Bull. Nov. Comp. Center, Num. Anal., 5 (1994), 35–53

RK-method of advanced accuracy: new point of view

Yu.I. Kuznetsov

In the paper new concepts of construction and analysis of numerical schemes of solving ODE systems are used for RK-method of advance accuracy. The analysis of compatibility of equations, the expression for the truncation error are given. The new analitic technics is based on the combinatorial identities. The constructed theory is closed.

1. Introduction

In the previous papers [1, 2] the new principle of construction and analysis of numerical schemes of solving ODE systems, leading to some modification of RK-method, was suggested. Usually an ordinary differential equation

$$\frac{\partial y}{\partial t} = f(t, y), \quad 0 \le t \le T, \quad y(0) = y_0,$$

on the interval $[t_n, t_{n+1}]$, $t_{n+1} = t_n + \tau$, is approximated of the nonlinear algebraic system

$$\begin{aligned} \eta &= y_n e + \tau B f, \\ y_{n+1} &= y_n + \tau b_{m+1} f, \end{aligned} \tag{1}$$

where $B = (\beta_{i,j})$, $i, j = 1(1)m+1$, is the matrix and $b_i = (\beta_{i1}, \dots, \beta_{im})$, $i = 1(1)m+1$, are the vector-rows, $\eta = (\eta_1, \dots, \eta_m)^T$, $f = (f_1, \dots, f_m)^T$, $e = (1, \dots, 1)^T$, $f_j = f(\xi_j, \eta_j)$, $\xi_i = t_n + \lambda_i \tau$, $i = 0(1)m+1$, $\lambda_0 = 0$, $\lambda_{m+1} = 1$, $\lambda_i \le \lambda_{i+1}$. It was proposed to use only the part of the fundamental equations, connecting the matrices B, b_{m+1}, $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_m)$:

$$\begin{aligned} &b_{m+1} B^{j-1} (kB - \Lambda) \Lambda^{k-1} e = 0, \\ &2 \le j + k \le N, \quad j = 1\,(1)\,N-1, \end{aligned} \tag{2}$$

and

$$\begin{aligned} &b_{m+1} B^j \Lambda^{k-1} e = \frac{(k-1)!}{(k+j)!}, \\ &k = 1(1)M - j, \quad j = 0\,(1)\,M-1. \end{aligned} \tag{3}$$

where M is the order of approximation and N is the order of weak approximation. The local truncation error l_{n+1} has the form

$$l_{n+1} = E_{M+1}\tau^{M+1}, \qquad E_j = 1 - (M+1)b_{m+1}\Lambda^M e, \tag{4}$$

but lacking equations are replaced by giving special properties of the characteristic polynomials of the matrix B,

$$Q_m(\lambda) = \lambda^m - \sum_{j=1}^{m} q_j\lambda^{m-j}, \tag{5}$$

and the matrix Λ,

$$P_m(\lambda) = \lambda^m - \sum_{j=1}^{m} p_j\lambda^{m-j}. \tag{6}$$

It was established that the property of nilpotency of the matrix B is ensuring for constructing the diagonal implicit and singly implicit RK-methods. In the present paper the truncated systems (2), (3) are used for constructing advanced accuracy RK-schemes too. The main result of it is the new technics of analysis, founded on the combinatorial identities, the analysis of compatibility of RK-schemes and the expression for the local truncation error as in the multistep methods.

2. Combinatorial identities

The neccesary combinatorial identities are founded on the fact that the solution of the system

$$\sum_{j=1}^{v} \binom{k}{j} x_j = 1, \quad k = V - v + 1\,(1)\,V, \tag{7}$$

with $V \leq 2v$, has the form

$$x_j = (-1)^{j-1}\binom{v}{j}\Big/\binom{V}{j}, \quad j = 1\,(1)\,v, \tag{8}$$

where

$$\binom{k}{i} = \frac{k!}{k!(k-i)!}.$$

Substitution of (8) to (7) brings about the first combinatorial identity

$$\sum_{j=0}^{v}(-1)^j\binom{k}{j}\binom{V-j}{v-j}=0, \quad k=V-v+1\,(1)\,V. \tag{9}$$

One may find from this

$$\sum_{j=0}^{v}(-1)^j\binom{v+i+l-1}{j}\binom{2v-j+l-1}{v+l-1}=(-1)^v\binom{i-1}{v}, \quad i>0; \tag{10}$$

$$\sum_{j=0}^{v}(-1)^j\binom{v-i+l-1}{j}\binom{2v-j+l-1}{v+l-1}=\binom{v+i}{v}, \tag{11}$$

$$i=0\,(1)\,v+l-1;$$

$$\sum_{l=0}^{k}\frac{(-1)^l}{i+l}\binom{k}{l}=\frac{1}{(k+1)\binom{k+i}{k+1}}; \tag{12}$$

$$\sigma_j(i,m)=\sum_{l=0}^{m}\frac{(-1)^l}{l+i+1}\binom{m}{l}\binom{i+l}{j+l}$$
$$=\frac{(-1)^{i+j}}{(m+1)\binom{m+i+1}{m+1}}+\sum_{l=1+j}^{i-m}(-1)^{m+l+j-1}\frac{1}{i+1-l}\binom{i}{m+l}, \tag{13}$$

$$j=0,1, \quad i\geq j;$$

$$\sum_{j=0}^{v}(-1)^j\binom{V+1}{j}\binom{V-j}{v-j}=(-1)^v; \tag{14}$$

$$\sum_{j=0}^{V-v}(-1)^j\binom{V-v}{j}\binom{V-j}{v-j}=1. \tag{15}$$

Direct verification helps to establish the validity of one more important identity

$$\binom{k-1}{i}\binom{k+i+1}{k}\binom{2k+1}{k+1}+\binom{k-1}{i-1}\binom{k+i}{k}\binom{2k+2}{k+1}$$
$$=\left[\binom{k-1}{i+1}\binom{k+i}{k}+\binom{k}{i}\binom{k+i+2}{k+1}\right]\binom{2k}{k}. \tag{16}$$

3. The Characteristic polynomials

It is well-known [3, 4] that stability function $R(\tau)$ has the form

$$R(\tau)=1+\tau b_{m+1}(E-\tau B)^{-1}e. \tag{17}$$

Using the Kayly-Hamilton theorem one may deduce the following representation [1]:

$$R(\tau) = \frac{1 + \sum_{k=1}^{m} \tau^k \left(\frac{1}{k!} - \sum_{j=1}^{k} \frac{q_j}{(k-j)!} \right)}{1 - \sum_{k=1}^{m} q_k \tau^k}. \tag{18}$$

Now let us apply the Kayly-Hamilton theorem to the polynomial $Q_m(\lambda)$ in the form

$$\sum_{j=1}^{m} q_j B^{m-j} = B^m.$$

Multiplying this equality by b_{m+1} on the left and by $\Lambda^i e$, $i = 1\,(1)\,M - m - 1$, on the right, find from (3) that

$$\sum_{j=1}^{m} \frac{q_j}{(k-j)!} = \frac{1}{k!}, \quad k = m+1\,(1)\,M.$$

Since for $M - m < m$ the number of equations in the system is not sufficient, let us complement it in the following way:

$$\sum_{j=1}^{k} \frac{q_j}{(k-j)!} = \frac{1}{k!}, \quad k = M - m + 1\,(1)\,m.$$

The choice of these equations annihilates the coefficients of the higher powers of the polynomial in the denominator of $R(\tau)$, which follows from (18). Admitting the matrix B to be degenerate, we also will assume

$$q_j = 0, \quad j = m - \theta + 1\,(1)\,m.$$

As a whole the system closed in this manner has the form

$$\sum_{j=1}^{m-\theta} \frac{q_j}{(k-j)!} = \frac{1}{k!}, \quad k = M - m + 1\,(1)\,M, \quad \theta = 0\,(1)\,2m - M, \tag{19}$$

if we count that

$$\frac{1}{i!} = 0 \quad \text{for} \quad i < 0.$$

Definition 1. The RK-method defined by (2) will be said to be the RK-method of advanced accuracy.

One should note that system (19) for $\theta = 0$ defines an extension of system (3) to negative powers of B

$$b_{m+1}B^{-1}e = 1, \quad b_{m+1}B^{-i}e = 0, \quad i = 2\,(1)\,2m - M.$$

These equalities may be generalized

$$\begin{aligned} &b_{m+1}B^{j-1}(kB - \Lambda)\Lambda^{k-1}e = 0, \\ &k = 1\,(1)\,M, \quad j = 0\,(-1)\,M - 2m + \omega + 1, \end{aligned} \tag{20}$$

where ω is some non-negative integer, or

$$\begin{aligned} &b_{m+1}B^{j}\Lambda^{k-1}e = \frac{(k-1)!}{(k+j)!}, \\ &j = M - 2m + \xi - 1, \quad k = 1\,(1)\,M. \end{aligned} \tag{21}$$

Solution of system (19), which coincides with system (7) with the replacements $v \to m - \theta$, $V \to M$, and $r_j \to j!q_j$, according to (8), has the form

$$q_j = (-1)^{j-1}\frac{1}{j!}\frac{\binom{m-\theta}{j}}{\binom{M}{j}}, \quad j = 1\,(1)\,m. \tag{22}$$

Substituting it to expression (18) for the stability function with allowance for the closure equations, we obtain

$$R(\tau) = \frac{N_{m-\theta,M-m+\theta}(\tau)}{D_{m-\theta,M-m+\theta}(\tau)},$$

where

$$\begin{aligned} D_{m-\theta,M-m+\theta}(\tau) &= \frac{1}{\binom{M}{m-\theta}}\sum_{k=0}^{m-\theta}\frac{(-\tau)^k}{k!}\binom{M-k}{M-m+\theta}, \\ N_{m-\theta,M-m+\theta}(\tau) &= \frac{1}{\binom{M}{m-\theta}}\sum_{k=0}^{M-m+\theta}\frac{\tau^k}{k!}\binom{M-k}{m-\theta}. \end{aligned}$$

The restriction

$$M \le 2(m-\theta) \tag{23}$$

follows from condition

$$|R(-\infty)| < 1.$$

4. The matrix Λ

Let us define the matrix Λ, using extension (21) of system (3). Assume $p_j = 0$, $j = m - \omega + 1\,(1)\,m$, $0 \le \omega \le 2m - M$, i.e., let us assume ω-fold degeneracy of the matrix Λ. Meanwhile, the number of equations required for definition of p_j reduces correspondingtly. Multiplying equality

$$\Lambda^m = \sum_{i=0}^{m-1} p_{m-i}\Lambda^i \tag{24}$$

by $b_{m+1}B^j$, where $j = M - 2m + \omega\ (1)\ M - m - 1$, from the left, and by vector e from the right, obtain

$$\sum_{j=1}^{m-\omega} p_j \frac{(m-j)!}{(k-j)!} = \frac{m!}{k!}, \quad k = M - m + \omega + 1\ (1)\ M. \tag{25}$$

Applying system (7) once again and making replacements $v \to m - \omega$, $V \to M$ and $r_j \to p_j(m-j)!/m!$, find

$$p_j = (-1)^{j-1}\binom{m}{j}\frac{\binom{m-\omega}{j}}{\binom{M}{j}}, \quad j = 1\ (1)\ m. \tag{26}$$

Theorem 1. *The characteristic polynomial of the matrix* Λ *is*

$$\begin{aligned}\mathcal{P}_m(\lambda) &= \frac{1}{\binom{M}{m-\omega}}\sum_{j=0}^{m-\omega}(-1)^j\binom{m}{j}\binom{M-j}{m-\omega-j}\lambda^{m-j} \\ &= \frac{m!}{M!}\frac{d^{M-m}}{d\lambda^{M-m}}(\lambda^{M-m+\omega}(\lambda-1)^{m-\omega}).\end{aligned} \tag{27}$$

For $\omega = 0$, $M = 2m$, these are the Legendre polynomials.

Theorem 2. *The Eucledean algorithm applied to the polynomials* $\psi_m(\lambda) = \mathcal{P}_m(\lambda)$ *and* $\psi_{m-1}(\lambda) = \mathcal{P}_m^{(1)}(\lambda)/m$, *brings about the sequence*

$$\psi_k(\lambda) = \sum_{j=0}^{k}(-1)^{k-j}\lambda^j\binom{k}{j}\binom{k+j-u+2}{j-\omega+1}\Big/\binom{2k-u+2}{k-\omega+1},$$
$$k = m-1\ (-1)\ 0, \quad u = 2m - M,$$

satisfying recursive relation

$$\psi_0(\lambda) = 1, \quad \psi_1(\lambda) = \lambda - \frac{2-\omega}{4-u},$$
$$\psi_k(\lambda) = (\lambda - U_k)\psi_{k-1}(\lambda) - V_k\psi_{k-2}(\lambda), \quad k = 2\ (1)\ m,$$

thereto,

$$U_k = \frac{2k(k-u+1)+\omega(u-2)}{(2k-u)(2k-u+2)},$$
$$V_k = \frac{(k-1)(k-\omega)(k-u+1)(k-u+\omega)}{(2k-u)^2((2k-u)^2-1)}, \quad k = m-1\,(-1)\,1,$$
$$U_m = \frac{m-\omega}{M}, \quad V_m = \frac{(m-1)(m-\omega)(M-m+\omega)}{(M-1)M^2}.$$

Proof. The validity of recursive relation follows from combinatorial identity (16) and may be directly verified. □

Since the values λ_j, $j = 1\,(1)\,m$, are the roots of $\mathcal{P}_m(\lambda)$, it is essential for them to belong to the interval $[0,1]$. The analysis of the Sturm system of the polynomial $\mathcal{P}_m(\lambda)$ with the combinatorical identities (9), (11), (14) leads to the following assertion.

Theorem 3. *All roots of the polynomial* $\mathcal{P}_m(\lambda)$ *are real, single and lie in the interval* $[0,1]$, *i.e.,*

$$0 \le \lambda_j < \lambda_{j+1} \le 1, \quad j = 1\,(1)\,m-1.$$

It follows from Theorem 3 and the second representation in (27) that

$$0 \le \omega \le 1, \quad 0 \le m - \omega - (M-m) \le 1.$$

These inequlities may be written in the form

$$\max(u-1,0) \le \omega \min(u,1), \quad u = 0,1,2. \tag{28}$$

Because the Vandermond matrix W,

$$W = \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{m-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \lambda_m & \lambda_m^2 & \dots & \lambda_m^{m-1} \end{pmatrix} \tag{29}$$

is nondegenerate one may use the eqalities (3) at $j = 0$,

$$b_{m+1}W = e^T h, \tag{30}$$

where $h = \mathrm{diag}(1, 1/2, \dots, 1/m)$, for locating b_{m+1}. (If the l-th row and k-th colomn of matrix A denote by $A_{l\bullet}$ and $A_{\bullet k}$, then $W_{\bullet k} = \Lambda^{k-1}e$.)

5. The Butcher matrix

So far, the above analysis has shown that the RK-schemes of advanced accuracy are characterized by the following parameters: 1) $M = 2m$, $\omega = 0$; 2) $M = 2m - 1$, $\omega = 0$, and $\omega = 1$; 3) $M = 2m - 2$, $\omega = 1$. By means of recursive relations of Theorem 2, one can calculate the values λ_j, $j = 1\,(1)\,m$, which enables us to define uniquely the matrices Λ and W (the ascending order $\lambda_j < \lambda_{j+1}$, $j = 1\,(1)\,m-1$, is assumed). Relation (30) defines the row b_{m+1}. It remains only to determine the Butcher matrix.

To start with this determination, let us note first that conditions (23) and (28) are compatible, if and only if θ assumes the following values:

$$\begin{aligned} \theta &= 0,1 \quad \text{for} \quad M = 2m-2, \\ \theta &= 0 \quad \text{for} \quad M = 2m-1, 2m. \end{aligned} \tag{31}$$

Therefore, all possible combinations of ω and u, consistent with restrictions (28), correspond to a non-degenerate matrix B ($\theta = 0$). The matrix B may be degenerate only for $u = 2$, $\omega = 1$.

Lemma 1. *The vectors* $r_j = b_{m+1}B^{j-1}$, $j = M-2m+\omega+\theta+1\,(1)\,M-m-\omega$, *are linearly independent.*

Proof. It is necessary to prove the linear independence of $m - \theta$ vectors. Suppose, *ad absurdum*, that there exists a linear combination

$$R = \sum_j \gamma_j r_j, \quad j = M - 2m + \omega + \theta + 1\,(1)\,M - m + \omega,$$

turning to zero, i.e., $R = 0$. Multiplying this equality from the right by the vector $\Lambda^s e$, $s = 0\,(1)\,m - \theta - 1$, by means of (3) and (21) obtain

$$R\Lambda^s e = \sum_j \gamma_j b_{m+1} B^{j-1} \Lambda^s e = \sum_j \gamma_j \frac{s!}{(s+j)!} = 0.$$

Introduce the replacement $s + j = k - t$, where $k = M - m + \omega + 1 + s$. Evidently, in this case $t = 1\,(1)\,m - \theta$, $k = M - m + \omega + 1\,(1)\,m + \omega - \theta$. Then

$$R\Lambda^s e = s! \sum_{t=1}^{m-\theta} \frac{\gamma_{k-s-t}}{(k-t)!} = \frac{s!}{k!} \sum_{t=1}^{m-\theta} \binom{k}{t} t!\gamma_{g-t} = 0,$$

$$g = M - m + \omega + 1, \quad k = M - m + \omega + 1\,(1)\,M + \omega - \theta.$$

The matrix of this system is the same as that of system (7), when $v = m-\theta$, $V = M + \omega - \theta$, hence, it is non-degenerate. The solution of such system is unique: $t!\gamma_{g-t} = 0$, i.e., $\gamma_{g-t} = 0$, $t = 1\,(1)\,m - \theta$. Therefore, the vectors r_j, $j = M - 2m + \omega + \theta + 1\,(1)\,M - m + \omega$, are linearly independent. □

Corollary 1. *If $\theta = 1$ then the vectors r_j, $j = 1\,(1)\,m$, are linearly dependent. Indeed, in this case the value γ_j, $j = 1\,(1)\,m$, will be determined by the nonhomogeneous system*

$$s!\sum_{j=1}^{m}\frac{\gamma_j}{(s+j)!} = \gamma\delta_{s,m-1}, \quad s = 0\,(1)\,m-1,$$

where $\gamma \neq 0$, $\delta_{s,m-1}$ – the Kronecker symbol. The matrix of this system is nonsingular.

Remark 1. If $\theta = 1$ one may deduce by analogy, that the vectors r_j, $j = 2\,(1)\,m$, are linearly independent.

Corollary 2. *Lemma 1 yields that the non-degenerate matrix B is complete.*

Lemma 2. *In the RK-methods of advanced accuracy, relations* (2) *still hold for $N = M$.*

Proof. Consider the case $N = M + 1$. Represent the left-hand side of (2) for $k + j = M + 1$, $k = m + 1$ in the form

$$Z = (m+1)b_{m+1}B^{M-m}\Lambda^m e - b_{m+1}B^{M-m-1}\Lambda^{m+1}e.$$

Expressing Λ^m through equality (6), with the help of (3) we obtain

$$\begin{aligned} Z &= \sum_{j=1}^{m} p_j \frac{(m-j)}{(M-j+1)!} = \frac{m!}{M!}\sum_{j=1}^{m-\omega}(-1)^{j-1}\binom{m-\omega}{j}\frac{j}{M-j+1} \\ &= \frac{m!}{M!}\sum_{j=0}^{m-\omega}(-1)^j\binom{m-\omega}{j} - \sum_{j=0}^{m-\omega}(-1)^j\binom{m-\omega}{j}\frac{M+1}{M-j+1}. \end{aligned}$$

The first sum on the right-hand side is, evidently, equal to zero. The second sum may be calculated through identity (12) and is not equal to zero. Therefore, $Z \neq 0$ and $N < M + 1$. But solvability of relations (12), (25) yields $N = M$. □

Now we can construct an algorithm of calculating of the matrix B.

Theorem 4. *For a non-degenerate matrix B ($\theta = 0$) there hold the equalities*

$$B\Lambda^{k-1}e = \frac{1}{k}\Lambda^k e, \quad k = 1\,(1)\,m - \omega, \tag{32}$$

i.e., the condition $C(m-\omega)$ is satisfied.

Proof. Relations (2) and (20) can be represented in the form

$$\begin{aligned} &b_{m+1}B^{j-1}(kB - \Lambda)\Lambda^{k-1}e = 0, \\ &j = M - 2m + \omega + 1\,(1)\,M - m + \omega, \quad k = 1\,(1)\,m - \omega. \end{aligned} \tag{33}$$

According to Lemma 1, it follows that

$$kB\Lambda^{k-1}e - \Lambda^k e = 0, \quad k = 1\,(1)\,m - \omega.$$

If $\omega = 0$, then the statement of the theorem is proved for this case. Suppose that $\omega = 1$, and let us prove that the condition $C(m)$ is impossible. Indeed, in the contrary it denotes that along with (25) the relation

$$\sum_{j=1}^{m} p_j \frac{(m-j)!}{(k-j)!} = \frac{m!}{k!}$$

also holds for $k = M - m + 1$, i.e.,

$$\sum_{j=0}^{m-1} (-1)^j \binom{k}{j} \frac{\binom{m-1}{j}}{\binom{M}{j}} = 0,$$

is valid, or the same,

$$\sum_{j=0}^{m-1} (-1)^j \binom{M - m + 1}{j} \binom{M - j}{m - 1 - j} = 0.$$

But according to (15), for $V = M$, $v = m - 1$, the expression on the left-hand side of this equality equals 1. Thus, the condition $C(m)$ should be omitted. □

Therefore, if $\omega = 0$, equalities (32) are transformed into the equality $BW = AWh$ and define the matrix B uniquely. In accordance with (27) and (28), there are only two RK-schemes, for which $\theta = \omega = 0$: $M = 2m$ is the Gaussian scheme and $M = 2m - 1$ is the Radau scheme-1, both are uniquely defined.

The case $\omega = 1$ corresponding to degenerate matrix Λ appears to be more complicated and requires additional analysis. Consider first a non-degenerate matrix B.

Theorem 5. *For a non-degenerate matrix B ($\theta = 0$) and a degenerate matrix Λ ($\omega = 1$) holds*

$$\beta_{i1} = \frac{1}{m(M - m + 1)}, \quad i = 1\,(1)\,m + 1. \tag{34}$$

Proof. The equation

$$BW = \Lambda W h + w e_m^T, \tag{35}$$

where w is some vector, follows from the condition $C(m-1)$, i.e., (32) at $\omega = 1$, therefore,

$$\tilde{B} = W^{-1} B W = (F^T + z e_m^T)^h = F_z^T h, \quad z = m W^{-1} w, \tag{36}$$

and

$$F = \begin{pmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ 0 & & & 0 & 1 \\ p_m & & \cdots & p_2 & p_1 \end{pmatrix}.$$

Using the expressions [2]

$$q_i = \frac{(m-i)!}{m!}(p_i + z_{m-i+1})$$

and (22), (26) we shall obtain

$$z_{m-i+1} = (-1)^{i-1} \frac{\binom{m}{i}}{\binom{M}{i}} \left(\binom{m-\theta}{i} - \binom{m-\omega}{i} \right). \tag{37}$$

Because $\theta = 0$, $\omega = 1$,

$$z = \sum_{i=1}^{m} (-1)^{i-1} \frac{\binom{m}{i}}{\binom{M}{i}} \binom{m-1}{i-1} e_{m-i+1}. \tag{38}$$

Let us take the vector w from equation (35)

$$w = BW_{\bullet m} - \frac{1}{m} \Lambda W_{\bullet m}. \tag{39}$$

It is evident that

$$W_{\bullet m} = \sum_{j=1}^{m-1} p_{m-j} W_{\bullet j} - p_{m-1} c_1, \tag{40}$$

since

$$\lambda_1 = 0, \quad p_m = 0.$$

Hence, applying the condition $C(m-1)$ one may obtain

$$w = \sum_{j=1}^{m-1} \frac{m-j}{mj} p_{m-j} \Lambda^j e - p_{m-1} B e_1. \tag{41}$$

Together with (36), (38) this equation gives the expression for the first column of the matrix B

$$\begin{aligned} Be_1 &= \frac{1}{mp_{m-1}} \left(\sum_{j=1}^{m-1} \frac{j}{m-j} (-1)^{j-1} \frac{\binom{m}{j}\binom{m-1}{j}}{\binom{M}{j}} \Lambda^{m-j} e - \right. \\ &\qquad \left. \sum_{j=1}^{m} (-1)^{j-1} \frac{\binom{m}{j}\binom{m-1}{j-1}}{\binom{M}{j}} \Lambda^{m-j} e \right) \\ &= \frac{(-1)^m}{mp_{m-1}\binom{M}{m}} e = \frac{\binom{M}{m-1}}{m\binom{M}{m} m} e = \frac{1}{m(M-m+1)} e. \end{aligned}$$

Thus, equations (34) are valid for $i = 1\ (1)\ m$. But they are valid for $i = m+1$ as well, since from (30) we have

$$\beta_{m+1,1} = e^T h W^{-1} e_1 = \frac{1}{p_{m-1}} \sum_{j=1}^{m} \frac{p_{m-j}}{j},$$

because using (5) we may get the equality

$$\beta_{m+1,1} = \frac{1}{m(M-m+1)} \left(\sum_{j=0}^{m} (-1)^{m-j+1} \binom{m}{j} \binom{M-j}{m-j} + 1 \right).$$

The expression in the brackets equals to unit with respect to identity (9). □

There are two RK-schemes satisfying condition (34), which possess the property $\theta = 0, \omega = 1$: $M = 2m-1$ – the Radau scheme-2 and $M = 2m-2$ – the Lobatto scheme-1. The last possible case, as follows from (28), consists in the choice $\omega = \theta = 1$. According to (31), here $M = 2m-2$. As well as in the Lobatto scheme-1, these RK-schemes use the end-points of the integration step $\lambda_1 = 0$, $\lambda_m = 1$ (see the proof of Theorem 1).

As the matrix B is degenerate, one cannot use Theorem 6. Thus, let us apply relations (2), assuming $N = 2m - 2$.

By Lemma 1, there are only $m - 1$ linearly independent vectors $r_j = b_{m+1}B^{j-1}$, $j = 1\,(1)\,m - 1$. These vectors, due to (2) are orthogonal to $m - 1$ vectors $v_k = (kB - \Lambda)\Lambda^{k-1}e$, $k = 1\,(1)\,m - 1$, and the vector v_m is orthogonal only to the first $m - 2$ vectors r_j.

Let us consider the case of the degenerate matrix V, when $\theta = \omega = 1$, $M = 2m - 2$. The equations (2) one may present in the form

$$\begin{aligned} r_j v_k &= 0, \quad j, k = 1\,(1)\,m - 1, \\ r_j v_m &= 0, \quad j = 1\,(1)\,m - 2, \end{aligned} \tag{42}$$

and so all, where $r_j = b_{m+1}B^{j-1}$, $v_k = (kB - \Lambda)\Lambda^{k-1}e$. With respect to Corollary 1 and the remark to Lemma 1

$$r_m = \sum_{j=1}^{m-1} \alpha_j r_j,$$

and $\alpha_1 \neq 0$. Multiplying this equation on the matrix B to right, expressing B^m by means of the theorem of the Kayly-Hamilton and using of the linear independence of the vectors r_j, $j = 2\,(1)\,m$, we find

$$r_m = \sum_{j=1}^{m-1} q_{m-j} r_j. \tag{43}$$

Let us introduce the left and right null-spaces of the matrix B,

$$xB = 0, \quad Bs = 0. \tag{44}$$

The vectors x and r_j, $j = 1\,(1)\,m - 1$, are the basis in R_m.

Lemma 3. *The equation*

$$BW = \Lambda W h + s u^T + w e_m^T \tag{45}$$

is valid for RK-scheme with $\theta = \omega = 1$, $M = 2m - 2$, if s is defined in (44) *and*

$$u^T e_m = 0, \quad b_{m+1} w = 0. \tag{46}$$

Proof. Let γ,

$$\gamma = \sum_{j=1}^{m-1} \alpha_j r_j + \alpha x,$$

be an arbitrary vector in R_m. Then according to (42), (43), (44)

$$\gamma B v_k = \sum_{j=1}^{m-1} \alpha_j r_j B v_k + \alpha x B v_k = 0, \quad k = 1\,(1)\,m-1.$$

The equalities

$$B v_k = 0, \quad k = 1\,(1)\,m-1,$$

follow from arbitrarity of vectors γ. Thus, the vectors v_k are orthogonal to the space of row of the matrix B and, therefore, are belonging to the right null-space of B. Hence,

$$B\Lambda^{k-1}e = \frac{1}{k}\Lambda^k e + \frac{\delta_k}{k} s, \quad k = 1\,(1)\,m-1,$$

or

$$BW = \Lambda W h + s \sum_{k=1}^{m-1} \frac{\delta_k}{k} e_k^T + (B - \frac{1}{m}\Lambda) W_{\bullet m} e_m^T. \tag{47}$$

Supposing

$$w = (B - \frac{1}{m}\Lambda) W_{\bullet m}$$

we get (45). □

So we dispose of $m-1$ equations from Lemma 3 and the condition of the rank of B being equal to $m-1$ ($\theta = 1$), used in this case for defining B. This is not sufficient for unique definition, thus let us specify also the degeneracy type. For instance, specification of the first row of the matrix B being equal to zero is consistent. Consider the case $b_1 = 0$.

Theorem 6. *If $\theta = \omega = 1$ and $M = 2m-2$, equality*

$$B\Lambda^{k-1}e = \frac{1}{k}\Lambda^k e, \quad k = 1\,(1)\,m, \tag{48}$$

takes place, if and only if the following conditions are satisfied:

1. *Matrix $\bar{B}$ of size $m-1$, obtained from B by crossing the first row and the first column out, is non-degenerate.*
2. *The first row of the matrix B is formed by zeroes.*

Proof. The necessity of Condition 2 is evident: (48) implies for the first row of the matrix B the following equality

$$b_1 W = 0$$

with a non-degenerate Vandermonde matrix. Hence, $x = e_1^T$. In this case the vectors $B_{\bullet k}$, $k = 2\,(1)\,m$, are linearly independent. Indeed, in contrary, $e_1^T s = 0$ and hence, $xs = 0$, what is impossible for $\theta = 1$.

In order to prove the sufficiency, let us assume that both conditions of the theorem are satisfied. It follows from this that $x = e_1^T$, $e_1^T s \neq 0$. But (47) implies

$$b_1 W = 0 + e_1^T s \sum_{k=1}^{m-1} \frac{\delta_k}{k} e_k^T + b_1 W_{\bullet m} e_m^T,$$

i.e.,

$$\sum_{k=1}^{m-1} \frac{\delta_k}{k} = 0.$$

The vectors e_k^T are linearly independent hence $\delta_k = 0$, $k = 1\,(1)\,m-1$ or $u = 0$. This is the condition $C(m-1)$. So the matrix B has been defined by (35). As $\theta = \omega$ we have $z = 0$ from (37) and $w = 0$. It implies the condition $C(m)$. □

The RK-scheme with $\theta = \omega = 1$, $M = 2m - 2$, for which condition (48) is satisfied, will be called the Lobatto scheme-2.

Corollary. *In the conditions of Theorem 6 RK-schemes with $\theta = \omega$ are collocation ones, i.e., the components of the vector η approximate the solution at the nodes ξ_i with order m (the condition $C(m)$). In the RK-schemes with $\theta = 0$ and $\omega = 1$ the solution at the nodes is approximated with order $m - 1$ (the condition $C(m-1)$).*

Before passing to another form of degeneracy of the matrix B, let us go back to those already considered.

Theorem 7. *In the Radau scheme-1 and both Lobatto schemes, there holds*

$$b_m = b_{m+1}.$$

Proof. For these schemes $\lambda_m = 1$. In the Radau scheme-1 and the Lobatto scheme-2 the vectors b_m and b_{m+1} are calculated actually from the same equation due to $\lambda_m = 1$ and condition $C(m)$. In the Lobatto scheme-1 the equation of (35) for the last row by (37) has the form

$$b_m W = e^T h + e_m^T w e_m^T = e^T h + m^{-1} e^T z e_m^T,$$

by means of the identity (10) with $v = m - 1$, $l = 0$, $i = 1$ and the expression (38) one may find that $e^T z = 0$. It leads to the statement of the theorem. □

Theorem 8. *The Lobatto scheme-2 is unique among RK-schemes with conditions $C(m-1)$, $\theta = \omega = 1$.*

Proof. Let us suppose, that in the representation (45) $u = 0$, i.e., the equalities (35), (36) take place. As $\theta = \omega = 1$, then the equations (17), (24) implay the equality

$$q_i \frac{m!}{(m-j)!} = p_j,$$

and identity $z = 0$, i.e., $w = 0$ follows from the equation

$$q_i = \frac{(m-i)!}{m!}(p_i + z_{m-i+1}), \quad i = 1\,(1)\,m.$$

Hence, the condition $C(m)$ takes place and the statement of this theorem is provided by Theorem 6. □

Theorem 9. *RK-scheme with the condition $Be_m = 0$, i.e., $s = e_m$ (the last column of the matrix B is equal to zero), does not exist among RK-scheme with $\theta = \omega = 1$.*

Proof. Equality $b_{m+1}s = 0$ follows from the orthogonality conditions (42). Thus $Bs = 0$, $s \neq 0$ and rank $B = m - 1$, then the row b_{m+1} belongs to row-space of matrix B,

$$b_{m+1} = \rho B, \tag{49}$$

where ρ is the vector-row. Suggesting $Be_m = 0$ implies the equality $\beta_{m+1,m} = 0$. However, it is not so. Indeed, from (30)

$$\beta_{m+1,m} = e^T h W^{-1} e_m.$$

Let us calculate this expression. Denote

$$\varphi_k(\lambda) = P_m(\lambda)/(\lambda - \lambda_k) = \sum_{j=1}^{m} \psi_{k,m-j} \lambda^{j-1}.$$

Then

$$W^{-1}e_m = \frac{1}{\varphi_m(1)} \sum_{j=1}^{m} \psi_{m,m-j} e_j. \tag{50}$$

The polynomial (27) with $M = 2m - 2$, $\theta = \omega = 1$ has the form

$$P_m(\lambda) = \frac{\lambda - 1}{\binom{2m-2}{m-1}} \sum_{j=0}^{m-1} (-1)^j \binom{m-2}{j} \binom{2m-2-j}{m-1} \lambda^{m-j-1},$$

hence,

$$\psi_{m,j} = (-1)^j \frac{\binom{m-2}{j}\binom{2m-2-j}{m-1}}{\binom{2m-2}{m-1}}, \quad j = 0\,(1)\,m-1. \tag{51}$$

Thus the identity (11) with $v = m - 1$, $i = l = 1$ implies

$$\varphi_m(1) = \frac{1}{\binom{2m-2}{m-1}} \sum_{j=0}^{m-1} (-1)^j \binom{m-2}{j} \binom{2m-2-j}{m-1} = \frac{m}{\binom{2m-2}{m-1}},$$

therefore,

$$\begin{aligned}
\beta_{m+1,m} &= \frac{1}{\varphi_m(1)} \sum_{j=1}^{m} \frac{1}{j} \psi_{m,m-j} \\
&= \frac{1}{m} \sum_{j=0}^{m-2} (-1)^j \binom{m-2}{j} \binom{2m-2-j}{m-1} \frac{1}{m-j} \\
&= \frac{1}{(m-1)m} \sum_{j=0}^{m} (-1)^j \binom{m-2}{j} \binom{2m-2-j}{m-2} = \frac{1}{(m-1)m}.
\end{aligned}$$

The last equality is the result of the identity (11) with $i = 0$, $l = -1$, $v = m$. □

6. Compatibility of equations

Relations (3) determine $M(M+1)/2$ equations for $m^2 + 2m$ unknown parameters β_{ij}, λ_j. It means that for any m and M system (3) is either under-defined, or over-defined. In the meantime the conclusions concerning the order of polynomial approximation are implicitly based on the compatibility of this system. Since a class of equations from (3) is not included to definition of characteristic polynomials, it should be separately analyzed. Two groups of equations were not considered above

$$b_{m+1}B^{m-i}\Lambda^k e = \frac{k!}{(m+k+i+1)!},$$
$$b_{m+1}B^{i}\Lambda^{m+k} e = \frac{(m+k)!}{(m+k+i+1)!}, \quad i+k = 0\,(1)\,M-m-1.$$

Let us show that these groups contain no new information. Let us express B^{m+i} and Λ^{m+k} through the corresponding Cayley-Hamilton identities. Both groups are transformed by this to the combinatorial identities

$$\sum_{j=0}^{m}(-1)^j\binom{M-j}{m-j}\binom{m+k+i+1}{j} = 0,$$
$$\sum_{j=0}^{m}(-1)^j\frac{\binom{m}{j}\binom{M-j}{m-j}}{\binom{m-j+k+i+1}{j+1}} = 0, \quad i+k = 0\,(1)\,M-m-1.$$

The first of these identities is system (9) and the second one is the k -th order difference by index i of the first identity, taken for $k=0$, if rewritten in the form

$$\sum_{j=0}^{m}(-1)^j\frac{\binom{m}{j}\binom{M-j}{m-j}}{\binom{m-j+i+1}{i+1}} = 0$$

it implies compatibility of relations (3) for $M \le 2m$.

Let us demonstrate that $M = 2m$ is the maximum admissible order of accuracy, i.e., for $M > 2m$ system (3) is incompatible. Since for $m > 2m$ the system of equations (19) should hold at least for $k = m+1\,(1)\,2m+1$, with the account of the coefficients structure on the right-hand side of equations (19), it proves the validity of equations

$$\sum_{j=0}^{m} q_j\frac{1}{(k-j)!} = 0, \quad k = m+1\,(1)\,2m+1,$$

where $q_0 = -1$. The matrix of this system is non-degenerate, since it coincides exactly with the matrix of system (19), where m is replaced with $m+1$. Thus, the coefficients q_j might be only zeroes, which contradicts to representation (22) and the definition of $q_0 = -1$. Hence, system (3) is incompatible for $M > 2m$.

7. Truncation error

Let us calculate the local truncation error, which is determined by (4). Using the Cayley-Hamilton identity for the matrix Λ, by means of (3) and (5) we obtain

$$\begin{aligned} E_{M+1} &= 1-(M+1)b_{m+1}\Lambda^M e = 1-(M+1)\sum_{j=1}^{m-\omega} p_j b_{m+1}\Lambda^{M-j}e \\ &= \sum_{j=0}^{m-\omega}(-1)^j\binom{m}{j}\frac{\binom{m-\omega}{j}}{\binom{M}{j}}\cdot\frac{M+1}{M-j+1} \\ &= \sum_{j=0}^{m-\omega}(-1)^{m-\omega-j}\binom{m-\omega}{j}\frac{\binom{M-m+\omega+j}{M-\omega}}{M-m+\omega+j+1}\cdot\frac{M+1}{\binom{M}{m}}. \end{aligned}$$

Applying relation (13), find (for given values of M, ω, the sum on the right-hand side vanishes)

$$E_{M+1} = \frac{(-1)^{M-\omega}}{\binom{M}{m}\binom{M}{m-\omega}}, \tag{52}$$

due to which the expression for the local truncation error has the form

$$l_{n+1} = (-1)^{M-\omega}\frac{\tau^{M+1}}{\binom{M}{m}\binom{M}{m-\omega}}. \tag{53}$$

It is easily seen from this expression that the m-stage Radau schemes-1 and 2 enclose the true solution from the both sides (in assumption of a constant sign of the $M+1$-th derivative of y on the segment $[t_n, t_{n+1}]$). The same property is shared by $m-1$-th stage Gauss scheme and m-stage Lobatto scheme.

References

[1] Yu. Kuznetsov, Modified Runge-Kutta method, Bulletin of the Novosibirsk Computing Center, series Num. Anal., issue 1, 1993, 17–40.

[2] Yu. Kuznetsov, Modified Runge-Kutta method II, Bulletin of the Novosibirsk Computing Center, series Num. Anal., issue 2, 1993, 45–63.

[3] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations. Runge-Kutta and General Linear Methods, Wiley, Chichester – New York – Brisbane – Toronto – Singapore, 1987.

[4] K. Dekker, J.G. Verwer, Stability of Runge-Kutta Methods for Stiff Nonlinear Differential Equations, North Holland, Amsterdam – New York – Oxford, 1984.