Phase-Fitted modified Runge-Kutta pairs of orders $6(5)$.

Ch. Tsitouras[*][1] and **I. Th. Famelis**[**][2]

[1] TEI of Chalkis, Department of Applied Sciences, GR34400 Psahna, Greece
[2] TEI of Athens, Department of Mathematics, GR12210 Athens, Greece

Modified Runge-Kutta methods are well suited for fulfilling properties that require coefficients depending on step-length. By a simple perturbation of very few coefficients we may produce various function-fitted methods and avoid the overhead of evaluating all of them in every step. In this paper we present the formula of determining the extra algebraic equations of condition generated by the major subcategory of these methods. Also phase-lag and phase-fitted properties are analyzed for this case. Finally a specific phase-fitted pair of orders $6(5)$ is given.

1 Introduction

We consider the numerical solution of the non-stiff initial value problem,

$$y' = f(x, y), \ y(x_0) = y_0 \in \Re^m, \ x \in [x_0, x_f] \tag{1}$$

where the function $f : \Re \times \Re^m \to \Re^m$ is assumed to be as smooth as necessary. The general $s-$stage embedded Runge-Kutta pair of orders $p(p-1)$, for the approximate solution of the problem (1) can be defined by the following Butcher scheme [2, 3]:

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

where $A \in \Re^{s \times s}$, is strictly lower triangular, b^T, $\hat{b}^T$, $c \in \Re^s$ with $c = A \cdot e$, $e = [1, \ 1, \cdots, 1]^T \in \Re^s$. The vectors $\hat{b}$, b define the coefficients of the $(p-1)-$th and $p-$th order approximations respectively.

Starting with a given value $y(x_0) = y_0$, this method produces approximations at the mesh points $x_0 < x_1 < x_2 < \cdots < x_f$. Throughout this paper, we assume that local extrapolation is applied, hence the integration is advanced using the $p-$th order approximation. For estimating the error, two approximations are evaluated at each step x_n to $x_{n+1} = x_n + h_n$. These are:

$$\hat{y}_{n+1} = y_n + h_n \sum_{j=1}^{s} \hat{b}_j f_j \ \text{ and } \ y_{n+1} = y_n + h_n \sum_{j=1}^{s} b_j f_j,$$

where $f_i = f(x_n + c_i h_n, \ y_n + h_n \sum_{j=1}^{i-1} a_{ij} f_j), \ i = 1, 2, \cdots, s.$

The local error estimate $E_n = \|y_n - \hat{y}_n\|$ of the $(p-1)-$th order Runge-Kutta pair is used for the automatic selection of the step size. Given a Tolerance $TOL > E_n$, the algorithm $h_{n+1} = 0.9 \cdot h_n \cdot (\frac{TOL}{E_n})^{\frac{1}{p}}$ furnishes the next step length. In case $TOL < E_n$ then we reject the current step and evaluate another smaller one using again the previous formula but with h_{n+1} being now h_n.

[*] E-mail: tsitoura@teihal.gr, URL-address: http://users.ntua.gr/tsitoura/, Phone: +30 6974 321 039, Fax: +30 2228 023 766
[**] E-mail: ifamelis@teiath.gr, URL-address: http://www.teiath.gr/stef/mathematics/ifamelis/, Phone: +30 210 538 5392

Let $y_n(x)$ be the solution of the local initial value problem $y'_n(x) = f(x, y_n(x))$, $x \geq x_n$, $y_n(x_n) = y_n$. Then E_{n+1} is an estimate of the error in the local solution $y_n(x)$ at $x = x_{n+1}$. The local truncation error t_{n+1} associated with the higher order method is

$$t_{n+1} = y_{n+1} - y_n(x_n + h_n) = \sum_{q=1}^{\infty} h_n^q \sum_{i=1}^{\lambda_q} T_{qi} P_{qi} = h_n^{p+1} \Phi(x_n, y_n) + O(h_n^{p+2})$$

where $T_{qi} = Q_{qi} - \xi_{qi}/q!$ with Q_{qi} algebraic functions of A, b, c and ξ_{qi} positive integers. P_{qi} are differentials of f evaluated at (x_n, y_n) and $T_{qi} = 0$ for $q = 1, 2, \cdots, p$ and $i = 1, 2, \cdots, \lambda_q$. The number of elementary differentials for each order is λ_q and coincides with the number of rooted trees of order q. It is known that $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 2, \lambda_4 = 4, \lambda_5 = 9, \lambda_6 = 20, \lambda_7 = 48 \cdots$, etc [1]. More details can be found in [4].

The set $T^{(q)} = \{T_{q1}, T_{q2}, \cdots, T_{q,\lambda_q}\}$ is formed by the $q-$th order truncation error coefficients. It is usual practice a $(q-1)-$th order method to have minimized $\|T^{(q)}\|_2 = \sqrt{\sum_{j=1}^{\lambda_q} T_{qj}^2}$.

2   Modified Runge-Kutta methods

Vanden Derghe et. al. [5] proposed the modified Runge-Kutta methods where the stages evaluated by:

$$f_i = f(x_n + c_i h_n, \ \gamma_i y_n + h_n \sum_{j=1}^{i-1} a_{ij} f_j), \ \ i = 1, 2, \cdots, s.$$

So the parameter vector $\gamma = [\gamma_1 \ \gamma_2 \cdots \gamma_s]^T$ is introduced. The $s-$stages modified Runge-Kutta method is given by the Butcher tableau:

$$
\begin{array}{c|c|cccc}
c_1 & \gamma_1 & & & & \\
c_2 & \gamma_2 & a_{21} & & O & \\
\vdots & \vdots & \vdots & \ddots & & \\
c_s & \gamma_s & a_{s1} & & a_{s,s-1} & \\
\hline
& & b_1 & & b_{s-1} & b_s
\end{array}
$$

If $\gamma_i \neq 1$ for some $1 \leq i \leq s$ then f enters in the expression for truncation error coefficients T's and little can be said about algebraic order conditions for this type of methods. Modified Runge-Kutta are used considering $\gamma_i = 1 + \gamma_{i2} v^2 + \gamma_{i4} v^4 + \cdots$, where $v = \omega h$ for some real parameter ω. In that case powers of h produce extra truncation error coefficients and the corresponding truncation error becomes:

$$t_{n+1} = \sum_{q=1}^{\infty} h_n^q \left(\sum_{i=1}^{\lambda_q} T_{qi} P_{qi} + \sum_{i=1}^{\widetilde{\lambda}_q} \widetilde{T}_{qi} \widetilde{P}_{qi} \right) = h_n^{p+1} \Phi(x_n, y_n) + O(h_n^{p+2})$$

where $\widetilde{T}_{qi} = \widetilde{Q}_{qi}$ with $\widetilde{Q}_{qi}$ algebraic functions of A, b, c and vectors $g_2 = [\gamma_{12}, \ \gamma_{22}, \ \gamma_{32}, \cdots]^T$, $g_4 = [\gamma_{14}, \ \gamma_{24}, \ \gamma_{34}, \cdots]^T$, etc. $\widetilde{P}_{qi}$ are differentials of f and $y(x)$ evaluated at (x_n, y_n) and $\widetilde{T}_{qi} = 0$ for $q = 1, 2, \cdots, p$ and $i = 1, 2, \cdots, \widetilde{\lambda}_q$. $\widetilde{\lambda}_q$ is the number of the additional elementary differentials for each order for the modified Runge-Kutta methods. We observed that $\widetilde{\lambda}_1 = 0, \widetilde{\lambda}_2 = 0, \widetilde{\lambda}_3 = 1, \widetilde{\lambda}_4 = 2, \widetilde{\lambda}_5 = 7, \widetilde{\lambda}_6 = 18, \widetilde{\lambda}_7 = 52$, etc. Franco [6] and Vyver [8] have already presented the additional equations of condition up to fifth order.

The sets $\widetilde{T}^{(q)} = \{\widetilde{T}_{q1}, \ \widetilde{T}_{q2}, \cdots, \widetilde{T}_{q,\widetilde{\lambda}_q}\}$ are formed as seen in Table- 1:

In this table operation "*" may understood as component-wise multiplication:

$$[b_1 \ b_2 \cdots b_s]^T * [\gamma_1 \ \gamma_2 \cdots \gamma_s]^T = [b_1 \gamma_1 \ b_2 \gamma_2 \cdots b_s \gamma_s]^T.$$

This operation has the less priority. Parentheses, powers and dot products are always evaluated before "*". Absence of an operation sign means that we use dot product.

The additional $\widetilde{T}'_q$s are evaluated from the original T'_qs in all possible combinations according to the following rules.

Table 1 The additional order conditions of orders 3 through 6.

$$\widetilde{T}_{3,1} = bg_2, \qquad \widetilde{T}_{4,1} = b\,(c * g_2), \qquad \widetilde{T}_{4,2} = bAg_2, \qquad \widetilde{T}_{5,1} = bA^2g_2, \qquad \widetilde{T}_{5,2} = bA(c * g_2),$$

$$\widetilde{T}_{5,3} = b(c * Ag_2), \quad \widetilde{T}_{5,4} = b(g_2 * Ac), \quad \widetilde{T}_{5,5} = bg_2^2, \qquad \widetilde{T}_{5,6} = b(g_2 * c^2) \qquad \widetilde{T}_{5,7} = bg_4$$

$$\widetilde{T}_{6,1} = bA^3g_2, \qquad \widetilde{T}_{6,2} = bA^2(c * g_2), \quad \widetilde{T}_{6,3} = bA(c * Ag_2) \quad \widetilde{T}_{6,4} = bA(g_2 * Ac) \quad \widetilde{T}_{6,5} = bAg_4$$

$$\widetilde{T}_{6,6} = bA(c^2 * g_2) \quad \widetilde{T}_{6,7} = bAg_2^2 \qquad \widetilde{T}_{6,8} = b(c * A^2g_2) \quad \widetilde{T}_{6,9} = b(c * A(c * g_2)) \quad \widetilde{T}_{6,10} = b(Ac * Ag_2)$$

$$\widetilde{T}_{6,11} = b(g_2 * A^2c) \quad \widetilde{T}_{6,12} = b(g_2 * Ac^2) \quad \widetilde{T}_{6,13} = b(g_2 * Ag_2) \quad \widetilde{T}_{6,14} = b(c^2 * Ag_2) \quad \widetilde{T}_{6,15} = b(c * g_2 * Ac)$$

$$\widetilde{T}_{6,16} = b(c * g_2^2) \qquad \widetilde{T}_{6,17} = b(c^3 * g_2) \qquad \widetilde{T}_{6,18} = b(c * g_4)$$

1. g_2 substitutes c^2, g_2^2 substitutes c^4, g_2^3 substitutes c^6, $\cdots$ etc.

2. g_4 substitutes c^4, g_4^2 substitutes c^8,... g_6 substitutes c^6, g_6^2 substitutes c^{12}, $\cdots$ etc.

3. Every possible combination of g_2, g_4, $\cdots$ may substitute the corresponding power of c, e.g. g_2g_4 substitutes c^6 or $g_2^2g_4g_6$ substitutes c^{14}.

4. It is not obligatory for a power of c to be entirely substituted, e. g. g_2c^2 may substitute c^4.

5. All combinations of substitution apply, e.g. original $b(c^2 * Ac^2)$ may substituted by $\widetilde{T}_{6,12}$, $\widetilde{T}_{6,13}$, $\widetilde{T}_{6,14}$.

As example observe that original truncation error coefficient $bc^5 - 1/6$, generates three additional equations. Namely $\widetilde{T}_{6,16} = b(g_2^2 * c) = 0$, $\widetilde{T}_{6,17} = b(g_2 * c^3) = 0$ and $\widetilde{T}_{6,18} = b(g4 * c) = 0$. On the other hand there are truncation error coefficients like $T_{2,1} = bc - \frac{1}{2}$ or $T_{3,1} = bAc - \frac{1}{6}$ that do not produce any $\widetilde{T}$'s.

3 Phase-Lag property and Phase-fitted modified Runge-Kutta methods

The application of a modified Runge-Kutta method to the test problem $y' = i\omega y$, $\omega \in \Re$, $i = \sqrt{-1}$, leads to the numerical scheme, $y_{n+1} = (1 - iv^2b \cdot (I_s + ivA)^{-1}\gamma)y_n = (Q(v^2) + iR(v^2))y_n$, where $v = \omega h$, h the step length, identity matrix $I_s \in \Re^{s \times s}$ and Q, P polynomials in v^2.

Actually we have

$$Q(v^2) = 1 - \tau_2 v^2 + \tau_4 v^4 - \tau_6 v^6 \pm \cdots, \quad R(v^2) = \tau_1 v - \tau_3 v^3 + \tau_5 v^5 \mp \cdots$$

with $\tau_0 = 1$, $\tau_1 = b\gamma$, $\tau_2 = bA\gamma$, $\tau_3 = bA^2\gamma$, $\tau_4 = bA^3\gamma$, $\cdots$ etc. This series is finite for explicit methods.

The phase lag of a modified Runge-Kutta method is the difference in the angles between theoretical and numerical solution. Thus it is defined as the argument of polynomial $Q(v^2) + iR(v^2)$, which is

$$\delta(v^2) = v - \arg(Q(v^2) + iR(v^2)).$$

A phase fitted method satisfies $\tan(v) = R(v^2)/Q(v^2)$ or $Q(v^2)\tan(v) = R(v^2)$. Every conventional Runge-Kutta method of $p-$th order can be modified entering just one γ_i (say γ_2) in order to solve the previous equation. We *conjecture* that this modification is of $p-$th order also satisfying by default all the additional order conditions.

Here we deal with the Runge-Kutta pair of orders 6(5) described in [7]. That pair was chosen because it had minimized the Euclidean norm of the principal truncation error $\|T^{(7)}\|_2 \approx 1.23 \cdot 10^{-5}$. The coefficients of this pair were not explicitly given in [7] so we present them here in Table 2.

We decided to alter only γ_3 and γ_4. Then we may solve simultaneously the following equations which are linear in these two coefficients:

$$Q(v^2)\tan(v) = R(v^2) \text{ and } \hat{Q}(v^2)\tan(v) = \hat{R}(v^2),$$

where $\hat{Q}(v^2) = 1 - \hat{\tau}_2 v^2 + \hat{\tau}_4 v^4 \mp \cdots$ and $\hat{R}(v^2) = \hat{\tau}_1 v - \hat{\tau}_3 v^3 \pm \cdots$ with $\hat{\tau}_1 = \hat{b}\gamma$, $\hat{\tau}_2 = \hat{b}A\gamma$,..etc. The expressions found are very lengthy and we present here a truncated form accurate to 16 digits.

$$\gamma_3 = \frac{\left\{ \begin{array}{c} v(-65305672138101957 + 8599340737185909279v^2 - 544643545858918291v^4 - 8331976784951199v^6 \\ +4863338667502587v^8 + 138457181971342v^{10} + 100256112768v^{12}) + (84534052330854 2510v \\ +2779876026560310441v^3 - 209128028368800712v^5 + 2430077755103021v^7 - 2586030383486v^9 + 793469990912v^{11} \\ -100256112768v^{13})\cos(2v) + (-390017425585220276 - 9042997967743938122v^2 + 687469882997434084v^4 \\ -50022310453042888v^6 - 2025204517188v^{10} + 140969132679v^{12})\sin(2v) \end{array} \right\}}{\left\{ \begin{array}{c} v^3(-3938718492122557254 + 2607869618740511300v^2 - 230139430387523620v^4 + 12091920773403244v^6 \\ +19185121135089v^8 + (-3938718492122557254 + 232511189651454343v^2)\cos(2v) - 15697577664141165720v\sin(2v)) \end{array} \right\}}$$

Table 2 The coefficients of the Runge-Kutta pair of orders $6(5)$ [7], accurate at 16 digits.

0								
$\frac{17}{183}$	$\frac{17}{183}$							
$\frac{12}{83}$	$\frac{3756}{117113}$	$\frac{13176}{117113}$						
$\frac{18}{83}$	$\frac{9}{166}$	0	$\frac{27}{166}$					
$\frac{71}{125}$	$\frac{55915731}{85159748}$	0	$-\frac{388019101}{155376874}$	$\frac{223573204}{92819845}$				
$\frac{42}{59}$	$-\frac{406585057}{236217205}$	0	$\frac{775681043}{107388827}$	$-\frac{410381131}{74670154}$	$\frac{77706261}{110079566}$			
$\frac{199}{200}$	$\frac{281572459}{68199282}$	0	$-\frac{1844127705}{109029499}$	$\frac{2749721557}{191899305}$	$-\frac{113931059}{73345148}$	$\frac{32727553}{32573572}$		
1	$\frac{276654081}{61910575}$	0	$-\frac{1175802683}{64092361}$	$\frac{439568282}{28315819}$	$-\frac{85495876}{49623813}$	$\frac{85908423}{79433356}$	$-\frac{580531}{104179841}$	
1	$\frac{24503}{381483}$	0	0	$\frac{46353896}{139258673}$	$\frac{19636650}{73309589}$	$\frac{11608951}{64542974}$	$\frac{38826028}{25699703}$	$-\frac{14933}{11016}$
6th	$\frac{24503}{381483}$	0	0	$\frac{46353896}{139258673}$	$\frac{19636650}{73309589}$	$\frac{11608951}{64542974}$	$\frac{38826028}{25699703}$	$-\frac{14933}{11016}$
5th	$\frac{7185863}{91275696}$	0	0	$\frac{10274196}{36984265}$	$\frac{34121257}{67323961}$	$-\frac{20245245}{160728943}$	$\frac{432688272}{102699917}$	$-\frac{296917782}{74219783}\ \frac{1}{20}$

$$\gamma_4 = \frac{\left\{\begin{array}{c} v(1032019641740741097473 - 189505731792181862952v^2 + 14399819581288050017v^4 \\ -220596858253066269v^6 + 23931701840512638v^8 + 30721404312284v^{10}) - v(-278413394724761389533 \\ +5655093876343691015v^2 + 6724253806806v^{10})\cos(2v) + (-98213015307892349568 + 35739924443797240974v^2 \\ -1842667837252740723v^4 - 6044485507690v^{10} + 1353457522376v^{12})\sin(2v) \end{array}\right\}}{\left\{\begin{array}{c} v(-326913634846172252154 + 216453178355462437970v^2 - 19101572722164460540v^4 + 1003629424192469311v^6 \\ +1592365054212460v^8 + (-326913634846172252154 + 19298428741070710483v^2)\cos(2v) - 130289886123716754818v\sin(2v)) \end{array}\right\}}$$

Expanding γ_3, γ_4 in series we have:

$$\gamma_3 \approx 1 - 2.479604820001983 \cdot 10^{-5}v^4 + O(v^6), \quad \gamma_4 \approx 1 + O(v^6).$$

Finally we form vectors $g_2 = [0, 0, 0, 0, 0, 0, 0, 0]^T$ and g_4 to find that the modification of this pair is of orders $6(5)$ indeed. This was verified checking only $\widetilde{T}_{5,7}$, $\widetilde{T}_{6,5}$ and $\widetilde{T}_{6,18}$ for sixth order formula and $\widetilde{\hat{T}}_{5,7}$ for the lower order one.

References

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

[2] J. C. Butcher, Implicit Runge-Kutta processes, Math. Comput., **18**, 50-64 (1964)

[3] J. C. Butcher, On Runge-Kutta processes of high order, J. Austral. Math. Soc., **4**, 179-194 (1964)

[4] J. C. Butcher, The Numerical Analysis of Ordinary Differential Equations. Runge-Kutta & General Linear methods, (John Wiley & sons, Chichester, 1987)

[5] G. Vanden Berghe, H. De Meyer, M. Van Daele and T. Van Hecke, Exponentially-fitted explicit Runge-Kutta methods, Comput. Phys. Commun. **123** 7-15 (1999)

[6] J. M. Franco, An embedded pair of exponentially fitted explicit RungeKutta methods, J. Comput. Appl. Math., **149**, 407-414 (2002)

[7] Ch. Tsitouras, A parameter study of explicit Runge-Kutta pairs of orders $6(5)$, Appl. Math. Lett., **11**, pp. 65-69, (1998)

[8] H Van de Vyver, An embedded phase-fitted modified RungeKutta method for the numerical integration of the radial Schrödinger equation, Phys. Lett. A, **352**, 278-285 (2006)