Monografías del Seminario Matemático García de Galdeano **31**, 179–189 (2004)

# CONSISTENCY OF A CLASS OF RK METHODS FOR INDEX-2 DAES

Inmaculada Higueras and Teo Roldán

**Abstract.** When index 2 semi-explicit differential algebraic equations (DAEs) are solved with a Runge-Kutta method $(\mathcal{A}, b)$, a standard assumption is the regularity of the matrix coefficient $\mathcal{A}$. However, Runge-Kutta methods with singular matrix coefficient $\mathcal{A}$ can also be used for index 2 DAEs if the matrix $\mathcal{A}$ has a special structure. In this case, the standard consistency analysis is not longer valid. In this paper we give conditions to ensure a certain order of consistency.

*Keywords:* Differential-algebraic equations, DAE, ESDIRK methods

*AMS classification:* 65L07

## §1. Introduction

We consider semi-explicit index-2 differential algebraic systems of the form

$$\begin{cases} y' = f(y,z) & y(x_0) = y_0 \\ 0 \ = g(y) & z(x_0) = z_0 \ , \end{cases} \tag{1}$$

where $f : \mathbb{R}^l \times \mathbb{R}^m \longrightarrow \mathbb{R}^l$ and $g : \mathbb{R}^l \longrightarrow \mathbb{R}^m$ are sufficiently smooth functions, and the matrix $g_y f_z$ is invertible in a neighborhood of the solution of (1).

If we consider an $s$-stage Runge-Kutta method $(\mathcal{A}, b)$ to solve (1), a standard assumption is the regularity of the matrix $\mathcal{A}$. Nevertheless we can also use methods with singular matrices of the form

$$\begin{array}{c|cc} 0 & 0 & 0^t \\ \bar{c} & a & \bar{\mathcal{A}} \\ \hline & b_1 & \bar{b}^t \end{array} \tag{2}$$

where $a \in \mathbb{R}^{s-1}$ and $\bar{\mathcal{A}}$ is an $(s-1) \times (s-1)$ regular matrix [9]. On the method $(\mathcal{A}, b)$ we assume that conditions $C(1)$ and $B(1)$ hold, i.e.

$$a + \bar{\mathcal{A}}\bar{e} = \bar{c} \tag{3}$$

$$b_1 + \bar{b}^t\bar{e} = 1 \tag{4}$$

where $\bar{e} = (1, \dots, 1)^t \in \mathbb{R}^{s-1}$. Moreover, in order to have $R(\infty)$ bounded, with $R(z)$ the stability function of the method, we impose

$$b_1 - \bar{b}^t \bar{\mathcal{A}}^{-1} a = 0\,. \tag{5}$$

In this way, (3)-(5) imply

$$\bar{b}^t \bar{\mathcal{A}}^{-1} \bar{c} = 1\,. \tag{6}$$

For these methods, the first internal stages are $Y_1 = y_n$, $Z_1 = z_n$, and the rest of the stages are given by the non-linear system

$$\begin{aligned} \bar{Y}_{n+1} &= \bar{e} \otimes y_n + h\, a \otimes f(y_n, z_n) + h(\bar{\mathcal{A}} \otimes I_l) f(\bar{Y}_{n+1}, \bar{Z}_{n+1})\,, && (7)\\ 0 &= g(\bar{Y}_{n+1})\,, && (8) \end{aligned}$$

where $\bar{Y}_{n+1} = (Y_2^t, \dots, Y_s^t)^t \in \mathbb{R}^{l(s-1)}$, $f(\bar{Y}_{n+1}, \bar{Z}_{n+1}) = (f(Y_2, Z_2)^t, \dots, f(Y_s, Z_s)^t)^t \in \mathbb{R}^{l(s-1)}$, and in an analogous way for $\bar{Z}_{n+1}$ and $g(\bar{Y}_{n+1})$. The symbol $\otimes$ denotes the Kronecker product. As the matrix $\bar{\mathcal{A}}$ is regular, system (7)-(8) can be solved for $\bar{Y}_{n+1}$ and $\bar{Z}_{n+1}$.

Once these values have been computed, with condition (5), we obtain

$$y_{n+1} = R_0(\infty) y_n + (\bar{b}^t \bar{\mathcal{A}}^{-1} \otimes I_l)\, \bar{Y}_{n+1}\,,$$

where $R_0(\infty) = 1 - \bar{b}^t \bar{\mathcal{A}}^{-1} \bar{e}$, and similarly we can compute

$$z_{n+1} = R_0(\infty) z_n + (\bar{b}^t \bar{\mathcal{A}}^{-1} \otimes I_m)\, \bar{Z}_{n+1}\,.$$

If the method is stiffly accurate, i.e. $a_{si} = b_i$, $i = 1, \dots, s$, we simply obtain

$$y_{n+1} = \bar{Y}_{n+1,s} \qquad z_{n+1} = \bar{Z}_{n+1,s}\,.$$

Observe that in this case the numerical solution satisfies $g(y_{n+1}) = 0$. If the method is not stiffly accurate, the numerical solution must be projected onto the constraint $g(y) = 0$ (see [1], [9]). Examples of methods of the form (2) are Lobatto IIIA methods and the ESDIRK methods considered for example in [4], [11], [2] and [10].

Runge-Kutta methods with singular matrix coefficient $\mathcal{A}$ of the form (2) have been studied in [9]. In [9, Theorem 5.1] local errors for stiffly accurate methods are given in terms of the simplifying conditions $B(p)$, $C(q)$ and $D(r)$. More precisely, $B(p)$, $C(q)$ and $D(r)$ ensure that the local errors are $\delta y_h(x_0) = \vartheta(h^{\min\{p,2q,q+r+1\}+1})$, for the differential component, and $\delta z_h(x_0) = \vartheta(h^q)$ for the algebraic one. The following example shows that this order bound is not sharp.

**Example 1.** We consider the family of four stage stiffly accurate methods satisfying $B(3)$ and $C(2)$ [4],

$$\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
2\lambda & \lambda & \lambda & 0 & 0 \\
c_3 & \dfrac{6c_3\lambda - 4\lambda^2 - c_3^2}{4\lambda} & \dfrac{c_3 u_1}{4\lambda} & \lambda & 0 \\
1 & \dfrac{12u_2\lambda^2 + 6u_3\lambda - u_3}{12c_3\lambda} & \dfrac{6\lambda u_2 + u_3}{12\lambda u_1} & \dfrac{6\lambda^2 - 6\lambda + 1}{3c_3u_1} & \lambda \\
\hline
\end{array} \tag{9}$$

with

$$u_1 = c_3 - 2\lambda\,, \qquad u_2 = 1 - c_3\,, \qquad u_3 = 3c_3 - 2\,.$$

We choose $\lambda \approx 0.43586652$ to get $R(\infty) = 0$. For this value of $\lambda$ and any $c_3$, Theorem 5.1 in [9] states that $\delta y_h(x_0) = \vartheta(h^4)$, and $\delta z_h(x_0) = \vartheta(h^2)$, and thus Theorem 5.2 in [9] ensures order of convergence 3 for the differential component $y$ and order 2 for the algebraic component $z$.

We have considered the problem

$$\begin{array}{rcl} y_1' &=& y_1\, y_2^2\, z^2 \\ y_2' &=& y_1^2\, y_2^2 - 3y_2^2\, z^2 \qquad\qquad t \in [1,2] \\ 0 &=& y_1^2\, y_2 - 1 \end{array}$$

and we have tested this method with two values of $c_3$, namely $c_3 = 0.75$ and $c_3 = 1.153799789$. In Table I we show the observed orders for the differential variable $y$ and the algebraic one $z$.

| | $c_3 = 0.75$ | | $c_3 = 1.153799789$ | |
|---|---|---|---|---|
| | $y$ | $z$ | $y$ | $z$ |
| $h = 0.01$ | 2.96 | 2.19 | 2.97 | 2.96 |
| $h = 0.005$ | 2.98 | 2.11 | 2.98 | 2.98 |
| $h = 0.0025$ | 3.00 | 2.06 | 3.00 | 2.99 |

Table I. Observed orders

We see that for $c_3 = 0.75$ the order is as stated in Theorem 5.2 in [9], but for $c_3 = 1.153799789$ the order for the algebraic variable $z$ is higher. □

In this paper we explore the order of consistency for Runge-Kutta methods of the form (2) and give sharp order conditions in terms of the rooted trees. We also ensure certain order of consistency with the help of some special simplifying assumptions. The rest of the paper is organized as follows. In Section 2, we review some results on Runge-Kutta methods with regular matrix coefficient $\mathcal{A}$. These results are extended in Section 3 for Runge-Kutta methods with matrix coefficient of the form (2).

## §2. Review on RK methods with regular matrix coefficient $\mathcal{A}$

In this section we briefly review some results for RK methods $(\mathcal{A}, b)$ with $\mathcal{A}$ regular [6, VII.4 and VII.5]. In [6, VII.4] estimations of the local error are given in terms of the simplifying conditions $B(p)$ and $C(q)$ whereas in [6, VII.5] they are given in terms of rooted trees. It is well known that the order of obtained from the simplifying assumptions is not optimal, and usually the observed order of convergence is greater than the predicted one. This is not the case when rooted trees theory is used. The drawback of the rooted trees theory is its complexity when high orders are desired.

For index-2 DAEs the $DA2$-series theory is used. We denote by $DAT2 = DAT2_y \cup DAT2_z$ the set of rooted trees with two type of vertex, meagre and fat. The expression $[t_1, \dots, t_\mu, u_1, \dots, u_\nu]_y$ denotes the tree which is obtained by joining the roots of $t_1, \dots, t_\mu, u_1, \dots, u_\nu$ to a meagre vertex whereas $[t_1, \dots, t_\mu]_z$ denotes the tree obtained by joining the roots of $t_1, \dots, t_\mu$ to a fat vertex, provided that $t_1 \neq [u]_y$ if $\mu = 1$. The letter $\tau$ denotes the tree consisting of a single meagre vertex. The order of a tree $t \in DAT2$, denoted by $\rho(t)$, is the difference between the number of meagre and fat vertices of that tree. Finally, given two vectors $u, v \in \mathbb{R}^s$, $u{\bullet}v$ denotes the product component by component. For further details, see [5], [6].

The internal stages of the Runge-Kutta method [5, Theorem 5.7] can be written as as $DA2$-series in terms of the coefficients $\Phi_y(t)$ and $\Phi_z(u)$ which are defined by

$$\Phi_y(\emptyset_y) = e\,, \quad \Phi_z(\emptyset_z) = e\,, \quad \Phi_y(\tau) = c\,,$$

$$\Phi_y(t) = \rho(t)\,\mathcal{A}\,[\Phi_y(t_1){\bullet}\dots{\bullet}\Phi_y(t_\mu){\bullet}\Phi_z(u_1){\bullet}\dots{\bullet}\Phi_z(u_\nu)]\,, \quad \text{if } t = [t_1,\dots,t_\mu,u_1,\dots,u_\nu]_y \in DAT2_y\,,$$

$$\Phi_z(u) = \frac{1}{\rho(u)+1}\,\mathcal{A}^{-1}[\Phi_y(t_1){\bullet}\dots{\bullet}\Phi_y(t_\mu)]\,, \quad \text{if } u = [t_1,\dots,t_\mu]_z \in DAT2_z\,.$$

We remark that these coefficients $\Phi_y(t)$ and $\Phi_z(u)$ are related to the coefficients $\phi_y(t)$ and $\phi_z(u)$ defined in [5] by $\Phi_y(t) = \gamma(t)\,\mathcal{A}\,\phi_y(t)$, $\Phi_z(u) = \gamma(u)\,\phi_z(u)$.

We introduce the notation $\delta y_h(x) = y_1 - y(x+h)$ and $\delta z_h(x) = z_1 - z(x+h)$ for the local error for the variable $y$ and $z$ respectively, and $P(x_0) = I - (f_z(g_y f_z)^{-1} g_y)(y_0, z_0)$. We have the following result.

**Theorem 1.** *[6, VII, Theorem 5.8]*

1. *The local error satisfies $\delta y_h(x_0) = \vartheta(h^p)$, $P(x_0)\delta y_h(x_0) = \vartheta(h^{p+1})$ if*

$$b^t \mathcal{A}^{-1} \Phi_y(t) = 1 \qquad \forall t \in DAT2_y\,,\ 1 \le \rho(t) \le p-1 \tag{10}$$

   *and those of order $\rho(t) = p$ which are not of the form $[u]_y$ with $u \in DAT2_z$.*

2. *The local error satisfies $\delta z_h(x_0) = \vartheta(h^r)$ if*

$$b^t \mathcal{A}^{-1} \Phi_z(u) = 1 \qquad \forall u \in DAT2_z\,,\ 1 \le \rho(u) \le r-1 \tag{11}$$

To apply the above result, the complete set of trees up to a given order must be constructed. As the number of trees increases considerably with the order, handling the set of trees for high orders is quite cumbersome. That is why in some cases it is preferred to get the order of consistency with the help of simplifying conditions, in spite of the fact that the order bounds obtained are not sharp.

**Theorem 2.** *Consider a Runge-Kutta method with coefficients $(\mathcal{A}, b)$ with $\mathcal{A}$ regular. Then*

1. *[6, VII, Theorem 5.10] If the method is stiffly accurate, then the conditions $B(p)$, $C(q)$, $D(\eta)$ with $p \le 2q$ and $p \le q + \eta + 1$ imply that the $y$-component of the local error satisfies $\delta y_h(x_0) = \vartheta(h^{p+1})$. Moreover if $f$ is linear in $z$, then the assumption $p \le 2q$ can be relaxed to $p \le 2q + 1$.*

2. *Conditions* $B(p)$ *and* $C(q)$ *with* $p \geq q$*, imply that* $\delta z_h(x_0) = \vartheta(h^q)$.

*Proof.* Part 2. It can be checked that $C(q)$ implies that for any tree $u \in DAT_z$ with $\rho(u) \leq q-1$, we get $\Phi_z(u) = c^{\rho(u)}$. Hence, using condition $B(p)$, with $p \geq q$, we obtain (11) for $r = q$. □

Although Theorem 2 is extremely useful for methods with high stage order $q$, it gives poor results for methods with low stage order. For example, with $C(2)$ and $B(3)$ we can only ensure $\delta y_h(x_0) = \vartheta(h^4)$ and $\delta z_h(x_0) = \vartheta(h^2)$. Theorem 2 can be improved if another set of conditions is considered [3],

$$\begin{aligned}
A_1(s):&\quad b^t\mathcal{A}^{-1}c^k = 1 && k = 1,\dots,s\\
A_2(s'):&\quad b^t\mathcal{A}^{-1}e = b^t\mathcal{A}^{-2}c &&\\
&\quad b^t\mathcal{A}^{-2}c^k = k && k = 1,\dots,s'\,.
\end{aligned}$$

The following result was proved in [7].

**Theorem 3.** *[7] If the coefficients of the Runge-Kutta method satisfy* $B(p)$*,* $C(q)$*,* $D(\eta)$ *and* $A_1(s)$*, with* $q \leq p \leq \min\{2q, q+2\}$*,* $p \leq q+\eta+1$*, and* $p \leq s+1$*, then* $\delta_h y(x_0) = \vartheta(h^p)$ *and* $P(x_0)\delta y_h(x_0) = \vartheta(h^{p+1})$*. Moreover if* $f$ *is linear in* $z$*, then the assumption* $p \leq 2q$ *can be relaxed to* $p \leq 2q+1$*.*

*If the coefficients of a Runge-Kutta method satisfy* $B(q)$*,* $C(q)$ *and* $A_2(q+1)$*, then* $\delta z_h(x_0) = \vartheta(h^{q+1})$*.*

Thus conditions $B(2)$ and $C(2)$ together with $A_2(3)$ ensure $\delta z_h(x_0) = \vartheta(h^3)$.

## §3. Results for RK methods with singular coefficient matrix $\mathcal{A}$

As it has been pointed out previously, the above results require the matrix $\mathcal{A}$ to be regular. A simple way to transfer them to methods of the form (2) is to embed the method (2) into

$$\begin{array}{c|cc}
\varepsilon & \varepsilon & 0^t\\
\bar{c} & a & \bar{\mathcal{A}}\\
\hline
 & b_1 & \bar{b}^t
\end{array}
\quad = \quad
\begin{array}{c|c}
c_\varepsilon & \mathcal{A}_\varepsilon\\
\hline
 & b^t
\end{array}\,. \tag{12}$$

If $\varepsilon \neq 0$, the coefficient matrix is regular and we can apply the above results. As the internal stages $Y_{n,\varepsilon}$, $Z_{n,\varepsilon}$ for this numerical method converge to $(y_{n-1}, \bar{Y}_n)$ and $(z_{n-1}, \bar{Z}_n)$ respectively when $\varepsilon$ tends to zero, the results can be transferred to the method (2).

## 3.1. Extension of Theorem 1

In order to apply Theorem 1, we simply have to ensure that $\Phi_{y,\varepsilon}$, $\Phi_{z,\varepsilon}$ are bounded when $\varepsilon$ tends to zero. Recall that the matrix $\mathcal{A}_\varepsilon^{-1}$, which contains the term $1/\varepsilon$, is involved in the definition of these functions. In [8] it was proved that for the $\varepsilon$-method, the functions $\Phi_{y,\varepsilon}$ and $\Phi_{z,\varepsilon}$ satisfy

$$\Phi_{*,\varepsilon}(t) = \begin{pmatrix} \vartheta\left(\varepsilon^{\rho(t)}\right) \\ \psi_*(t) + \vartheta(\varepsilon) \end{pmatrix} \qquad \forall\, t \in DAT2\,, \rho(t) \geq 1\,,$$

where the functions $\psi_y(t) : DAT2_y \to \mathbb{R}^{s-1}$ and $\psi_z(u) : DAT2_z \to \mathbb{R}^{s-1}$ are defined recursively for the coefficients

$$\begin{array}{c|c} \bar{c} & \bar{\mathcal{A}} \\ \hline \end{array}$$

as

$$\begin{aligned}
&\psi_y(\emptyset_y) = \bar{\mathcal{A}}^{-1}\bar{c}\,, \quad \psi_z(\emptyset_z) = \bar{\mathcal{A}}^{-1}\bar{c}\,, \quad \psi_y(\tau) = \bar{c}\,,\\
&\psi_y(t) = \rho(t)\,\bar{\mathcal{A}}\,[\psi_y(t_1)\bullet\ldots\bullet\psi_y(t_\mu)\bullet\psi_z(u_1)\bullet\ldots\bullet\psi_z(u_\nu)]\,, \quad \text{if } t = [t_1,\ldots,t_\mu,u_1,\ldots,u_\nu]_y \in DAT2_y\,,\\
&\psi_z(u) = \frac{1}{\rho(u)+1}\,\bar{\mathcal{A}}^{-1}[\psi_y(t_1)\bullet\ldots\bullet\psi_y(t_\mu)]\,, \quad \text{if } u = [t_1,\ldots,t_\mu]_z \in DAT2_z\,.
\end{aligned}$$

In particular, for the order one trees it holds

$$\begin{aligned}
\Phi_{y,\varepsilon}(\tau) &= \begin{pmatrix} \varepsilon \\ \psi_y(\tau) \end{pmatrix}\,,\\
\Phi_{z,\varepsilon}(u_{1,1}) &= \begin{pmatrix} \frac{1}{2}\varepsilon \\ \psi(u_{1,1}) - \frac{1}{2}\varepsilon\bar{\mathcal{A}}^{-1}a \end{pmatrix}\,, \qquad \Phi_{z,\varepsilon}(u_{1,2}) = \begin{pmatrix} \varepsilon \\ \psi_z(u_{1,2}) \end{pmatrix}\,.
\end{aligned}$$

with $u_{1,1} = [\tau,\tau]_z$ and $u_{1,2} = [[\tau]_y]_z$. Next we extend Theorem 1.

**Theorem 4.** *Consider a Runge-Kutta method of the form (2). Assume that condition (5) holds. Then*

1. *The local error for the differential component $\delta y_h(x_0)$ satisfies $\delta y_h(x_0) = \vartheta(h^p)$, $P(x_0)\delta y_h(x_0) = \vartheta(h^{p+1})$ if*

$$\bar{b}^t\bar{\mathcal{A}}^{-1}\psi_y(t) = 1 \qquad \forall t \in DAT2_y\,,\ 1 \leq \rho(t) \leq p-1 \tag{13}$$

   *and those trees of order $p$ which are not of the form $[u]_y$ with $u \in DAT2_z$.*

2. *The local error $\delta z_h(x_0)$ for the algebraic component satisfies $\delta z_h(x_0) = \vartheta(h^r)$ if*

$$\bar{b}^t\bar{\mathcal{A}}^{-1}\psi_z(u) = 1 \qquad \forall u \in DAT2_z\,,\ 1 \leq \rho(u) \leq r-1\,. \tag{14}$$

*Proof.* We simply have apply Theorem 1 to the $\varepsilon$-method, and take the limit as $\varepsilon$ tends to zero. A simple computation gives

$$b^t\mathcal{A}_\varepsilon^{-1}\Phi_{*,\varepsilon}(t) = \bar{b}^t\bar{\mathcal{A}}^{-1}\psi_*(t) + \vartheta(\varepsilon^{\min\{\rho(t)-1,1\}})\,.$$

Hence, if $\rho(t) \geq 2$, when $\varepsilon$ tends to zero we obtain (13) and (14). It remains to prove conditions (13) and (14) for the order one trees. There is one tree in $DAT_y$ with $\rho(t) = 1$, namely $t = \tau$. In this case,

$$b^t \mathcal{A}_\varepsilon^{-1} \Phi_{y,\varepsilon}(\tau) = b_1 - \bar{b}^t \bar{\mathcal{A}}^{-1} a + \bar{b}^t \bar{\mathcal{A}}^{-1} \psi_y(\tau) \,.$$

Thus, using condition (5), we obtain (13). In $DAT_z$ there are two trees with $\rho(u) = 1$, namely $u_{1,1} = [\tau, \tau]_z$ and $u_{1,2} = [[\tau]_y]_z$. We compute

$$\begin{aligned} b^t \mathcal{A}_\varepsilon^{-1} \Phi_{z,\varepsilon}(u_{1,1}) &= \frac{1}{2} \left( b_1 - \bar{b}^t \bar{\mathcal{A}}^{-1} a \right) - \frac{1}{2} \, \varepsilon \, \bar{b}^t \bar{\mathcal{A}}^{-2} a + \bar{b}^t \bar{\mathcal{A}}^{-1} \psi_z(u_{1,1}) \,, \\ b^t \mathcal{A}_\varepsilon^{-1} \Phi_{z,\varepsilon}(u_{1,2}) &= b_1 - \bar{b}^t \bar{\mathcal{A}}^{-1} a + \bar{b}^t \bar{\mathcal{A}}^{-1} \psi_z(u_{1,2}) \,. \end{aligned}$$

In this way, using condition (5), when $\varepsilon$ tends to zero we obtain (14). □

In Table II we give the order conditions for the trees in $DAT_y$ with $1 \leq \rho(t) \leq 3$, and those of order 4 which are not of the form $[u]_y$, with $u \in DAT_z$. For the trees with $\rho(t) = 2, 3$, we show the conditions associated to trees of the form $[u]_y$, with $u \in DAT_z$. In Table III we give the trees in $DAT_z$ with $1 \leq \rho(t) \leq 2$.

| $\rho(t)$ | Conditions | |
|---|---|---|
| 1 | $\bar{b}^t \, \mathcal{A}^{-1} \, \bar{c} = 1$ | |
| 2 | $\bar{b}^t \, \bar{c} = \frac{1}{2}$ | |
| … | ………………………………………………… | |
| $[u]_y$ | $\bar{b}^t \bar{\mathcal{A}}^{-1} \, \bar{c}^2 = 1$ | |
| 3 | $\bar{b}^t \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \bullet \bar{c} \right) = \frac{2}{3}$ | $\bar{b}^t \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \right)^2 = \frac{4}{3}$ |
| | $\bar{b}^t \, \bar{c}^2 = \frac{1}{3}$ | $\bar{b}^t \, \bar{\mathcal{A}} \, \bar{c} = \frac{1}{6}$ |
| … | ………………………………………………… | |
| $[u]_y$ | $\bar{b}^t \, \bar{\mathcal{A}}^{-1} \, \bar{c}^3 = 1$ | $\bar{b}^t \, \bar{\mathcal{A}}^{-1} \left( \bar{c} \bullet \mathcal{A} \, \bar{c} \right) = \frac{1}{2}$ |
| 4 | $\bar{b}^t \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \bullet \bar{c}^2 \right) = \frac{1}{2}$ | $\bar{b}^t \left( \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \right)^2 \bullet \bar{c} \right) = 1$ |
| | $\bar{b}^t \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \right)^3 = 2$ | $\bar{b}^t \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^3 \bullet \bar{c} \right) = \frac{3}{4}$ |
| | $\bar{b}^t \left( \bar{\mathcal{A}}^{-1} \left( \bar{c} \bullet \bar{\mathcal{A}} \, \bar{c} \right) \bullet \bar{c} \right) = \frac{3}{8}$ | $\bar{b}^t \left( \bar{\mathcal{A}}^{-1} \left( \bar{c} \bullet \bar{\mathcal{A}} \, \bar{c} \right) \bullet \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \right) = \frac{3}{4}$ |
| | $\bar{b}^t \left( \bar{\mathcal{A}} \, \bar{c} \bullet \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \right) = \frac{1}{4}$ | $\bar{b}^t \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^3 \bullet \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \right) = \frac{3}{2}$ |
| | $\bar{b}^t \, \bar{\mathcal{A}} \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \bullet \bar{c} \right) = \frac{1}{6}$ | $\bar{b}^t \, \bar{\mathcal{A}} \left( \bar{\mathcal{A}}^{-1} \, \bar{c}^2 \right)^2 = \frac{1}{3}$ |
| | $\bar{b}^t \, \bar{c}^3 = \frac{1}{4}$ | $\bar{b}^t \left( \bar{c} \bullet \bar{\mathcal{A}} \, \bar{c} \right) = \frac{1}{8}$ |
| | $\bar{b}^t \, \bar{\mathcal{A}}^2 \, \bar{c} = \frac{1}{24}$ | $\bar{b}^t \, \bar{\mathcal{A}} \, \bar{c}^2 = \frac{1}{12}$ |

Table II. Order conditions for $t \in DAT_y$

| $\rho(u)$ | Conditions | |
|---|---|---|
| 1 | $\bar{b}^t\, \mathcal{A}^{-1}\, \bar{c} = 1$ | $\bar{b}^t \bar{\mathcal{A}}^{-2}\, \bar{c}^2 = 2$ |
| 2 | $\bar{b}^t\, \bar{c} = \frac{1}{2}$ | $\bar{b}^t \bar{\mathcal{A}}^{-1}\, \bar{c}^2 = 1$ |
| | $\bar{b}^t\, \bar{\mathcal{A}}^{-1} \left( \bar{c} \bullet \bar{\mathcal{A}}^{-1}\, \bar{c}^2 \right) = 2$ | $\bar{b}^t\, \bar{\mathcal{A}}^{-1} \left( \bar{\mathcal{A}}^{-1}\, \bar{c}^2 \right)^2 = 4$ |
| | $\bar{b}^t\, \bar{\mathcal{A}}^{-2}\, \bar{c}^3 = 3$ | $\bar{b}^t\, \bar{\mathcal{A}}^{-2} \left( \bar{c} \bullet \bar{\mathcal{A}}\, \bar{c} \right) = \frac{3}{2}$ |

Table III. Order conditions for $u \in DAT_z$

**Example 2.** We consider again the method (9). For any $c_3$, the conditions for $\rho(t) \leq 3$ in Table I and $\rho(u) = 1$ in Table II are satisfied. It can be checked that for $c_3 = 1.153799789$ all the conditions in Table II for $\rho(u) = 2$ are also satisfied. However, for $c_3 = 0.75$, conditions $\bar{b}^t\, \bar{\mathcal{A}}^{-2}\, \bar{c}^3 = 3$ and $\bar{b}^t\, \bar{\mathcal{A}}^{-2} \left( \bar{c} \bullet \bar{\mathcal{A}}\, \bar{c} \right) = \frac{3}{2}$ in Table II are not fulfilled. $\square$

## 3.2. Extension of Theorem 2

Observe that in general, the simplifying conditions are not transferred from the original method (2) to the $\varepsilon$-method. For example, the $\varepsilon$-method only satisfies $C(1)$ with independence of the $C(q)$ condition satisfied by (2). This fact is not a drawback because as we will take the limit when $\varepsilon$ tends to zero, it is enough to consider the simplifying assumptions in the limit case. In [8] the simplifying assumptions for the $\varepsilon$-method were defined as

$$\begin{aligned}
B_\varepsilon(p):\quad & \lim_{\varepsilon \to 0} \left( b^t c_\varepsilon^{k-1} - \frac{1}{k} \right) = 0\,, && k = 1, \dots, p \\
C_\varepsilon(q):\quad & \lim_{\varepsilon \to 0} \left( \mathcal{A}_\varepsilon c_\varepsilon^{k-1} - \frac{c_\varepsilon^k}{k} \right) = 0\,, && k = 1, \dots, q \\
D_\varepsilon(r):\quad & \lim_{\varepsilon \to 0} \left( (b \bullet c_\varepsilon^{k-1})^t \mathcal{A}_\varepsilon - \frac{1}{k} \left[ b^t - (b \bullet c_\varepsilon^k)^t \right] \right) = 0\,, && k = 1, \dots, r\,.
\end{aligned}$$

It can be proved [8, Proposition 6] that the method (2) satisfies $B(p)$, $C(q)$, $D(r)$ if and only if the $\varepsilon$-method satisfies $B_\varepsilon(p)$, $C_\varepsilon(q)$, $D_\varepsilon(r)$ respectively. Thus we can use for the $\varepsilon$-method the same simplifying conditions as for the method (2).

Applying Theorem 2 to the $\varepsilon$-method and taking the limit when $\varepsilon$ tends to zero, we obtain the following result.

**Theorem 5.** *Consider a Runge-Kutta method of the form (2).*

1. *If $b_i = a_{si}$, $i = 1, \dots, s$, then the conditions $B(p)$, $C(q)$, $D(\eta)$ with $p \leq 2q$ and $p \leq q + \eta + 1$ imply that the $y$-component of the local error satisfies $\delta y_h(x_0) = \vartheta(h^{p+1})$. Moreover if $f$ is linear in $z$, then the assumption $p \leq 2q$ can be relaxed to $p \leq 2q + 1$.*

2. *Conditions $B(p)$ and $C(q)$ with $p \geq q$, imply that $\delta z_h(x_0) = \vartheta(h^q)$.*

Observe that this is precisely Theorem 5.1 in [9].

## 3.3. Extension of Theorem 3

Conditions $A_1(s)$ and $A_2(s')$ make no sense if the coefficient matrix $\mathcal{A}$ is singular, but they can be imposed to the $\varepsilon$-method and take the limit when $\varepsilon$ tends to zero. We give the following definition.

**Definition 1.** We will say that the RK method (2) satisfies the condition $\bar{A}_1(s)$ if $s$ is the greatest integer such that

$$\bar{A}_1(s): \quad \bar{b}^t \bar{\mathcal{A}}^{-1} \bar{c}^k = 1 \qquad k = 1, \ldots, s$$

holds. We will say that the RK method (2) satisfies the condition $\bar{A}_2(s')$ if $s'$ is the greatest integer such that

$$\begin{aligned} \bar{A}_2(s'): \quad & \bar{b}^t \bar{\mathcal{A}}^{-1} \bar{e} = -\bar{b}^t \bar{\mathcal{A}}^{-2} a + \bar{b}^t \bar{\mathcal{A}}^{-2} \bar{c} && k = 1 \\ & \bar{b}^t \bar{\mathcal{A}}^{-2} \bar{c}^k = k && k = 2, \ldots, s' \end{aligned}$$

holds.

This definition is justified by the following result whose proof is straightforward.

**Proposition 6.** *If $A_{1,\varepsilon}(s)$ and $A_{2,\varepsilon}(s')$ denote respectively the conditions*

$$\begin{aligned} A_{1,\varepsilon}(s): \quad & \lim_{\varepsilon \to 0} \left( b^t \mathcal{A}_\varepsilon^{-1} c_\varepsilon^k - 1 \right) = 0\,, && k = 1, \ldots, s \\ A_{2,\varepsilon}(s'): \quad & \lim_{\varepsilon \to 0} \left( b^t \mathcal{A}_\varepsilon^{-1} e - b^t \mathcal{A}_\varepsilon^{-2} c_\varepsilon \right) = 0 && k = 1 \\ & \lim_{\varepsilon \to 0} \left( b^t \mathcal{A}_\varepsilon^{-2} c_\varepsilon^k - k \right) = 0\,, && k = 2, \ldots, s' \end{aligned}$$

*for the $\varepsilon$-method (12), then*

1. *The method (2) satisfies $\bar{A}_1(s)$ if and only if the $\varepsilon$-method satisfies $A_{1,\varepsilon}(s)$.*
2. *The method (2) satisfies $\bar{A}_2(s')$ if and only if the $\varepsilon$-method satisfies $A_{2,\varepsilon}(s')$.*

Applying now Theorem 3 to the $\varepsilon$-method and taking the limit when $\varepsilon$ tends to zero, we obtain the following result.

**Theorem 7.** *Consider a Runge-Kutta method of the form (2). If the conditions $B(p)$, $C(q)$, $D(\eta)$ and $\bar{A}_1(s)$ hold, with $q \leq p \leq \min\{2q, q+2\}$, $p \leq q+\eta+1$ and $p \leq s+1$. Then $\delta y_h(x_0) = \vartheta(h^p)$ and $P(x_0)\delta y_h(x_0) = \vartheta(h^{p+1})$. Moreover if $f$ is linear in $z$, then the assumption $p \leq 2q$ can be relaxed to $p \leq 2q+1$.*

*If the coefficients of a Runge-Kutta method of the form (2) satisfy $B(q)$, $C(q)$ and $\bar{A}_2(q+1)$, then $\delta z_h(x_0) = \vartheta(h^{q+1})$.*

**Example 3.** We consider again the method (9). For any $c_3$, the method satisfies $B(3)$, $C(2)$, $\bar{A}_1(\infty)$ and $\bar{A}_2(2)$. Therefore for any $c_3$ we get $\delta y_h(x_0) = \vartheta(h^3)$, $P(x_0)\delta y_h(x_0) = \vartheta(h^4)$ and $\delta z_h(x_0) = \vartheta(h^2)$. Condition $\bar{A}_2(3)$ gives us the value $c_3 = 1.153799789$, and hence for this value $\delta z_h(x_0) = \vartheta(h^3)$. □

## Acknowledgements

Research supported by the Ministerio the Ciencia y Tecnología, Project BFM2001-2188.

## References

[1] ASCHER, U., AND PETZOLD, L. Projected implicit Runge-Kutta methods for differential-algebraic equations. *SIAM J. Numer. Anal. 28* (1991), 1097–1120.

[2] BIJL, H., CARPENTER, M, VATSA, V., AND KENNEDY, C. Implicit time integration schemes for the unsteady compressible Navier-Stokes equations: laminar flow, *J. Comput. Phys. 179* (2002), 313–329.

[3] BRENAN, K., CAMPBELL, S., AND PETZOLD, L. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. North-Holland, New York, 1989.

[4] CAMERON, F. A class of low order DIRK methods for a class of DAEs *Applied Numerical Mathematics* 31 (1999), 1–16.

[5] HAIRER, E., LUBICH, C., AND ROCHE, M. *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*. Lecture Notes in Mathematics 1409, Springer, Berlin, 1989.

[6] HAIRER, E., AND WANNER, G. *Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems*. Springer, Berlin, 1996.

[7] HIGUERAS, I. On simplifying assumptions of Runge-Kutta methods for index 2 differential algebraic problems. *Computing 54* (1995), 185–190.

[8] HIGUERAS, I., AND ROLDÁN, T. Starting algorithms for a class of RK methods for index-2 DAEs, *Preprint Departamento de Matemática e Informática, Sección 1, n. 4 (2001). Accepted for publication in Computers and Mathematics with Applications*.

[9] JAY, L. Convergence of a class of runge-kutta methods for differential-algebraic systems of index 2, *BIT 33* (1993), 137–150.

[10] KENNEDY, C., AND CARPENTER, M. Additive Runge-Kutta schemes for convection-diffusion-reaction equations, *Appl. Numer. Math. 44* (2003), 139–181.

[11] WILLIAMS, R., BURRAGE, K., CAMERON, I. , AND M. KERR A four-stage index 2 diagonally implicit Runge-Kutta method, *Appl. Numer. Math. 40* (2002), 415–432.

Inmaculada Higueras, Teo Roldán
Departamento de Matemática e Informática,
Universidad Pública the Navarra,
31006 Pamplona (SPAIN)
`higueras@unavarra.es` and `teo@unavarra.es`