Numer. Math. 53, 423–438 (1988)

Numerische Mathematik

Natural Runge-Kutta and Projection Methods*

Marino Zennaro
Dipartimento di Matematica e Informatica, Universita' degli Studi di Udine, I-33100 Udine, Italy

Summary. Recently the author defined the class of natural Runge-Kutta methods and observed that it includes all the collocation methods. The present paper is devoted to a complete characterization of this class and it is shown that it coincides with the class of the projection methods in some polynomial spaces.

Subject Classifications: AMS(MOS) 65L05; CR: G1.7.

1. Introduction

Recently the author [7] introduced the notion of Natural Continuous Extension (NCE) of a v-stage Runge-Kutta (RK) method of order p

$$K_i = f(t_0 + c_i h, y_0 + h \sum_{j=1}^{v} a_{ij} K_j), \qquad i = 1, \ldots, v, \tag{1.1a}$$

$$y_1 = y_0 + h \sum_{i=1}^{v} b_i K_i \tag{1.1b}$$

for the numerical solution of the initial value problem (IVP)

$$\begin{aligned} y'(t) &= f(t, y(t)), \qquad t \geqq t_0 \\ y(t_0) &= y_0 \end{aligned} \tag{1.2}$$

in $\mathbb{R}^m$ at the point $t_1 = t_0 + h$, where $f(t, y)$ is as smooth as necessary.

The RK method (1.1) has order $p (\geqq 1)$ if

$$|y(t_1) - y_1| = O(h^{p+1}), \tag{1.3}$$

where $|\cdot|$ stands for any norm on $\mathbb{R}^m$.

* This work was supported by the Italian Ministero della Pubblica Istruzione, funds 40%

A NCE of degree d is determined by (and therefore can be identified with) a v-vector polynomial $\mathbf{b}(\theta)=(b_1(\theta), \ldots, b_v(\theta))^T$, where $b_i(\theta)\in\Pi_d$ (the space of polynomials of degree $\leqq d$), $i=1, \ldots, v$.

With

$$u(t_0+\theta h)=y_0+h\sum_{i=1}^{v} b_i(\theta)\, K_i, \tag{1.4}$$

the following properties are satisfied by $u(t)$:

$$u(t_0)=y_0 \quad \text{and} \quad u(t_1)=y_1\,; \tag{1.5}$$

$$\max_{t_0\leqq t\leqq t_1} |y'(t)-u'(t)|=O(h^d); \tag{1.6}$$

$$\left|\int_{t_0}^{t_1} G(t)[y'(t)-u'(t)]\,dt\right|=O(h^{p+1}) \tag{1.7}$$

for every sufficiently smooth matrix-valued function $G(t)$.

The NCEs were introduced in order to get a continuous approximation of the solution $y(t)$ of (1.2) without any extra function evaluations. Condition (1.5) assures that the NCE interpolates the nodal values given by the RK formula (1.1). Condition (1.6) assures the highest order of uniform approximation which is possible by means of a polynomial of degree d (possibly less than p). Finally, the asymptotic orthogonality condition (1.7) allows to preserve the nodal order p when the NCE is used, for instance, while solving an IVP with driving equation or a delay differential equation (DDE).

In [8] the author investigated some stability properties of RK methods for DDEs. In that setting he gave the definition of *natural RK methods*, which was suggested merely by technical necessities. In fact, as we shall see, the description of the so called P-stability region is much simplified for them.

In order to recall this definition, note that a RK method (1.1) is defined by the triplet $A=\|a_{ij}\|$, $\mathbf{b}=(b_1, \ldots, b_v)^T$ and $\mathbf{c}=(c_1, \ldots, c_v)^T$ (the so called Butcher notation) and that NCE determines a $v\times v$-matrix $B=\|b_j(c_i)\|$.

Definition 1.1. The RK method (1.1) is *natural* if it has a NCE such that $B=A$.

The P-stability region of a RK method for DDEs is the set S_P of the pairs of complex numbers (α, β), $\alpha:=ah$, $\beta:=bh$, such that the numerical solution of

$$y'(t)=a\,y(t)+b\,y(t-\tau), \qquad t>0$$
$$y(t)=g(t) \qquad \text{for } -\tau\leqq t\leqq 0$$

($a, b\in\mathbb{C}$, $\tau>0$ and $g(t)$ is continuous and complex valued) asymptotically vanishes for every delay τ and for every step-length h satisfying

$$h=\tau/m, \qquad m \text{ positive integer.}$$

Theorem 13 in [8] states that the interior of the P-stability region is

$$S_{P'}=\{(\alpha, \beta)\in\mathbb{C}\times\mathbb{C}\,|\,\alpha\in S_A \quad \text{and} \quad |\beta|<\sigma_\alpha\},$$

where S_A is the A-stability region of the underlying RK method,

$$\sigma_\alpha = \inf\{|z| \,|\, z \in \mathbb{C} \quad \text{and} \quad |r_\alpha(z)| = 1\}$$
$$r_\alpha(z) = 1 + (\alpha + z)\,\mathbf{b}^T (I - \alpha A - zB)^{-1}\mathbf{e}, \quad \mathbf{e} = (1, \ldots, 1)^T.$$

In general, the computation of σ_α is not easy. To see this it is sufficient to look at the Heun method ($v = p = 2$) with linear interpolation for which

$$r_\alpha(z) = [1 + \alpha + \alpha^2/2 + (\alpha + 1)\,z/2]/(1 - z/2).$$

In this case the analytical computation of σ_α takes about two pages!

On the other hand, if the method is natural, i.e. $B = A$, then it immediately follows that

$$r_\alpha(z) = R(\alpha + z),$$

where R is the absolute stability function of the RK method (see also Section 4), and hence that (Theorem 16 in [8])

$$\sigma_\alpha = d(\alpha, S_A)$$

for every $\alpha \in S_A$, where $d(.,.)$ is the Euclidean distance in the complex plane.

Therefore we have a drastic simplification. For example, for $v = p = 2$, we can consider the trapezoidal rule with linear interpolation, which is a natural RK method. We have

$$r_\alpha(z) = R(\alpha + z) = [1 + (\alpha + z)/2]/[1 - (\alpha + z)/2]$$

and hence straightway

$$\sigma_\alpha = -\operatorname{Re}(\alpha)$$

for every $\alpha \in S_A = \{\alpha \in \mathbb{C} \,|\, \operatorname{Re}(\alpha) < 0\}$. In this case we have a so called P-stable method.

Similar simplifications hold also when natural RK methods are applied to neutral DDEs (see Bellen, Jackiewicz and Zennaro [1]).

The main purpose of this paper is to characterize the class of the natural RK methods. In [8] it was noticed that any collocation method is natural, since the collocation polynomial is a NCE such that $B = A$, where A is the matrix of the equivalent RK formula (1.1).

In general, we shall prove that a nonconfluent RK method (1.1) (for which $c_i \neq c_j$ if $i \neq j$, see Dekker and Verwer [2]) is natural if and only if it is equivalent to a projection method in some polynomial space. More precisely, it is necessary and sufficient that there exists a linear projector Q_d of $C^0([t_0, t_1], \mathbb{R}^m)$ onto $\Pi_{d-1}([t_0, t_1], \mathbb{R}^m)$ (d is the degree of the NCE such that $B = A$), depending on A, $\mathbf{b}$ and $\mathbf{c}$, such that

$$\begin{aligned} &u(t_0) = y_0 \\ &u'(t) = Q_d f(t, u(t)), \qquad u \in \Pi_d([t_0, t_1], \mathbb{R}^n) \\ &y_1 = u(t_1). \end{aligned} \tag{1.8}$$

In view of this result we can say that the word "natural" in Definition 1.1 is justified. In fact it turns out that the RK formula (1.1) is naturally derived from a polynomial method which yields a really natural continuous approximation.

The paper ends with an analysis of the absolute stability properties of the natural RK methods, generalizing what was proved by Norsett and Wanner [4] for collocation.

2. Some General Facts about NCEs

Given a nonconfluent RK method (that is $c_i \neq c_j$ if $i \neq j$) of the form (1.1), we fix our attention on the v-vectors $\mathbf{b}$ and $\mathbf{c}$ and on the integer $p(\geqq 1)$, i.e. the order of the method.

As it is shown in [7], if a v-vector polynomial $\mathbf{b}(\theta)$ is a NCE, then it satisfies the following properties:

$$\mathbf{b}(0)=\mathbf{0} \quad \text{and} \quad \mathbf{b}(1)=\mathbf{b}; \tag{2.1}$$

$$\mathbf{b}'(\theta)^T \mathbf{c}^{r-1} \equiv \theta^{r-1}, \quad r=1, \ldots, d; \tag{2.2}$$

$$\int_0^1 \theta^s \mathbf{b}'(\theta)^T \mathbf{c}^{r-1} d\theta = 1/(r+s), \quad r=d+1, \ldots, p \quad \text{and} \quad s=0, \ldots, p-r, \tag{2.3}$$

where $\mathbf{c}^{r-1} := (c_1^{r-1}, \ldots, c_v^{r-1})^T$.

Note that (2.3) makes sense only if $d<p$.

Moreover, the degree d of a NCE is subject to the following inequality:

$$q \leqq d \leqq \mu, \tag{2.4}$$

where $q := [(p+1)/2]$ and $\mu := \min\{v, p\}$.

Definition 2.1. A v-vector polynomial $\mathbf{b}(\theta)$ of degree d satisfying (2.4) is a *feasible interpolant* for the RK method (1.1) if it verifies conditions (2.1), (2.2) and (2.3).

It is not true that every feasible interpolant is also a NCE. This fact can be understood by noticing that (2.2) and (2.3) are conditions corresponding only to some special elementary differentials of the Butcher series (see [7]). In general, only NCEs of minimal degree q are assured to exist. We can conclude that different RK methods of the same order p based on the same $\mathbf{b}$ and $\mathbf{c}$ can be more or less rich of NCEs. Obviously, this depends on the matrix A.

Now observe that, since $\mu \leqq p$, (2.3) implies

$$\int_0^1 \theta^s \mathbf{b}'(\theta)^T \mathbf{c}^{r-1} d\theta = 1/(r+s), \quad r=d+1, \ldots, \mu \quad \text{and} \quad s=0, \ldots, p-r. \tag{2.5}$$

Remark 2.2. Looking at the proof of (2.4) in [7] (Theorem 5), we observe that (2.2) and (2.5) for $r=d+1$ are sufficient to prove the inequality $q \leqq d \leqq v$. To get $d \leqq p$, one must instead consider the condition (1.6), which implies (2.2) as

a particular case, the condition (1.5), which is equivalent to (2.1), and the order p of the method, i.e. (1.3).

Now consider a v-vector polynomial $\mathbf{b}(\theta)$ of degree d satisfying (2.4), (2.1), (2.2) and (2.5) and, for any polynomial $\pi \in \Pi_{v-1}$, define

$$P_{v,d}\,\pi(\theta) := \sum_{i=1}^{v} b_i'(\theta)\,\pi(c_i). \tag{2.6}$$

We can easily conclude that $P_{v,d}$ is a linear projector of Π_{v-1} onto Π_{d-1} such that

$$\int_0^1 \theta^s P_{v,d}\,\pi(\theta)\,d\theta = \int_0^1 \theta^s \pi(\theta)\,d\theta \qquad \text{for all } s \geqq 0 \tag{2.7}$$

and $\pi \in \Pi_{\mu-1}$ such that $s + \deg(\pi) \leqq p-1$.

Moreover,

$$b_i'(\theta) = P_{v,d}\,l_i(\theta), \qquad i = 1, \ldots, v, \tag{2.8}$$

where the $l_i(\theta)$'s are the Lagrange coefficients

$$l_i(\theta) = \prod_{j \neq i}^{1,v} \frac{\theta - c_j}{c_i - c_j}, \qquad i = 1, \ldots, v. \tag{2.9}$$

This is yield by (2.6), since $l_i \in \Pi_{v-1}$ and $l_i(c_j) = \delta_{ij}$ (the Kronecker δ).

Finally, remark that

$$\int_0^1 P_{v,d}\,l_i(\theta)\,d\theta = b_i, \qquad i = 1, \ldots, v. \tag{2.10}$$

Definition 2.3. Given an integer d satisfying (2.4), a linear projector $P_{v,d}$ of Π_{v-1} onto Π_{d-1} is a *feasible projector* for the RK method (1.1) if it verifies conditions (2.7) and (2.10).

Lemma 2.4. *A v-vector polynomial $\mathbf{b}(\theta)$ of degree d satisfying* (2.4), (2.1), (2.2) *and* (2.5) *is a feasible interpolant.*

Proof. It is sufficient to prove that (2.5) implies (2.3) in the case $\mu = v$. To this aim, assume $\mu = v < p$ and let $v+1 \leqq r \leqq p$ and $0 \leqq s \leqq p-r$. Furthermore, consider the v-vector polynomial $\mathbf{l}(\theta) := (l_1(\theta), \ldots, l_v(\theta))^T$ of degree $v-1$ (see (2.9)).

Then, since $0 \leqq s \leqq p-r \leqq p-v-1$, by (2.8) and (2.7) we get

$$\int_0^1 \theta^s \mathbf{b}'(\theta)^T \mathbf{c}^{r-1}\,d\theta = \int_0^1 \theta^s \left(\sum_{i=1}^{v} P_{v,d}\,l_i(\theta)\,c_i^{r-1} \right) d\theta = \int_0^1 \theta^s \mathbf{l}(\theta)^T \mathbf{c}^{r-1}\,d\theta. \tag{2.11}$$

Now observe that, since the RK method (1.1) has order p, the quadrature formula with nodes c_i and weights b_i has polynomial order $\geqq p-1$ and therefore the polynomial $M(\theta) = (\theta - c_1)\ldots(\theta - c_v)$ is orthogonal to Π_{p-v-1}. Thus, since

$\theta^{r-1} - \mathbf{l}(\theta)^T \mathbf{c}^{r-1} = \pi(\theta)\, M(\theta)$ with $\pi \in \Pi_{r-\nu-1}$ and since $s+r-\nu-1 \leqq p-\nu-1$, we obtain

$$\int_0^1 \theta^s \mathbf{l}(\theta)^T \mathbf{c}^{r-1}\, d\theta = 1/(r+s), \tag{2.12}$$

which, together with (2.11), completes the proof. □

Theorem 2.5. *The feasible interpolants of degree d form an affine manifold F_d of $(\Pi_d)^\nu$ whose dimension is*

$$\delta_{\nu,d} = d(\nu-d) - (\mu-d)(2p+1-\mu-d)/2 - (\nu-\mu). \tag{2.13}$$

Moreover, formulas (2.6) *and* (2.8) *with* $\mathbf{b}(0)=\mathbf{0}$ *define a bilinear correspondence between F_d and the set $\mathscr{P}_{\nu,d}$ of the feasible projectors $P_{\nu,d}$, which then is an affine manifold of $\mathscr{L}(\Pi_{\nu-1}, \Pi_{d-1})$.*

Proof. Assume that $\mathbf{b}_{(1)}(\theta), \ldots, \mathbf{b}_{(k)}(\theta) \in F_d$ and choose $\rho_1, \ldots, \rho_k \in \mathbb{R}$ such that $\rho_1 + \ldots + \rho_k = 1$. Then it is quite easy to see that $\mathbf{b}(\theta) := \rho_1 \mathbf{b}_{(1)}(\theta) + \ldots + \rho_k \mathbf{b}_{(k)}(\theta)$ satisfies (2.1), (2.2) and (2.3). Thus F_d is an affine manifold.

In order to compute its dimension $\delta_{\nu,d}$, note that the general ν-vector polynomial $\mathbf{b}(\theta)$ such that $\mathbf{b}(0)=\mathbf{0}$ (see (2.1)) is determined by νd coefficients. On the other hand, (2.2) imposes d^2 linearly independent conditions and there are other $(\mu-d)(2p+1-\mu-d)/2$ linearly independent conditions given by (2.5), which, by Lemma 2.4, is equivalent to (2.3).

If $\mu=\nu\leqq p$, then the weights b_i of the quadrature formula (1.1 b) of the RK method are uniquely determined and hence $\mathbf{b}(1)=\mathbf{b}$ (see (2.1)) is already included in (2.5) for $s=0$. On the contrary, if $\mu=p<\nu$, then the vector $\mathbf{b}$ of the weights can be chosen in an affine manifold of $\mathbb{R}^\nu$ whose dimension is $\nu-p$ and hence, in this case, (2.5) for $s=0$ includes only p of the ν conditions $\mathbf{b}(1)=\mathbf{b}$. We can conclude that, in any case, other $\nu-\mu$ conditions are to be counted to get formula (2.13).

To prove the last part of the theorem we have only to observe that, if $P_{\nu,d}$ is a feasible projector for the RK method (1.1), then (2.8) with $\mathbf{b}(0)=\mathbf{0}$ defines a ν-vector polynomial $\mathbf{b}(\theta)$ of degree d satisfying (2.4), (2.1), (2.2) and (2.5). Hence, by Lemma 2.4, it is a feasible interpolant. □

Theorem 2.6. *The NCEs of degree d form an affine submanifold N_d of F_d.*

Proof. It is obvious that $N_d \subseteq F_d$. If $\mathbf{b}_{(1)}(\theta), \ldots, \mathbf{b}_{(k)}(\theta) \in N_d$, then $u_{(1)}(t), \ldots, u_{(k)}(t)$ given by (1.4) satisfy (1.5), (1.6) and (1.7). Choose $\rho_1, \ldots, \rho_k \in \mathbb{R}$ such that $\rho_1 + \ldots + \rho_k = 1$ and consider the function $u(t) := \rho_1 u_{(1)}(t) + \ldots + \rho_k u_{(k)}(t)$ which, by (1.4), corresponds to $\mathbf{b}(\theta) = \rho_1 \mathbf{b}_{(1)}(\theta) + \ldots + \rho_k \mathbf{b}_{(k)}(\theta)$. It is quite easy to see that also $u(t)$ verifies (1.5), (1.6) and (1.7) and thus $\mathbf{b}(\theta) \in N_d$. □

These theorems yield an upper bound to the number of NCEs of degree d that a RK method may possess.

Note that for $d=\nu\leqq p$ we get $\delta_{\nu,\nu}=0$. This means that only one NCE of degree ν can exist, as it was already proved in [7]. Furthermore, observe that

in this case the corresponding feasible projector $P_{v,v}$ is the identity map on Π_{v-1}.

The last part of this section is devoted to study the relationships among the various manifolds F_d and N_d, where d varies in the interval $[q, \mu]$. Obviously, this makes sense only for methods such that $q<\mu$.

To this purpose, consider two integers $d, d' \in [q, \mu]$ with $d'<d$ and then a linear projector $P_{d,d'}$ of Π_{d-1} onto $\Pi_{d'-1}$ such that

$$\int_0^1 \theta^s P_{d,d'}\pi(\theta)\,d\theta = \int_0^1 \theta^s \pi(\theta)\,d\theta \qquad \text{for all } s \geqq 0 \tag{2.14}$$

and $\pi \in \Pi_{d-1}$ such that $s+\deg(\pi) \leqq p-1$.

Following the same line of the previous proofs, we can easily get the following result.

Theorem 2.7. *The set $\mathscr{P}_{d,d'}$ of the projectors $P_{d,d'}$ is an affine manifold of $\mathscr{L}(\Pi_{d-1}, \Pi_{d'-1})$ whose dimension is*

$$\delta_{d,d'} = (d-d')(d+3d'-2p-1)/2. \tag{2.15}$$

The next theorem gives the relationships between $\mathscr{P}_{v,d}$ and $\mathscr{P}_{v,d'}$, where $d'<d$.

Theorem 2.8. *If $P_{v,d} \in \mathscr{P}_{v,d}$ and $P_{d,d'} \in \mathscr{P}_{d,d'}$ with $d'<d$, then the linear operator $P_{v,d'}$ defined by*

$$P_{v,d'} = P_{d,d'} \cdot P_{v,d} \tag{2.16}$$

belongs to $\mathscr{P}_{v,d'}$.

Conversely, if $P_{v,d'} \in \mathscr{P}_{v,d'}$ with $q \leqq d' \leqq \mu-1$, then for every $d>d'$, $d \leqq \mu$, there exist $P_{v,d} \in \mathscr{P}_{v,d}$ and $P_{d,d'} \in \mathscr{P}_{d,d'}$ such that (2.16) holds.

In both cases $P_{d,d'}$ is the restriction of $P_{v,d'}$ to Π_{d-1}.

Proof. The direct implication is easily got by using the properties of $P_{d,d'}$. The fact that $P_{d,d'}$ is the restriction of $P_{v,d'}$ to Π_{d-1} is yielded by (2.16), since $P_{v,d}$ is the identity map on Π_{d-1}.

The converse is trivial if $d=v=\mu$. Otherwise, given $P_{v,d'}$ and $d<v$ such that $d'<d \leqq \mu$, note that the restriction of $P_{v,d'}$ to Π_{d-1} is a projector $P_{d,d'} \in \mathscr{P}_{d,d'}$. Then consider the linear projector $P_{v,d}$ of Π_{v-1} onto Π_{d-1} such that $P_{v,d}\theta^{r-1} = P_{v,d'}\theta^{r-1}$, $r=d+1, \ldots, v$.

It is immediately seen that $P_{v,d}$ satisfies (2.7) and that (2.16) holds. Finally, the validity of (2.10) for $P_{v,d'}$, (2.16) and (2.14) for $s=0$ yield the validity of (2.10) also for $P_{v,d}$. Hence $P_{v,d} \in \mathscr{P}_{v,d}$. □

It is interesting to remark that our construction of $P_{v,d}$ satisfying (2.16) for a given $P_{v,d'}$ is not necessarily the unique one possible. In other words, the dimension $\delta_{v,d'}$ of $\mathscr{P}_{v,d'}$ may be strictly less than $\delta_{v,d}+\delta_{d,d'}$ (see (2.13) and (2.15)).

As a consequence of Theorem 2.8 and Theorem 2.5 we have the following corollary.

Corollary 2.9. *A v-vector polynomial $\boldsymbol{\beta}(\theta)$ of degree d' such that $q \leqq d' \leqq \mu-1$ and $\boldsymbol{\beta}(0)=\mathbf{0}$ is a feasible interpolant for the RK method (1.1) if and only if for*

every $d > d'$, $d \leq \mu$, *there exist a feasible interpolant* $\mathbf{b}(\theta)$ *of degree* d *and a projector* $P_{d,d'} \in \mathscr{P}_{d,d'}$ *such that*

$$\beta_i'(\theta) = P_{d,d'}\, b_i'(\theta), \qquad i = 1, \ldots, v. \tag{2.17}$$

We conclude this section by improving the result given by Theorem 6 in [7].

Theorem 2.10. *If the RK method* (1.1) *has a NCE* $\mathbf{b}(\theta)$ *of degree* $d > q$, *then every* $\boldsymbol{\beta}(\theta)$ *defined by* (2.17) *is also a NCE of degree* d' *for every* $d' = q, \ldots, d-1$.

Moreover, if $N_d = F_d$ *for* $d > q$, *then we have also* $N_{d'} = F_{d'}$ *for every* $d' = q, \ldots, d-1$.

Proof. The latter part of the theorem is clearly a consequence of the former part and of Corollary 2.9.

To prove the former part, one can look at the proof of Theorem 6 in [7] and just observe that it is still valid by using a projector $P_{d,d'} \in \mathscr{P}_{d,d'}$ instead of the projector used there. We do not do this here for the sake of brevity, since a lot of notation should be introduced to this aim. □

3. Projection Methods and Equivalence Theorem

Consider v distinct abscissae $c_1, \ldots, c_v \in [0, 1]$, an integer d satisfying (2.4) and a projector $P_{v,d}$ of Π_{v-1} onto Π_{d-1} satisfying (2.7). Then we can define the linear projector Q_d of $C^0([t_0, t_1], \mathbb{R}^m)$ onto $\Pi_{d-1}([t_0, t_1], \mathbb{R}^m)$ by

$$Q_d\, \varphi(t) := \sum_{i=1}^{v} P_{v,d}\, l_i(\theta)\, \varphi(t_0 + c_i h), \qquad \theta := (t - t_0)/h, \tag{3.1}$$

where the $l_i(\theta)$'s are the Lagrange coefficients (2.9).

Lemma 3.1. *Assume that for* $h \in (0, h_0]$, $h_0 > 0$, *all the derivatives of* $\varphi_h \in C^p([t_0, t_1], \mathbb{R}^m)$ *and* $F_h \in C^p([t_0, t_1], \mathscr{L}(\mathbb{R}^m))$ *are uniformly bounded with respect to* h. *Then the projector* Q_d *defined by* (3.1) *is such that*

$$\max_{t_0 \leq t \leq t_1} |\varphi_h(x) - Q_d\, \varphi_h(x)| = O(h^d) \tag{3.2}$$

and

$$\left| \int_{t_0}^{t_1} F_h(x)[\varphi_h(x) - Q_d\, \varphi_h(x)]\, dx \right| = O(h^{p+1}). \tag{3.3}$$

Proof. By the hypotheses on $F_h(x)$ and $\varphi_h(x)$, we can expand them in Taylor polynomials at the point t_0, whose coefficients are uniformly bounded as $h \to 0$. Then, since Q_d is a projector onto $\Pi_{d-1}([t_0, t_1], \mathbb{R}^m)$, we immediately get (3.2). In order to obtain (3.3), we must instead make use of the orthogonality properties (2.7) and (2.11)–(2.12) of $P_{v,d}$. □

Theorem 3.2. *A projection method* (1.8) *in which* Q_d *is given by* (3.1) *is equivalent to a nonconfluent RK method* (1.1) *of order* p, *and* $P_{v,d}$ *is a feasible projector of it.*

Proof. As done by Norsett and Wanner [4, 5] for collocation, with

$$K_i = f(t_0 + c_i h, u(t_0 + c_i h)), \qquad i = 1, \ldots, v, \tag{3.4a}$$

$$a_{ij} = \int_0^{c_i} P_{v,d}\, l_j(\theta)\, d\theta, \qquad i, j = 1, \ldots, v, \tag{3.4b}$$

$$b_i = \int_0^1 P_{v,d}\, l_i(\theta)\, d\theta, \qquad i = 1, \ldots, v, \tag{3.4c}$$

the equivalence with a v-stage RK method (1.1) is easily obtained.

To see that the order is p, consider the Grobner-Alekseev nonlinear variation-of-constants formula for the IVP (1.2)

$$y(t) - u(t) = \int_{t_0}^{t} K(t, x)[f(x, u(x)) - u'(x)]\, dx, \tag{3.5}$$

where $K(t, x)$ is a variational matrix depending on $u(t)$ (see again [4] or [5]).

Then, by (1.8), we get

$$y(t_1) - y_1 = \int_{t_0}^{t_1} K(t_1, x)[f(x, u(x)) - Q_d f(x, u(x))]\, dx. \tag{3.6}$$

By the smoothness of $f(t, y)$, some standard analysis yields the uniform boundedness of the derivatives of $u(t)$ and $K(t, x)$ as $h \to 0$. Therefore the hypotheses of Lemma 3.1 are satisfied and (3.3) yields (1.3).

Finally, by (2.7) and (3.4c), we have that $P_{v,d}$ is a feasible projector. □

The next theorem is a kind of converse of Theorem 3.2.

Theorem 3.3. *If a nonconfluent RK method* (1.1) *of order* p *is equivalent to a projection method* (1.8), *then the projector* Q_d *is defined by* (3.1), *where* $P_{v,d}$ *is a feasible projector.*

Proof. In order to be equivalent to a v-stage RK method of the form (1.1), a projection method (1.8) must be determined by a projector Q_d of $C^0([t_0, t_1], \mathbb{R}^m)$ onto $\Pi_{d-1}([t_0, t_1], \mathbb{R}^m)$ which makes use only of function evaluations at the points $t_0 + c_i h$, $i = 1, \ldots, v$. Thus we have

$$Q_d \varphi(t) = \sum_{i=1}^{v} \gamma_i(\theta)\, \varphi(t_0 + c_i h), \qquad \theta := (t - t_0)/h, \tag{3.7}$$

where $\gamma_i(\theta) \in \Pi_{d-1}$, $i = 1, \ldots, v$.

By the equivalence of the methods applied to the simple IVP $y'(t)=f(t)$, $y(t_0)=t_0$, by (1.1 b) and (3.7), we immediately get

$$\int_0^1 \gamma_i(\theta)\,d\theta = b_i, \qquad i=1, \ldots, v. \tag{3.8}$$

The method being of order p, (1.3) must hold for every IVP (1.2) and thus, by (3.7) and (3.6) and by expanding $K(t_1, x)$ and $f(x, u(x))$ in Taylor polynomials at the point t_0, we can easily conclude that

$$\int_0^1 \theta^s \boldsymbol{\gamma}(\theta)^T \mathbf{c}^{r-1}\,d\theta = 1/(r+s), \qquad r=1, \ldots, p \quad \text{and} \quad s=0, \ldots, p-r, \tag{3.9}$$

where $\boldsymbol{\gamma}(\theta) := (\gamma_1(\theta), \ldots, \gamma_v(\theta))^T$.

If we define $\mathbf{b}'(\theta) := \boldsymbol{\gamma}(\theta)$, $\mathbf{b}(0) := \mathbf{0}$, by (3.7), (3.8) and (3.9), we have that $\mathbf{b}(\theta)$ satisfies (2.1), (2.2) and (2.3). Hence, by Remark 2.2, also the inequality $q \leqq d \leqq v$ holds. Moreover, since Q_d is a projector onto $\Pi_{d-1}([t_0, t_1], \mathbb{R}^m)$, by (3.6) and (1.3), it is $d \leqq p$ and thus d satisfies (2.4).

Finally, (2.8) completes the proof. □

We have characterized the projectors Q_d which define a RK method. Now we are in a position to state the main result of this paper.

Theorem 3.4. *A nonconfluent v-stage RK method* (1.1) *is natural if and only if it is equivalent to a projection method* (1.8).

Proof. First assume that the RK method (1.1) is equivalent to a projection method (1.8). In order to prove that it is natural, by Theorem 3.3 and by (3.4 b), we have only to show that the polynomial $u(t)$, solution of (1.8), satisfies (1.6) and (1.7).

To do this, consider once again (3.5) and, consequently,

$$y'(t)-u'(t) = \int_{t_0}^{t} \frac{\partial K}{\partial t}(t, x)[f(x, u(x)) - Q_d f(x, u(x))]\,dx + f(t, u(t)) - Q_d f(t, u(t))$$

and

$$\int_{t_0}^{t_1} G(t)[y'(t)-u'(t)]\,dt = \int_{t_0}^{t_1} H(x)[f(x, u(x)) - Q_d f(x, u(x))]\,dx,$$

where

$$H(x) = G(x) + \int_x^{t_1} G(\xi)\cdot\frac{\partial K}{\partial t}(\xi, x)\,d\xi.$$

Then Lemma 3.1 furnishes the desired result.

Conversely, assume that the nonconfluent RK method (1.1) is natural. This means that it has a NCE $\mathbf{b}(\theta)$ of some degree d such that (3.4 b) and (3.4 c)

hold, where $P_{v,d}$ is the corresponding feasible projector. Then, by Theorem 3.2, we can conclude that it is equivalent to the projection method (1.8) in which Q_d is given by (3.1). □

The above characterization allows us to classify as natural some other classes of RK methods besides collocation. Incidentally, note that for collocation methods the feasible projector $P_{v,v}$ in (3.1) is the identity map on Π_{v-1}.

A simple but popular example is given by the so called θ-methods, for which $v=2$, $d=1$ and $\mu=p$ ($=1$ or $=2$). The feasible projector $P_{2,1}$ in (3.1) is defined by

$$P_{2,1}\,\pi(\theta)=(1-\rho)\,\pi(0)+\rho\,\pi(1).$$

(Here the usual parameter θ has been replaced by ρ not to change the notation used in this paper).

Moreover, we quote the *one-step subregion methods* studied by Vermiglio [6], for which the abscissae $c_1, \ldots, c_v$ are the v Lobatto points in [0, 1], $d=q=v-1$, $p=2v-2$ and $\mu=v$. The feasible projector $P_{v,v-1}$ is the only one possible for these data (in fact (2.13) yields $\delta_{v,v-1}=0$), that is the interpolation projector at the $v-1$ Gaussian points in [0, 1].

Finally, we quote the *fully implicit methods* studied by Jackiewicz [3] (see also Bellen, Jackiewicz and Zennaro [1]), in which $d=v-1$ and $p\geqq v-1$. The feasible projector $P_{v,v-1}$ is the interpolation projector at $v-1$ distinct points, other than the c_i's.

The next theorem completely characterizes the NCEs of a natural RK method.

Theorem 3.5. *Assume that the nonconfluent RK method* (1.1) *is natural and let* (1.8) *be the equivalent projection method. Then* $N_k=F_k$ *for* $k=q, \ldots, \sigma$, $\sigma:=\min\{d+1, v\}$, *and, even if* $v>d+1$, $N_k=\emptyset$ *for* $k>d+1$.

Proof. Consider a feasible interpolant $\boldsymbol{\beta}(\theta)$ of degree σ. To prove that it is a NCE, in view of (3.4a), we must show that the polynomial

$$v(t_0+\theta h):=y_0+h\sum_{i=1}^{v}\beta_i(\theta)f(t_0+c_i h, u(t_0+c_i h)) \tag{3.10}$$

is such that

$$\max_{t_0\leq t\leq t_1}|y'(t)-v'(t)|=O(h^\sigma) \tag{3.11}$$

and

$$\left|\int_{t_0}^{t_1} G(t)[y'(t)-v'(t)]\,dt\right|=O(h^{p+1}) \tag{3.12}$$

for every sufficiently smooth matrix-valued function $G(t)$.

If $P_{v,\sigma}$ is the feasible projector corresponding to $\boldsymbol{\beta}(\theta)$ and if Q_σ is the projector consequently defined by (3.1), then (3.10) yields

$$\begin{aligned} y'(t)-v'(t)&=f(t, y(t))-Q_\sigma f(t, u(t))\\ &=(I-Q_\sigma)f(t, y(t))\\ &\quad-(I-Q_\sigma)[f(t, y(t))-f(t, u(t))]\\ &\quad+f(t, y(t))-f(t, u(t)), \end{aligned} \tag{3.13}$$

where I is the identity map on $C^0([t_0, t_1], \mathbb{R}^m)$.

Since u satisfies (1.6), we have

$$\max_{t_0 \leqq t \leqq t_1} |f(t, y(t)) - f(t, u(t))| = O(h^{d+1}).$$

Therefore, since $\sigma \leqq d+1$, (3.2) (in which d is replaced by σ) yields (3.11).

The validity of (3.12) is a consequence of (3.13), (3.3) and (1.7).

We can conclude that $N_\sigma = F_\sigma$ and therefore, by Theorem 2.10, that $N_k = F_k$ for $k = q, \ldots, \sigma$.

To complete the proof, assume $\nu > d+1$ and suppose that a NCE $\boldsymbol{\beta}(\theta)$ of degree $k > d+1$ exists. Then consider the IVP $y'(t) = y(t)$, $y(0) = 1$, whose solution is $\exp(t)$. Reasoning as above, since $f(t, u(t)) = u(t)$ is a polynomial of degree $\leqq d$ and since $k-1 > d$, for the function $v(t)$ defined by (3.10) we get $v'(t) = Q_k u(t) = u(t)$ and thus, by (3.11) in which σ is replaced by k,

$$\max_{t_0 \leqq t \leqq t_1} |\exp(t) - u(t)| = O(h^k).$$

This fact is clearly an absurde. In fact a polynomial of degree $\leqq d \leqq k-2$ cannot be uniformly a k-th order approximation to the exponential. □

At the light of this result we can say that the natural RK (or projection) methods are the richest of NCEs. In particular, for a collocation method every feasible interpolant is also a NCE.

We conclude this section with some remarks about the implementation of a ν-stage natural RK method.

Although at first sight (1.1) seems to lead to the solution of a $\nu \times m$-dimensional nonlinear system each step, the true dimension of the problem is instead $d \cdot m$ only, where d is the degree of the polynomial $u(t)$ in (1.8). In other words, although the method makes use of function evaluations at ν points, its "difficulty" is given by the degree d rather than by the number of stages ν. Alternatively, we could say that d is the *degree of implicitness* of the method.

To see this, we can write the general polynomial $u \in \Pi_d([t_0, t_1], \mathbb{R}^m)$ such that $u(t_0) = y_0$ in the form

$$u(t) = y_0 + h \sum_{j=1}^{d} \pi_j(\theta) H_j, \qquad \theta := (t - t_0)/h, \tag{3.14}$$

where $\pi_j(0) = 0$, $j = 1, \ldots, d$, $\{\pi'_1, \ldots, \pi'_d\}$ is a suitable basis of Π_{d-1} and $H_1, \ldots, H_d \in \mathbb{R}^m$.

Then, by (3.1) and (2.8), substitution of (3.14) into (1.8) yields

$$\sum_{j=1}^{d} \pi'_j(\theta) H_j = \sum_{k=1}^{\nu} b'_k(\theta) f(t_0 + c_k h, y_0 + h \sum_{j=1}^{d} \pi_j(c_k) H_j), \tag{3.15}$$

which is a system whose unknowns are the m-vectors $H_1, \ldots, H_d$.

In order to simplify (3.15), we recall that there are exactly d polynomials among the $b_k(\theta)$'s which are linearly independent. We can suppose, without

any restriction, that these polynomials are $b_1(\theta), \ldots, b_d(\theta)$. Therefore, if $d<v$, there exist coefficients α_{jk} such that

$$b_k(\theta)=\sum_{j=1}^{d} \alpha_{jk}\, b_j(\theta), \qquad k=d+1, \ldots, v. \tag{3.16}$$

By choosing

$$\pi_i'(\theta):=b_i'(\theta), \qquad i=1, \ldots, d,$$

since they are linearly independent, (3.15) and (3.16) furnish

$$H_i=f(t_0+c_i h, y_0+h \sum_{j=1}^{d} b_j(c_i)\, H_j) + \sum_{k=d+1}^{v} \alpha_{ik} f(t_0+c_k h, y_0+h \sum_{j=1}^{d} b_j(c_k)\, H_j), \qquad i=1, \ldots, d. \tag{3.17}$$

If $d=v$ (collocation methods), then we find again (1.1) with $H_i=K_i$, $i=1, \ldots, v$.

Otherwise, if $d<v$, then

$$H_i=K_i+\sum_{k=d+1}^{v} \alpha_{ik}\, K_k, \qquad i=1, \ldots, d.$$

The form of (3.17), as well as that of (1.1), suggests, for example, the application of a direct iterative scheme. However, for any d, the number of function evaluations per iteration always is equal to v.

4. Stability Analysis

In this section we briefly analyze the stability properties of the natural RK (or projection) methods with respect to the usual linear test equation

$$\begin{aligned} y'(t)&=\lambda y(t), \qquad \lambda=\mathbb{C} \\ y(0)&=1. \end{aligned} \tag{4.1}$$

The main appearance is that the result given by Norsett and Wanner [4] for collocation similarly extends to all projection methods.

In order to study the *absolute stability function* $R(z)$, $z:=h\lambda$, which is such that $y_1=R(z)\, y_0$ when the projection method (1.8) is applied to (4.1), without any restriction we can fix $h=1$. Moreover, since the case $d=v$ leads to collocation, we assume $d\leqq v-1$.

Therefore Theorem 3.3 yields

$$\begin{aligned} u'(\theta)&=zP_{v,d}\, u(\theta) \\ u(0)&=1, \end{aligned} \tag{4.2}$$

where $P_{v,d}$ is the feasible projector in (3.1).

Remark that $R(z)=y_1=u(1)$.

Since u is a polynomial of degree $\leqq d$, (4.2) is equivalent to

$$\begin{aligned} u'(\theta) &= zP_{d+1,d}\,u(\theta) \\ u(0) &= 1, \end{aligned} \tag{4.3}$$

where $P_{d+1,d}$ is the restriction of $P_{v,d}$ to Π_d.

Theorem 4.1. *The solution of* (4.2) *is the polynomial*

$$u(\theta)=\left(\sum_{j=0}^{d}\rho^{(d-j)}(\theta)\,z^j\right)\Big/\left(\sum_{j=0}^{d}\rho^{(d-j)}(0)\,z^j\right), \tag{4.4}$$

where

$$\rho(\theta)=\theta^d-P_{d+1,d}\,\theta^d \tag{4.5}$$

is a monic polynomial of degree d.

Proof. The polynomial $u(\theta)$ can be written as

$$u(\theta)=u_d\,\theta^d+\text{terms of lower degree.} \tag{4.6}$$

We have $u_d\neq 0$. In fact, if it were not so, then we should have $u\in\Pi_{d-1}$ and therefore, by (4.3), $u'(\theta)=zu(\theta)$, $u(0)=1$, that would yield the absurde equality $u(\theta)=\exp(z\theta)$.

Now define

$$\rho(\theta):=[u(\theta)-P_{d+1,d}\,u(\theta)]/u_d, \tag{4.7}$$

which, by (4.6), yields (4.5).

By (4.3) and (4.7) we get

$$u'(\theta)-zu(\theta)=K\rho(\theta), \qquad K:=-zu_d.$$

Reasoning as in the proof of Theorem 4 in [4], the application of the variation-of-constants formula and repeated partial integration lead to (4.4). □

In particular, for $\theta=1$ formula (4.4) furnishes the form of the absolute stability function

$$R(z)=\left(\sum_{j=0}^{d}\rho^{(d-j)}(1)\,z^j\right)\Big/\left(\sum_{j=0}^{d}\rho^{(d-j)}(0)\,z^j\right). \tag{4.8}$$

We can conclude that the absolute stability properties of a projection method (1.8) are determined by the polynomial $\rho(\theta)$. On the other hand, formula (4.5) defines a one-to-one correspondence between the polynomials $\rho(\theta)$ and the projectors $P_{d+1,d}$ and consequently, for fixed d, formula (4.8) between the absolute stability functions $R(z)$ and the projectors $P_{d+1,d}$.

Therefore the number of possible absolute stability functions for a projection method (1.8) is given by Theorem 2.7, which states that the projectors $P_{d+1,d}$ form an affine manifold whose dimension is $\delta_{d+1,d}=2d-p$. It is remarkable

that this number is independent of the number v of stages of the equivalent natural RK method.

A consequence of this is that different projection methods may have the same absolute stability function $R(z)$.

For instance, if the polynomial $\rho(\theta)$ has d distinct roots $\xi_1, \ldots, \xi_d$ in [0, 1], then the projection method (1.8) has the same absolute stability function of the collocation method at the points $\xi_1, \ldots, \xi_d$ (compare once again [4]).

However, this is not a general occurence.

Theorem 4.2. *The polynomial $\rho(\theta)$ defined by* (4.5) *has at least $p-d$ distinct roots in* (0, 1) *where it changes its sign.*

In particular, if $d=q$ and $p=2q$, then $\rho(\theta)$ is uniquely determined and coincides with the (monic) Legendre polynomial of degree q in [0, 1].

If $d=q$ and $p=2q-1$, then $\delta_{d+1,d}=1$ and, in any case, $\rho(\theta)$ has q distinct real roots. Moreover, at least $q-1$ of them are in (0, 1).

Proof. By (2.14) and (4.5), we have

$$\int_0^1 \theta^s \rho(\theta)\, d\theta = 0, \qquad s=0, \ldots, p-d-1. \tag{4.9}$$

If we suppose that $\rho(\theta)$ changes its sign only at r points $\xi_1, \ldots, \xi_r \in (0, 1)$, $r \leqq p-d-1$, then we should have

$$\int_0^1 (\theta-\xi_1)\ldots(\theta-\xi_r)\, \rho(\theta)\, d\theta = 0.$$

Since the integrand has the same sign in [0, 1], we should get $\rho(\theta) \equiv 0$, which makes the absurde.

Now assume $d=q$ and $p=2q$. Then (2.15) yields $\delta_{d+1,d}=0$ and therefore $\rho(\theta)$ is uniquely determined. Moreover, (4.9) states that it is orthogonal to Π_{q-1} in [0, 1] and thus it must coincide with the (monic) Legendre polynomial of degree q in [0, 1].

If $d=q$ and $p=2q-1$, then (2.15) yields $\delta_{d+1,d}=1$ and therefore $\rho(\theta)$ is not uniquely determined. In any case, it has $q-1$ distinct roots in (0, 1) where it changes its sign. It is a real polynomial and thus the remaining root is real. If this last root coincided with another root, it would be a double root and the sign of $\rho(\theta)$ would not change. Therefore all the roots are distinct. □

We conclude the paper with a corollary of the above Theorem 4.2, which characterizes the absolute stability function $R(z)$ of a projection method (1.8) in which $d=q$.

Corollary 4.3. *If $d=q$, then the projection method* (1.8) *has the same absolute stability function $R(z)$ of a collocation method at q points (possibly one of them lies outside* [0, 1]).

In particular, if $p=2q$, then these points are the q Gaussian points in [0, 1] *and thus $R(z)$ is A-acceptable.*

References

1. Bellen, A., Jackiewicz, Z., Zennaro, M.: Stability analysis of one-step methods for neutral delay differential equations Numer. Math. **52**, 605–619
2. Dekker, K., Verwer, J.G.: Stability of Runge-Kutta methods for stiff nonlinear differential equations. Amsterdam: North-Holland 1984
3. Jackiewicz, Z.: One-step methods of any order for neutral functional differential equations. SIAM J. Numer. Anal. **21**, 486–511 (1984)
4. Norsett, S.P., Wanner, G.: The real-pole sandwich for rational approximations and oscillation equations. BIT **19**, 79–94 (1979)
5. Norsett, S.P., Wanner, G.: Perturbed collocation and Runge-Kutta methods. Numer. Math. **38**, 193–208 (1981)
6. Vermiglio, R.: A one-step subregion method for delay differential equations. Calcolo **22**, 429–455 (1985)
7. Zennaro, M.: Natural continuous extensions of Runge-Kutta methods. Math. Comput. **46**, 119–133 (1986)
8. Zennaro, M.: *P*-stability properties of Runge-Kutta methods for delay differential equations. Numer. Math. **49**, 305–318 (1986)

Received March 17, 1987 / November 27, 1987