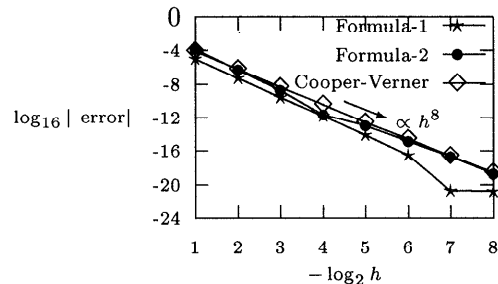


Fig. 4 Largest errors in the numerical solution of Example 1 at the last step.

results are shown in Table 5. Quadruple-precision arithmetic was also used. From Table 5, we see that Formula 1 works well for  $h \leq 0.04$ , but that for  $h \geq 0.05$  the computation fails. The computation by Formula 2 can be continued for larger values of  $h$  than 0.04, but it breaks down for  $h \geq 0.07$ .

In this study, we have shown the existence of nine-stage eighth-order limiting formulas, as  $c_2 \rightarrow 0$  and  $c_8 \rightarrow c_9 = 1$ . The derivatives involved in our formulas are not individual partial derivatives, but the products of the Jacobian matrix and some vector. These products can be calculated very easily by automatic differen-

**Table 4** Errors in the numerical solution of Example 1 at the last step.

Step size		Formula 1	Formula 2	Cooper and Verner
2 <sup>-1</sup>	y <sub>1</sub>	.109 × 10 <sup>-5</sup>	-.238 × 10 <sup>-4</sup>	-.139 × 10 <sup>-4</sup>
	y <sub>2</sub>	-.759 × 10 <sup>-6</sup>	.923 × 10 <sup>-5</sup>	.428 × 10 <sup>-5</sup>
	y <sub>3</sub>	-.281 × 10 <sup>-6</sup>	.403 × 10 <sup>-5</sup>	.206 × 10 <sup>-5</sup>
2 <sup>-2</sup>	y <sub>1</sub>	.183 × 10 <sup>-8</sup>	-.311 × 10 <sup>-7</sup>	-.357 × 10 <sup>-7</sup>
	y <sub>2</sub>	-.139 × 10 <sup>-8</sup>	.127 × 10 <sup>-7</sup>	.115 × 10 <sup>-7</sup>
	y <sub>3</sub>	-.497 × 10 <sup>-9</sup>	.543 × 10 <sup>-8</sup>	.560 × 10 <sup>-8</sup>
2 <sup>-3</sup>	y <sub>1</sub>	.332 × 10 <sup>-11</sup>	-.346 × 10 <sup>-10</sup>	-.948 × 10 <sup>-10</sup>
	y <sub>2</sub>	-.260 × 10 <sup>-11</sup>	.148 × 10 <sup>-10</sup>	.326 × 10 <sup>-10</sup>
	y <sub>3</sub>	-.893 × 10 <sup>-12</sup>	.634 × 10 <sup>-11</sup>	.159 × 10 <sup>-10</sup>
2 <sup>-4</sup>	y <sub>1</sub>	.600 × 10 <sup>-14</sup>	.118 × 10 <sup>-13</sup>	-.278 × 10 <sup>-12</sup>
	y <sub>2</sub>	-.479 × 10 <sup>-14</sup>	-.231 × 10 <sup>-14</sup>	.102 × 10 <sup>-12</sup>
	y <sub>3</sub>	-.151 × 10 <sup>-14</sup>	-.477 × 10 <sup>-15</sup>	.496 × 10 <sup>-13</sup>
2 <sup>-5</sup>	y <sub>1</sub>	.100 × 10 <sup>-16</sup>	.312 × 10 <sup>-15</sup>	-.902 × 10 <sup>-15</sup>
	y <sub>2</sub>	-.831 × 10 <sup>-17</sup>	-.119 × 10 <sup>-15</sup>	.345 × 10 <sup>-15</sup>
	y <sub>3</sub>	-.206 × 10 <sup>-17</sup>	-.476 × 10 <sup>-16</sup>	.169 × 10 <sup>-15</sup>

**Table 5** Relative errors in numerical solution of Example 2 (E<sub>1</sub>: first step, E<sub>100</sub>: last step).

Step size		Formula 1	Formula 2	Cooper and Verner
0.02	E <sub>1</sub>	-.365 × 10 <sup>-3</sup>	.270 × 10 <sup>-3</sup>	.134 × 10 <sup>-2</sup>
	E <sub>100</sub>	.391 × 10 <sup>-9</sup>	-.190 × 10 <sup>-9</sup>	-.339 × 10 <sup>-6</sup>
0.03	E <sub>1</sub>	-.952 × 10 <sup>-2</sup>	.401 × 10 <sup>-2</sup>	.271 × 10 <sup>-1</sup>
	E <sub>100</sub>	.239 × 10 <sup>-6</sup>	-.768 × 10 <sup>-7</sup>	-.155 × 10 <sup>-4</sup>
0.04	E <sub>1</sub>	-.997 × 10 <sup>-1</sup>	.227 × 10 <sup>-1</sup>	.238
	E <sub>100</sub>	-.383 × 10 <sup>-6</sup>	.991 × 10 <sup>-7</sup>	-.557 × 10 <sup>-3</sup>
0.05	E <sub>1</sub>	-.626	.613 × 10 <sup>-1</sup>	1.331
	E <sub>100</sub>	-.644 × 10 <sup>38</sup>	-.110 × 10 <sup>-6</sup>	-.578 × 10 <sup>71</sup>
0.06	E <sub>1</sub>	-2.826	.141 × 10 <sup>-1</sup>	5.589
	E <sub>100</sub>	—	.367 × 10 <sup>-9</sup>	—
0.07	E <sub>1</sub>	—	-.658	—
	E <sub>100</sub>	—	.632 × 10 <sup>58</sup>	—

tiation, using the intermediate values obtained during the function evaluation. Thus the total cost is about twice the number of intermediate variables and in general is far less than the function evaluation. Moreover, since automatic methods for simultaneous computation of functions and partial derivatives are now available<sup>1),6),15)</sup>, our methods can be easily calculated.

In conclusion, we can state that Formula 1 is efficient for non-stiff problems. Simple parameters are preferred, because explicit Runge-Kutta formulas are not suitable for stiff systems.

**Acknowledgments** The author would like to thank Professor M. Tanaka for his kind suggestions. She also would like to thank Professor T. Mitsui for his helpful comments, and the anonymous referees for their constructive re-

marks. She is grateful to Professor M. Hoshi for his support.

## References

- 1) Bischof, C., Carle, A., Corliss, G., Griewank, A. and Hovland, P.: ADIFOR - Generating Derivative Codes from Fortran Programs, *Scientific Programming*, Vol.1, pp.11-29 (1992).
- 2) Bulirsch, R. and Stoer, J.: Numerical Treatment of Ordinary Differential Equations by Extrapolation Methods, *Num. Math.*, Vol.8, pp.1-13 (1966).
- 3) Butcher, J.C.: *The Numerical Analysis of Ordinary Differential Equations*, John Wiley & Sons, New York (1987).
- 4) Cooper, G.J. and Verner, J.H.: Some Explicit Runge-Kutta Methods of High Order, *SIAM J. Num. Anal.*, Vol.9, pp.389-405 (1972).
- 5) Iri, M.: Simultaneous Computation of Functions, Partial Derivatives and Estimates of Rounding Errors - Complexity and Practicality, *Japan. J. Appl. Math.*, Vol.1, pp.223-252 (1984).
- 6) Kubota, K: PADRE2, A FORTRAN Pre-compiler Yielding Error Estimates and Second Derivatives, *Proc. SIAM Workshop on Automatic Differentiation of Algorithms - Theory, Implementation, and Application* (1991).
- 7) Ono, H.: Five- and Six-stage Runge-Kutta-Type Formulas of Orders Numerically Five and Six, *J. Inf. Process.*, Vol.12, No.3, pp.251-260 (1989).
- 8) Ono, H.: Limiting Formulas of Eight-stage Explicit Runge-Kutta Methods of Order Seven, *Numerical Analysis of Ordinary Differential Equations and Its Applications*, Mitsui, T. and Shinohara, Y. (Eds.), pp.1-14, World Scientific, Singapore (1996).
- 9) Ono, H. and Toda, H.: Runge-Kutta-Type Seventh-order Limiting Formula, *J. Inf. Process.*, Vol.12, No.3, pp.286-298 (1989).
- 10) Rall, L.B.: *Automatic Differentiation - Techniques and Applications*, Lecture Notes in Computer Science, Vol.120, Springer-Verlag, Berlin (1981).
- 11) Ralston, A. and Rabinowitz, P.: *A First Course in Numerical Analysis*, 2nd edition, McGraw-Hill, New York (1987).
- 12) Tanaka, M., Muramatsu, S. and Yamashita, S.: On the Optimization of Some Nine-stage Seventh-order Runge-Kutta Methods, *Trans. IPS Japan*, Vol.33, No.12, pp.1512-1526 (1992) (in Japanese).
- 13) Tanaka, M., Yamashita, S., Kubo, E. and Nozaki, Y.: On Nine-stage Seventh-order Explicit Runge-Kutta Methods with Extended Region of Stability, *Trans. IPS Japan*, Vol.34,

No.1, pp.52–61 (1993) (in Japanese).

- 14) Toda, H.: On the Truncation Error of a Limiting Formula of Runge–Kutta Methods, *Trans. IPS Japan*, Vol.21, No.4, pp.285–296 (1980) (in Japanese).
- 15) Yoshida, T.: Automatic Derivative Derivation System, *Trans. IPS Japan*, Vol.30, No.7, pp.799–806 (1989) (in Japanese).

**Appendix: Outline of the Derivation of Eqs. (29), (30) and  $a_{54}$**

Using the notation  $\rho_i$ , we see that Eqs. (12), (13), (14), (16), and (20) can be written in the form

$$\sum_{i=4}^8 \rho_i c_i^k = \frac{1}{\prod_{m=1}^2 (k+m)} \quad (k = 2, 3, \dots, 6).$$

From this system, we obtain the solutions  $\rho_i$ . By using the last equation of (2) and (6), Eq. (13) can be rewritten in terms of  $\sigma_i$ . From this resultant equation, together with Eqs. (15), (17), and (21), we obtain the system

$$\sum_{i=4}^7 \sigma_i c_i^k = \frac{1}{\prod_{m=1}^3 (k+m)} \quad (k = 2, 3, 4, 5),$$

and the solutions  $\sigma_i$ .

If all the equations used above hold, then four equations, (19), (24), (25), and (26), can be omitted, as shown below. Replacing  $\sigma_i$  in Eq. (15) with  $\rho_i$  in accordance with definition (28), subtracting Eq. (19) from this equation, and substituting  $2\sigma_i$  for  $\rho_i(1-c_i)$  following from Eq. (29), we see that Eq. (19) becomes Eq. (18). In a very similar way, using Eqs. (17), (19), and (18), we can see that Eqs. (24), (25), and (26) are identical to Eqs. (22), (27), and (23), respectively.

Rewriting  $c_i^3$  in Eq. (15) using the last equation of (2), and then using Eq. (7), together with Eqs. (18) and (22), we obtain

$$\sum_{i=4}^6 \tau_i c_i^k = \frac{1}{\prod_{m=1}^4 (k+m)} \quad (k = 2, 3, 4),$$

and the solutions  $\tau_i$ . Similarly, rewriting Eq. (18) by using the last equation of (2) and (8), together with Eq. (23), we obtain the system

$$\sum_{i=4}^5 \phi_i c_i^k = \frac{1}{\prod_{m=1}^5 (k+m)} \quad (k = 2, 3),$$

and the solutions  $\phi_i$ .

The relation (30) is obtained from Eq. (27) as follows:

(18)  $\times c_7 - (27)$  yields

$$(\sigma_5(c_7 - c_5)a_{54} + \sigma_6(c_7 - c_6)a_{64})c_4^3$$

$$+ \sigma_6(c_7 - c_6)a_{65}c_5^3 = \frac{c_7}{840} - \frac{1}{1344}. \quad (39)$$

(15)  $\times c_7 - (17)$  yields

$$\begin{aligned} \sigma_4 c_4^3(c_7 - c_4) + \sigma_5 c_5^3(c_7 - c_5) \\ + \sigma_6 c_6^3(c_7 - c_6) = \frac{c_7}{120} - \frac{1}{210}. \end{aligned}$$

In the latter equation, rewriting  $c_i^3$  by using the last equation of (2), and then using Eqs. (7) and (9), we obtain

$$\begin{aligned} (\sigma_5(c_7 - c_5)a_{54} + \sigma_6(c_7 - c_6)a_{64})c_4^2 \\ + \sigma_6(c_7 - c_6)a_{65}c_5^2 = \frac{c_7}{360} - \frac{1}{630}. \quad (40) \end{aligned}$$

From the system of Eqs. (39) and (40) we obtain

$$a_{65} = \frac{-56c_4c_7 + 32c_4 + 24c_7 - 15}{20160\sigma_6c_5^2(c_7 - c_6)(c_5 - c_4)},$$

while, from Eqs. (28) and (29) we find

$$a_{65} = \frac{\phi_5}{\tau_6} = \frac{5040c_6^2(c_6 - c_4)(c_6 - c_5)(-8c_4 + 3)}{20160c_5^2(c_5 - c_4)(14c_4c_5 - 6(c_4 + c_5) + 3)}.$$

Equating these two values of  $a_{65}$ , we get Eq. (30).

The parameter  $a_{54}$  in the form shown in the solutions (33) is derived as follows:

From the system of Eqs. (39) and (40), we obtain

$$\begin{aligned} (\sigma_5(c_7 - c_5)a_{54} + \sigma_6(c_7 - c_6)a_{64})c_4^2(c_5 - c_4) \\ = \frac{56c_5c_7 - 32c_5 - 24c_7 + 15}{20160}, \quad (41) \end{aligned}$$

and from Eqs. (28) and (29), we obtain

$$\phi_4 = \tau_4 a_{54} + \tau_5 a_{64} = \frac{8c_5 - 3}{20160c_4^2(c_5 - c_4)}. \quad (42)$$

The solution  $a_{54}$  is obtained from the system of Eqs. (41) and (42) as

$$a_{54} = \frac{-c_5^2(c_4(56c_5^2 - 42c_5 + 9) - 3c_5)}{12c_4^2(14c_4^2 - 14c_4 + 3)}$$

and can be written by using Eq. (30) in the form

$$a_{54} = \frac{-54c_4(28c_4^2 - 21c_4 + 3)}{(56c_4^2 - 42c_4 + 9)^4} = \frac{c_5^3(c_5 - c_4)}{c_4^3}.$$

(Received March 17, 1997)

(Accepted July 1, 1997)



**Harumi Ono** was born in 1932. She received her B.S. degree in mathematics from Ochanomizu University in 1954, and her D.E. degree from University of Tokyo in 1985. She had been in Chiba University as an associate professor until March 1997. Her main research interest is numerical analysis. She is a member of IPSJ, JSIAM, and JSAS.