Journal of Computational and Applied Mathematics 279 (2015) 159–172

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Comparison of boundedness and monotonicity properties of one-leg and linear multistep methods

A. Mozartova, I. Savostianov, W. Hundsdorfer*

CWI, P.O. Box 94079, 1090-GB Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 28 December 2012
Received in revised form 19 August 2014

MSC:
65L05
65L06
65L20
65M20

Keywords:
Initial value problem
Method of lines (MOL)
Multistep methods
Monotonicity
Boundedness
Strong-stability-preserving (SSP)

ABSTRACT

One-leg multistep methods have some advantage over linear multistep methods with respect to storage of the past results. In this paper boundedness and monotonicity properties with arbitrary (semi-)norms or convex functionals are analyzed for such multistep methods. The maximal stepsize coefficient for boundedness and monotonicity of a one-leg method is the same as for the associated linear multistep method when arbitrary starting values are considered. It will be shown, however, that combinations of one-leg methods and Runge–Kutta starting procedures may give very different stepsize coefficients for monotonicity than the linear multistep methods with the same starting procedures. Detailed results are presented for explicit two-step methods.

1. Introduction

The ODE systems and basic assumption

We consider systems of ordinary differential equations (ODEs) with given initial value in a vector space $\mathbb{V}$,

$$u'(t) = F(u(t)), \qquad u(0) = u_0, \tag{1.1}$$

where $F : \mathbb{V} \to \mathbb{V}$ and $u_0 \in \mathbb{V}$. Throughout this paper we will make the following basic assumption: there is a constant $\tau_0 > 0$ such that

$$\|v + \tau_0 F(v)\| \le \|v\| \quad \text{for all } v \in \mathbb{V}, \tag{1.2}$$

where $\|\cdot\|$ denotes a norm, seminorm, or convex functional on $\mathbb{V}$ (cf. Section 2).

It is easy to see that (1.2) implies $\|v + \Delta t\, F(v)\| \le \|v\|$ for all $\Delta t \in (0, \tau_0]$. Consequently, applying the forward Euler method $u_n = u_{n-1} + \Delta t\, F(u_{n-1})$, $n \ge 1$, with stepsize $\Delta t > 0$ to compute approximations $u_n \approx u(t_n)$ at $t_n = n\Delta t$, we obtain $\|u_n\| \le \|u_0\|$ for $n \ge 1$ under the stepsize restriction $\Delta t \le \tau_0$. For general one-step methods, this property under a stepsize restriction $\Delta t \le \gamma\, \tau_0$ is often referred to as *monotonicity* or *strong stability preservation* (SSP). For multistep methods this can be generalized in several ways, which will be addressed in this paper.

* Corresponding author.
E-mail addresses: mozartovaas@yahoo.com (A. Mozartova), igor.savostianov@gmail.com (I. Savostianov), willem.hundsdorfer@cwi.nl (W. Hundsdorfer).

http://dx.doi.org/10.1016/j.cam.2014.10.025
0377-0427/

Linear multistep and one-leg methods

In this paper we will consider one-leg and linear multistep methods for finding the approximations $u_n \approx u(t_n)$ at the step points $t_n = n\Delta t$, $n \geq 1$. It is supposed that starting vectors $u_0, u_1, \ldots, u_{k-1} \in \mathbb{V}$ are known.

A linear multistep (LM) method applied to (1.1) reads

$$u_n = \sum_{j=1}^{k} a_j u_{n-j} + \Delta t \sum_{j=0}^{k} b_j F(u_{n-j}) \tag{1.3}$$

for $n \geq k$. The parameters a_j, b_j and $k \in \mathbb{N}$ define the method. Along with this linear multistep method, we also consider the corresponding k-step one-leg (OL) method

$$u_n = \sum_{j=1}^{k} a_j u_{n-j} + \Delta t \beta F(v_n), \qquad v_n = \sum_{j=0}^{k} \hat{b}_j u_{n-j} \tag{1.4}$$

for $n \geq k$, where $\hat{b}_j = b_j/\beta$ and $\beta = \sum_{j=0}^{k} b_j \neq 0$. If $b_0 = 0$ these multistep methods are called explicit, and if $b_0 \neq 0$ they are called implicit.

One-leg methods were introduced by Dahlquist [1], originally only to facilitate the analysis of linear multistep methods. Subsequently, it was realized that one-leg methods might be useful on their own, not just as an analysis tool. It is known that the conditions for consistency of order p are the same if $p = 1, 2$, but for larger p the one-leg method has to satisfy more order conditions than the corresponding linear multistep method; cf. [2], for instance.

On the other hand, one-leg methods have an advantage over the corresponding linear multistep methods with respect to storage, which is often important for large-scale problems when function evaluations of F are expensive. If, for example, $b_0 = 0$ but $a_k, b_k \neq 0$, then for a step (1.3) with the linear multistep method we need storage of the vectors $u_{n-1}, \ldots, u_{n-k}$ and $F(u_{n-2}), \ldots, F(u_{n-k})$, together with an evaluation of $F(u_{n-1})$. For a step (1.4) with the one-leg method only storage of $u_{n-1}, \ldots, u_{n-k}$ is needed, together with evaluation of $F(v_n)$.

Scope of the paper

In this paper we will first consider the property

$$\|u_n\| \leq \mu \cdot \max_{0 \leq j < k} \|u_j\| \quad \text{for all } n \geq k \text{ and } 0 < \Delta t \leq \gamma\, \tau_0, \tag{1.5}$$

whenever the basic assumption (1.2) is satisfied. Here the factor $\mu \geq 1$ and the stepsize coefficient $\gamma \geq 0$ are determined by the multistep method, and we are interested in having $\gamma > 0$ as large as possible. If (1.5) holds with $\mu = 1$, then this property will be called *monotonicity*. For many interesting methods, this property (1.5) will only hold with some $\mu > 1$, in which case we refer to it as *boundedness*.

It is known, see e.g. [3,4], that the condition for monotonicity for either the linear multistep method (1.3) or the one-leg method (1.4) reads

$$a_j \geq \gamma\, b_j \geq 0 \quad (1 \leq j \leq k). \tag{1.6}$$

This requires that all coefficients of the method are non-negative, which severely restricts the class of methods. It is therefore of interest to study more relaxed properties.

The boundedness property (1.5) with some $\mu \geq 1$, has been studied for linear multistep methods. Sufficient stepsize conditions $\Delta t \leq \gamma \tau_0$ were derived in [5,6] for having (1.5) with arbitrary seminorms under the basic assumption (1.2). More simple conditions were found in [7], and these conditions were shown be necessary as well as sufficient.

In (1.5) the starting values $u_1, \ldots, u_{k-1}$ are arbitrary. In practice these starting values will be computed from the given initial value u_0, for instance by a Runge–Kutta method. For such combinations of multistep methods and Runge–Kutta starting procedures the following monotonicity property

$$\|u_n\| \leq \|u_0\| \quad \text{for all } n \geq 1 \text{ and } 0 < \Delta t \leq \gamma\, \tau_0, \tag{1.7}$$

can still be valid, even if the multistep method itself is not monotone, but only bounded for arbitrary starting values, that is, (1.5) is valid with $\mu > 1$, not with $\mu = 1$.

For some combinations of linear multistep methods and Runge–Kutta starting procedures, the monotonicity property (1.7) was studied in [7], where conditions were derived with arbitrary seminorms and nonnegative sublinear functionals. Earlier, for some two-step methods, sufficient conditions with seminorms were found in [5].

In this paper we will first describe in Section 2 a general framework for having boundedness with arbitrary starting vectors, or monotonicity with starting procedures. This framework, which is valid for general multistep multistage methods, will be based on the approach of Spijker [4] for monotonicity, and of Hundsdorfer, Mozartova and Spijker [8] for boundedness. The results will then be applied to linear multistep methods and one-leg methods. For this, the methods will be formulated in Section 3 in terms of input and output processes, so that the general framework is applicable.

Section 4 contains a discussion for these two classes of multistep methods on the boundedness property (1.5) with arbitrary starting vectors. It will be seen that the maximal stepsize coefficient for boundedness of a one-leg method is the same as for the associated linear multistep method. In view of the close connection between one-leg and linear multistep methods, this result is not surprising.

Next, in Section 5, we will derive conditions for having monotonicity (1.7) for these multistep methods with a starting procedure. It will be seen that then the conditions for linear multistep methods are different from the ones for one-leg methods. In Section 6, a detailed study of the conditions will reveal that, for the class of explicit two-step methods, combinations of the one-leg methods with natural Runge–Kutta starting procedures can give monotonicity with much larger stepsizes than for the corresponding linear multistep methods with the same starting procedures.

2. General framework

We will primarily study ODE systems (1.1) with the basic assumption (1.2) for arbitrary (semi-)norms. In applications, the vector space $\mathbb{V}$ for the ODE systems is usually $\mathbb{R}^M$ with arbitrary $M \geq 1$. Interesting (semi-)norms are for example the maximum norm $\|v\|_\infty = \max_{1\leq j\leq M} |v_j|$, for $v = (v_j) \in \mathbb{R}^M$, and the total variation semi-norm $\|v\|_{\mathrm{TV}} = \sum_{j=1}^M |v_{j-1} - v_j|$, where $v_0 = v_M$ for problems obtained by spatial discretization from partial differential equations with periodic boundary conditions. Norms that are generated by an inner product, such as the Euclidean norm $\|v\|_2 = (\sum_{j=1}^M |v_j|^2)^{1/2}$, are not a focus in this paper; for such norms other boundedness results exist, under more relaxed stepsize conditions, related to G-stability, see [1,2] for example.

To include related properties, such as maximum principles (as in [4]) and positivity preservation (as in [9]), it can be useful to consider more general functionals. Recall that $\varphi : \mathbb{V} \to \mathbb{R}$ is a *convex functional* on $\mathbb{V}$ if

$$\varphi(\lambda v + (1-\lambda)w) \leq \lambda\,\varphi(v) + (1-\lambda)\,\varphi(w) \quad (\text{for all } 0 \leq \lambda \leq 1 \text{ and } v, w \in \mathbb{V}).$$

It is called a *nonnegative sublinear functional* if $\varphi(v + w) \leq \varphi(v) + \varphi(w)$ and $\varphi(cv) = c\varphi(v) \geq 0$ for all real $c \geq 0$ and $v, w \in \mathbb{V}$. It is a *seminorm* if we have in addition $\varphi(-v) = \varphi(v) \geq 0$ for all $v \in \mathbb{V}$. Finally, if it also holds that $\varphi(v) = 0$ only if $v = 0$, then φ is a *norm*.

In the remainder, the notations of [8,4] will be followed as far as possible.

For any given $m \geq 1$, let $e_1, e_2, \ldots, e_m$ stand for the unit basis vectors in $\mathbb{R}^m$, that is, the jth element of e_i equals one if $i = j$ and zero otherwise. The $m \times m$ identity matrix is denoted by I. Furthermore, $e = e_1 + e_2 + \cdots + e_m$ is the vector in $\mathbb{R}^m$ with all components equal to one. If it is necessary to specify the dimension we will denote these unit vectors by $e^{[m]}$; usually the proper dimension will be clear from the context.

If $K = (\kappa_{ij})$ is an $m \times l$ matrix, then $|K|$ stands for the matrix with entries $|\kappa_{ij}|$, and we will denote by the boldface symbol $\boldsymbol{K}$ the associated linear mapping from $\mathbb{V}^l$ to $\mathbb{V}^m$, that is, $\eta = \boldsymbol{K}\xi$ for $\eta = [\eta_i] \in \mathbb{V}^m, \xi = [\xi_i] \in \mathbb{V}^l$ if $\eta_i = \sum_{j=1}^l \kappa_{ij}\xi_j \in \mathbb{V}$ $(1 \leq i \leq m)$. Inequalities for vectors or matrices are to be understood component-wise. In particular, we will use the notation $K \geq 0$ when all entries κ_{ij} of this matrix are non-negative.

2.1. *The generic form*

Application of a multistep method with a fixed number of steps leads to a process of the generic form

$$y_i = \sum_{j=1}^k s_{ij}x_j + \Delta t \sum_{j=1}^m t_{ij}F(y_j) \quad (1 \leq i \leq m), \tag{2.1}$$

producing the output vectors $y_1, y_2, \ldots, y_m$ from the input data $x_1, \ldots, x_k$ in $\mathbb{V}$. Typically, the set of output vectors will contain approximations u_n, $n \geq k$, whereas the input vectors x_j will consist of linear combinations of the starting vectors $u_0, u_1, \ldots, u_{k-1}$ and their function values $F(u_0), F(u_1), \ldots, F(u_{k-1})$.

Let $y = [y_i] \in \mathbb{V}^m$, $x = [x_i] \in \mathbb{V}^k$, and denote $\boldsymbol{F}(y) = [F(y_i)] \in \mathbb{V}^m$. The coefficient matrices for the process (2.1) are $S = (s_{ij}) \in \mathbb{R}^{m\times k}$ and $T = (t_{ij}) \in \mathbb{R}^{m\times m}$. With the above notations, the generic process (2.1) can be written in a compact way as

$$y = \boldsymbol{S}\,x + \Delta t\boldsymbol{T}\,\boldsymbol{F}(y). \tag{2.2}$$

Let $[S\ T]$ be the $m \times (k+m)$ matrix whose first k columns equal those of S and whose last m columns are equal to those of T. As we will see, the generic processes that are generated by the multistep methods will be such that all rows of S are not zero and all rows of $[S\ T]$ are different from each other. With unit basis vectors $e_i \in \mathbb{R}^m$, $1 \leq i \leq m$, this can be expressed as

$$e_i^T S \neq 0 \quad \text{for all } i, \tag{2.3a}$$

$$e_i^T[\,S\ T\,] \neq e_j^T[\,S\ T\,] \quad \text{if } i \neq j. \tag{2.3b}$$

It is obvious that two identical rows in $[S\ T]$ lead to two output vectors y_i and y_j, $i \neq j$, with $y_i = y_j$ for any function F and arbitrary input vectors x_l. This was called reducibility in [4]. In this paper we will refer to such a scheme as *reducible* (in the sense of Spijker), and a scheme for which all rows of $[S\ T]$ are different from each other is called *irreducible* (in the sense of Spijker).

2.2. Boundedness for arbitrary starting vectors

If $\gamma > 0$ is such that $I + \gamma T$ is not singular, we can write the process (2.2) also in the form

$$y = \boldsymbol{R}\,x + \boldsymbol{P}\left(y + \frac{\Delta t}{\gamma}\boldsymbol{F}(y)\right), \tag{2.4}$$

where $R \in \mathbb{R}^{m\times k}$ and $P \in \mathbb{R}^{m\times m}$ are given by

$$R = (I+\gamma T)^{-1}S, \qquad P = (I+\gamma T)^{-1}\gamma T. \tag{2.5}$$

The number of steps with the multistep methods will be arbitrary, so the number m will allowed to be arbitrarily large as well. Consider, for given vector space $\mathbb{V}$ and seminorm $\|\cdot\|$, the boundedness property

$$\begin{cases} \max\limits_{1\le i\le m}\|y_i\| \le \mu\cdot\max\limits_{1\le j\le k}\|x_j\| \quad \text{whenever (1.2) is valid, } \Delta t \le \gamma\,\tau_0, \text{ and } x, y \text{ satisfy (2.2), } m \ge 1, \end{cases} \tag{2.6}$$

with a stepsize coefficient $\gamma > 0$ and boundedness factor $\mu \ge 1$. Note that this bound holds uniformly for all initial value problems (1.1) under the basic assumption (1.2) with given $\tau_0 > 0$.

For any $m \times m$ matrix $K = (\kappa_{ij})$, let spr(K) be the spectral radius of K, and let $\|K\|_\infty = \max_i \sum_j |\kappa_{ij}|$ stand for the induced maximum norm of K. From Theorem 2.4 in [8] we have the following result:

Theorem 2.1. *Assume $I+\gamma T$ is not singular, and spr$(|P|) < 1$. Then, for any vector space $\mathbb{V}$ with seminorm $\|\cdot\|$, the boundedness property (2.6) is valid provided that*

$$\|(I-|P|)^{-1}|R|\,\|_\infty \le \mu \quad \text{for all } m. \tag{2.7}$$

Moreover, if (2.3) holds, then the condition (2.7) is necessary for (2.6) to be valid for the class of spaces $\mathbb{V} = \mathbb{R}^M$, $M \ge 1$, with the maximum norm.

For the multistep methods considered in this paper, the assumptions that $I + \gamma T$ is not singular and spr$(|P|) < 1$ will hold trivially. Furthermore we note that boundedness as in (2.6), that is, boundedness with respect to the input vectors x_j, can be considered for functionals that are more general than seminorms. However this does not lead to boundedness results with respect to the starting vectors $u_0, \ldots, u_{k-1}$, as in (1.5), unless additional constraints on the methods are imposed. For example, as pointed out in [7], for linear multistep methods that would lead again to the very strict conditions (1.6).

2.3. Monotonicity with starting procedures

Instead of arbitrary starting vectors $u_0, u_1, \ldots, u_{k-1}$ for the multistep methods, we will consider Runge–Kutta starting procedures to generate these vectors from u_0. Assume this starting procedure produces a vector $w = [w_i] \in \mathbb{V}^{m_0}$, $m_0 \ge k$, where $u_j = w_{i_j}$ for $j = 0, 1, \ldots, k-1$ and the remaining w_i are internal stage vectors of the starting procedure.

The whole starting procedure, which may consist of several steps of a Runge–Kutta method, can be conveniently written as a single step

$$w = \boldsymbol{e_0}u_0 + \Delta t\boldsymbol{K_0}\boldsymbol{F}(w), \tag{2.8}$$

where $e_0 = e^{[m_0]} = (1,\ldots,1)^T \in \mathbb{R}^{m_0}$, and $K_0 \in \mathbb{R}^{m_0\times m_0}$ is the coefficient matrix of this Runge–Kutta starting procedure; examples are given in Section 6. As is well known, see e.g. [4], the conditions of Kraaijevanger [10]

$$(I+\gamma K_0)^{-1}e_0 \ge 0, \qquad (I+\gamma K_0)^{-1}\gamma K_0 \ge 0 \tag{2.9}$$

guarantee that the starting procedure itself is monotone with the stepsize coefficient γ, that is, $\|w_j\| \le \|u_0\|(1 \le j \le m_0)$ whenever (1.2) is valid, $\Delta t \le \gamma\,\tau_0$, for any vector space $\mathbb{V}$ and convex functional $\|\cdot\|$.

The above Runge–Kutta starting procedure will give an input vector of the form

$$x = \boldsymbol{S_0}u_0 + \Delta t\boldsymbol{T_0}\boldsymbol{F}(w) \tag{2.10}$$

with $S_0 \in \mathbb{R}^{k\times 1}$, $T_0 \in \mathbb{R}^{k\times m_0}$. The total scheme, consisting of the multistep method and starting procedure can therefore be written as

$$\begin{cases} w = \boldsymbol{e_0}u_0 + \Delta t\boldsymbol{K_0}F(w), \\ y = \boldsymbol{SS_0}u_0 + \Delta t\boldsymbol{ST_0}\boldsymbol{F}(w) + \Delta t\boldsymbol{TF}(y). \end{cases} \tag{2.11}$$

For the multistep methods considered in this paper, the output vectors y_i will be consistent approximations to $u(t_n)$ for some $n \ge 0$. By considering $F \equiv 0$ it then follows that

$$S\,S_0 = e, \tag{2.12}$$

where $e = (1, 1, \ldots, 1)^T \in \mathbb{R}^m$. Therefore, for any fixed m, the total scheme (2.11) is then just an $(m_0 + m)$-stage Runge–Kutta method, with an $(m_0 + m) \times (m_0 + m)$ coefficient matrix

$$K = \begin{pmatrix} K_0 & O \\ S\,T_0 & T \end{pmatrix}. \tag{2.13}$$

To obtain monotonicity results we substitute $\gamma(v+\frac{\Delta t}{\gamma}\boldsymbol{F}(v))-\gamma v$ for the terms $\Delta t\boldsymbol{F}(v)$ in (2.11). This gives, after a little manipulation,

$$\begin{cases} w = (\boldsymbol{I}+\gamma\boldsymbol{K_0})^{-1}\boldsymbol{e_0}\,u_0 + (\boldsymbol{I}+\gamma\boldsymbol{K_0})^{-1}\gamma\boldsymbol{K_0}\left(w+\dfrac{\Delta t}{\gamma}\boldsymbol{F}(w)\right), \\ y = \boldsymbol{R}\,\boldsymbol{R_0}u_0 + \boldsymbol{R}\,\boldsymbol{P_0}\left(w+\dfrac{\Delta t}{\gamma}\boldsymbol{F}(w)\right) + \boldsymbol{P}\left(y+\dfrac{\Delta t}{\gamma}\boldsymbol{F}(y)\right), \end{cases} \tag{2.14}$$

with matrices R, P as before and

$$R_0 = S_0 - \gamma T_0(I+\gamma K_0)^{-1}e_0, \qquad P_0 = \gamma T_0(I+\gamma K_0)^{-1}. \tag{2.15}$$

These expressions arise in a natural way by writing $x = \boldsymbol{R_0}u_0 + \boldsymbol{P_0}\big(w+\frac{\Delta t}{\gamma}\boldsymbol{F}(w)\big)$, together with relation (2.4).

We will consider, for m arbitrarily large, and a given vector space $\mathbb{V}$ with convex functional $\|\cdot\|$, the following monotonicity property with stepsize coefficient $\gamma > 0$:

$$\Big\{ \max_{1\le n\le m} \|y_n\| \le \|u_0\| \quad \text{whenever (1.2) is valid, } \Delta t \le \gamma\,\tau_0, \text{ and } x, y \text{ satisfy (2.2), (2.8), (2.10), } m \ge 1. \tag{2.16}$$

As we will see next, this type of monotonicity of the multistep methods with starting procedures will hold under the condition

$$R\,R_0 \ge 0, \qquad R\,P_0 \ge 0, \qquad P \ge 0 \quad (\text{for all } m \ge 1), \tag{2.17}$$

where R, P are defined by (2.5). The following result is similar to Theorem 4.4 in [7], where sufficiency of condition (2.17) was proven for nonnegative sublinear functionals.

Theorem 2.2. *Assume $I+\gamma T$ is not singular, $spr(|P|) < 1$, and the starting procedure is such that (2.9) holds. Let $\|\cdot\|$ be a convex functional on a vector space $\mathbb{V}$. Then (2.17) implies the monotonicity property (2.16). Moreover, if all rows of the matrix K in (2.13) are different from each other and (2.12) holds, then (2.17) is also necessary for this monotonicity property to hold for the class of spaces $\mathbb{V} = \mathbb{R}^M$, $M \ge 1$, with the maximum norm.*

Proof. Assume (2.9). Let $\eta = (\eta_i) \in \mathbb{R}^m$ with $\eta_i = \|y_i\|$. Since we have $\|w_j + \frac{\Delta t}{\gamma}F(w_j)\| \le \|w_j\| \le \|u_0\|$ for $1 \le j \le m_0$, it follows from the second equality in (2.14) that

$$\eta \le R\,R_0 \cdot \|u_0\| + R\,P_0\,e_0 \cdot \|u_0\| + P\,\eta.$$

In case $F \equiv 0$, all vectors w_j, y_i will be equal to u_0, which implies $e = R\,R_0 1 + R\,P_0 e_0 + Pe$. Hence

$$(I-P)\eta \le (I-P)\,e \cdot \|u_0\|.$$

Since $\mathrm{spr}(P) < 1$, we have $(I-P)^{-1} = \sum_{j\ge 0} P^j \ge 0$, and therefore $\eta \le e \cdot \|u_0\|$, that is, $\|y_i\| \le \|u_0\|$ for $1 \le i \le m$.

If we have $S\,S_0 = e$, then the total scheme (2.11) can be viewed as a step with a Runge–Kutta method with coefficient matrix K, and if all rows of K are different from each other, then this method is irreducible (in the sense of Spijker). The necessity of (2.17) therefore follows from [4, Theorem 2.7]. □

We note that, because (2.14) has the form of a Runge–Kutta coefficient matrix, the sufficiency of (2.17) for monotonicity could also be derived from the results in [11,12] if $S\,S_0 = e$.

3. Reformulations of the linear multistep and one-leg methods

For the linear multistep methods (1.3) and the one-leg methods (1.4), it will be assumed throughout this paper that

$$\sum_{j=1}^{k} a_j = 1, \qquad \sum_{j=1}^{k} j\,a_j = \sum_{j=0}^{k} b_j > 0, \qquad b_0 \ge 0. \tag{3.1a}$$

Here, the two equalities give the conditions for consistency of order one. Having $\sum_{j=0}^{k} b_j > 0$ is then necessary for zero-stability of the methods. The assumption $b_0 \ge 0$ will be convenient, and it holds for all well-known methods. Furthermore, for the generating polynomials $\rho(\zeta) = \zeta^k - \sum_{j=1}^{k} a_j\zeta^{k-j}$ and $\sigma(\zeta) = \sum_{j=0}^{k} b_j\zeta^{k-j}$ it will be assumed that

$$\rho(\zeta) \text{ and } \sigma(\zeta) \text{ have no common factor.} \tag{3.1b}$$

Methods that do not satisfy this last condition are said to be reducible (in the sense of Dahlquist), and these are essentially equivalent to a $(k-1)$-step method.

As before, $e_1, e_2, \ldots, e_m$ stand for the unit basis vectors in $\mathbb{R}^m$, and I is the $m \times m$ identity matrix. Further, $E = [e_2, \ldots, e_m, 0]$ will denote the $m \times m$ backward shift matrix, that is, all entries of E are zero except the entries on the first lower diagonal, which are 1.

Let $A, B \in \mathbb{R}^{m\times m}$ be defined by

$$A = \sum_{j=1}^{k} a_j E^j, \qquad B = \sum_{j=0}^{k} b_j E^j, \tag{3.2}$$

where $E^0 = I$. These are lower triangular Toeplitz matrices, with coefficients a_j, b_j on the jth lower diagonal. For $m \geq k$ we also introduce $J = [e_1, \ldots, e_k] \in \mathbb{R}^{m\times k}$, containing the first k columns of the identity matrix I. To make the notations fitting for any $m \geq 1$, we can define $J = [e_1, \ldots, e_m, O]$ for $1 \leq m < k$, with O being the $m \times (k-m)$ zero matrix. Finally, $A_0, B_0 \in \mathbb{R}^{k\times k}$ are given by

$$A_0 = \begin{pmatrix} a_k & \cdots & a_2 & a_1 \\ & a_k & & a_2 \\ & & \ddots & \vdots \\ & & & a_k \end{pmatrix}, \qquad B_0 = \begin{pmatrix} b_k & \cdots & b_2 & b_1 \\ & b_k & & b_2 \\ & & \ddots & \vdots \\ & & & b_k \end{pmatrix}. \tag{3.3}$$

3.1. Formulation of linear multistep methods with input vectors

In order to apply the general results on boundedness and monotonicity we will formulate the multistep methods in terms of input and output vectors. For the linear multistep methods this was done in [7], and this is repeated here to keep the paper self-sufficient and to make comparison with the subsequent results for the one-leg methods easier.

The output vectors of the linear multistep scheme (1.3) are $y_n = u_{k-1+n}$, $n \geq 1$. The starting values $u_0, u_1, \ldots, u_{k-1}$ will enter the scheme in the first k steps in the combinations

$$x_l = \sum_{j=l}^{k} a_j u_{k-1+l-j} + \Delta t \sum_{j=l}^{k} b_j F(u_{k-1+l-j}) \quad \text{for } 1 \leq l \leq k. \tag{3.4}$$

The multistep scheme (1.3) can now be written as

$$y_n = x_n + \sum_{j=1}^{n-1} a_j y_{n-j} + \Delta t \sum_{j=0}^{n-1} b_j F(y_{n-j}) \quad \text{for } 1 \leq n \leq k, \tag{3.5a}$$

$$y_n = \sum_{j=1}^{k} a_j y_{n-j} + \Delta t \sum_{j=0}^{k} b_j F(y_{n-j}) \quad \text{for } n > k, \tag{3.5b}$$

where the starting values are contained within the source terms in the first k steps. The vectors $x_1, \ldots, x_k \in \mathbb{V}$ are the input vectors for the scheme.

Considering m steps of the multistep scheme, $m \geq k$, leading to (3.5) with $n = 1, 2, \ldots, m$, the resulting scheme can now be written compactly as

$$y = \boldsymbol{J}x + \boldsymbol{A}y + \Delta t \boldsymbol{BF}(y). \tag{3.6}$$

Clearly this is of the form (2.2) with

$$S = (I - A)^{-1} J, \qquad T = (I - A)^{-1} B, \tag{3.7}$$

which gives (2.4) with

$$R = (I - A + \gamma B)^{-1} J, \qquad P = (I - A + \gamma B)^{-1} \gamma B. \tag{3.8}$$

If we consider the problem (1.1) with $F \equiv \beta$ and solution $u(t) = \alpha + \beta t$, then exact starting values $u_j = u(t_j)$ $(0 \leq j < k)$ will give $u_n = u(t_n)$ (for all $n \geq k$) because of consistency of the methods. From this it is easily seen that (2.3) holds, and therefore the scheme is irreducible (in the sense of Spijker). It should be remarked that this is not directly related to the Dahlquist irreducibility condition (3.1b) for the multistep methods.

The matrix $I - A + \gamma B$ is invertible for any $\gamma > 0$, because $b_0 \geq 0$. The matrix P is again a lower triangular Toeplitz matrix, and it has the entry $\pi_0 = \gamma b_0/(1 + \gamma b_0) \in [0, 1)$ on the main diagonal. The spectral radius $\mathrm{spr}(|P|)$ of the matrix $|P|$ is therefore less than one.

The coefficients of the matrices R and P are easily found recursively. Let $\rho_j = 0$ for $j < 0$. If we set $(I - A + \gamma B)^{-1} = \sum_{n\geq 0} \rho_n E^n$ and $P = \sum_{n\geq 0} \pi_n E^n$, then these Toeplitz coefficients ρ_n, π_n are given by $\rho_0 = 1/(1 + \gamma b_0)$ and

$$\rho_n = \sum_{j=1}^{k} a_j \rho_{n-j} - \gamma \sum_{j=0}^{k} b_j \rho_{n-j} \quad \text{for } n \geq 1, \tag{3.9a}$$

$$\pi_n = \gamma \sum_{j=0}^{k} b_j \rho_{n-j} \quad \text{for } n \geq 0. \tag{3.9b}$$

An inequality of the type $Rv \geq 0$ for all $m \geq 1$, with a vector $v = (v_0, v_1, \dots, v_k)^T$, is now equivalent to having $\sum_{j=0}^{k} v_j \rho_{n-j} \geq 0$ for all $n \geq 0$.

3.2. Formulation of one-leg methods with input vectors

To derive results for one-leg methods, we will proceed in a similar way, using an input–output formulation. To distinguish the arising vectors and associated matrices for the one-leg methods from those of the linear multistep methods, we will use the upper bar symbol for the one-leg vectors and matrices. In particular, the matrices S, T, R, P will be as in (3.7) and (3.8) for the linear multistep methods, and the corresponding matrices for the one-leg methods will be denoted by $\bar{S}, \bar{T}, \bar{R}$ and $\bar{P}$. Likewise, in the generic form (2.1), (2.2) the dimensions m, k will now read $\bar{m}$ and $\bar{k}$.

Consider m steps of the one-leg method (1.4), and let $\bar{y} = [\bar{y}_i] \in \mathbb{V}^{\bar{m}}$, $\bar{m} = 2m$, with

$$\bar{y}_i = u_{k-1+i}, \qquad \bar{y}_{m+i} = v_{k-1+i} \quad \text{for } 1 \leq i \leq m. \tag{3.10}$$

As input we have $\bar{x} = [\bar{x}_j] \in \mathbb{V}^{\bar{k}}$, $\bar{k} = 2k$, with

$$\bar{x}_i = \sum_{j=i}^{k} a_j u_{k-1+i-j}, \qquad \bar{x}_{i+k} = \sum_{j=i}^{k} \hat{b}_j u_{k-1+i-j} \quad i = 1, \dots, k. \tag{3.11}$$

Let $J \in \mathbb{R}^{m\times k}$ and $A, B \in \mathbb{R}^{m\times m}$ be as before. Then the m steps of the one-leg method can be written as

$$\bar{y} = \bar{\boldsymbol{J}}\bar{x} + \bar{\boldsymbol{A}}\bar{y} + \Delta t\bar{\boldsymbol{B}}\boldsymbol{F}(\bar{y}), \tag{3.12}$$

where $\bar{J} \in \mathbb{R}^{\bar{m}\times\bar{k}}$ and $\bar{A}, \bar{B} \in \mathbb{R}^{\bar{m}\times\bar{m}}$ are given by

$$\bar{J} = \begin{pmatrix} J & 0 \\ 0 & J \end{pmatrix}, \qquad \bar{A} = \begin{pmatrix} A & O \\ \frac{1}{\beta}B & O \end{pmatrix}, \qquad \bar{B} = \begin{pmatrix} O & \beta I \\ O & O \end{pmatrix}, \tag{3.13}$$

with zero matrices $O \in \mathbb{R}^{m\times m}$ and $0 \in \mathbb{R}^{m\times k}$. We can rewrite (3.12) in the following form, comparable to (2.2),

$$\bar{y} = \bar{\boldsymbol{S}}\bar{x} + \Delta t\, \bar{\boldsymbol{T}}\boldsymbol{F}(\bar{y}) \tag{3.14}$$

with $\bar{S} \in \mathbb{R}^{\bar{m}\times\bar{k}}$ and $\bar{T} \in \mathbb{R}^{\bar{m}\times\bar{m}}$ defined by

$$\bar{S} = (\bar{I} - \bar{A})^{-1}\bar{J}, \qquad \bar{T} = (\bar{I} - \bar{A})^{-1}\bar{B}, \tag{3.15}$$

with $\bar{m} \times \bar{m}$ identity matrix $\bar{I}$. Working out these matrices, in terms of $S = (I - A)^{-1}J$ and $T = (I - A)^{-1}B$, gives

$$\bar{S} = \begin{pmatrix} S & 0 \\ \frac{1}{\beta}BS & J \end{pmatrix}, \qquad \bar{T} = \begin{pmatrix} O & \beta(I - A)^{-1} \\ O & T \end{pmatrix}. \tag{3.16}$$

This can be further rewritten, for example with $BS = TJ$.

If there is only one index j, $0 \leq j \leq k$, with $b_j \neq 0$, then the one-leg method is the same as the linear multistep method. For genuine one-leg methods, with $b_j \neq 0$ for at least two indices j, it will now be shown that the scheme is irreducible (in the sense of Spijker).

Lemma 3.1. *Suppose $b_j \neq 0$ for two or more indices $0 \leq j \leq k$. Then all rows of $[\bar{S}\ \bar{T}]$ are different from each other.*

Proof. The u_n, v_n are consistent approximations to $u(t_n), u(\bar{t}_n)$, respectively, with $\bar{t}_n = \sum_{j=0}^{k} \hat{b}_j t_{n-j}$. Using the same arguments as for the linear multistep methods, it follows that the first m rows of $[\bar{S}\ \bar{T}]$ are different from each other, and the same holds for the last m rows.

It remains to show that none of the first m rows can be equal to any of the last m rows. For this it is sufficient to show that the $m \times (m + k)$ matrices

$$C_1 = \big(0\beta(I - A)^{-1}\big) \quad \text{and} \quad C_2 = \big(JT\big)$$

have no common rows. Because of the entry J in C_2, it is clear that the first k rows of C_2 cannot coincide with any of the rows of C_1.

Since the lower triangular Toeplitz matrices B and $I - A$ commute, we have $T = B(I - A)^{-1}$. Equal rows of $\beta(I - A)^{-1}$ and T can therefore only happen if $\beta e_i^T(I - A)^{-1} = e_j^T B(I - A)^{-1}$ with $i \geq j$. But then $\beta e_i^T = e_j^T B$, which implies that either $i, j \leq k$ or that only the coefficient b_{i-j} is not zero.

Consequently, if the matrices C_1 and C_2 have common rows, then there is only one index j, $0 \leq j \leq k$, with $b_j \neq 0$. □

As for the linear multistep methods [7], it follows from consistency of the one-leg method that $\bar{S}$ has no zero rows. The conditions (2.3) are therefore fulfilled with the matrices $\bar{S}, \bar{T}$ instead of S, T, provided the one-leg method is not a linear multistep method.

Since $I + \gamma T$ is invertible for any $\gamma > 0$, the same holds for $I + \gamma \bar{T}$. From (3.12) we therefore obtain the transformed form, comparable to (2.4),

$$\bar{y} = \bar{\boldsymbol{R}}\bar{x} + \bar{\boldsymbol{P}}\Big(\bar{y} + \frac{\Delta t}{\gamma}\boldsymbol{F}(\bar{y})\Big), \tag{3.17}$$

where $\bar{R} \in \mathbb{R}^{\bar{m}\times\bar{k}}$ and $\bar{P} \in \mathbb{R}^{\bar{m}\times\bar{m}}$ are given by

$$\bar{R} = (\bar{I} - \bar{A} + \gamma\bar{B})^{-1}\bar{J}, \qquad \bar{P} = (\bar{I} - \bar{A} + \gamma\bar{B})^{-1}\gamma\bar{B}. \tag{3.18}$$

These matrices $\bar{R}$ and $\bar{P}$ in (3.18) for the one-leg method will be expressed in terms of the $m \times m$ matrices R and P for the linear multistep method, as given by (3.8). Let us here denote $L = (I - A + \gamma B)^{-1}$. Then it is found that

$$(\bar{I} - \bar{A} + \gamma\bar{B})^{-1} = \begin{pmatrix} I - A & \beta\gamma I \\ -\frac{1}{\beta}B & I \end{pmatrix}^{-1} = \begin{pmatrix} L & -\beta\gamma L \\ \frac{1}{\beta}BL & (I - A)L \end{pmatrix}.$$

The blocks consist of products of Toeplitz matrices that commute. Since $LJ = R$, it follows that

$$\bar{R} = \begin{pmatrix} R & -\beta\gamma R \\ \frac{1}{\beta}BR & (I - A)R \end{pmatrix}, \qquad \bar{P} = \begin{pmatrix} O & \beta\gamma L \\ O & P \end{pmatrix}. \tag{3.19}$$

Note that since $\mathrm{spr}(|P|) < 1$, we also have $\mathrm{spr}(|\bar{P}|) < 1$. Furthermore we see that $\bar{P} \geq 0$ iff $P \geq 0$ and $R \geq 0$.

4. Boundedness for arbitrary starting vectors

In this section conditions are given for boundedness of the linear multistep methods (1.3) and the one-leg methods (1.4). It will always be assumed that (3.1) is satisfied. Furthermore, in this section, boundedness is understood in the sense of property (2.6) for any seminorm, with some $\mu \geq 1$, and with y_i, x_j, m, k replaced by $\bar{y}_i$, $\bar{x}_j$, $\bar{m} = 2m$ and $\bar{k} = 2k$, respectively, for the one-leg methods.

To formulate the results we will use some standard linear stability concepts for multistep methods, as given in [13,14,2] for instance. We denote the *stability region* of the methods by $\mathcal{S}$, and its interior by $\mathrm{int}(\mathcal{S})$. If $0 \in \mathcal{S}$ the method is said to be *zero-stable*.

It was shown in [7] that for a zero-stable linear multistep method satisfying (3.1), the condition

$$-\gamma \in \mathrm{int}(\mathcal{S}), \qquad P \geq 0 \quad (\text{for all } m), \tag{4.1}$$

is necessary and sufficient for the boundedness property (2.6) to hold with some $\mu \geq 1$. As we will see, the same result is valid for the one-leg methods. This can be shown using relations between a linear multistep method and the corresponding one-leg method, as given in [1] or [2, Sect. V.6]. It is also possible to prove this directly from Theorem 2.1, which will be done here.

For this, we consider the matrix

$$\bar{M} = (\bar{I} - |\bar{P}|)^{-1}|\bar{R}| = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \tag{4.2}$$

with matrices $\bar{P}$, $\bar{R}$ for the one-leg method given by (3.19). Using the fact that $(I - A)R = (I - P)J$ and $BR = \frac{1}{\gamma}PJ$, it follows by some calculations that the blocks $M_{ij} \in \mathbb{R}^{m\times k}$ can be written as

$$\begin{cases} M_{11} = (I - |P|)^{-1}|R|, \\ M_{12} = \beta\gamma\left((I - |P|)^{-1}|I - P| + I\right)|R|, \\ M_{21} = \frac{1}{\beta\gamma}(I - |P|)^{-1}|P|J, \\ M_{22} = (I - |P|)^{-1}|I - P|J. \end{cases} \tag{4.3}$$

According to Theorem 2.1, boundedness of the one-leg method is equivalent to having a bound $\|\bar{M}\|_\infty \leq \mu$ uniformly for $m \geq 1$. By considering the M_{11} block, we therefore see that boundedness of the one-leg method implies boundedness of the linear multistep method.

On the other hand, suppose the linear multistep method is bounded. Then we know that $P \geq 0$. Zero-stability implies that $\|S\|_\infty$ is bounded uniformly in m. Therefore, the maximum norms of the matrices R and

$$(I - |P|)^{-1}J = (I - A + \gamma B)S$$

are also bounded uniformly in m. Using the relations $(I - P)^{-1}P = (I - P)^{-1} - I$ and $(I - P)^{-1}|I - P| \leq (I - P)^{-1}(I + P)$, it follows that the maximum norms of all the blocks M_{ij} in (4.3) are bounded uniformly in m.

In conclusion, we have the following result on boundedness for our multistep methods (1.3) and (1.4).

Theorem 4.1. *Consider a one-leg or linear multistep method, satisfying (3.1). Assume the method is zero-stable, and let $\gamma > 0$. Then there is a $\mu \geq 1$ such that the boundedness property (2.6) is valid for any vector space $\mathbb{V}$ and seminorm $\|\cdot\|$ if and only if condition (4.1) holds.*

The above result, with equal stepsize coefficients for a one-leg method and its linear multistep counterpart, is hardly surprising, given the close connection between one-leg methods and linear multistep methods. We will see, however, that the allowable stepsizes for one-leg and linear multistep methods can be very different if we require monotonicity with starting procedures.

The boundedness property (2.6) is expressed in terms of the input vectors x_j. However, with seminorms this is easily translated into boundedness with respect to the starting values $u_0, \ldots, u_{k-1}$ as in (1.5), and likewise reversely; see [7, Sect. 4.2].

5. Monotonicity with starting procedures

5.1. Linear multistep methods with starting procedures

Consider the Runge–Kutta starting procedure (2.8), producing the vector $w = [w_i] \in \mathbb{V}^{m_0}$. Let $J_0 \in \mathbb{R}^{k\times m_0}$ be the matrix, with columns $e_1, \ldots, e_k \in \mathbb{R}^k$ interceded by zero columns, which selects those components w_i that correspond to a starting value u_j of the multistep method, that is,

$$\boldsymbol{J_0} w = (u_0, \ldots, u_{k-1})^T, \qquad \boldsymbol{J_0}\boldsymbol{F}(w) = (F(u_0), \ldots, F(u_{k-1}))^T.$$

Further, let A_0, B_0 be as in (3.3). Then it follows from (3.4) that

$$x = \boldsymbol{A_0}\boldsymbol{J_0} w + \Delta t \boldsymbol{B_0}\boldsymbol{J_0}\boldsymbol{F}(w). \tag{5.1}$$

This gives the representation $x = \boldsymbol{S_0} u_0 + \Delta t \boldsymbol{T_0}\boldsymbol{F}(w)$, as in (2.10), with matrices $S_0 \in \mathbb{R}^{k\times 1}$, $T_0 \in \mathbb{R}^{k\times m_0}$ given by

$$S_0 = A_0 J_0 e_0, \qquad T_0 = A_0 J_0 K_0 + B_0 J_0. \tag{5.2}$$

To obtain monotonicity results, the scheme is now written in the form (2.14) with matrices R, P given by (3.8) and with $R_0 \in \mathbb{R}^{k\times 1}$, $P_0 \in \mathbb{R}^{k\times m_0}$ given by

$$R_0 = A_0 J_0 e_0 - P_0 e_0, \qquad P_0 = \gamma (A_0 J_0 K_0 + B_0 J_0)(I + \gamma K_0)^{-1}. \tag{5.3}$$

These matrices R_0, P_0 can be further rewritten as

$$R_0 = (A_0 - \gamma B_0) J_0 (I + \gamma K_0)^{-1} e_0, \tag{5.4a}$$

$$P_0 = (A_0 - \gamma B_0) J_0 (I + \gamma K_0)^{-1} \gamma K_0 + \gamma B_0 J_0. \tag{5.4b}$$

Consequently, the conditions (2.17) for monotonicity of the total scheme are:

$$\begin{cases} P \geq 0, \\ R\,(A_0 - \gamma B_0) J_0 (I + \gamma K_0)^{-1} e_0 \geq 0, \\ R\left((A_0 - \gamma B_0) J_0 (I + \gamma K_0)^{-1} \gamma K_0 + \gamma B_0 J_0\right) \geq 0. \end{cases} \tag{5.5}$$

Note that the first inequality, $P \geq 0$, that appears here is the essential condition for boundedness of the linear multistep method (cf. (4.1)).

5.2. One-leg methods with starting procedures

As for the linear multistep methods, we now consider the formulas that are obtained if a Runge–Kutta starting procedure is used for a one-leg method. It is assumed, as before, that this starting procedure is of the form $w = \boldsymbol{e_0} u_0 + \Delta t \boldsymbol{K_0}\boldsymbol{F}(w)$ and $\boldsymbol{J_0} w = (u_0, \ldots, u_{k-1})^T$. From (3.11) it is then seen that

$$\bar{x} = \begin{pmatrix} \boldsymbol{A_0} \\ \frac{1}{\beta}\boldsymbol{B_0} \end{pmatrix} \boldsymbol{J_0} w = \begin{pmatrix} \boldsymbol{A_0} \\ \frac{1}{\beta}\boldsymbol{B_0} \end{pmatrix} \boldsymbol{J_0}\big(\boldsymbol{e_0} u_0 + \Delta t \boldsymbol{K_0}\boldsymbol{F}(w)\big).$$

This can be written as

$$\bar{x} = \bar{\boldsymbol{S}}_{\boldsymbol{0}} u_0 + \Delta t \bar{\boldsymbol{T}}_{\boldsymbol{0}} \boldsymbol{F}(w), \tag{5.6}$$

with

$$\bar{S}_0 = \begin{pmatrix} A_0 \\ \frac{1}{\beta} B_0 \end{pmatrix} J_0 e_0, \qquad \bar{T}_0 = \begin{pmatrix} A_0 \\ \frac{1}{\beta} B_0 \end{pmatrix} J_0 K_0. \tag{5.7}$$

For any fixed number of steps m, the total scheme can be viewed as an $(m_0 + \bar{m})$-stage Runge–Kutta method, $\bar{m} = 2m$, with an $(m_0 + \bar{m}) \times (m_0 + \bar{m})$ coefficient matrix

$$\bar{K} = \begin{pmatrix} K_0 & O \\ \bar{S}\,\bar{T}_0 & \bar{T} \end{pmatrix}. \tag{5.8}$$

To derive monotonicity results we substitute, similar as before, $\gamma(v + \frac{\Delta t}{\gamma}\boldsymbol{F}(v)) - \gamma v$ for all terms $\Delta t\boldsymbol{F}(v)$. This gives, as in (2.14), the total scheme in the form

$$\begin{cases} w = (I + \gamma\boldsymbol{K_0})^{-1}\boldsymbol{e_0}\, u_0 + (I + \gamma\boldsymbol{K_0})^{-1}\gamma\boldsymbol{K_0}\Big(w + \dfrac{\Delta t}{\gamma}\boldsymbol{F}(w)\Big), \\ \bar{y} = \bar{\boldsymbol{R}}\,\bar{\boldsymbol{R}}_{\boldsymbol{0}} u_0 + \bar{\boldsymbol{R}}\,\bar{\boldsymbol{P}}_{\boldsymbol{0}}\Big(w + \dfrac{\Delta t}{\gamma}\boldsymbol{F}(w)\Big) + \bar{\boldsymbol{P}}\Big(\bar{y} + \dfrac{\Delta t}{\gamma}\boldsymbol{F}(\bar{y})\Big), \end{cases}$$

with matrices $\bar{R}$, $\bar{P}$ given by (3.19) and

$$\bar{R}_0 = \bar{S}_0 - \bar{P}_0 e_0, \qquad \bar{P}_0 = \gamma\bar{T}_0(I + \gamma K_0)^{-1}. \tag{5.9}$$

Inserting the expressions (5.7) into (5.9) gives

$$\bar{R}_0 = \begin{pmatrix} A_0 \\ \frac{1}{\beta}B_0 \end{pmatrix} J_0(I + \gamma K_0)^{-1}e_0, \qquad \bar{P}_0 = \begin{pmatrix} A_0 \\ \frac{1}{\beta}B_0 \end{pmatrix} J_0(I + \gamma K_0)^{-1}\gamma K_0. \tag{5.10}$$

To compare the occurring matrices in the monotonicity requirement for the one-leg method with those of the linear multistep method, we note that

$$\bar{R}\begin{pmatrix} A_0 \\ \frac{1}{\beta}B_0 \end{pmatrix} = \begin{pmatrix} R(A_0 - \gamma B_0) \\ \frac{1}{\beta}(BRA_0 + (I - A)RB_0) \end{pmatrix} = \begin{pmatrix} R(A_0 - \gamma B_0) \\ \frac{1}{\beta\gamma}(PJ(A_0 - \gamma B_0) + \gamma JB_0) \end{pmatrix}.$$

The conditions for monotonicity of the total scheme, $\bar{R}\bar{R}_0 \geq 0$, $\bar{R}\bar{P}_0 \geq 0$ and $\bar{P} \geq 0$, can therefore be written as

$$\begin{cases} P \geq 0, \qquad R \geq 0, \\ R\,(A_0 - \gamma B_0)J_0(I + \gamma K_0)^{-1}e_0 \geq 0, \\ R\,(A_0 - \gamma B_0)J_0(I + \gamma K_0)^{-1}\gamma K_0 \geq 0, \\ \big(PJ(A_0 - \gamma B_0) + \gamma JB_0\big)J_0(I + \gamma K_0)^{-1}e_0 \geq 0, \\ \big(PJ(A_0 - \gamma B_0) + \gamma JB_0\big)J_0(I + \gamma K_0)^{-1}\gamma K_0 \geq 0. \end{cases} \tag{5.11}$$

These conditions differ from the ones given in (5.5) for the corresponding linear multistep method. We will see in the next section that for explicit two-step methods the conditions (5.11) for the one-leg method are easier to fulfill than (5.5).

6. Application for explicit two-step methods

As an application of the general formulas derived in the previous section, we will give here detailed results for the class of explicit two-step methods of order one. With this class of methods we can take a_1 and b_1 as free parameters, and set $a_2 = 1 - a_1$, $b_2 = 2 - a_1 - b_1$. The methods have order two if $b_1 = 2 - \frac{1}{2}a_1$, and they are zero-stable if $0 \leq a_1 < 2$. The methods with $b_1 = 1$ or $a_1 = 2$ do not satisfy the Dahlquist irreducibility condition (3.1b). Furthermore, if $b_1 = 0$ or $b_2 = 0$, then the one-leg (OL) methods coincide with the corresponding linear multistep (LM) methods. For the one-leg methods it can therefore be assumed that $b_1 \neq 0$ and $b_2 \neq 0$.

For this family of methods, with free parameters a_1, b_1, we will display in contour plots the maximal values of γ such that we have monotonicity or boundedness with arbitrary starting vectors (for seminorms) or monotonicity with starting procedures (for convex functionals). For the 'white' areas in the contour plots, there is no positive γ.

Fig. 1 shows the maximal values of γ for which we have monotonicity with arbitrary starting vectors (left panel), or boundedness with arbitrary starting vectors (right panel).

These maximal stepsize coefficients are often called *threshold values*. The threshold values for monotonicity and boundedness are the same for the one-leg and linear multistep methods, and therefore the right panel in this figure coincides with [7, Fig. 4]. We see that for boundedness the area of nonzero thresholds is much larger than for monotonicity and it includes many interesting methods, for example the second-order methods with $b_1 = 2 - \frac{1}{2}a_1$.

These threshold values for monotonicity were obtained directly from condition (1.6). To find the threshold values for boundedness in the figure, condition $P \geq 0$ was verified numerically with $m = 1000$. Inspection with larger m showed that the results do not differ anymore visually; in fact, for most methods a much smaller value of m would have been sufficient. The conditions for monotonicity with starting procedures – such as (5.5) for the linear multistep methods and the related conditions (5.11) for the one-leg methods – will be verified in the same way, using the recursive formulas (3.9) for the coefficients of the relevant Toeplitz matrices.

In these plots, $b_1 = 1$ is a special case: starting with forward Euler, the whole linear two-step scheme reduces to an application of the forward Euler method, so then we have monotonicity with $\gamma = 1$. For the one-leg methods, $a_1 + b_1 = 2$ is also a special case: we then have $b_2 = 0$, so the one-leg method is then a linear multistep method, written in a reducible form.

6.1. Starting with the explicit Euler method

Consider explicit two-step methods, and suppose u_1 is computed by the forward Euler method,

$$u_1 = u_0 + \Delta tF(u_0). \tag{6.1}$$

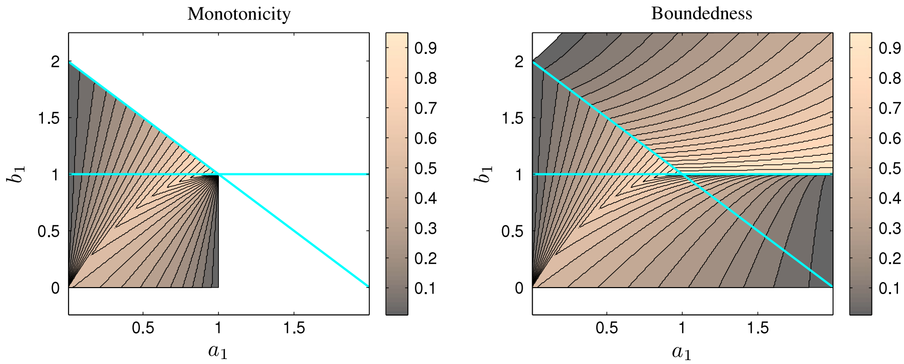

Fig. 1. Explicit two-step methods of order one, with parameters $a_1 \in [0, 2)$ horizontally and $b_1 \in [-\frac{1}{4}, \frac{9}{4}]$, $b_1 \neq 1$, vertically. The contour plot shows the optimal $\gamma > 0$ for monotonicity (left picture) and for boundedness (right picture) with arbitrary starting vectors. The contour levels are at $j/20$, $j = 0, 1, \ldots$; for the 'white' areas, there is no positive γ. The lines $b_1 = 1$ and $a_1 + b_1 = 2$ are special (reducible) cases.

This is of the form (2.8) with $m_0 = 2$, $w = (w_1, w_2)^T = (u_0, u_1) \in \mathbb{V}^2$ and we get

$$K_0 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad J_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{6.2}$$

For the linear multistep methods we then obtain from (5.4),

$$R_0 = \begin{pmatrix} c_2 + (1-\gamma)c_1 \\ (1-\gamma)c_2 \end{pmatrix}, \qquad P_0 = \begin{pmatrix} \gamma c_1 + \gamma b_2 & \gamma b_1 \\ \gamma c_2 & \gamma b_2 \end{pmatrix}, \tag{6.3}$$

with $c_j = a_j - \gamma b_j$ for $j = 1, 2$. For the one-leg methods this Euler starting procedure leads to (5.10) with

$$\bar{R}_0 = \begin{pmatrix} a_2 + (1-\gamma)a_1 \\ (1-\gamma)a_2 \\ \hat{b}_2 + (1-\gamma)\hat{b}_1 \\ (1-\gamma)\hat{b}_2 \end{pmatrix}, \qquad \bar{P}_0 = \begin{pmatrix} \gamma a_1 & 0 \\ \gamma a_2 & 0 \\ \gamma \hat{b}_1 & 0 \\ \gamma \hat{b}_2 & 0 \end{pmatrix}. \tag{6.4}$$

The total schemes with the linear two-step methods are irreducible (in the sense of Spijker) because all u_n are consistent approximations to $u(t)$ at different time levels. The combinations of the two-step one-leg methods and the forward Euler starting procedure are also irreducible (in the sense of Spijker) if $b_j \neq 0$ for $j = 1, 2$. To show this we consider this total scheme, written as one step of a big Runge–Kutta method with coefficient matrix (cf. (5.8))

$$\bar{K} = \begin{pmatrix} K_0 & O \\ \bar{S}\bar{T}_0 & \bar{T} \end{pmatrix}$$

where $\bar{S} \in \mathbb{R}^{\bar{m}\times 4}$, $\bar{T}_0 \in \mathbb{R}^{4\times 2}$ and $\bar{T} \in \mathbb{R}^{\bar{m}\times\bar{m}}$ are given by

$$\bar{S} = \begin{pmatrix} (I-A)^{-1}J & 0 \\ \frac{1}{\beta}B(I-A)^{-1}J & J \end{pmatrix}, \qquad \bar{T}_0 = \begin{pmatrix} a_1 & 0 \\ a_2 & 0 \\ \hat{b}_1 & 0 \\ \hat{b}_2 & 0 \end{pmatrix}, \qquad \bar{T} = \begin{pmatrix} O & \beta(I-A)^{-1} \\ O & B(I-A)^{-1} \end{pmatrix}.$$

Assume $b_1, b_2 \neq 0$. This implies $\hat{b}_1 \neq 0$, $\hat{b}_1 \neq 1$, and it also easily follows that all rows f $\bar{T}$ are then different from each other (which is easier than in the proof of Lemma 3.1 because we now have $k = 2$). Further it is clear that $\beta(I-A)^{-1}$ has no zero row. The first row of the lower triangular Toeplitz matrix $B(I-A)^{-1}$ is its only zero row, but the first row of $(\frac{1}{\beta}B(I-A)^{-1}J \;\; J)\bar{T}_0$ equals $(\hat{b}_1 \;\; 0)$, which is not equal to one of the rows of K_0. Therefore the coefficient matrix $\bar{K}$ of the total scheme has no equal rows.

In Fig. 2 the maximal values of γ are shown for which we have monotonicity with the forward Euler starting procedure for the explicit linear two-step methods (left picture) and the explicit one-leg methods (right picture). From this figure we conclude that the monotonicity properties with forward Euler starting procedure are better for these explicit one-leg methods than for the corresponding linear multistep methods if $b_1 > 1$, $a_1 + b_1 > 2$.

6.2. Starting with the explicit trapezoidal rule

Now suppose that u_1 is computed by the explicit trapezoidal rule, also known as the modified Euler method,

$$v_1 = u_0 + \Delta t F(u_0), \qquad u_1 = u_0 + \frac{1}{2}\Delta t F(u_0) + \frac{1}{2}\Delta t F(v_1). \tag{6.5}$$

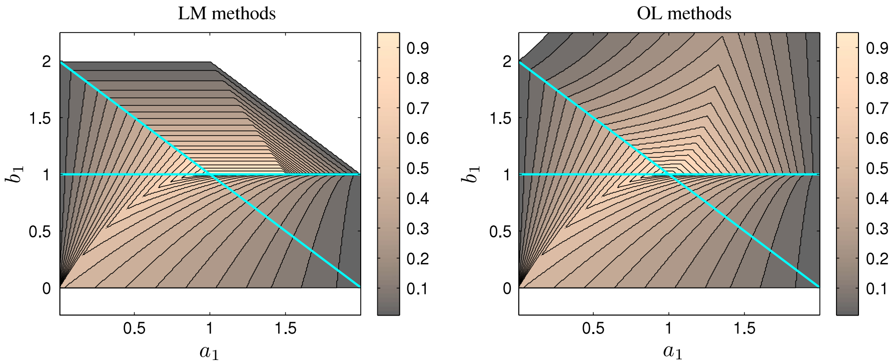

Fig. 2. Optimal $\gamma > 0$ for monotonicity of the explicit two-step methods with explicit Euler starting method. Left picture: linear multistep methods; right picture: one-leg methods. Explanations as in Fig. 1.

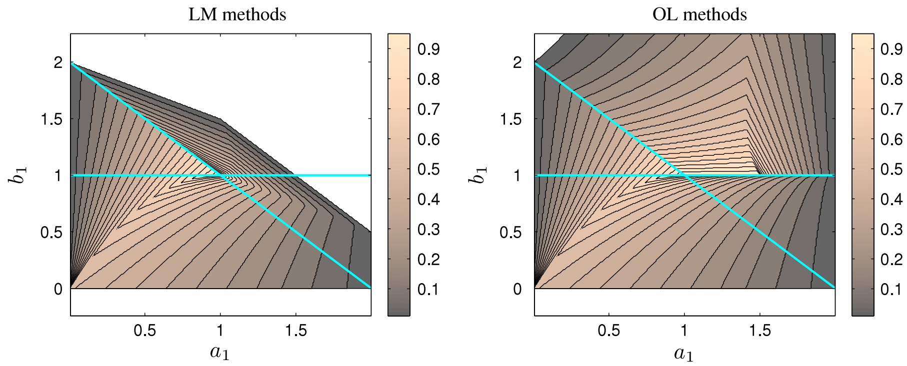

Fig. 3. Optimal $\gamma > 0$ for monotonicity of the explicit two-step methods with the explicit trapezoidal rule as starting method. Left picture: linear multistep methods; right picture: one-leg methods. Explanations as in Fig. 1.

This fits in our general form with $m_0 = 3$, $w = (w_1, w_2, w_3)^T = (u_0, v_1, u_1)^T$ and

$$K_0 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, \qquad J_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.6}$$

Here we have $\|w_j\| \le \|u_0\| (1 \le j \le m_0)$, whenever (1.2) is valid and $\Delta t \le \tau_0$.

For the linear multistep method this gives, as in [7],

$$R_0 = \begin{pmatrix} c_2 + c_1 r_0 \\ c_2 r_0 \end{pmatrix}, \qquad P_0 = \begin{pmatrix} c_1 q_0 + \gamma b_2 & c_1 q_1 & \gamma b_1 \\ c_2 q_0 & c_2 q_1 & \gamma b_2 \end{pmatrix}, \tag{6.7}$$

with $c_j = a_j - \gamma b_j$ for $j = 1, 2$, and $r_0 = 1 - \gamma + \frac{1}{2}\gamma^2$, $q_0 = \frac{1}{2}\gamma(1-\gamma)$, $q_1 = \frac{1}{2}\gamma$. For the one-leg methods we obtain, with the same r_0, q_0, q_1, the formulas

$$\bar{R_0} = \begin{pmatrix} a_2 + a_1 r_0 \\ a_2 r_0 \\ \hat{b}_2 + \hat{b}_1 r_0 \\ \hat{b}_2 r_0 \end{pmatrix}, \qquad \bar{P_0} = \begin{pmatrix} a_1 q_0 & a_1 q_1 & 0 \\ a_2 q_0 & a_2 q_1 & 0 \\ \hat{b}_1 q_0 & \hat{b}_1 q_1 & 0 \\ \hat{b}_2 q_0 & \hat{b}_2 q_1 & 0 \end{pmatrix}. \tag{6.8}$$

In the same way as with the explicit Euler starting procedure, it can be verified that the total schemes are irreducible (in the sense of Spijker), under the assumption $b_j \neq 0\,(j = 1, 2)$ for the one-leg methods. The monotonicity conditions (5.5) and (5.11) are therefore not only sufficient but also necessary.

The maximal values of $\gamma > 0$ for monotonicity with this explicit trapezoidal rule starting procedure are shown in Fig. 3; in the left picture for the linear two-step methods, and in the right picture for the one-leg methods.

For the linear multistep methods monotonicity with the explicit trapezoidal rule as starting procedure leads to monotonicity thresholds that are less than or equal to those with the forward Euler method.

The one-leg methods with the explicit trapezoidal rule as starting method give here almost the same thresholds as with the forward Euler method, except for a parameter region with $a_1, b_1 > 1$. There the thresholds are somewhat improved

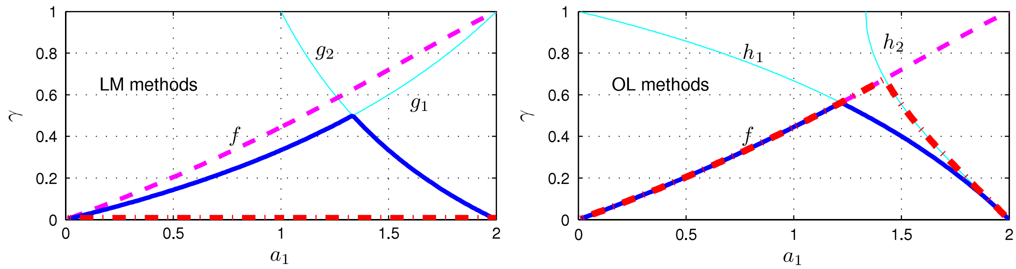

Fig. 4. Explicit two-step methods of order two, with parameter $a_1 \in [0, 2)$. Vertical axis: thresholds for boundedness or monotonicity. Left picture: linear multistep methods; right picture: one-leg methods. Dashed line for boundedness; solid lines for monotonicity with forward Euler start; dash-dotted lines for monotonicity with explicit trapezoidal rule as starting method (where $\gamma = 0$ for the linear two-step methods).

with the explicit trapezoidal rule. This is in marked contrast to the situation for the corresponding linear multistep methods, where the thresholds deteriorate with this starting method in comparison with the explicit Euler start.

6.3. *Explicit two-step methods of order two*

The most interesting explicit two-step methods are of course those with order two, given by $b_1 = 2 - \frac{1}{2}a_1$. This is a one-parameter family with a_1 as free parameter. Here a more clear picture is provided in Fig. 4, where thresholds are plotted for boundedness and for monotonicity with forward Euler and the explicit trapezoidal rule as starting methods for the linear two-step methods and the corresponding one-leg methods. This gives a cut of the previous figures along the line $\frac{1}{2}a_1 + b_1 = 2$.

The curves in these figures describing the thresholds are actually quite simple. The condition for boundedness is $\gamma \le f(a_1)$ with $f(z) = 2z(3-z)/(4-z)^2$; this value was shown in [5] to be sufficient and from the requirement $\pi_2 = \gamma(a_1 b_1 + a_2) - \gamma^2 b_1^2 \ge 0$, it directly follows that it is also necessary.

For the linear multistep methods with forward Euler start it can be shown, in the same way as in [15], that a sufficient condition for monotonicity is given by $\gamma \le \min\{g_1(a_1), g_2(a_1)\}$ with $g_1(z) = z/(4-z)$, $g_2(z) = (2-z)/z$; see also results in [5] for $a_1 \le \frac{4}{3}$. We see in the figure that these sufficient conditions for monotonicity are also necessary. With an explicit trapezoidal rule start there is no positive threshold for monotonicity; this can be shown by considering the Runge–Kutta method that arises with $m = 1$, giving the coefficient matrix

$$K = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ \frac{1}{2} & \frac{1}{2} & 0 & \\ \beta_1 & \beta_2 & \beta_3 & 0 \end{pmatrix} \tag{6.9}$$

with $\beta_1 = b_2 + \frac{1}{2}a_1$, $\beta_2 = \frac{1}{2}a_1$ and $\beta_3 = b_1$ being the weights of this explicit 3-stage method. The conditions for order two imply that $\beta_1 = 0$, and using [10, Thm. 4.2] it can be shown that this Runge–Kutta method does not have a positive stepsize coefficient for monotonicity.

For the one-leg methods with forward Euler starting procedure we have $\gamma \le \min\{f(a_1), h_1(a_1)\}$ as sufficient condition with $h_1(z) = 2(2-z)/(4-z)$. This follows from the results in [15] on positivity. With an explicit trapezoidal rule start the sufficient condition becomes $\gamma \le \min\{f(a_1), h_2(a_1)\}$ with $h_2(z) = 1 - \sqrt{(3z-4)/(4-z)}$, $z \ge \frac{4}{3}$. This sufficient condition can be derived in a similar way as in [15]. Again, we see from Fig. 4 that these sufficient conditions for monotonicity are necessary as well. This necessity can actually be proven by directly considering the Runge–Kutta methods that arise with $m = 1$ and $m = 2$, that is, after one or two steps of the one-leg method with these starting procedures.

7. Concluding remarks

In view of the reduced storage requirements, compared to linear multistep methods, one-leg methods are interesting for large-scale computations. In this paper results have been presented for having boundedness with arbitrary starting values, as well as for monotonicity with Runge–Kutta starting procedures.

It was seen that the stepsize restriction $\Delta t \le \gamma\tau_0$ for the boundedness property (1.5) with seminorms is the same for a one-leg method as for the associated linear multistep method. Differences between one-leg methods and linear multistep methods arise when we consider the monotonicity property (1.7) with starting procedures.

For explicit two-step methods it was seen that the monotonicity properties with standard starting procedures are better for the one-leg methods than for the linear multistep methods. However, no general conclusions are to be drawn from this. For the implicit two-step methods of order two it was found that the requirements for monotonicity, starting with backward Euler or with the θ-method, $\theta = b_0$, were not always better for the one-leg methods than for the corresponding linear multistep methods. These implicit methods turn out to have thresholds less than or equal to two. Since this is not very much larger than for the explicit methods, the implicit methods are not recommended if monotonicity is important, and therefore the results for these implicit methods have not been discussed here in detail.

A detailed study and comparison of the monotonicity conditions (5.5) and (5.11) for higher order linear multistep and one-leg methods, with suitable Runge–Kutta starting procedures, is part of our research plans.

Finally we mention that some numerical tests were performed with the second-order explicit two-step methods for scalar conservation laws $u_t + f(u)_x = 0$ in one spatial dimension with $f(u) = u$ (linear advection) and $f(u) = \frac{1}{2}u^2$ (Burgers equation). For spatial discretization a van Leer type flux-limited scheme was used, for which it is known that the resulting ODE system satisfies the basic assumption (1.2) with τ_0 proportional to the mesh width in space Δx. However, no significant difference was found in these tests between the explicit two-step one-leg methods and the corresponding linear multistep methods with the various starting procedures, even though the theoretical properties of the one-leg methods are more favorable than those of the linear multistep methods. Consequently, the practical relevance of the differences between one-leg and linear multistep methods found in this paper are not yet fully established. However, for practical computations the theoretical findings do give a foundation for the explicit two-step one-leg methods with explicit trapezoidal start that is more solid than for the linear multistep methods.

Acknowledgments

The work of A. Mozartova has been supported by a grant from the Netherlands Organization for Scientific Research NWO. The work of I. Savostianov and W. Hundsdorfer for this publication has been supported by Award No. FIC/2010/05 from the King Abdullah University of Science and Technology (KAUST).

References

[1] G. Dahlquist, Error analysis for a class of methods for stiff nonlinear initial value problems, in: G.A. Watson (Ed.), Procs. Dundee Conf. 1975, in: Lecture Notes in Math., vol. 506, Springer, 1976, pp. 60–74.

[2] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II — Stiff and Differential-Algebraic Problems, second ed., in: Springer Series in Comput. Math., vol. 14, Springer, 1996.

[3] S. Gottlieb, D. Ketcheson, C.-W. Shu, Strong Stability Preserving Runge–Kutta and Multistep Time Discretizations, World Scientific, 2011.

[4] M.N. Spijker, Stepsize restrictions for general monotonicity in numerical initial value problems, SIAM J. Numer. Anal. 45 (2007) 1226–1245.

[5] W. Hundsdorfer, S.J. Ruuth, R.J. Spiteri, Monotonicity-preserving linear multistep methods, SIAM J. Numer. Anal. 41 (2003) 605–623.

[6] W. Hundsdorfer, S.J. Ruuth, On monotonicity and boundedness properties of linear multistep methods, Math. Comp. 75 (2006) 655–672.

[7] W. Hundsdorfer, A. Mozartova, M.N. Spijker, Stepsize restrictions for boundedness and monotonicity of multistep methods, J. Sci. Comput. 50 (2012) 265–286.

[8] W. Hundsdorfer, A. Mozartova, M.N. Spijker, Stepsize conditions for boundedness in numerical initial value problems, SIAM J. Numer. Anal. 47 (2009) 3797–3819.

[9] W. Hundsdorfer, A. Mozartova, M.N. Spijker, Special boundedness properties in numerical initial value problems, BIT 51 (2011) 909–936.

[10] J.F.B.M. Kraaijevanger, Contractivity of Runge–Kutta methods, BIT 31 (1991) 482–528.

[11] L. Ferracina, M.N. Spijker, An extension and analysis of the Shu–Osher representation of Runge–Kutta methods, Math. Comp. 74 (2005) 201–219.

[12] I. Higueras, Representations of Runge–Kutta methods and strong stability preserving methods, SIAM J. Numer. Anal. 43 (2005) 924–948.

[13] J.C. Butcher, Numerical Methods for Ordinary Differential Equations, Wiley, 2003.

[14] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I — Nonstiff Problems, second ed., in: Springer Series Comput. Math., vol. 8, Springer, 1993.

[15] N.N. PhamThi, W. Hundsdorfer, B.P. Sommeijer, Positivity for explicit two-step methods in linear multistep and one-leg form, BIT 46 (2006) 875–882.