ON RUNGE-KUTTA PROCESSES OF HIGH ORDER

J. C. BUTCHER

(received 28 October 1963)

Introduction

An (explicit) Runge-Kutta process is a means of numerically solving the differential equation

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

at the point $x = x_0+h$, where y, f may be vectors.

The equations defining a ν stage Runge-Kutta process are

$$\begin{aligned}
g_1 &= f(x_0, y_0),\\
g_2 &= f(x_0+c_2 h, y_0+ha_{21} g_1),\\
g_3 &= f(x_0+c_3 h, y_0+h(a_{31} g_1+a_{32} g_2)),\\
&\vdots \qquad\qquad \vdots\\
g_\nu &= f(x_0+c_\nu h, y_0+h(a_{\nu 1} g_1+a_{\nu 2} g_2+\cdots+a_{\nu,\nu-1} g_{\nu-1})),\\
y &= y_0+h(b_1 g_1+b_2 g_2+\cdots+b_\nu g_\nu)
\end{aligned}$$

where

$$\begin{aligned}
c_2 &= a_{21},\\
c_3 &= a_{31}+a_{32},\\
&\vdots \qquad\quad \vdots\\
c_\nu &= a_{\nu 1}+a_{\nu 2}+\cdots a_{\nu,\nu-1},
\end{aligned} \tag{1}$$

and a_{21}, a_{31}, a_{32}, $\cdots$, $a_{\nu,\nu-1}$, b_1, b_2, $\cdots b_\nu$ are a set of parameters which characterize the process. For convenience, we will introduce additional symbols a_{11}, a_{12}, $\ldots$, $a_{\nu\nu}$, c_1 all equal to zero, so that the notation of [1] may be used. In [1] it was shown that the condition that a Runge-Kutta process be accurate to terms of order ρ is

$$\Phi = 1/\gamma,\ r \leqq \rho, \tag{2}$$

where the Φ are certain polynomial expressions in the numbers a_{11}, a_{12}, $\cdots$, b_1, b_2, $\cdots$, b_ν and r, γ are certain integers depending on the form of the particular Φ to which they correspond. r is called the order of Φ.

For example, if

$$\Phi = [_{r-1}\phi]_{r-1} \equiv \sum_{i_1, i_2, \cdots, i_{r-1}=1}^{\nu} b_{i_1} a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_{r-2} i_{r-1}} c_{i_{r-1}}$$

then $\gamma = r!$ and if

$$\Phi = [\phi^{r-1}] \equiv \sum_{i=1}^{\nu} b_i c_i^{r-1}$$

then $\gamma = r$.

As we are assuming $a_{ij} = 0$ if $i \leqq j$, the only terms contributing to $[{}_{r-1}\phi]_{r-1}$ are those in which $i_1 > i_2 > \cdots > i_{r-1} > 1$ so that this expression vanishes identically unless $r \leqq \nu$. Hence we have the well known result that a Runge-Kutta process of order ρ has at least ρ stages. For $\rho = 1, 2, 3, 4$ it is a simple matter to find processes of order ρ with exactly ρ stages. For $\rho = 5$ there exists a process originally due to Kutta [2] but corrected by Nyström [3] with 6 stages; while for $\rho = 6$ there are processes due to Huta [4, 5] (see also [6]) with 8 stages. If $N(\rho)$ is the minimum number of stages necessary for a process of order ρ then we can write

$$N(\rho) = \rho, \ \rho \leqq 4,$$

$$5 \leqq N(5) \leqq 6,$$

$$6 \leqq N(6) \leqq 8.$$

The limits for $N(5)$ and $N(6)$ will each be improved in this paper. We shall show firstly that $N(5) > 5$ (implying $N(5) = 6$) and secondly that $N(6) \leqq 7$. This second result is shown by actually constructing suitable processes.

Although the possibility that $N(6) = 6$ is still open it happens that many details of the proof that $N(5) > 5$ readily generalize and might take a future investigator some way towards deciding this question.

The non-existence of a 5 stage, 5th order process

In this section we shall suppose that a 5 stage, 5th order Runge-Kutta process exists and show that this leads to a contradiction. Unfortunately, a great number of different cases arise and it appears to be necessary to consider these separately, even though the treatments of the different cases have many similarities. However, it is possible to avoid the repetition of many trivial details in treating some cases, by referring back to cases already treated for which the argument is similar.

To clarify the procedure of classification and to make cross-referencing easier, a decimal system of numbering the different cases will be used. For example, case 1 is broken into cases 1.1 and 1.2 and the first of these is broken into cases 1.11 and 1.12, and so on.

We now list the explicit form of the 17 equations of the form (2) when $\nu = \rho = 5$. The range of each summed subscript is from 1 to 5.

(3) $\phi \equiv \sum_i b_i = 1$,

(4) $[\phi] \equiv \sum_i b_i c_i = \frac{1}{2}$,

(5) $[\phi^2] \equiv \sum_i b_i c_i^2 = \frac{1}{3}$,

(6) $[{}_2\phi]_2 \equiv \sum_{i,j} b_i a_{ij} c_j = \frac{1}{6}$,

(7) $[\phi^3] \equiv \sum_i b_i c_i^3 = \frac{1}{4}$,

(8) $[[\phi]\phi] \equiv \sum_{i,j} b_i c_i a_{ij} c_j = \frac{1}{8}$,

(9) $[{}_2\phi^2]_2 \equiv \sum_{i,j} b_i a_{ij} c_j^2 = \frac{1}{12}$,

(10) $[{}_3\phi]_3 \equiv \sum_{i,j,k} b_i a_{ij} a_{jk} c_k = \frac{1}{24}$,

(11) $[\phi^4] \equiv \sum_i b_i c_i^4 = \frac{1}{5}$,

(12) $[[\phi]\phi^2] \equiv \sum_{i,j} b_i c_i^2 a_{ij} c_j = \frac{1}{10}$,

(13) $[[\phi]^2] \equiv \sum_i b_i (\sum_j a_{ij} c_j)^2 = \frac{1}{20}$,

(14) $[[\phi^2]\phi] \equiv \sum_{i,j} b_i c_i a_{ij} c_j^2 = \frac{1}{15}$,

(15) $[[{}_2\phi]_2\phi] \equiv \sum_{i,j,k} b_i c_i a_{ij} a_{jk} c_k = \frac{1}{30}$,

(16) $[{}_2\phi^3]_2 \equiv \sum_{i,j} b_i a_{ij} c_j^3 = \frac{1}{20}$,

(17) $[{}_2[\phi]\phi]_2 \equiv \sum_{i,j,k} b_i a_{ij} c_j a_{jk} c_k = \frac{1}{40}$,

(18) $[{}_3\phi^2]_3 \equiv \sum_{i,j,k} b_i a_{ij} a_{jk} c_k^2 = \frac{1}{60}$,

(19) $[{}_4\phi]_4 \equiv \sum_{i,j,k,l} b_i a_{ij} a_{jk} a_{kl} c_l = \frac{1}{120}$.

We note that (19) can be written in the form

$$b_5 a_{54} a_{43} a_{32} c_2 = \frac{1}{120},$$

so that none of the factors on the left can vanish.

It may happen that some of $b_1, b_2, \cdots, b_5$ may vanish or that some of $c_1, c_2, \cdots, c_5$ may be equal to the same number. In such cases let us suppose that (3), (4), (5), (7), (11) can be written together in the form

$$\sum_{i=1}^{\bar{\nu}} \bar{b}_i \bar{c}_i^{k-1} = \frac{1}{k}, \qquad k = 1, 2, 3, 4, 5,$$

where $\bar{c}_1, \bar{c}_2, \cdots, \bar{c}_{\bar{\nu}}$ are all different and $\bar{b}_1, \bar{b}_2, \cdots, \bar{b}_{\bar{\nu}}$ do not vanish. We will now show that

(a) $\bar{\nu} > 2$, and

(b) if $\bar{\nu} = 3$, $\bar{c}_1 = 0$,

then $\bar{c}_2, \bar{c}_3$ are the pair of numbers $(6-\sqrt{6})/10$, $(6+\sqrt{6})/10$. To prove (a) we suppose that $\bar{\nu} = 2$ so that we have

$$\begin{aligned} 0 &= \sum_{i=1}^{2} \bar{b}_i (\bar{c}_i - \bar{c}_1)^2 (\bar{c}_i - \bar{c}_2)^2 \\ &= \int_0^1 (t-\bar{c}_1)^2 (t-\bar{c}_2)^2 \mathrm{d}t > 0, \end{aligned}$$

a contradiction.

To prove (b) we find in the same sort of way that

$$\int_0^1 t(t-\bar{c}_2)(t-\bar{c}_3)\,\mathrm{d}t = 0,$$

$$\int_0^1 t^2(t-\bar{c}_2)(t-\bar{c}_3)\,\mathrm{d}t = 0,$$

a pair of linear equations in $(\bar{c}_2+\bar{c}_3)$ and $\bar{c}_2\bar{c}_3$. The solution turns out to be $\bar{c}_2+\bar{c}_3 = 6/5$, $\bar{c}_2\bar{c}_3 = 3/10$ and the result follows.

The statements (a), (b) will be referred to by these letters in a number of places in the following.

It is convenient to define

$$p_i = \sum_j a_{ij}c_j - \tfrac{1}{2}c_i^2, \tag{20}$$

$$q_i = \sum_j a_{ij}c_j^2 - \tfrac{1}{3}c_i^3, \tag{21}$$

$$r_j = \sum_i b_i a_{ij} - b_j(1-c_j), \tag{22}$$

$$s_j = \sum_i b_i c_i a_{ij} - \tfrac{1}{2}b_j(1-c_j^2), \tag{23}$$

so that p_1, q_1 both vanish while $p_2 = -\frac{1}{2}c_2^2$ and $q_2 = -\frac{1}{3}c_2^3$ do not vanish. The following equations are now easily verified making use of (3)—(19),

$$\sum_i b_i p_i = 0, \tag{24}$$

$$\sum_i b_i c_i p_i = 0, \tag{25}$$

$$\sum_i b_i c_i^2 p_i = 0, \tag{26}$$

$$\sum_i b_i p_i^2 = 0, \tag{27}$$

$$\sum_i b_i q_i = 0, \tag{28}$$

$$\sum_i b_i c_i q_i = 0, \tag{29}$$

$$\sum_i r_i = 0, \tag{30}$$

$$\sum_i r_i c_i = 0, \tag{31}$$

$$\sum_i r_i c_i^2 = 0, \tag{32}$$

$$\sum_i r_i c_i^3 = 0, \tag{33}$$

$$\sum_i r_i p_i = 0, \tag{34}$$

$$\sum_i r_i c_i p_i = 0, \tag{35}$$

$$\sum_{i,j} r_i a_{ij} p_j = 0, \tag{36}$$

$$\sum_i s_i = 0, \tag{37}$$

$$\sum_i s_i c_i = 0, \tag{38}$$

$$\sum_i s_i c_i^2 = 0, \tag{39}$$

$$\sum_i s_i p_i = 0. \tag{40}$$

For example, to verify (36) we have

$$\begin{aligned}\sum_{i,j} r_i a_{ij} p_j &= \sum_{i,j,k,l} (b_i a_{ij} - b_j(1-c_j)) a_{jk}(a_{kl}c_l - \tfrac{1}{2}c_k^2) \\ &= [_4\phi]_4 - \tfrac{1}{2}[_3\phi^2]_3 - [_3\phi]_3 + \tfrac{1}{2}[_2\phi^2]_2 + [\phi[_2\phi]_2] - \tfrac{1}{2}[\phi[\phi^2]] \\ &= \tfrac{1}{120} - \tfrac{1}{2}\cdot\tfrac{1}{60} - \tfrac{1}{24} + \tfrac{1}{2}\cdot\tfrac{1}{12} + \tfrac{1}{30} - \tfrac{1}{2}\cdot\tfrac{1}{15} = 0.\end{aligned}$$

We now distinguish two cases according as no two of c_2, c_3, c_4, c_5 are equal; or otherwise,

Case 1: No two of c_2, c_3, c_4, c_5 are equal. This case will be broken into two subcases:

Case 1.1: Not all of b_2p_2, b_3p_3, b_4p_4, b_5p_5 vanish. Regarding (24)—(27) as a set of linear equations in these variables we see that the following determinantal equation holds

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ c_2 & c_3 & c_4 & c_5 \\ c_2^2 & c_3^2 & c_4^2 & c_5^2 \\ p_2 & p_3 & p_4 & p_5 \end{vmatrix} = 0.$$

Since c_2, c_3, c_4, c_5 are all different it follows that

$$p_i = tc_i^2 + uc_i + v, \qquad i = 2, 3, 4, 5, \tag{41}$$

where t, u, v are certain constants.

Substituting (41) into (25), (26) and making use of (4), (5), (7), (11) we see that

$$\frac{t}{4} + \frac{u}{3} + \frac{v}{2} = 0,$$

$$\frac{t}{5} + \frac{u}{4} + \frac{v}{3} = 0,$$

so that v does not vanish, for if it did, so would t, u and hence p_2.

We now distinguish two further cases,

Case 1.11: Not all of r_2, r_3, r_4, r_5 vanish. Making use of (31)—(34) we see that

$$p_i = t'c_i^3 + u'c_i^2 + v'c_i, \qquad i = 2, 3, 4, 5,$$

so that using also (41) we see that c_2, c_3, c_4, c_5 satisfy a cubic equation. This is possible only if the coefficients of this equation vanish. In particular this implies $v = 0$ and leads to a contradiction.

Case 1.12: $r_2 = r_3 = r_4 = r_5 = 0$. Since $r_5 = -b_5(1-c_5)$ and $b_5 \neq 0$ (19) it follows that $c_5 = 1$. Hence $s_5 = -\frac{1}{2}b_5(1-c_5^2) = 0$. We now distinguish two further cases,

Case 1.121: Not all of s_2, s_3, s_4 vanish. Using (38)—(40) we see that

$$p_i = t''c_i^2 + u''c_i, \qquad i = 2, 3, 4,$$

so that using (41) (compare case 1.11) we have a contradiction.

Case 1.122: $s_2 = s_3 = s_4 = 0$. Since $r_4 = b_5a_{54} - b_4(1-c_4)$, $s_4 = b_5c_5a_{54} - \frac{1}{2}b_4(1-c_4^2)$ and since $c_5 = 1$ we see that $b_4(1-c_4) = \frac{1}{2}b_4(1-c_4^2)$. Hence, since $b_4(1-c_4) = b_5a_{54} \neq 0$ it follows that $c_4 = 1$, a contradiction. We now return to

Case 1.2: $b_2p_2 = b_3p_3 = b_4p_4 = b_5p_5 = 0$. Since $p_2 \neq 0$, $b_2 = 0$. However, $b_5 \neq 0$ and (a) not both of b_3, b_4 are zero. We distinguish 3 cases,

Case 1.21: $b_3 = 0$, $b_4 \neq 0$. In this case $p_1 = p_4 = p_5 = 0$ so that (34), (40) take the forms

$$b_4(a_{43}p_3+a_{42}p_2)+b_5(a_{53}p_3+a_{52}p_2) = 0,$$
$$b_4c_4(a_{43}p_3+a_{42}p_2)+b_5c_5(a_{53}p_3+a_{52}p_2) = 0.$$

Since $c_4 \neq c_5$, it follows that $a_{43}p_3+a_{42}p_2 = a_{53}p_3+a_{52}p_2 = 0$ so that (36) reduces to

$$r_3a_{32}p_2 = 0$$

implying $r_3 = 0$. Hence from (31), (32), (33) it follows that $r_2 = r_4 = r_5 = 0$ which is impossible (compare case 1.12) as $c_5 \neq 1$ (b).

Case 1.22: $b_3 \neq 0$, $b_4 = 0$. In this case $p_1 = p_3 = p_5 = 0$ and (using (34), (40); compare case 1.21) $b_3a_{32}p_2 = 0$ which is impossible.

Case 1.23: $b_3 \neq 0$, $b_4 \neq 0$. $p_1 = p_3 = p_4 = p_5 = 0$ so that from (34), $r_2 = 0$. Hence (compare case 1.21), $r_3 = r_4 = r_5 = 0$. We now deduce (40) that $s_2 = 0$ and (compare case 1.12) that $s_5 = 0$. Using (38), (39) we see that $s_3 = s_4 = 0$ leading (compare case 1.122) to a contradiction.

This concludes the treatment of case 1. We now consider

Case 2: Two of $c_2, \cdots, c_5$ are equal. First we distinguish the cases where c_5 is or is not different from the other c_i.

Case 2.1: c_5 does not equal another c_i; but suppose $c_I = c_J$. Using (30)—(33) we see that $r_5 = 0$ (hence, compare case 1.12, $c_5 = 1$, $s_5 = 0$) and by (a), (b) no two c_i can be equal other than $c_I = c_J$. Hence, using (30)—(32), (37)—(39) we see that $r_I+r_J = s_I+s_J = 0$. We now distinguish two further cases:

Case 2.11: $r_I = r_J = s_I = s_J = 0$. In this case it would follow (compare case 1.122) that $c_4 = 1 = c_5$, a contradiction.

Case 2.12: $r_I \neq 0$ or $s_I \neq 0$. However, $r_i = s_i = 0$ if $i \neq I$, $i \neq J$ so that either (34) of (40) implies $p_I = p_J$. If

Case 2.121: $p_I = p_J = 0$; it follows (making use of (24)—(26), (b)) that $p_2 = 0$, a contradiction. On the other hand, if

Case 2.122: $p_I = p_J \neq 0$; we see that neither of I, J is 1. It follows from (24)—(26) that $b_I+b_J = 0$ which by (b) is impossible. We now return to

Case 2.2: $c_5 = c_I (I < 5)$. We distinguish 3 cases (by (a) there are no more).

Case 2.21: No other equality exists between $c_1, c_2, \cdots, c_5$. In this case we may assume $I \neq 1$ so that we find, using (24)—(27) that $b_Ip_I+b_5p_5 = b_Ip_I^2+b_5p_5^2 = 0$. Hence, either $p_I = p_5$ and $b_I+b_5 = 0$ or $b_Ip_I = b_5p_5 = 0$.

Case 2.211: $p_I = p_5$, $b_I+b_5 = 0$. Either

Case 2.2111: $I \neq 2$; in which case $b_2 = 0$ and by (a) we have a contradiction or

Case 2.2112: $I = 2$; from (24)—(26), $p_3 = p_4 = 0$. Thus from (34), (40) $r_2 + r_5 = s_2 + s_5 = 0$ and from (31), (32), (38), (39) it follows that $r_3 = r_4 = s_3 = s_4 = 0$. Hence, (36) reduces to

$$r_5 a_{52} p_2 = 0.$$

Since $p_2 \neq 0$ and $r_5 \neq 0$ (compare case 1.122) it follows that $a_{52} = 0$. The equations $r_2 + r_5 = s_2 + s_5 = 0$ reduce to the forms

$$b_3 a_{32} + b_4 a_{42} = 0,$$
$$b_3 c_3 a_{32} + b_4 c_4 a_{42} = 0,$$

so that $b_3 a_{32} = 0$; a contradiction by (a) and (19). We now return to

Case 2.212: $b_I p_I = b_5 p_5 = 0$. Hence $p_5 = 0$. From (30)—(33) we see that $r_I + r_5 = 0$. We now consider the cases $r_5 = 0$, $r_5 \neq 0$ separately.

Case 2.2121: $r_5 = 0$. It follows (compare case 1.12) that $c_5 = 1$, $s_5 = 0$. We consider the cases $I = 2$, $I \neq 2$ separately.

Case 2.21211: $I = 2$; in which case, by (24)—(26) and (b), $p_1 = p_3 = p_4 = p_5 = 0$; so that using (40), followed by (37)—(39) we can deduce that $s_1 = s_2 = s_3 = s_4 = s_5 = 0$, leading (compare case 1.122) to a contradiction.

Case 2.21212: $I \neq 2$. By (24)—(26) it follows that $b_2 p_2 = 0$ and hence $b_2 = 0$. This is impossible (b). We now return to

Case 2.2122: $r_5 \neq 0$. Since $r_i = 0$ $(i \neq I, i \neq 5)$ follows from (30)—(33); (34) simplifies to the form $p_I = p_5$ so that $p_I = 0$ (thus excluding the possibility $I = 2$). It now follows from (40) that $s_2 = 0$ so that (37)—(39) imply $s_I + s_5 = 0$. We now consider two alternatives:

Case 2.21221: $I \neq 4$. In this case $r_4 = s_4 = 0$ so that $c_5 = \frac{1}{2}(1 + c_4)$. However, by (b) c_4, c_5 are two of the numbers 0, $(6 - \sqrt{6})/10$, $(6 + \sqrt{6})/10$ so we have a contradiction.

Case 2.21222: $I = 4$. We have $r_4 + r_5 = s_4 + s_5 = 0$ leading to $c_4 = c_5 = 1$, a contradiction as $r_5 \neq 0$. We now return to the second subcase of case 2.2.

Case 2.22: $c_5 = c_I$, $c_J = c_K$ where I, J, K, 5 are all different. It is clear that one of I, J, K must be 2, for otherwise we get the same contradiction as in case 2.2111. Let us suppose that $c_2 = c_L$ $(L > 2)$. We have (compare case 2.21) that either

Case 2.221: $b_2 + b_L = 0$, which is impossible (a), or

Case 2.222: $b_2 p_2 = b_L p_L = 0$. We have $b_2 = 0$ and (a) $b_L \neq 0$ so that $p_L = 0$. We distinguish three further cases:

Case 2.2221: $L = 3$. $c_2 = c_3$ and in addition $c_5 = c_1$ or $c_5 = c_4$. We treat these cases separately:

Case 2.22211: $c_5 = c_1 = 0$. In this case we deduce from (30)—(32), (37)—(39) that $r_4 = s_4 = 0$; hence $b_4(1-c_4) = b_5 a_{54} \neq 0$, $\frac{1}{2}b_4(1-c_4^2) = b_5 c_5 a_{54} = 0$. It follows that $c_4 = -1$ which is impossible (b).

Case 2.22212: $c_5 = c_4$. Using (28), (29) we deduce $q_3 = 0$. Since also $p_3 = 0$, we have $a_{32}c_2 = \frac{1}{2}c_3^2$, $a_{32}c_2^2 = \frac{1}{3}c_3^3$ which is impossible as $c_2 = c_3 \neq 0$. We now consider the second subcase of case 2.222.

Case 2.2222: $L = 4$. $c_2 = c_4$ and in addition $c_5 = c_1$ or $c_5 = c_3$.

Case 2.22221: $c_5 = c_1 = 0$. From (24)—(27) and (a) we deduce $b_2 = p_3 = p_4 = p_5 = 0$. From (34), (40) we deduce $r_2 = s_2 = 0$ and hence, from (30)—(32), (37)—(39) we deduce $r_4 = s_4 = 0$ leading (compare case 2.22211) to a contradiction.

Case 2.22222: $c_5 = c_3$. We proceed as in case 2.21 (the possibilities corresponding to cases 2.211, 2.2121 are immediately rejected by (a) and (b)) until we reach the situation corresponding to case 2.2122. We now have $b_3 p_3 = p_5 = 0$, $r_3 + r_5 = 0$, $r_5 \neq 0$. From (34), (35) we now deduce $p_3 = p_5 = 0$. Using (30)—(34), (37)—(40) we can now deduce that $r_4 = s_4 = 0$ giving (compare case 2.21221) a contradiction. We now return to the third and last sub-case of case 2.222.

Case 2.2223: $L = 5$. Making use of (30)—(32) we see that $r_2 = r_5 = 0$ (which (b) is impossible) or, using also (34), (35) that $p_2 = p_5$ and using also (24)—(26) that $b_2 + b_5 = 0$ which is impossible by (a). We now return to the final subcase of case 2.2.

Case 2.23: $c_5 = c_I = c_J$. Clearly one of I, J (say I) equals 2, for otherwise we could deduce from (24)—(26) that $b_2 = 0$. We consider the three possible values of J separately.

Case 2.231: $J = 1$. This implies $c_2 = 0$ which is impossible.

Case 2.232: $J = 3$. A contradiction results (compare case 2.21221).

Case 2.233: $J = 4$. From (24), (25), (28), (29) and (a) it follows that $p_3 = q_3 = 0$. Thus $a_{32}c_2 = \frac{1}{2}c_3^2$, $a_{32}c_2^2 = \frac{1}{3}c_3^3$ implying $3c_2 = 2c_3$. However, by (b) c_2, c_3 are each one of the numbers $(6-\sqrt{6})/10$, $(6+\sqrt{6})/10$, a contradiction.

This completes the treatment of the different cases. Each case leads to a contradiction, so the result is proved that there is no 5 stage, 5th order Runge-Kutta process.

We remark finally that case 1.1 in this proof generalizes almost immediately to the corresponding case of a proof that $N(\rho) > \rho$ for $\rho > 5$. Unfortunately, cases 1.2 and 2 as they are handled here do not generalize in the same sort of way, so the more general question remains open.

7 stage, 6th order processes

In searching for such processes we are guided by the analysis of the previous section to make the following assumptions:

$$\sum_{i=1}^{7} b_i a_{ij} = b_j(1-c_j), \qquad j = 1, 2, \cdots, 7, \tag{42}$$

$$\sum_{j=1}^{7} a_{ij} c_j = \tfrac{1}{2} c_i^2, \qquad i = 3, 4, \cdots, 7, \tag{43}$$

$$b_2 = 0. \tag{44}$$

Assuming (42) has the effect of ensuring that all first degree (in the sense of [1]) elementary weights have the correct values if the same is true for other elementary weights. This result is proved in [6]. Similarly, we need not concern ourselves with elementary weights of the form $\Phi = [[\phi]\Phi_1\Phi_2 \cdots \Phi_s]$ as the corresponding γ is given by $\gamma = 2r\gamma_1\gamma_2 \cdots \gamma_s$ (by (31) of [1]) and for $\Phi' = [\phi^2\Phi_1\Phi_2 \cdots \Phi_s]$, $\gamma' = r\gamma_1\gamma_2 \cdots \gamma_s = \frac{1}{2}\gamma$, where r is the order of Φ and Φ' and $\gamma_1, \gamma_2, \cdots, \gamma_s$ correspond to $\Phi_1, \Phi_2, \cdots, \Phi_s$. If $\Phi_i = \sum_{j=1}^{7} b_j \chi_{ij}$ where χ_{ij} is independent of $b_1, b_2, \cdots, b_7$ we have (if (43) and (44) are assumed)

$$\begin{aligned}
\Phi &= \sum_{i,j,k_1,\cdots,k_s=1}^{7} b_i a_{ij} c_j a_{ik_1} \chi_{1k_1} a_{ik_2} \chi_{2k_2} \cdots a_{ik_s} \chi_{sk_s} \\
&= \tfrac{1}{2} \sum_{i,k_1,\cdots,k_s=1}^{7} b_i c_i^2 a_{ik_1} \chi_{1k_1} a_{ik_2} \chi_{2k_2} \cdots a_{ik_s} \chi_{sk_s} \\
&= \tfrac{1}{2}\Phi' \\
&= \frac{1}{2\gamma'} \\
&= \frac{1}{\gamma}, \text{ the correct value.}
\end{aligned}$$

The equations that still remain to be solved will now be listed. The range of each summed subscript is from 1 to 7.

(45) $\phi \equiv \sum_i b_i = 1$,

(46) $[\phi] \equiv \sum_i b_i c_i = \frac{1}{2}$,

(47) $[\phi^2] \equiv \sum_i b_i c_i^2 = \frac{1}{3}$,

(48) $[\phi^3] \equiv \sum_i b_i c_i^3 = \frac{1}{4}$,

(49) $[\phi^4] \equiv \sum_i b_i c_i^4 = \frac{1}{5}$,

(50) $[\phi^5] \equiv \sum_i b_i c_i^5 = \frac{1}{6}$,

$$[[_2\phi]_2\phi] \equiv \sum_{i,j,k} b_i c_i a_{ij} a_{jk} c_k = \tfrac{1}{30}, \tag{51}$$

$$[[\phi^2]\phi] \equiv \sum_{i,j} b_i c_i a_{ij} c_j^2 = \tfrac{1}{15}, \tag{52}$$

$$[[[\phi]\phi]\phi] \equiv \sum_{i,j,k} b_i c_i a_{ij} c_j a_{jk} c_k = \tfrac{1}{48}, \tag{53}$$

$$[[\phi^3]\phi] \equiv \sum_{i,j} b_i c_i a_{ij} c_j^3 = \tfrac{1}{24}, \tag{54}$$

$$[[_2\phi]_2\phi^2] \equiv \sum_{i,j,k} b_i c_i^2 a_{ij} a_{jk} c_k = \tfrac{1}{36}, \tag{55}$$

$$[[\phi^2]\phi^2] \equiv \sum_{i,j} b_i c_i^2 a_{ij} c_j^2 = \tfrac{1}{18}, \tag{56}$$

$$[[_3\phi]_3\phi] \equiv \sum_{i,j,k,l} b_i c_i a_{ij} a_{jk} a_{kl} c_l = \tfrac{1}{144}, \tag{57}$$

$$[[_2\phi^2]_2\phi] \equiv \sum_{i,j,k} b_i c_i a_{ij} a_{jk} c_k^2 = \tfrac{1}{72}. \tag{58}$$

Making use of (43) and observing also that $\sum_{j=1}^{7} a_{ij} c_j = \frac{1}{2} c_i^2$ in the case $i = 1$ but not in the case $i = 2$ (we assume, of course, that $c_2 \neq 0$ for otherwise we would in effect be searching for a process with only 6 stages; for the same reason we shall assume that $b_7 \neq 0$ so that, by (42), $c_7 = 1$) it follows that (51), (53), (55), (57) are equivalent respectively to (52), (54), (56), (58) if (and only if) the following equations are satisfied:

$$\sum_i b_i c_i a_{i2} = 0, \tag{59}$$

$$\sum_i b_i c_i^2 a_{i2} = 0, \tag{60}$$

$$\sum_{i,j} b_i c_i a_{ij} a_{j2} = 0. \tag{61}$$

Before attempting to solve (42)—(50), (52), (54), (56), (58)—(61) we might consider what appears to be gained by making the assumptions (42), (43), (44). For a 7 stage process there are 28 parameters a_{21}, a_{31}, a_{32}, $\cdots$, a_{76}, b_1, b_2, $\cdots$, b_7 to be chosen and there are 37 conditions for them to satisfy if the process is to be of order 6. On the other hand, there are 6 conditions to be satisfied from equations (43), (44) and 13 more in the equations which follow. (42) contains only 5 independent conditions since (making use of (43)—(47))

$$\sum_j \left(\sum_i b_i a_{ij} - b_j(1-c_j)\right) = \sum_i b_i c_i - \sum_j b_j + \sum_j b_j c_j = 0,$$

and

$$\sum_j c_j \left(\sum_i b_i a_{ij} - b_j(1-c_j)\right) = \tfrac{1}{2}\sum_i b_i c_i^2 - \sum_j b_j c_j + \sum b_j c_j^2 = 0.$$

Hence, we now have only 24 conditions to satisfy in the 28 parameters so that, far from suspecting that no solution exists, we may hope for some flexibility in the choice of a solution.

We now replace the equations (52), (54), (56), (58), (59), (60), (61) by a more convenient set such that the 24 equations we have to satisfy are replaced by an equivalent set. The modified versions of the 7 equations are

$$(62)\qquad \sum_{i,j} b_i(1-c_i)a_{ij}c_j(c_j-c_3) = \tfrac{1}{60} - \frac{c_3}{24},$$

$$(63)\qquad \sum_{i,j} b_i(1-c_i)a_{ij}c_j(c_j-c_3)(c_j-c_4) = \tfrac{1}{120} - \frac{c_3+c_4}{60} + \frac{c_3c_4}{24},$$

$$(64)\qquad \sum_{i,j} b_i(1-c_i)(c_6-c_i)a_{ij}c_j(c_j-c_3) = -\tfrac{1}{90} + \frac{c_3}{40} + \frac{c_6}{60} - \frac{c_3c_6}{24},$$

$$(65)\qquad \sum_{i,j,k} b_i(1-c_i)a_{ij}a_{jk}c_k(c_k-c_3) = \tfrac{1}{360} - \frac{c_3}{120},$$

$$(66)\qquad \sum_{i} b_i(1-c_i)a_{i2} = 0,$$

$$(67)\qquad \sum_{i} b_i(1-c_i)(c_6-c_i)a_{i2} = 0$$

$$(68)\qquad \sum_{i,j} b_i(1-c_i)a_{ij}a_{j2} = 0.$$

These equations are easily verified. For example, the left hand side of (64) is

$$\begin{aligned}[\phi^2[\phi^2]] - (1+c_6)[\phi[\phi^2]] + c_6[{}_2\phi^2]_2 - c_3[\phi^2[\phi]] + c_3(1+c_6)[\phi[\phi]] - c_3c_6[{}_2\phi]_2 \\ = \tfrac{1}{18} - (1+c_6)\tfrac{1}{15} + c_6\tfrac{1}{12} - c_3\tfrac{1}{10} + c_3(1+c_6)\tfrac{1}{8} - c_3c_6\tfrac{1}{6} \\ = -\tfrac{1}{90} + \frac{c_3}{40} + \frac{c_6}{60} - \frac{c_3c_6}{24}.\end{aligned}$$

Furthermore, it can be seen that the original equations can be deduced from these ones. The advantage of these transformed equations is that the numbers of terms actually occurring on the left hand sides are reduced. Bearing in mind (66), (67), (68) we can write (62), (63), (64), (65) in the forms

$$(69)\quad b_6(1-c_6)[a_{65}c_5(c_5-c_3)+a_{64}c_4(c_4-c_3)]+b_5(1-c_5)a_{54}c_4(c_4-c_3) = \tfrac{1}{60} - \frac{c_3}{24},$$

$$(70)\quad b_6(1-c_6)a_{65}c_5(c_5-c_3)(c_5-c_4) = \tfrac{1}{120} - \frac{c_3+c_4}{60} + \frac{c_3c_4}{24},$$

$$(71)\quad b_5(1-c_5)(c_6-c_5)a_{54}c_4(c_4-c_3) = -\tfrac{1}{90} + \frac{c_3}{40} + \frac{c_6}{60} - \frac{c_3c_6}{24},$$

$$(72)\quad b_6(1-c_6)a_{65}a_{54}c_4(c_4-c_3) = \tfrac{1}{360} - \frac{c_3}{120}.$$

We shall use (72) with (70), (71) to obtain a condition on the numbers $c_2, c_3, \cdots, c_6$ and then solve for a_{65}, a_{54} from (70), (71), (72). After this a_{64} can be found from (69) and then the remaining parameters from the remaining equations.

We write

$$\theta = b_5(c_5-1)(c_5-c_6)(c_5-c_4)(c_5-c_3)c_5$$

so that the product of the left hand sides of (70), (71) is the product of θ with the left hand side of (72). We have therefore,

$$(73)\quad \left(\tfrac{1}{360} - \frac{c_3}{120}\right)\theta = \left(\tfrac{1}{120} - \frac{c_3+c_4}{60} + \frac{c_3c_4}{24}\right)\left(-\tfrac{1}{90} + \frac{c_3}{40} + \frac{c_6}{60} + \frac{c_3c_6}{24}\right).$$

To evaluate θ we solve the system (45)—(50) to obtain

$$\theta = \frac{c_3c_4c_6}{6} - \frac{c_3c_4+c_4c_6+c_3c_6}{12} + \frac{c_3+c_4+c_6}{20} - \tfrac{1}{30}.$$

We substitute this into (73) and simplify to obtain

$$(74)\qquad (1-c_6)[2c_4-(1+10c_4)c_3+15c_4c_3^2] = 0.$$

The possibility $c_6 = 1$ must be rejected as it leads, by (72) to the value $c_3 = \frac{1}{3}$, so that (69), (71) take the forms

$$b_5(1-c_5)a_{54}c_4(c_4-\tfrac{1}{3}) = \tfrac{1}{360},$$
$$b_5(1-c_5)^2a_{54}c_4(c_4-\tfrac{1}{3}) = 0,$$

and clearly, these cannot be satisfied.

Since c_4 occurs linearly in the other factor, it is convenient to select a value for c_3 and then to solve for c_4. On grounds of simplicity, the most acceptable value of c_3 is $c_3 = \frac{2}{3}$ leading to the value $c_4 = \frac{1}{3}$. However, we shall also consider the values $c_3 = (5\pm\sqrt{5})/10$, $c_4 = (5\mp\sqrt{5})/10$ as these are capable of yielding processes of interest.

With $c_3 = \frac{2}{3}$, $c_4 = \frac{1}{3}$, there is very little restriction on the values of c_2, c_5, c_6. If we assume $c_5 = c_6$ then from (71) we see that $c_6 = \frac{1}{2}$. We shall examine this case in some detail.

From (45)—(50) we find, $b_1 = b_7 = \frac{11}{120}$, $b_3 = b_4 = \frac{27}{40}$, $b_5+b_6 = -\frac{8}{15}$. We shall write $b_6 = -\lambda/15$ so that $b_5 = (\lambda-8)/15$, and $c_2 = \mu$. In what follows it will be assumed that $\lambda \neq 0$, $\mu \neq 0$. We solve (70) for a_{65}, (72) for a_{54} and (69) for a_{64}. The results are

$$a_{54} = -\tfrac{3}{8}, \quad a_{64} = -\tfrac{3}{8} - \frac{3}{2\lambda}, \quad a_{65} = \frac{2}{\lambda}.$$

Using (43) with $i = 3$ followed by (67) yields the values

$$a_{32} = \frac{2}{9\mu}, \quad a_{42} = \frac{1}{9\mu}.$$

We now write $a_{52} = u/\mu$, $a_{62} = v/\mu$ and use (43) with $i = 4, 5, 6$, to find

$$a_{43} = -\tfrac{1}{12}, \quad a_{53} = \tfrac{3}{8} - \frac{3u}{2}, \quad a_{63} = \tfrac{3}{8} - \frac{3v}{2} - \frac{3}{4\lambda},$$

to find the values of u, v we use (66), (68) to obtain the equations

$$(8-\lambda)u+\lambda v = 3,$$
$$(8-4\lambda)u+4\lambda v = 3,$$

with solution $u = v = \frac{3}{8}$.

These values can now be substituted into the expressions for a_{52}, a_{62}, a_{53}, a_{63}. Finally, we use (1) and (42) to find $a_{21}, a_{31}, \cdots, a_{71}$, $a_{72}, \cdots, a_{76}$.

It is convenient to specify a Runge-Kutta process by writing the parameters in an array as follows

$$\begin{array}{c|cccccc} 0 & & & & & & \\ c_2 & a_{21} & & & & & \\ c_3 & a_{31} & a_{32} & & & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \\ c_\nu & a_{\nu 1} & a_{\nu 2} & \cdots & a_{\nu,\nu-1} & & \\ \hline & b_1 & b_2 & \cdots & b_{\nu-1} & b_\nu & \end{array}.$$

In the present case the array is

$$\begin{array}{c|cccccccc} 0 & & & & & & & & \\ \mu & \mu & & & & & & & \\ \frac{2}{3} & \frac{2}{3}-\frac{2}{9\mu} & \frac{2}{9\mu} & & & & & & \\ \frac{1}{3} & \frac{5}{12}-\frac{1}{9\mu} & \frac{1}{9\mu} & -\frac{1}{12} & & & & & \\ \frac{1}{2} & \frac{17}{16}-\frac{3}{8\mu} & \frac{3}{8\mu} & -\frac{3}{16} & -\frac{3}{8} & & & & \\ \frac{1}{2} & \frac{17}{16}-\frac{3}{8\mu}+\frac{1}{4\lambda} & \frac{3}{8\mu} & -\frac{3}{16}-\frac{3}{4\lambda} & -\frac{3}{8}-\frac{3}{2\lambda} & \frac{2}{\lambda} & & \\ 1 & -\frac{27}{44}+\frac{3}{11\mu} & -\frac{3}{11\mu} & \frac{63}{44} & \frac{18}{11} & \frac{4\lambda-16}{11} & -\frac{4\lambda}{11} & \\ \hline & \frac{11}{120} & 0 & \frac{27}{40} & \frac{27}{40} & \frac{\lambda-8}{15} & -\frac{\lambda}{15} & \frac{11}{120} \end{array}.$$

For example, if we choose $\lambda = 4$, $\mu = \frac{1}{3}$ the array is

0						
$\frac{1}{3}$	$\frac{1}{3}$					
$\frac{2}{3}$	0	$\frac{2}{3}$				
$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{3}$	$-\frac{1}{12}$			
$\frac{1}{2}$	$-\frac{1}{16}$	$\frac{9}{8}$	$-\frac{3}{16}$	$-\frac{3}{8}$		
$\frac{1}{2}$	0	$\frac{9}{8}$	$-\frac{3}{8}$	$-\frac{3}{4}$	$\frac{1}{2}$	
1	$\frac{9}{44}$	$-\frac{9}{11}$	$\frac{63}{44}$	$\frac{18}{11}$	0	$-\frac{16}{11}$
	$\frac{11}{120}$	0	$\frac{27}{40}$	$\frac{27}{40}$	$-\frac{4}{15}$	$-\frac{4}{15}$	$\frac{11}{120}$

and we have a process which seems to be suitable for practical use.

Other choices of c_2, c_5, c_6 lead to processes just as suitable for practical computations and we content ourselves with writing down the arrays for just two of them. They are

0							
1	1						
$\frac{2}{3}$	$\frac{4}{9}$	$\frac{2}{9}$					
$\frac{1}{3}$	$\frac{11}{36}$	$\frac{1}{9}$	$-\frac{1}{12}$				
$-\frac{1}{3}$	$\frac{151}{36}$	$\frac{29}{9}$	$-\frac{7}{4}$	-6			
$\frac{4}{3}$	$-\frac{112}{9}$	$-\frac{116}{9}$	$\frac{32}{3}$	18	-2		
1	$-\frac{5}{4}$	$-\frac{29}{23}$	$\frac{397}{276}$	$\frac{152}{69}$	$-\frac{10}{69}$	$\frac{1}{69}$	
	$\frac{23}{160}$	0	$\frac{29}{80}$	$\frac{29}{80}$	$-\frac{1}{160}$	$-\frac{1}{160}$	$\frac{23}{160}$,

0							
$\frac{1}{2}$	$\frac{1}{2}$						
$\frac{2}{3}$	$\frac{2}{9}$	$\frac{4}{9}$					
$\frac{1}{3}$	$\frac{7}{36}$	$\frac{2}{9}$	$-\frac{1}{12}$				
$\frac{5}{6}$	$-\frac{35}{144}$	$-\frac{55}{36}$	$\frac{35}{48}$	$\frac{15}{8}$			
$\frac{1}{6}$	$-\frac{1}{360}$	$-\frac{11}{36}$	$-\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{10}$		
1	$-\frac{41}{260}$	$\frac{22}{13}$	$\frac{43}{156}$	$-\frac{118}{39}$	$\frac{32}{195}$	$\frac{80}{39}$	
	$\frac{13}{200}$	0	$\frac{11}{40}$	$\frac{11}{40}$	$\frac{4}{25}$	$\frac{4}{25}$	$\frac{13}{200}$.

We now consider the choice of values, $c_3 = (5 \pm \sqrt{5})/10$, $c_4 = (5 \mp \sqrt{5})/10$. We also choose $c_5 = c_3$, $c_2 = c_6 = c_4$ and we notice that both sides of (70) vanish. From (45)$-$(50), we find $b_1 = b_7 = \frac{1}{12}$, $b_3 + b_5 = b_4 + b_6 = \frac{5}{12}$. Although there is some freedom in the choice of b_5, b_6 we assume each of these equals $\frac{5}{12}$ so that $b_3 = b_4 = b_2 = 0$. We now find a_{54} from (71), a_{65}

from (72) and the rest of the parameters as for the previous cases. The final array is

$$
\begin{array}{c|cccccc}
0 & & & & & & \\
\frac{5\mp\sqrt{5}}{10} & \frac{5\mp\sqrt{5}}{10} & & & & & \\
\frac{5\pm\sqrt{5}}{10} & \mp\frac{\sqrt{5}}{10} & \frac{5\pm 2\sqrt{5}}{10} & & & & \\
\frac{5\mp\sqrt{5}}{10} & \frac{-15\pm 7\sqrt{5}}{20} & \frac{-1\pm\sqrt{5}}{4} & \frac{15\mp 7\sqrt{5}}{10} & & & \\
\frac{5\pm\sqrt{5}}{10} & \frac{5\mp\sqrt{5}}{60} & 0 & \frac{1}{6} & \frac{15\pm 7\sqrt{5}}{60} & & \\
\frac{5\mp\sqrt{5}}{10} & \frac{5\pm\sqrt{5}}{60} & 0 & \frac{9\mp 5\sqrt{5}}{12} & \frac{1}{6} & \frac{-5\pm 3\sqrt{5}}{10} & \\
1 & \frac{1}{6} & 0 & \frac{-55\pm 25\sqrt{5}}{12} & \frac{-25\mp 7\sqrt{5}}{12} & 5\mp 2\sqrt{5} & \frac{5\pm\sqrt{5}}{2} \\
\hline
 & \frac{1}{12} & 0 & 0 & 0 & \frac{5}{12} & \frac{5}{12} \quad \frac{1}{12}.
\end{array}
$$

This process is interesting in that it is an explicit counterpart of a certain 4 stage implicit process [7, table III]. Since the classical 4th order Runge-Kutta process

$$
\begin{array}{c|cccc}
0 & & & & \\
\frac{1}{2} & \frac{1}{2} & & & \\
\frac{1}{2} & 0 & \frac{1}{2} & & \\
1 & 0 & 0 & 1 & \\
\hline
 & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}
\end{array}
$$

may be thought of as a counterpart of the corresponding 3 stage implicit process, the possibility is naturally suggested that counterparts exist for higher order implicit processes as well. Since implicit processes can be found of arbitrarily high order [7, 8] such a relationship would be valuable in yielding high order explicit processes.

References

[1] Butcher, J. C., Coefficients for the study of Runge-Kutta integration processes, This *Journal* 3 (1963), 185–201.

[2] Kutta, W., Beitrag zur näherungsweisen Integration totaler Differentialgleichungen, *Zeit. Math. Physik* 46 (1901), 435–452.

[3] Nyström, E. J., Über die numerische Integration von Differentialgleichungen, *Acta Soc. Sci. Fennicae* 50, No. 13 (1925).

[4] Huťa, A., Une amélioration de la méthode de Runge-Kutta-Nyström pour la résolution numérique des équations différentielles du premier ordre, *Acta Fac. Nat. Univ. Comenian. Math.* 1 (1956) 201–224.

[5] Huta, A., Contribution à la formule de sixième ordre dans la méthode de Runge-Kutta-Nyström, *Acta Fac. Nat. Univ. Comenian. Math.* 2 (1957), 21—24.
[6] Butcher, J. C., On the integration processes of A. Huťa, This *Journal* 3 (1963), 202—206.
[7] Butcher, J. C., Integration processes based on Radau quadrature formulas, *Math. Comp.*, 18 (April, 1964).
[8] Butcher, J. C., Implicit Runge-Kutta processes, *Math. Comp.*, 18 (1964), 50—64.

University of Canterbury,
New Zealand.