Journal of Computational and Applied Mathematics 261 (2014) 72–84

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

# Construction of algebraically stable DIMSIMs

Giuseppe Izzo [a,*], Zdzislaw Jackiewicz [b,c]

[a] Dipartimento di Matematica e Applicazioni "R.Caccioppoli", Università di Napoli "Federico II", 80126 Napoli, Italy
[b] Department of Mathematics, Arizona State University, Tempe, AZ 85287, United States
[c] AGH University of Science and Technology, Kraków, Poland

ARTICLE INFO

*Article history:*
Received 12 May 2011
Received in revised form 17 October 2013

*Keywords:*
General linear methods
Order conditions
Algebraic stability

ABSTRACT

The class of general linear methods for ordinary differential equations combines the advantages of linear multistep methods (high efficiency) and Runge–Kutta methods (good stability properties such as $A$-, $L$-, or algebraic stability), while at the same time avoiding the disadvantages of these methods (poor stability of linear multistep methods, high cost for Runge–Kutta methods). In this paper we describe the construction of algebraically stable general linear methods based on the criteria proposed recently by Hewitt and Hill. We also introduce the new concept of $\epsilon$-algebraic stability and investigate its consequences. Examples of $\epsilon$-algebraically stable methods are given up to order $p = 4$.



## 1. Introduction

Consider the solution of an initial value problem for a system of differential equations

$$\begin{cases} y'(t) = f\big(y(t)\big), & t \in [t_0, T], \\ y(t_0) = y_0, \end{cases} \tag{1.1}$$

$f : \mathbb{R}^m \to \mathbb{R}^m$, by the general linear method (GLM) of the form

$$\begin{cases} Y_i^{[n]} = h\displaystyle\sum_{j=1}^{s} a_{ij} f(Y_j^{[n]}) + \sum_{j=1}^{r} u_{ij} y_j^{[n-1]}, & i = 1, 2, \ldots, s, \\ y_i^{[n]} = h\displaystyle\sum_{j=1}^{s} b_{ij} f(Y_j^{[n]}) + \sum_{j=1}^{r} v_{ij} y_j^{[n-1]}, & i = 1, 2, \ldots, r, \end{cases} \tag{1.2}$$

$i = 1, 2, \ldots, N$. Here, $N$ is a positive integer, $h = (T - t_0)/N$ is a fixed stepsize, $t_n = t_0 + nh$, $Y_i^{[n]}$ is an approximation of stage order $q$ to $y(t_{n-1} + c_i h)$, that is

$$Y_i^{[n]} = y(t_{n-1} + c_i h) + O(h^{q+1}), \quad i = 1, 2, \ldots, s, \tag{1.3}$$

and $y_i^{[n]}$ is an approximation of order $p$ to a linear combination of scaled derivatives of the solution $y$ to (1.1), i.e., $y_i^{[n]}$ satisfy the relations

$$y_i^{[n]} = q_{i,0} y(t_n) + q_{i,1} h y'(t_n) + \cdots + q_{i,p} h^p y^{(p)}(t_n) + O(h^{p+1}), \quad i = 1, 2, \ldots, r, \tag{1.4}$$

* Corresponding author. Tel.: +39 3294167433.
*E-mail addresses:* giuseppe.izzo@unina.it (G. Izzo), jackiewicz@asu.edu (Z. Jackiewicz).

0377-0427/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cam.2013.10.037

with some scalars $q_{i,j}$. Such methods are characterized by the abscissa vector $\mathbf{c} = [c_1, \ldots, c_s]^T$, four coefficient matrices $\mathbf{A} \in \mathbb{R}^{s\times s}$, $\mathbf{U} \in \mathbb{R}^{s\times r}$, $\mathbf{B} \in \mathbb{R}^{r\times s}$, $\mathbf{V} \in \mathbb{R}^{r\times r}$, the vectors $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_p$ given by

$$\mathbf{q}_0 = \begin{bmatrix} q_{1,0} \\ \vdots \\ q_{r,0} \end{bmatrix}, \quad \mathbf{q}_1 = \begin{bmatrix} q_{1,1} \\ \vdots \\ q_{r,1} \end{bmatrix}, \ldots, \mathbf{q}_p = \begin{bmatrix} q_{1,p} \\ \vdots \\ q_{r,p} \end{bmatrix},$$

and four integers: $p$—the order, $q$—the stage order, $r$—the number of external approximations, and $s$—the number of stages or internal approximations. We assume throughout the paper that $p = q = r = s$ and that the coefficient matrices $\mathbf{A}$ and $\mathbf{V}$ have the form

$$\mathbf{A} = \begin{bmatrix} \lambda & & & & \\ a_{2,1} & \lambda & & & \\ a_{3,1} & a_{3,2} & \lambda & & \\ \vdots & \vdots & \ddots & \ddots & \\ a_{s,1} & a_{s,2} & \cdots & a_{s,s-1} & \lambda \end{bmatrix}, \qquad \mathbf{V} = \begin{bmatrix} 1 & v_{1,2} & v_{1,3} & \cdots & v_{1,r} \\ 0 & v_{2,2} & v_{2,3} & \cdots & v_{2,r} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & v_{r-1,r-1} & v_{r-1,r} \\ 0 & 0 & 0 & \cdots & v_{r,r} \end{bmatrix},$$

$\lambda > 0$, $-1 \le v_{j,j} < 1$, $j = 2, 3, \ldots, r$. This representation of $\mathbf{V}$ implies that the GLM (1.2) is zero-stable, i.e., that the matrix $\mathbf{V}$ is power bounded (cf. [1]). The above representation for $\mathbf{A}$ implies that the methods we will consider are of type known as diagonally-implicit multi-stage integration methods (DIMSIMs). These methods were first introduced in [2] and further investigated in [3–6]. It is the purpose of this paper to construct new classes of algebraically stable DIMSIMs using the recent criteria proposed by Hewitt and Hill [7]. These criteria are described in Section 4.

We introduce next the notions of method equivalence and reducibility following the presentation in [8] (compare also [1]). We say that the GLMs defined by the coefficient matrices

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] \quad \text{and} \quad \left[\begin{array}{c|c} \widetilde{\mathbf{A}} & \widetilde{\mathbf{U}} \\ \hline \widetilde{\mathbf{B}} & \widetilde{\mathbf{V}} \end{array}\right]$$

are equivalent if there exists a permutation matrix $\mathbf{P} \in \mathbb{R}^{s\times s}$ and a nonsingular matrix $\mathbf{Q} \in \mathbb{R}^{r\times r}$ such that

$$\left[\begin{array}{c|c} \widetilde{\mathbf{A}} & \widetilde{\mathbf{U}} \\ \hline \widetilde{\mathbf{B}} & \widetilde{\mathbf{V}} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{P}^T\mathbf{A}\mathbf{P} & \mathbf{P}^T\mathbf{U}\mathbf{Q} \\ \hline \mathbf{Q}^{-1}\mathbf{B}\mathbf{P} & \mathbf{Q}^{-1}\mathbf{V}\mathbf{Q} \end{array}\right].$$

The GLM (1.2) is said to be reducible if $s = s_1 + s_2$ and $r = r_1 + r_2 + r_3$ with $s_2 + r_2 + r_3 > 0$, and there exists an equivalent GLM with a sparsity pattern of the form

$$\begin{array}{c} \\ s_1 \\ s_2 \\ r_1 \\ r_2 \\ r_3 \end{array} \begin{array}{c} \begin{array}{ccccc} s_1 & s_2 & r_1 & r_2 & r_3 \end{array} \\ \left[\begin{array}{cc|ccc} \mathbf{A}_{11} & \mathbf{0} & \mathbf{U}_{11} & \mathbf{0} & \mathbf{U}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{U}_{21} & \mathbf{U}_{22} & \mathbf{0} \\ \hline \mathbf{B}_{11} & \mathbf{0} & \mathbf{V}_{11} & \mathbf{0} & \mathbf{V}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{V}_{21} & \mathbf{V}_{22} & \mathbf{V}_{23} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{V}_{33} \end{array}\right] \end{array}.$$

In this case the GLM (1.2) can be reduced to a GLM with coefficient matrices

$$\left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{U}_{11} \\ \hline \mathbf{B}_{11} & \mathbf{V}_{11} \end{array}\right]$$

with $s_1$ internal stages and $r_1$ external stages. The method is said to be irreducible if it is not reducible.

The organization of this paper is as follows. In Section 2 we review the stage order and order conditions for GLMs (1.2) following the presentation in [2,1], and derive representation formulas for the coefficient matrices $\mathbf{U}$, $\mathbf{B}$, and the vector $\mathbf{q}_p$. In Section 3 we survey algebraic stability of GLMs, introduce a new concept of $\epsilon$-algebraic stability, and investigate its consequences. In Section 4 we review some technical tools to investigate algebraic and $\epsilon$-algebraic stability. In Section 5 we use the criteria proposed by Hewitt and Hill [7] to search for algebraically and $\epsilon$-algebraically stable DIMSIMs with $p = q = r = s = 2$. In Sections 6 and 7 this search is extended to higher order methods with $p = q = r = s = 3$ and $p = q = r = s = 4$, respectively. In Section 8 we present the results of some numerical experiments which demonstrate that $\epsilon$-algebraically stable methods constructed in Sections 5–7 do not suffer from order reduction. Finally, in Section 9 some concluding remarks are given and plans for future work are outlined.

## 2. Stage order and order conditions

In this section we review the conditions on the abscissa vector $\mathbf{c}$, coefficient matrices $\mathbf{A}, \mathbf{U}, \mathbf{B}, \mathbf{V}$, and the vectors $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_p$ which guarantee that the method (1.2) has stage order $q = p$ and order $p$. To formulate these conditions, taking into account (1.3) and (1.4), we assume that the components $y_i^{[n-1]}$ of the input vector $y^{[n-1]}$ for the step from $t_{n-1}$ to $t_n$ satisfy the relations

$$y_i^{[n-1]} = \sum_{k=0}^{p} q_{i,k} h^k y^{(k)}(t_{n-1}) + O(h^{p+1}), \quad i = 1, 2, \ldots, r. \tag{2.1}$$

Then the method (1.2) has stage order $q = p$ and order $p$ if

$$Y_i^{[n]} = y(t_{n-1} + c_i h) + O(h^{p+1}), \quad i = 1, 2, \ldots, s, \tag{2.2}$$

and

$$y_i^{[n]} = \sum_{k=0}^{p} q_{i,k} h^k y^{(k)}(t_n) + O(h^{p+1}), \quad i = 1, 2, \ldots, r, \tag{2.3}$$

for scalars $c_i$ and for the same scalars $q_{i,k}$. Define the vector $\mathbf{w}(z)$ by

$$\mathbf{w}(z) = \sum_{k=0}^{p} \mathbf{q}_k z^k.$$

We have the following theorem expressing the order and stage order conditions (2.2) and (2.3) in terms of the coefficients of the GLM.

**Theorem 2.1** (*Compare [2,9]*). *The method (1.2) has stage order $q = p$ and order $p$, i.e., the relation (2.1) implies (2.2) and (2.3), if and only if*

$$e^{\mathbf{c}z} = z\mathbf{A}e^{\mathbf{c}z} + \mathbf{U}\mathbf{w}(z) + O(z^{p+1}), \tag{2.4}$$

*and*

$$e^{z}\mathbf{w}(z) = z\mathbf{B}e^{\mathbf{c}z} + \mathbf{V}\mathbf{w}(z) + O(z^{p+1}). \tag{2.5}$$

*Here, $e^{\mathbf{c}z} = [e^{c_1 z}, e^{c_2 z}, \ldots, e^{c_s z}]^T$.*

This theorem is very convenient in a symbolic manipulation environment. Comparing the free terms in (2.4) and (2.5) leads to the preconsistency conditions

$$\mathbf{U}\mathbf{q}_0 = \mathbf{e}, \qquad \mathbf{V}\mathbf{q}_0 = \mathbf{q_0},$$

where $\mathbf{e} = [1, 1, \ldots, 1]^T \in \mathbb{R}^s$. The vector $\mathbf{q}_0$ is called the preconsistency vector. For DIMSIMs this vector may be taken as $\mathbf{q}_0 = [1, 0, \ldots, 0]^T$. Comparing terms of the first order in (2.4) and (2.5) leads to stage-consistency and consistency conditions

$$\mathbf{A}\mathbf{e} + \mathbf{U}\mathbf{q}_1 = \mathbf{c}, \qquad \mathbf{B}\mathbf{e} + \mathbf{V}\mathbf{q}_1 = \mathbf{q}_0 + \mathbf{q}_1.$$

The vector $\mathbf{q}_1$ is called the consistency vector.

We derive next the representation formulas for the coefficient matrix $\mathbf{U}$ in terms of $\mathbf{c}$, $\mathbf{A}$ and $\mathbf{q}_0, \ldots, \mathbf{q}_{p-1}$, for the vector $\mathbf{q}_p$ in terms of $\mathbf{c}$, $\mathbf{A}$ and $\mathbf{U}$, and for the coefficient matrix $\mathbf{B}$ in terms of $\mathbf{c}$, $\mathbf{V}$ and $\mathbf{q}_1, \ldots, \mathbf{q}_p$. Put

$$\mathbf{C} = \begin{bmatrix} \mathbf{e} & \mathbf{c} & \dfrac{\mathbf{c}^2}{2!} & \cdots & \dfrac{\mathbf{c}^{p-1}}{(p-1)!} \end{bmatrix} \in \mathbb{R}^{s\times s},$$

$$\mathbf{K} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{s\times s}, \qquad \mathbf{F} = \begin{bmatrix} 1 & \dfrac{1}{2!} & \cdots & \dfrac{1}{(p-1)!} & \dfrac{1}{p!} \\ 1 & 1 & \cdots & \dfrac{1}{(p-2)!} & \dfrac{1}{(p-1)!} \\ 0 & 1 & \cdots & \dfrac{1}{(p-3)!} & \dfrac{1}{(p-2)!} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{(s+1)\times s},$$

$$\mathbf{W} = \begin{bmatrix} \mathbf{q}_0 & \mathbf{q}_1 & \cdots & \mathbf{q}_p \end{bmatrix} \in \mathbb{R}^{s\times(s+1)},$$

and partition the matrix $\mathbf{W}$ as

$$\mathbf{W} = \left[\begin{array}{c|c} \mathbf{q}_0 & \mathbf{W}_0 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{W}_p & \mathbf{q}_p \end{array}\right], \quad \mathbf{W}_0, \mathbf{W}_p \in \mathbb{R}^{s\times s}.$$

We have the following theorem.

**Theorem 2.2.** *Assume that $p = q = r = s$. Moreover, assume that $c_i \neq c_j$ for $i \neq j$ and that the matrix $\mathbf{W}_p$ is nonsingular. Assume also the definitions for $\mathbf{C}$, $\mathbf{K}$, $\mathbf{F}$, $\mathbf{W}$, $\mathbf{W}_0$, and $\mathbf{W}_p$ as given above. Then if the method (1.2) has stage order $q = p$ and order $p$ we have*

$$\mathbf{U} = \big(\mathbf{C} - \mathbf{ACK}\big)\mathbf{W}_p^{-1}, \tag{2.6}$$

*and if* $\mathbf{U}$ *is nonsingular then*

$$\mathbf{q}_p = \mathbf{U}^{-1}\left(\frac{\mathbf{c}^p}{p!} - \mathbf{A}\frac{\mathbf{c}^{p-1}}{(p-1)!}\right). \tag{2.7}$$

*Moreover, the matrix* $\mathbf{B}$ *is given by*

$$\mathbf{B} = \big(\mathbf{WF} - \mathbf{VW}_0\big)\mathbf{C}^{-1}. \tag{2.8}$$

**Proof.** Expanding (2.4) into power series with respect to $z$ and comparing the powers of $z$ from zero to $p$ leads to the system of equations

$$\begin{cases} \mathbf{Uq}_0 = \mathbf{e}, \\ \mathbf{Uq}_1 + \mathbf{Ae} = \mathbf{c}, \\ \mathbf{Uq}_2 + \mathbf{Ac} = \dfrac{\mathbf{c}^2}{2!}, \\ \vdots \\ \mathbf{Uq}_{p-1} + \mathbf{A}\dfrac{\mathbf{c}^{p-2}}{(p-2)!} = \dfrac{\mathbf{c}^{p-1}}{(p-1)!}, \\ \mathbf{Uq}_p + \mathbf{A}\dfrac{\mathbf{c}^{p-1}}{(p-1)!} = \dfrac{\mathbf{c}^p}{p!}. \end{cases} \tag{2.9}$$

The first $p$ equations of (2.9) can be written in a vector form as

$$\mathbf{U}\begin{bmatrix}\mathbf{q}_0 & \mathbf{q}_1 & \cdots & \mathbf{q}_{p-1}\end{bmatrix} + \mathbf{A}\begin{bmatrix}\mathbf{0} & \mathbf{e} & \mathbf{c} & \cdots & \dfrac{\mathbf{c}^{p-2}}{(p-2)!}\end{bmatrix} = \begin{bmatrix}\mathbf{e} & \mathbf{c} & \dfrac{\mathbf{c}^2}{2!} & \cdots & \dfrac{\mathbf{c}^{p-1}}{(p-1)!}\end{bmatrix},$$

which is equivalent to (2.6) and the last equation of (2.9) is equivalent to (2.7). Expanding next (2.5) into power series with respect to $z$ and comparing the powers of $z$ from zero to $p$ we obtain the relation $\mathbf{q}_0 = \mathbf{Vq}_0$ and the system of equations

$$\begin{cases} \mathbf{q}_0 + \mathbf{q}_1 = \mathbf{Be} + \mathbf{Vq}_1, \\ \dfrac{1}{2!}\mathbf{q}_0 + \mathbf{q}_1 + \mathbf{q}_2 = \mathbf{Bc} + \mathbf{Vq}_2, \\ \vdots \\ \dfrac{1}{(p-1)!}\mathbf{q}_0 + \dfrac{1}{(p-2)!}\mathbf{q}_1 + \cdots + \mathbf{q}_{p-1} = \mathbf{B}\dfrac{\mathbf{c}^{p-2}}{(p-2)!} + \mathbf{Vq}_{p-1}, \\ \dfrac{1}{p!}\mathbf{q}_0 + \dfrac{1}{(p-1)!}\mathbf{q}_1 + \cdots + \mathbf{q}_p = \mathbf{B}\dfrac{\mathbf{c}^{p-1}}{(p-1)!} + \mathbf{Vq}_p. \end{cases}$$

This system can be written in vector form as

$$\mathbf{WF} = \mathbf{BC} + \mathbf{VW}_0,$$

which is equivalent to (2.8). This completes the proof. □

## 3. Algebraic stability

The notion of algebraic stability of GLMs (1.2) was introduced by Burrage and Butcher [10] and further investigated in [11–13,8,14].

**Definition 3.1.** The GLM (1.2) is said to be algebraically stable if there exist a real, symmetric, and positive definite matrix $\mathbf{G} \in \mathbb{R}^{r\times r}$ and a real, diagonal, and positive definite matrix $\mathbf{D} \in \mathbb{R}^{s\times s}$ such that the matrix $\mathbf{M}$ defined by

$$\mathbf{M} := \left[\begin{array}{c|c} \mathbf{DA} + \mathbf{A}^T\mathbf{D} - \mathbf{B}^T\mathbf{GB} & \mathbf{DU} - \mathbf{B}^T\mathbf{GV} \\ \hline \mathbf{U}^T\mathbf{D} - \mathbf{V}^T\mathbf{GB} & \mathbf{G} - \mathbf{V}^T\mathbf{GV} \end{array}\right] \tag{3.1}$$

is nonnegative definite.

We will write $\mathbf{M} \geq 0$ if the matrix is nonnegative definite and $\mathbf{M} > 0$ if $\mathbf{M}$ is positive definite. Note that the above definition requires $\mathbf{G} > 0$ and $\mathbf{D} > 0$.

We would like to observe that the original definition of algebraic stability presented in [10] required only that both matrices $\mathbf{G}$ and $\mathbf{D}$ be nonnegative definite (i.e., $\mathbf{G} \geq 0$ and $\mathbf{D} \geq 0$), while the definition presented in the monograph [15] requires only that $\mathbf{G}$ is positive definite and $\mathbf{D}$ is nonnegative definite (i.e., $\mathbf{G} > 0$ and $\mathbf{D} \geq 0$). The recent result of Hewitt and Hill [16] shows that all these definitions are equivalent for a large class of GLMs (1.2). This result reads as follows.

**Theorem 3.1.** *Assume that for an irreducible GLM (1.2) with coefficients* $\mathbf{c}$, $\mathbf{A}$, $\mathbf{U}$, $\mathbf{B}$, *and* $\mathbf{V}$, *the matrix* $\mathbf{M}$ *defined by (3.1) is nonnegative definite for some real, symmetric and nonnegative definite matrix* $\mathbf{G}$ *and a real, diagonal and nonnegative definite matrix* $\mathbf{D}$. *Then* $\mathbf{G}$ *and* $\mathbf{D}$ *are positive definite.*

An interesting illustration of this theorem will be given in Section 3.

It was demonstrated in [10] that the matrices $\mathbf{G}$ and $\mathbf{D}$ appearing in the definition of algebraic stability are not independent but rather are related by the equation

$$\mathbf{De} = \mathbf{B}^T\mathbf{G}\mathbf{q}_0,$$

where $\mathbf{q}_0$ is the preconsistency vector. Moreover, it follows from Lemma 9.5 in [15] that $\mathbf{G}\mathbf{q}_0$ is a left eigenvector of the matrix $\mathbf{V}$ corresponding to the eigenvalue 1, i.e.,

$$(\mathbf{I} - \mathbf{V}^T)\mathbf{G}\mathbf{q}_0 = \mathbf{0}.$$

These relationships between $\mathbf{G}$ and $\mathbf{D}$ are useful in investigations of algebraic stability.

We will now discuss some implications of algebraic stability of GLMs. To this end consider the test problem

$$y'(t) = f\big(y(t)\big), \quad t \geq 0, \tag{3.2}$$

$f : \mathbb{R}^m \to \mathbb{R}^m$, such that for some inner product $\langle\cdot,\cdot\rangle$ on $\mathbb{R}^m$ the function $f$ satisfies the monotonicity condition

$$\big\langle y, f(y)\big\rangle \leq 0, \tag{3.3}$$

for all $y \in \mathbb{R}^m$. It can be demonstrated that any solution $y$ to (3.2) with $f$ satisfying (3.3) satisfies the condition

$$\big\|y(t+\tau)\big\| \leq \big\|y(t)\big\| \tag{3.4}$$

for all $t \geq 0$ and $\tau > 0$, compare [17]. Here, $\|\cdot\|$ is the corresponding inner product norm $\|y\| = \sqrt{\langle y, y\rangle}$. We would like to point out again that in the original paper [10] it was assumed only that $\langle\cdot,\cdot\rangle$ is a semi-inner product on $\mathbb{R}^m$ and $\|\cdot\|$ the corresponding semi-norm on $\mathbb{R}^m$.

A GLM (1.2) which reflect the non-increasing nature of the solution $y$ to (3.2) as given by (3.4) will be called monotonic. To give a precise formulation of this property we have to choose a norm on the space $\mathbb{R}^{mr}$ (since the computed solution $y^{[n]}$ is represented by a vector in this space). Let

$$u = \begin{bmatrix} u_1^T & u_2^T & \cdots & u_r^T \end{bmatrix}^T \in \mathbb{R}^{mr}, \qquad v = \begin{bmatrix} v_1^T & v_2^T & \cdots & v_r^T \end{bmatrix}^T \in \mathbb{R}^{mr},$$

$u_i, v_j \in \mathbb{R}^r$, and for a symmetric and positive definite matrix $\mathbf{G} = [g_{ij}] \in \mathbb{R}^{r\times r}$ we define the inner product $\langle\cdot,\cdot\rangle_{\mathbf{G}}$ on $\mathbb{R}^{mr}$ and the corresponding inner product norm $\|\cdot\|_{\mathbf{G}}$ by the formulas

$$\big\langle u, v\big\rangle_{\mathbf{G}} = \sum_{i=1}^{r}\sum_{j=1}^{r} g_{ij}\big\langle u_i, v_j\big\rangle, \qquad \big\|u\big\|_{\mathbf{G}} = \sqrt{\big\langle u, u\big\rangle_{\mathbf{G}}}.$$

Then the GLM (1.2) is said to be monotonic if

$$\big\|y^{[n]}\big\|_{\mathbf{G}} \leq \big\|y^{[n-1]}\big\|_{\mathbf{G}}, \tag{3.5}$$

$n = 1, 2, \ldots$. It was proved in [10] that if the GLM (1.2) is algebraically stable then

$$\big\|y^{[n]}\big\|_{\mathbf{G}}^2 - \big\|y^{[n-1]}\big\|_{\mathbf{G}}^2 = 2\sum_{i=1}^{s} d_i\big\langle Y_i^{[n]}, hf(Y_i^{[n]})\big\rangle - \sum_{i=1}^{r+s}\sum_{j=1}^{r+s} m_{ij}\big\langle \alpha_i, \alpha_j\big\rangle, \tag{3.6}$$

where $d_i$ are diagonal elements of the matrix $\mathbf{D}$, $m_{ij}$ are elements of the matrix $\mathbf{M}$ given by (3.1), and the vector $\alpha \in \mathbb{R}^{m(r+s)}$ is defined by

$$\alpha = \begin{bmatrix} (y_1^{[n-1]})^T & \cdots & (y_r^{[n-1]})^T & hf(Y_1^{[n]})^T & \cdots & hf(Y_s^{[n]})^T \end{bmatrix}^T.$$

If the GLM (1.2) is algebraically stable then both terms on the right hand side of (3.6) are non-positive, so the GLM (1.2) is also monotonic, i.e., the numerical solution $y^{[n]}$ satisfies the monotonicity condition (3.5).

It will be demonstrated in Sections 5–7, that it is difficult to satisfy exactly conditions for algebraic stability, especially for high order methods, and for $\epsilon > 0$ we define a weaker property of $\epsilon$-algebraic stability.

**Definition 3.2.** The GLM (1.2) is said to be $\epsilon$-algebraically stable if there exist a real, symmetric, and positive definite matrix $\mathbf{G} \in \mathbb{R}^{r\times r}$ and a real, diagonal, and positive definite matrix $\mathbf{D} \in \mathbb{R}^{s\times s}$ such that the matrix $\mathbf{M}$ defined by (3.1) satisfies the relation

$$\beta^T \mathbf{M}\,\beta \geq -\epsilon\beta^T\beta \tag{3.7}$$

for any $\beta \in \mathbb{R}^{m(r+s)}$.

Clearly, for $\epsilon = 0$ this definition reduces to Definition 3.1 of algebraic stability. Taking into account that if (3.7) holds then

$$-\sum_{i=1}^{r+s}\sum_{j=1}^{r+s} m_{ij}\langle\alpha_i, \alpha_j\rangle = -\alpha^T\mathbf{M}\alpha \leq \epsilon\left\|y^{[n-1]}\right\|_{\mathbf{G}}^2 + h^2\epsilon\left\|f(Y^{[n]})\right\|_{\mathbf{G}}^2,$$

the next result is a consequence of Eq. (3.6).

**Theorem 3.2.** *Assume that the GLM (1.2) is $\epsilon$-algebraically stable. Then the numerical solution $y^{[n]}$ resulting from application of (1.2) to (3.2) with $f$ subject to (3.3) satisfies the relation*

$$\left\|y^{[n]}\right\|_{\mathbf{G}}^2 \leq (1+\epsilon)\left\|y^{[n-1]}\right\|_{\mathbf{G}}^2 + h^2\epsilon\left\|f(Y^{[n]})\right\|_{\mathbf{G}}^2, \tag{3.8}$$

$n = 1, 2, \ldots.$

Hence, if the method (1.2) is $\epsilon$-algebraically stable then it is no longer necessarily contractive, but the growth of the square of the norm $\|y^{[n]}\|_{\mathbf{G}}$ may still be acceptable for small values of $\epsilon$ and moderate stepsizes $h$.

## 4. Technical tools to investigate algebraic and $\epsilon$-algebraic stability

The construction of algebraically stable GLMs is a highly nontrivial task and so far only a few examples of such methods are known in the literature on this subject. Perhaps the most interesting family of such methods is presented by Burrage [11], who found algebraically stable formulas in the special subclass of GLMs, namely the class of multistep Runge–Kutta methods of order $p = 2s$. This construction utilizes an elegant extension of the collocation approach for Runge–Kutta methods. An early example of GLM with $p = 4$, $q = 3$, and $s = r = 2$ was constructed by Dekker [18,17]. Other isolated examples of algebraically stable methods are given in [16,7,19–21] and examples of $\epsilon$-algebraically stable GLMs of order up to four are given in [20,21], although the terminology of $\epsilon$-algebraic stability was not used in these papers.

The search for coefficients $\mathbf{c}$, $\mathbf{A}$, $\mathbf{U}$, $\mathbf{B}$, $\mathbf{V}$ and the vectors $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_p$ of GLMs, symmetric positive definite matrix $\mathbf{G}$ and diagonal positive definite matrix $\mathbf{D}$ such the matrix $\mathbf{M}$ is nonnegative definite is quite tedious and so far only successful for some classes of GLMs with a relatively small number of external stages $r$ and internal stages $s$. In this search the verification of the condition $\mathbf{M} \geq 0$ can be somewhat simplified by the observation by Hewitt and Hill [16] which is based on the Albert theorem [22]. Assume that $\mathbf{M} \in \mathbb{R}^{(s+r)\times(s+r)}$ is partitioned as

$$\mathbf{M} = \left[\begin{array}{c|c} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \hline \mathbf{M}_{12}^T & \mathbf{M}_{22} \end{array}\right],$$

where $\mathbf{M}_{11} \in \mathbb{R}^{s\times s}$, $\mathbf{M}_{12} \in \mathbb{R}^{s\times r}$, $\mathbf{M}_{22} \in \mathbb{R}^{r\times r}$. Then it follows from the Albert result [22] that $\mathbf{M} \geq 0$ if and only if

$$\mathbf{M}_{11} \geq 0, \quad \mathbf{M}_{22} - \mathbf{M}_{12}^T\mathbf{M}_{11}^+\mathbf{M}_{12} \geq 0, \qquad \mathbf{M}_{11}\mathbf{M}_{11}^+\mathbf{M}_{12} = \mathbf{M}_{12}, \tag{4.1}$$

or

$$\mathbf{M}_{22} \geq 0, \quad \mathbf{M}_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^+\mathbf{M}_{12}^T \geq 0, \qquad \mathbf{M}_{22}\mathbf{M}_{22}^+\mathbf{M}_{12}^T = \mathbf{M}_{12}^T. \tag{4.2}$$

Here, $\mathbf{A}^+$ stands for the Moore–Penrose pseudoinverse of the matrix $\mathbf{A}$ [23,24]. Hence, we can reduce the problem of checking if the matrix $\mathbf{M}$ of dimension $r+s$ is nonnegative definite to two smaller problems with matrices of dimension $s$ and $r$ or $r$ and $s$, respectively, and some extra equality constraints, as given in (4.1) or (4.2). The examples of algebraically stable GLMs found using this technique and some additional observations about the structure of the matrix $\mathbf{G}$, the so called minimality property, are given in [16]. In [19] a technique from control theory is proposed where the candidate $\mathbf{G}$ matrices for an algebraically stable GLM may be obtained in terms of the generalized eigenvectors of a generalized eigenproblem related to the condition $\mathbf{M} \geq 0$ and defined in terms of the coefficient matrices $\mathbf{A}$, $\mathbf{U}$, $\mathbf{B}$, and $\mathbf{V}$. The detailed description of this algorithm is also presented in [19] and its applicability illustrated on some known algebraically stable GLMs from [25,16,18,17].

Due to the limitations of symbolic manipulation packages the criteria based on (4.1) or (4.2) may be very difficult to apply for GLMs with larger number of external and internal stages, say, when $s \geq 3$ and $r \geq 3$. Perhaps the more practical approach, where the search for algebraically stable methods can be done numerically, is based on the Nyquist stability function defined by

$$\mathbf{N}(\xi) = \mathbf{A} + \mathbf{U}(\xi\mathbf{I} - \mathbf{V})^{-1}\mathbf{B}, \quad \xi \in \mathbb{C} - \sigma(\mathbf{V}), \tag{4.3}$$

(compare [8,26]). Here, $\sigma(\mathbf{V})$ stands for the spectrum of the matrix $\mathbf{V}$. Following [26] define the diagonal matrix $\mathbf{D}$ by $\mathbf{D} = \mathrm{diag}(\mathbf{B}^T\mathbf{G}\mathbf{q}_0)$. Define also by $\mathrm{He}(\mathbf{Q}) = (\mathbf{Q}+\mathbf{Q}^*)/2$ the Hermitian part of a complex square matrix $\mathbf{Q}$, where $\mathbf{Q}^*$ stands for the conjugate transpose of $\mathbf{Q}$. Then it was demonstrated in [26] (see also [8]) that a consistent GLM (1.2) is algebraically stable if the following conditions are satisfied:

1. The coefficient matrix $\mathbf{V}$ is power-bounded.
2. $\mathbf{U}\mathbf{x} \neq \mathbf{0}$ for all right eigenvectors of $\mathbf{V}$ and $\mathbf{B}^T\mathbf{x} \neq \mathbf{0}$ for all left eigenvectors of $\mathbf{V}$.
3. $\mathbf{D} > 0$ and $\mathrm{He}(\mathbf{D}\mathbf{A}) \geq 0$.
4. $\mathrm{He}\big(\mathbf{D}\mathbf{N}(\xi)\big) \geq 0$ for all $\xi$ such that $|\xi| = 1$ and $\xi \in \mathbb{C} - \sigma(\mathbf{V})$.

Examples of algebraically and $\epsilon$-algebraically stable GLMs in Nordsieck form and two-step Runge–Kutta methods up to order $p = 4$ which were found using the above criteria are presented in [20,21].

Another interesting recent idea of Hewitt and Hill [16,7] is to consider the equivalent GLMs with a simple structure of the symmetric and positive definite matrix $\mathbf{G}$, for example $\mathbf{G} = \mathbf{I}$. This search can be based on the observation that if the methods $(\mathbf{A},\mathbf{U},\mathbf{B},\mathbf{V})$ and $(\widetilde{\mathbf{A}},\widetilde{\mathbf{U}},\widetilde{\mathbf{B}},\widetilde{\mathbf{V}})$ are equivalent then the method $(\mathbf{A},\mathbf{U},\mathbf{B},\mathbf{V})$ is algebraically stable if and only if the method $(\widetilde{\mathbf{A}},\widetilde{\mathbf{U}},\widetilde{\mathbf{B}},\widetilde{\mathbf{V}})$ is algebraically stable [16], as well as on the following result.

**Lemma 4.1** *([16]). An algebraically stable GLM (1.2) with coefficient matrices $(\mathbf{A},\mathbf{U},\mathbf{B},\mathbf{V})$ is equivalent to an algebraically stable method with coefficients $(\widetilde{\mathbf{A}},\widetilde{\mathbf{U}},\widetilde{\mathbf{B}},\widetilde{\mathbf{V}})$ for which $\widetilde{\mathbf{G}} = \mathbf{I}$. Furthermore, if $\mathbf{D} > 0$ is such that the matrix $\mathbf{M}$ defined by (3.1) satisfies $\mathbf{M} \geq 0$, then we have $\widetilde{\mathbf{M}} \geq 0$, where the matrix $\widetilde{\mathbf{M}}$ is defined by*

$$\widetilde{\mathbf{M}} := \left[\begin{array}{c|c} \mathbf{D}\widetilde{\mathbf{A}} + \widetilde{\mathbf{A}}^T\mathbf{D} - \widetilde{\mathbf{B}}^T\widetilde{\mathbf{B}} & \mathbf{D}\widetilde{\mathbf{U}} - \widetilde{\mathbf{B}}^T\widetilde{\mathbf{V}} \\ \hline \widetilde{\mathbf{U}}^T\mathbf{D} - \widetilde{\mathbf{V}}^T\widetilde{\mathbf{B}} & \mathbf{I} - \widetilde{\mathbf{V}}^T\widetilde{\mathbf{V}} \end{array}\right].$$

Consider again the GLM (1.2) with coefficients $(\mathbf{A},\mathbf{U},\mathbf{B},\mathbf{V})$ and the matrix $\mathbf{M}$ defined by (3.1), which assuming that $\mathbf{G} = \mathbf{I}$, takes the form

$$\mathbf{M} = \left[\begin{array}{c|c} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \hline \mathbf{M}_{12}^T & \mathbf{M}_{22} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{D}\mathbf{A} + \mathbf{A}^T\mathbf{D} - \mathbf{B}^T\mathbf{B} & \mathbf{D}\mathbf{U} - \mathbf{B}^T\mathbf{V} \\ \hline \mathbf{U}^T\mathbf{D} - \mathbf{V}^T\mathbf{B} & \mathbf{I} - \mathbf{V}^T\mathbf{V} \end{array}\right].$$

Then, as discussed above, it follows from Albert theorem [22] that $\mathbf{M} \geq 0$ if and only if $\mathbf{M}_{22} \geq 0$, $\mathbf{R} \geq 0$, and $\mathbf{M}_{22}\mathbf{M}_{22}^{+}\mathbf{M}_{12}^T = \mathbf{M}_{12}^T$, where $\mathbf{R}$ is defined by

$$\begin{aligned}\mathbf{R} &:= \mathbf{M}_{11} - \mathbf{M}_{12}\mathbf{M}_{22}^{+}\mathbf{M}_{12}^T \\ &= \mathbf{D}\mathbf{A} + \mathbf{A}^T\mathbf{D} - \mathbf{B}^T\mathbf{B} - (\mathbf{D}\mathbf{U} - \mathbf{B}^T\mathbf{V})(\mathbf{I} - \mathbf{V}^T\mathbf{V})^{+}(\mathbf{D}\mathbf{U} - \mathbf{B}^T\mathbf{V})^T.\end{aligned}$$

Lemma 4.1 and the choice $\mathbf{R} = 0$, lead to the following algorithm for the construction of GLMs with coefficients $(\mathbf{A},\mathbf{U},\mathbf{B},\mathbf{V})$ for which $\mathbf{G} = \mathbf{I}$. This algorithm, which was recently proposed in [7], consists of the following steps.

1. Choose the matrix $\mathbf{G} = \mathbf{I}$.
2. Ensure that $\mathbf{D} = \mathrm{diag}(\mathbf{B}^T\mathbf{q}_0) > 0$.
3. Ensure that $\mathbf{M}_{22} = \mathbf{I} - \mathbf{V}^T\mathbf{V} \geq 0$.
4. Enforce the condition $\mathbf{R} = 0$.

Once the methods are found following the above algorithm we have also to verify the condition

$$\mathbf{M}_{22}\mathbf{M}_{22}^{+}\mathbf{M}_{12}^T = \mathbf{M}_{12}^T$$

appearing in (4.2). In our search for algebraically stable methods described in Sections 5–7 this condition was usually automatically satisfied. In fact, the verification of this condition was not even mentioned in [7].

**Remark.** Steps 1 to 3 do not cause loss of generality, while step 4 is a sufficient condition to enforce the algebraic stability of the methods (cf. [7]).

Carrying out this algorithm with the steps 3 and 4 above replaced by $\min\sigma(\mathbf{M}_{22}) \geq -\epsilon$ and $\|\mathbf{R}\| = O(\epsilon)$, where $\epsilon$ is small and positive, leads to methods which are $\epsilon$-algebraically stable.

We conclude this section with a technical result which places some restrictions on the structure of the matrix $\mathbf{V}$ such that $\mathbf{M}_{22} = \mathbf{I} - \mathbf{V}^T\mathbf{V} \geq 0$.

**Lemma 4.2.** *Assume that the matrix $\mathbf{V}$ has the form as described in Section 1. Then $\mathbf{M}_{22} = \mathbf{I} - \mathbf{V}^T\mathbf{V} \geq 0$ implies that $v_{1,j} = 0, j = 2, 3, \ldots, r$.*

**Proof.** It can be verified that the matrix $\mathbf{M}_{22}$ has the form

$$\mathbf{M}_{22} = \begin{bmatrix} 0 & -v_{1,2} & \cdots & -v_{1,r} \\ -v_{1,2} & \times & \cdots & \times \\ \vdots & \vdots & \ddots & \vdots \\ -v_{1,r} & \times & \cdots & \times \end{bmatrix},$$

where $\times$'s designate arbitrary entries. Then $\mathbf{M}_{22} \geq 0$ implies that

$$\det \begin{bmatrix} 0 & -v_{1,j} \\ -v_{1,j} & \times \end{bmatrix} = -v_{1,j}^2 \geq 0,$$

which can be only satisfied if $v_{1,j} = 0, j = 2, 3, \ldots, r$. $\square$

## 5. Methods with $p = q = r = s = 2$

Solving stage order and order conditions (2.4) and (2.5) or using the representation formulas (2.6)–(2.8) for $\mathbf{U}$, $\mathbf{q}_2$, $\mathbf{B}$, and using Lemma 4.2 leads to a seven parameter family of methods of stage order $q = 2$ and order $p = 2$ depending on the parameters $c_1, c_2, \lambda, a_{21}, v_{22}, q_{11}$, and $q_{21}$, for which

$$M_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 - v_{22} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \dfrac{1 - 2c_2}{2(c_1 - c_2)} & 0 \\ 0 & \dfrac{2c_1 - 1}{2(c_1 - c_2)} \end{bmatrix}.$$

Observe that the matrix $D$ is positive definite if and only if

$$\left(c_1 < \frac{1}{2} \text{ and } c_2 > \frac{1}{2}\right) \quad \text{or} \quad \left(c_1 > \frac{1}{2} \text{ and } c_2 < \frac{1}{2}\right).$$

Trying to construct methods which are algebraically stable we will make several simplifying assumptions choosing some parameters of the methods in advance. We assume first that $v_{22} = 0$ and $\mathbf{q}_1 = [q_{11}, q_{21}]^T = [0, 1]^T$, and then try to enforce the condition $\mathbf{R} = 0$. This leads to the system of three equations in four unknowns. Assuming in addition that $c_2 = 1$ this system has four real solutions including one solution with rational coefficients which is given by $c_1 = 1/2$, $\lambda = 3/2$, and $a_{21} = -1/8$. The coefficients of the resulting method are

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{cc|cc} \frac{3}{2} & 0 & 1 & -1 \\ -\frac{1}{8} & \frac{3}{2} & 1 & -\frac{3}{8} \\ \hline 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right].$$

For this method the spectrum of the matrix $\mathbf{M}$ defined by (3.1) is $\sigma(\mathbf{M}) = \{2, 0, 0, 0\}$. The matrix $\mathbf{D}$ takes the form $\mathbf{D} = \text{diag}(1, 0)$ and is not positive definite. Hence, it follows from Theorem 3.1 that this method must be reducible. It is easy to verify that the reduced method is

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{c|cc} \frac{3}{2} & 1 & -1 \\ \hline 1 & 1 & 0 \\ 1 & 0 & 0 \end{array}\right],$$

for which $\sigma(\mathbf{M}) = \{2, 0, 0\}$ and $\mathbf{D} = [1]$.

The other three solutions to the system $\mathbf{R} = 0$ using a finite precision arithmetic provide only approximations to algebraically stable methods. One such approximation in rational format obtained in Mathematica using the function `Rationalize`$[x, tol]$ with $tol = 10^{-16}$ is given by

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{U} \\ \hline \mathbf{B} & \mathbf{V} \end{array}\right] = \left[\begin{array}{cc|cc} \frac{235875500}{76532157} & 0 & 1 & -\frac{74457421}{26595157} \\ -\frac{317908253}{60937223} & \frac{235875500}{76532157} & 1 & \frac{248088986}{79136875} \\ \hline \frac{35834268}{51430529} & \frac{42749463}{140971448} & 1 & 0 \\ \frac{4321111}{66099453} & \frac{54070473}{57852454} & 0 & 0 \end{array}\right]. \tag{5.1}$$

For this method $\sigma(\mathbf{M}) = \{5.71, -2.09 \cdot 10^{-13}, 8.82 \cdot 10^{-14}, -3.56 \cdot 10^{-17}\}$ and $\|\mathbf{R}\| = 4.19 \cdot 10^{-13}$ so the method is $\epsilon$-algebraically stable with $\epsilon$ equal to $2 \cdot 10^{-13}$. Increasing the precision of the computations leads to $\epsilon$-algebraically stable methods with decreasing values of $\epsilon$. This is illustrated in Table 1, where we have listed the precision of the computations set by the parameter WorkingPrecision in Mathematica versus $\min\{\sigma(\mathbf{M})\}$, $\min\{\sigma(\mathbf{R})\}$, and $\|\mathbf{R}\|_F$, where MachinePrecision corresponds to about 16 correct decimal digits and $\|\cdot\|_F$ stands for Frobenius norm.

We can also construct methods for which the coefficient matrix $\mathbf{A}$ is diagonal. Putting $\mathbf{q}_1 = [0, 1]^T$ and $v_{22} = 0$ and solving the system $\mathbf{R} = 0$ with respect to $c_1$, $c_2$, and $\lambda$ we obtain $c_1 = (5 - \sqrt{22})/2$, $c_2 = (5 + \sqrt{22})/2$, $\lambda = 3/2$, and the

**Table 1**
WorkingPrecision versus $\min\{\sigma(\mathbf{M})\}$, $\min\{\sigma(\mathbf{R})\}$, and $\|\mathbf{R}\|_F$.

| WorkingPrecision | $\min\{\sigma(\mathbf{M})\}$ | $\min\{\sigma(\mathbf{R})\}$ | $\|\mathbf{R}\|_F$ |
|---|---|---|---|
| MachinePrecision | $-1.022\cdot 10^{-13}$ | $-1.216\cdot 10^{-13}$ | $2.454\cdot 10^{-13}$ |
| 20 | $-1.771\cdot 10^{-20}$ | $-8.080\cdot 10^{-20}$ | $9.988\cdot 10^{-20}$ |
| 30 | $-3.860\cdot 10^{-30}$ | $-1.579\cdot 10^{-29}$ | $2.434\cdot 10^{-29}$ |
| 40 | $-1.364\cdot 10^{-40}$ | $-6.500\cdot 10^{-40}$ | $6.501\cdot 10^{-40}$ |
| 50 | $-1.026\cdot 10^{-50}$ | $-2.412\cdot 10^{-50}$ | $3.315\cdot 10^{-49}$ |
| 100 | $-2.763\cdot 10^{-100}$ | $-8.439\cdot 10^{-100}$ | $8.440\cdot 10^{-100}$ |
| 200 | $-2.538\cdot 10^{-167}$ | $-2.591\cdot 10^{-167}$ | $1.323\cdot 10^{-167}$ |
| 500 | $-8.995\cdot 10^{-467}$ | $-4.400\cdot 10^{-466}$ | $4.400\cdot 10^{-466}$ |

coefficients of the resulting method are

$$\left[\begin{array}{c|c}\mathbf{A} & \mathbf{U}\\ \hline \mathbf{B} & \mathbf{V}\end{array}\right]=\left[\begin{array}{cc|cc}\frac{3}{2} & 0 & 1 & \frac{2-\sqrt{22}}{2}\\ 0 & \frac{3}{2} & 1 & \frac{2+\sqrt{22}}{2}\\ \hline \frac{11+\sqrt{22}}{22} & \frac{11-\sqrt{22}}{22} & 1 & 0\\ \frac{22+\sqrt{22}}{44} & \frac{22-\sqrt{22}}{44} & 0 & 0\end{array}\right].$$

For this method $\mathbf{D}=\operatorname{diag}((11+2\sqrt{22})/22,\ (11-2\sqrt{22})/22)$ and $\sigma(\mathbf{M})=\{115/44, 0, 0, 0\}$.

We relax next the condition on $\mathbf{q}_1$ and assume only that $q_{11}=0$. In this case we do not assume that $\mathbf{A}$ is diagonal. Assuming that the abscissa vector is given by $\mathbf{c}=[1/3, 2/3]^T$ and solving the system $\mathbf{R}=0$ we obtain $\lambda=2/3$, $a_{21}=-1/6$, and $q_{21}=\pm\sqrt{3}/6$. Choosing $q_{21}=\sqrt{3}/6$ leads to the method

$$\left[\begin{array}{c|c}\mathbf{A} & \mathbf{U}\\ \hline \mathbf{B} & \mathbf{V}\end{array}\right]=\left[\begin{array}{cc|cc}\frac{2}{3} & 0 & 1 & -\frac{2\sqrt{3}}{3}\\ -\frac{1}{6} & \frac{2}{3} & 1 & \frac{\sqrt{3}}{3}\\ \hline \frac{1}{2} & \frac{1}{2} & 1 & 0\\ -\frac{\sqrt{3}}{6} & \frac{\sqrt{3}}{3} & 0 & 0\end{array}\right],$$

which was obtained before in [7]. For this method $\sigma(\mathbf{M})=\{17/12, 0, 0, 0\}$ and $\mathbf{D}=\operatorname{diag}(1/2, 1/2)$.

## 6. Methods with $p=q=r=s=3$

Using the representation formulas (2.6)–(2.8) for $\mathbf{U}$, $\mathbf{q}_3$, $\mathbf{B}$, and taking into account Lemma 4.2, we obtain a 16 parameter family of methods of stage order $q=3$ and order $p=3$ depending on the parameters $c_1, c_2, c_3, \lambda, a_{21}, a_{31}, a_{32}, v_{22}, v_{23}, v_{33}, q_{11}, q_{21}, q_{31}, q_{12}, q_{22}$ and $q_{32}$. In this case it is not easy to force the condition $\mathbf{R}=0$, because it constitutes a nonlinear system of six equations in sixteen unknowns, but numerical approximations for algebraically stable methods can be found by minimizing the quantity $\|\mathbf{R}\|_F$. To this aim we have used the `NMinimize` function of Mathematica. An example of such a method obtained with WorkingPrecision equal to MachinePrecision is given by

$$\mathbf{c}=\left[-\frac{101945136}{112612663}, -\frac{4581335}{62678469}, \frac{31271554}{33364359}\right]^T,$$

$$\mathbf{A}=\begin{bmatrix}\frac{46917000}{104218099} & 0 & 0\\ \frac{143886276}{141018391} & \frac{46917000}{104218099} & 0\\ \frac{26519721}{51255491} & \frac{32428625}{49353386} & \frac{46917000}{104218099}\end{bmatrix},\qquad \mathbf{U}=\begin{bmatrix}1 & \frac{72449071}{83980704} & \frac{23278421}{81652661}\\ 1 & \frac{71510487}{58991348} & \frac{21000658}{151963191}\\ 1 & \frac{3714629}{143661548} & -\frac{28549797}{77302997}\end{bmatrix},$$

$$\mathbf{B}=\begin{bmatrix}\frac{24899124}{51799415} & \frac{29578637}{98768301} & \frac{24291430}{110495197}\\ -\frac{28254283}{168393484} & -\frac{23259899}{90332233} & -\frac{22315183}{65037990}\\ \frac{5762224}{48449435} & -\frac{6769993}{73337137} & -\frac{21501955}{134501824}\end{bmatrix},\qquad \mathbf{V}=\begin{bmatrix}1 & 0 & 0\\ 0 & -\frac{16561402}{68584587} & \frac{5012623}{85021484}\\ 0 & 0 & \frac{19299246}{50267989}\end{bmatrix},$$

$$\mathbf{q}_1=\left[-\frac{66758655}{88894537}, -\frac{70240814}{111634289}, -\frac{45996736}{212672177}\right]^T,\qquad \mathbf{q}_2=\left[\frac{41739183}{82954951}, \frac{9734946}{25441867}, -\frac{4987490}{89126563}\right]^T.$$

**Table 2**
WorkingPrecision versus $\min\{\sigma(\mathbf{M})\}$, $\min\{\sigma(\mathbf{R})\}$ and $\|\mathbf{R}\|_F$.

| WorkingPrecision | $\min\{\sigma(\mathbf{M})\}$ | $\min\{\sigma(\mathbf{R})\}$ | $\|\mathbf{R}\|_F$ |
|---|---|---|---|
| MachinePrecision | $-1.972\cdot 10^{-6}$ | $-2.228\cdot 10^{-6}$ | $2.228\cdot 10^{-6}$ |
| 20 | $-3.371\cdot 10^{-8}$ | $-3.808\cdot 10^{-8}$ | $3.808\cdot 10^{-8}$ |
| 30 | $-1.572\cdot 10^{-11}$ | $-1.776\cdot 10^{-11}$ | $1.776\cdot 10^{-11}$ |
| 40 | $-3.024\cdot 10^{-14}$ | $-3.417\cdot 10^{-14}$ | $3.417\cdot 10^{-14}$ |
| 50 | $-5.644\cdot 10^{-18}$ | $-6.378\cdot 10^{-18}$ | $6.378\cdot 10^{-18}$ |
| 100 | $-8.319\cdot 10^{-35}$ | $-9.400\cdot 10^{-35}$ | $9.400\cdot 10^{-35}$ |
| 200 | $-3.591\cdot 10^{-68}$ | $-4.058\cdot 10^{-68}$ | $4.058\cdot 10^{-68}$ |
| 500 | $-1.435\cdot 10^{-167}$ | $-1.621\cdot 10^{-167}$ | $1.621\cdot 10^{-167}$ |

It can be verified that for this method

$$\mathbf{D} = \mathrm{diag}(0.481, 0.299, 0.220),$$

and

$$\sigma(\mathbf{M}) = \{1.213, 0.849, -1.972\cdot 10^{-6}, -1.966\cdot 10^{-8}, -3.201\cdot 10^{-9}, 1.124\cdot 10^{-16}\}.$$

Similarly as in Section 5, increasing the precision of the computations leads to $\epsilon$-algebraically stable methods with decreasing values of $\epsilon$. This is illustrated in Table 2, where we have listed the precision of the computations set by the parameter WorkingPrecision in Mathematica versus $\min\{\sigma(\mathbf{M})\}$, $\min\{\sigma(\mathbf{R})\}$, and $\|\mathbf{R}\|_F$.

## 7. Methods with $p = q = r = s = 4$

Using the representation formulas (2.6)–(2.8) for $\mathbf{U}$, $\mathbf{q}_4$, $\mathbf{B}$, and taking into account Lemma 4.2, we obtain a 29 parameter family of methods of stage order $q = 4$ and order $p = 4$ depending on the parameters $c_1, c_2, c_3, c_4, \lambda, a_{21}, a_{31}, a_{32}, a_{41}, a_{42}, a_{44}, v_{22}, v_{23}, v_{24}, v_{33}, v_{34}, v_{44}, q_{11}, q_{21}, q_{31}, q_{41}, q_{12}, q_{22}, q_{32}, q_{42}, q_{13}, q_{23}, q_{33}$ and $q_{43}$.

As in the previous section we look for numerical approximations for algebraically stable methods minimizing the quantity $\|\mathbf{R}\|_F$ by means of the `NMinimize` Mathematica function. An example of such a method obtained with WorkingPrecision equal to MachinePrecision is given by

$$\mathbf{c} = \left[-\frac{136691984}{56181129}, -\frac{50483434}{97001607}, \frac{67774851}{79681789}, \frac{113679233}{29070903}\right]^T,$$

$$\mathbf{A} = \begin{bmatrix} \frac{72790755}{100862638} & 0 & 0 & 0 \\ \frac{476785}{87049992} & \frac{72790755}{100862638} & 0 & 0 \\ \frac{20050367}{136240340} & \frac{45511759}{88582495} & \frac{72790755}{100862638} & 0 \\ -\frac{47594845}{103765783} & \frac{106930055}{278211234} & \frac{27285361}{74242319} & \frac{72790755}{100862638} \end{bmatrix},$$

$$\mathbf{U} = \begin{bmatrix} 1 & -\frac{451634987}{140827502} & -\frac{152906573}{43743542} & \frac{145488433}{59069284} \\ 1 & -\frac{28380214}{46050151} & \frac{111117249}{128311658} & \frac{8094106}{111485217} \\ 1 & \frac{80259071}{139207756} & -\frac{23821648}{70030981} & \frac{93027797}{86529613} \\ 1 & -\frac{113721374}{32649629} & -\frac{117707786}{26849261} & -\frac{352780594}{16109787} \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} \frac{8055613}{173594494} & \frac{50692033}{76932071} & \frac{31238859}{106472948} & \frac{25987}{20323048} \\ -\frac{8452327}{97477693} & \frac{15801488}{88238837} & \frac{44127692}{96094939} & -\frac{2443177}{332935081} \\ -\frac{4049179}{33492836} & \frac{32497415}{81434453} & -\frac{29137108}{100267221} & -\frac{1159458}{119366215} \\ \frac{135447}{64189342} & -\frac{4040557}{130316218} & -\frac{3386596}{40127799} & -\frac{633671}{16754328} \end{bmatrix},$$

**Table 3**
WorkingPrecision versus $\min\{\sigma(\mathbf{M})\}$, $\min\{\sigma(\mathbf{R})\}$ and $\|\mathbf{R}\|_F$.

| WorkingPrecision | $\min\{\sigma(\mathbf{M})\}$ | $\min\{\sigma(\mathbf{R})\}$ | $\|\mathbf{R}\|_F$ |
|---|---|---|---|
| MachinePrecision | $-9.290 \cdot 10^{-8}$ | $-1.042 \cdot 10^{-7}$ | $1.158 \cdot 10^{-7}$ |
| 20 | $-2.532 \cdot 10^{-8}$ | $-2.864 \cdot 10^{-8}$ | $2.888 \cdot 10^{-8}$ |
| 30 | $-6.271 \cdot 10^{-9}$ | $-7.038 \cdot 10^{-9}$ | $7.041 \cdot 10^{-9}$ |
| 40 | $-8.294 \cdot 10^{-11}$ | $-9.301 \cdot 10^{-11}$ | $9.304 \cdot 10^{-11}$ |
| 50 | $-7.775 \cdot 10^{-12}$ | $-8.732 \cdot 10^{-12}$ | $8.733 \cdot 10^{-12}$ |
| 100 | $-5.419 \cdot 10^{-19}$ | $-6.099 \cdot 10^{-19}$ | $6.099 \cdot 10^{-19}$ |

$$\mathbf{V} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \dfrac{18257081}{41519760} & \dfrac{38469238}{94411247} & \dfrac{29985486}{70219573} \\ 0 & 0 & \dfrac{189820571}{228231862} & -\dfrac{1274917}{110945271} \\ 0 & 0 & 0 & \dfrac{1075320273}{2684824064} \end{bmatrix},$$

$$\mathbf{q}_1 = \begin{bmatrix} -\dfrac{36042078}{51394831} & \dfrac{38765702}{55691309} & -\dfrac{13608701}{119003375} & -\dfrac{28697948}{113848499} \end{bmatrix}^T,$$

$$\mathbf{q}_2 = \begin{bmatrix} \dfrac{107717787}{204603878} & -\dfrac{87950752}{125801249} & -\dfrac{31179097}{61611950} & \dfrac{3157174}{43616961} \end{bmatrix}^T,$$

$$\mathbf{q}_3 = \begin{bmatrix} -\dfrac{40924885}{150423523} & \dfrac{50875109}{115541270} & \dfrac{55776045}{110288264} & -\dfrac{96101253}{218223736} \end{bmatrix}^T.$$

It can be verified that for this method

$$\mathbf{D} = \mathrm{diag}(0.046, 0.659, 0.293, 0.001),$$

and

$$\sigma(\mathbf{M}) = \{1.208, 0.769, 0.031, -9.290 \cdot 10^{-8}, -4.116 \cdot 10^{-8}, 2.272 \cdot 10^{-8}, -3.284 \cdot 10^{-9}, -1.358 \cdot 10^{-18}\}.$$

Once again, increasing the precision of the computations leads to $\epsilon$-algebraically stable methods with decreasing values of $\epsilon$. This is illustrated in Table 3, where we have listed the precision of the computations set by the parameter WorkingPrecision in Mathematica versus $\min\{\sigma(\mathbf{M})\}$, $\min\{\sigma(\mathbf{R})\}$, and $\|\mathbf{R}\|_F$.

## 8. Numerical experiments

In this section we will illustrate that some of the DIMSIMs of order $p$ and stage order $q = p$ derived in this paper do not suffer from order reduction, unlike standard classical Runge–Kutta formulas. To illustrate this we have applied the Runge–Kutta–Gauss method of order $p = 4$ and stage order $q = 2$ and DIMSIMs of order $p = 2, 3$ and 4 and stage order $q = p$ given in Section 5 (method (5.1)), 6 and 7 to the van der Pol oscillator (see VDPOL problem in [15])

$$\begin{cases} y_1' = y_2, & y_1(0) = 2, \\ y_2' = ((1 - y_1^2)y_2 - y_1)/\varepsilon, & y_2(0) = -2/3, \end{cases} \tag{8.1}$$

$t \in [0, T]$, with a stiffness parameter $\varepsilon$. We have implemented all methods with a fixed stepsize $h$, and observed the order of convergence of the numerical approximations to the slowly varying parts of the solution, where the problem is stiff for small values of the parameter $\varepsilon$ (the problem is not stiff on the interval where the solution is changing rapidly). We compare the numerical results for the solution at the final time with a reference solution computed by the Matlab function *ode15s* with very tight tolerances $atol = 2 * 10^{-16}$ and $rtol = 10^{-14}$. The errors are measured in the $\|\cdot\|_2$ norm. Plots of error versus time step size for the Runge–Kutta–Gauss method of order $p = 4$ and stage order $q = 2$ and DIMSIMs of order $p = 2, 3, 4$ and stage order $q = p$ applied to problem (8.1) with $T = 3/4$, $h = T/N$, and $N = 256, 512, 1024$ and 2048 are reported in Figs. 1, 2, 3 and 4, respectively. The observed orders reported in Figs. 2–4 match the theoretical predictions, while the one reported in Fig. 1 shows an order reduction phenomenon.

## 9. Concluding remarks

We introduced the definition of $\epsilon$-algebraic stable method and investigated algebraically and $\epsilon$-algebraically stable DIMSIMs of order $p$, for which the stage order $q$, the number of internal stages $s$ and the number of external stages $r$ are

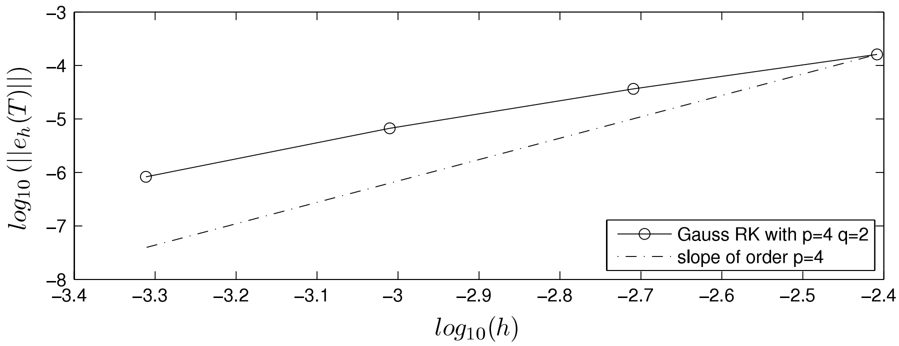


**Fig. 1.** Numerical results for the Runge–Kutta–Gauss formula with $p = 4$ and $q = 2$ on the van der Pol problem (8.1) with $\epsilon = 10^{-6}$ and $T = 3/4$.

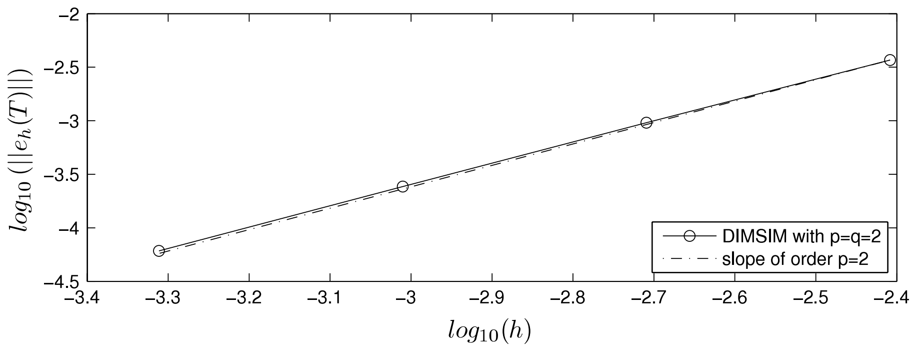


**Fig. 2.** Numerical results for the DIMSIM (5.1) derived in Section 5 on the van der Pol problem (8.1) with $\epsilon = 10^{-6}$ and $T = 3/4$.

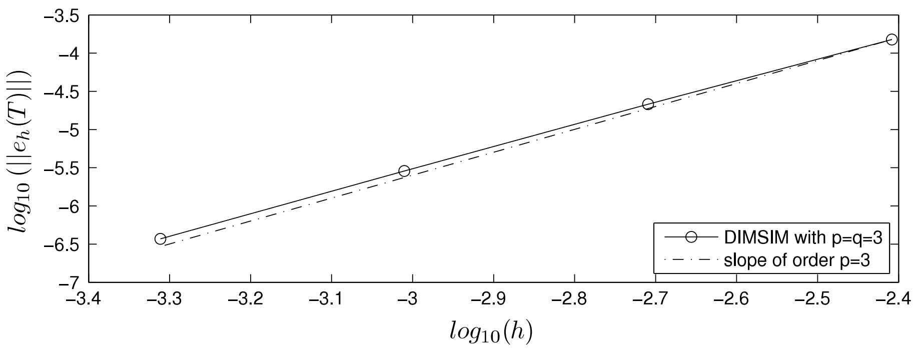


**Fig. 3.** Numerical results for the DIMSIM derived in Section 6 on the van der Pol problem (8.1) with $\epsilon = 10^{-6}$ and $T = 3/4$.

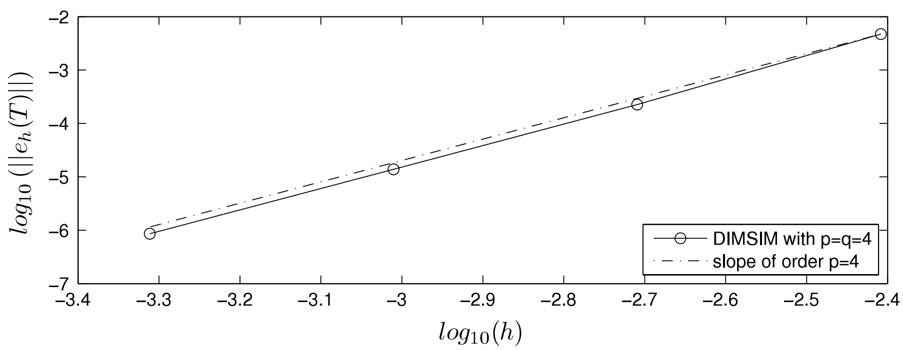


**Fig. 4.** Numerical results for the DIMSIM derived in Section 7 on the van der Pol problem (8.1) with $\epsilon = 10^{-6}$ and $T = 3/4$.

all equal to $p$. For this particular class of DIMSIMs we gave explicit representation formulas for matrices $\mathbf{B}$ and $\mathbf{U}$ and for the vector $\mathbf{q}_p$. Using a criteria proposed recently by Hewitt and Hill [7] we found algebraically and $\epsilon$-algebraically stable methods with $p = q = r = s = 2$, and $\epsilon$-algebraically stable methods for small values of the parameters $\epsilon$, with $p = q = r = s = 3$ and $p = q = r = s = 4$. Finally we showed by means of numerical tests that the DIMSIMs constructed in this paper do not suffer from order reduction, unlike standard classical Runge–Kutta methods.

Future work will address various implementation issues related to these methods such as choice of initial stepsize, the construction of starting procedures, local error estimation, stepsize and order changing strategies, strategies for updating the vector of external approximations, and construction of continuous interpolants of uniform order $p$, as well as developing variable stepsize variable order software based on these methods for stiff differential systems.

## Acknowledgments

The results reported in this paper were obtained during the visit of the first author (GI) to the Arizona State University in March and April of 2011. This author wishes to express his gratitude to the Department of Mathematics for hospitality during this visit.

## References

[1] Z. Jackiewicz, General Linear Methods for Ordinary Differential Equations, John Wiley & Sons, Hoboken, New Jersey, 2009.
[2] J.C. Butcher, Diagonally-implicit multi-stage integration methods, Appl. Numer. Math. 11 (1993) 347–363.
[3] J.C. Butcher, Z. Jackiewicz, Diagonally implicit general linear methods for ordinary differential equations, BIT 33 (1993) 452–472.
[4] J.C. Butcher, Z. Jackiewicz, Construction of diagonally implicit general linear methods of type 1 and 2 for ordinary differential equations, Appl. Numer. Math. 21 (1996) 385–415.
[5] J.C. Butcher, Z. Jackiewicz, Implementation of diagonally implicit multistage integration methods for ordinary differential equations, SIAM J. Numer. Anal. 34 (1997) 2119–2141.
[6] J.C. Butcher, Z. Jackiewicz, Construction of high order diagonally implicit multistage integration methods for ordinary differential equations, Appl. Numer. Math. 27 (1998) 1–12.
[7] L.L. Hewitt, A.T. Hill, Algebraically stable diagonally implicit general linear methods, Appl. Numer. Math. 60 (2010) 629–636.
[8] J.C. Butcher, The equivalence of algebraic stability and *AN*-stability, BIT 27 (1987) 510–533.
[9] J.C. Butcher, Numerical Methods for Ordinary Differential Equations, second ed., John Wiley & Sons, Chichester, 2008.
[10] K. Burrage, J.C. Butcher, Non-linear stability of a general class of differential equation methods, BIT 20 (1980) 185–203.
[11] K. Burrage, High order algebraically stable multistep Runge–Kutta methods, SIAM J. Numer. Anal. 24 (1987) 106–115.
[12] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations. Runge–Kutta and General Linear Methods, John Wiley & Sons, Chichester, New York, 1987.
[13] J.C. Butcher, Linear and non-linear stability for general linear methods, BIT 27 (1987) 182–189.
[14] J.C. Butcher, Thirty years of *G*-stability, BIT 46 (2006) 479–489.
[15] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems, Springer-Verlag, Berlin, Heidelberg, New York, 1996.
[16] L.L. Hewitt, A.T. Hill, Algebraically stable general linear methods and the *G*-matrix, BIT 49 (2009) 93–111.
[17] K. Dekker, J.G. Verwer, Stability of Runge–Kutta Methods for Stiff Nonlinear Differential Equations, North-Holland, Amsterdam, New York, Oxford, 1984.
[18] K. Dekker, Algebraic stability of general linear methods, Tech. Rep. 25, Computer Science Department, University of Auckland, New Zealand, 1981.
[19] A.T. Hill, *G*-matrices for algebraically stable general linear methods, Numer. Algorithms 53 (2010) 281–292.
[20] M. Braś, Z. Jackiewicz, Nordsieck methods with computationally verified algebraic stability, Appl. Math. Comput. 217 (2011) 8598–8610.
[21] R. D'Ambrossio, G. Izzo, Z. Jackiewicz, Search for highly stable two-step Runge–Kutta methods, Appl. Numer. Math. 62 (2012) 1361–1379.
[22] A. Albert, Conditions for positive and non-negative definiteness in terms of pseudoinverses, SIAM J. Appl. Math. 17 (1969) 434–440.
[23] J.W. Demmel, Applied Numerical Linear Algebra, SIAM, Philadelphia, 1997.
[24] G.H. Golub, C.F. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, London, 1996.
[25] J.C. Butcher, A.T. Hill, Linear multistep methods as irreducible general linear methods, BIT 46 (2006) 5–19.
[26] A.T. Hill, Nonlinear stability of general linear methods, Numer. Math. 103 (2006) 611–629.