SIAM J. NUMER. ANAL.
Vol. 54, No. 5, pp. 2799–2832
© 2016 Society for Industrial and Applied Mathematics

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

STRONG STABILITY PRESERVING EXPLICIT LINEAR MULTISTEP METHODS WITH VARIABLE STEP SIZE*

YIANNIS HADJIMICHAEL[†], DAVID I. KETCHESON[‡], LAJOS LÓCZI[†], AND ADRIÁN NÉMETH[§]

Abstract. Strong stability preserving (SSP) methods are designed primarily for time integration of nonlinear hyperbolic PDEs, for which the permissible SSP step size varies from one step to the next. We develop the first SSP linear multistep methods (of order two and three) with variable step size, and prove their optimality, stability, and convergence. The choice of step size for multistep SSP methods is an interesting problem because the allowable step size depends on the SSP coefficient, which in turn depends on the chosen step sizes. The description of the methods includes an optimal step-size strategy. We prove sharp upper bounds on the allowable step size for explicit SSP linear multistep methods and show the existence of methods with arbitrarily high order of accuracy. The effectiveness of the methods is demonstrated through numerical examples.

Key words. strong stability preservation, monotonicity, linear multistep methods, variable step size, time integration

AMS subject classifications. Primary, 65M20; Secondary, 65L06

DOI. 10.1137/15M101717X

1. Introduction. Strong stability preserving (SSP) linear multistep methods (LMMs) with uniform step size have been studied by several authors [9, 10, 4, 12, 3, 5]. In this work, we develop the first variable step-size (VSS) SSP multistep methods.

The principal area of application of SPP methods is the integration of nonlinear systems of hyperbolic conservation laws. In such applications, the allowable step size $h_{\max}$ is usually determined by a CFL-like condition, and in particular is inversely proportional to the fastest wave speed. This wave speed may vary significantly during the course of the integration, and its variation cannot generally be predicted in advance. Thus a fixed-step-size code may be inefficient (if $h_{\max}$ increases) or may fail completely (if $h_{\max}$ decreases).

SSP Runge–Kutta methods are used much more widely than SSP LMMs. Indeed, it is difficult to find examples of SSP LMMs used in realistic applications; this may be due to the lack of a VSS formulation. Trade-offs between Runge–Kutta and LMMs have been discussed at length elsewhere, but one reason for preferring LMMs over their Runge–Kutta counterparts in the context of SPP stems from the recent development of a high-order positivity preserving limiter [16]. Use of Runge–Kutta methods in conjunction with the limiter can lead to order reduction, so LMMs are recommended [1, 16].

*Received by the editors April 16, 2015; accepted for publication (in revised form) June 21, 2016; published electronically September 8, 2016.
http://www.siam.org/journals/sinum/54-5/M101717.html
Funding: This work was supported by the King Abdullah University of Science and Technology (KAUST), 4700 Thuwal, 23955-6900, Saudi Arabia. The work of the fourth author was partially supported by the grant TÁMOP-4.2.2.A-11/1/KONV-2012-0012.
[†]Computer, Electrical, aand Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Jeddah 23955-6900, Saudi Arabia (yiannis.hadjimichael@kaust.edu.sa, lajos.loczi@kaust.edu.sa).
[‡]Corresponding author. Computer, Electrical, aand Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Jeddah 23955-6900, Saudi Arabia (david.ketcheson@kaust.edu.sa).
[§]Széchenyi István University, Győr, H-9026, Hungary (nemetha@sze.hu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

There exist two approaches to VSS multistep methods [2]. In the first, polynomial interpolation is used to find values at equally spaced points, and then the fixed-step-size method is used. In the second, the method coefficients are varied to maintain high-order accuracy based on the solution values at the given (nonuniform) step intervals. Both approaches are problematic for SSP methods; the first, because the interpolation step itself may violate the SSP property, and the second, because the SSP step size depends on the method coefficients. Herein we pursue the second strategy.

The main contributions of this work are

- sharp bounds on the SSP coefficient of VSS SSP LMMs (Theorems 2–4);
- optimal methods of orders two and three (sections 3.2 and 3.3);
- existence of methods of arbitrary order (Theorem 5);
- analysis of the greedy step-size strategy (section 4);
- proof of stability and convergence of the optimal methods (Theorems 8–10).

The rest of the paper is organized as follows. In section 2.1, we review the theory of SSP LMMs while recognizing that the forward Euler permissible step size may change from step to step. The main result, Theorem 1, is a slight refinement of the standard one. In section 2.2 we show how an optimal SSP multistep formula may be chosen at each step, given the sequence of previous steps. In section 2.3 we provide for convenience a description of two of the simplest and most useful methods in this work. In sections 3.2 and 3.3 we derive and prove the optimality of several second- and third-order methods. In section 4 we investigate the relation between the SSP step size, the method coefficients, and the step-size sequence. We develop step-size strategies that ensure the SSP property under mild assumptions on the problem. In section 5 we prove that the methods, with the prescribed step-size strategies, are stable and convergent. In section 6 we demonstrate the efficiency of the methods with some numerical examples. Finally, sections 8 and 9 contain the proofs of the more technical theorems and lemmas.

Two topics that might be pursued in the future based on this work are

- VSS SSP LMMs of order higher than three;
- VSS of SSP methods with multiple steps and multiple stages.

2. SSP LMMs.

We consider the numerical solution of the initial value problem

$$u'(t) = f(u(t)), \qquad u(t_0) = u_0, \tag{2.1}$$

for $t \in [t_0, t_0 + T]$ by an explicit LMM. If a fixed numerical step size h is used, the method takes the form

$$u_n = \sum_{j=0}^{k-1} \left(\alpha_j u_{n-k+j} + h\beta_j f(u_{n-k+j})\right), \qquad n \geq k. \tag{2.2}$$

Here k is the number of steps and u_n is an approximation to the solution $u(nh)$.

Now let the step size vary from step to step so that $t_n = t_{n-1} + h_n$. In order to achieve the same order of accuracy, the coefficients α, β must also vary from step to step:

$$u_n = \sum_{j=0}^{k-1} \left(\alpha_{j,n} u_{n-k+j} + h_n\beta_{j,n} f(u_{n-k+j})\right). \tag{2.3}$$

At this point, it is helpful to establish the following terminology. We use the term *multistep formula*, or just *formula*, to refer to a set of coefficients α_j $(= \alpha_{j,n})$, β_j

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

($= \beta_{j,n}$) that may be used at step n. We use the term *multistep method* to refer to a full time-stepping algorithm that includes a prescription of how to choose a formula $(\alpha_{j,n}, \beta_{j,n})$ and the step size h_n at step n.

For $1 \le j \le k$ let

$$\omega_j := \frac{h_{n-k+j}}{h_n} > 0 \tag{2.4}$$

denote the step-size ratios and

$$\begin{cases} \Omega_0 :=0, \\ \Omega_j := \displaystyle\sum_{i=1}^{j} \omega_i \quad \text{for} \quad 1 \le j \le k. \end{cases} \tag{2.5}$$

Note that the values ω and Ω depend on n, but we often suppress that dependence since we are considering a single step.[1] It is useful to keep in mind that $\Omega_j = j$ if the step size is fixed. Also the simple relation $\Omega_k = \Omega_{k-1} + 1$ will often be used.

2.1. SPP. We are interested in initial value problems (2.1) whose solution satisfies a monotonicity condition

$$\|u(t+h)\| \le \|u(t)\| \qquad \text{for } h \ge 0, \tag{2.6}$$

where $\|\cdot\|$ represents any convex functional (for instance, a norm). We assume that f satisfies the (stronger) forward Euler condition

$$\|u + hf(u)\| \le \|u\| \qquad \text{for } 0 \le h \le h_{\mathrm{FE}}(u). \tag{2.7}$$

The discrete monotonicity condition (2.7) implies the continuous monotonicity condition (2.6).

The primary application of SSP methods is in the time integration of nonlinear hyperbolic PDEs. In such applications, h_{FE} is proportional to the CFL number

$$\nu = h\frac{a(u)}{\Delta x}, \tag{2.8}$$

where $a(u)$ is the largest wave speed appearing in the problem. This speed depends on u. For instance, in the case of Burgers' equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0, \tag{2.9}$$

we have $a(u) = \max_x |u|$. For scalar conservation laws like Burgers' equation, it is possible to determine a value of h_{FE}, based on the initial and boundary data, that is valid for all time. But for general systems of conservation laws, $a(u)$ can grow in time and so the minimum value of $h_{\mathrm{FE}}(u)$ cannot be determined without solving the initial value problem. We will often write just h_{FE} for brevity, but the dependence of h_{FE} on u should be remembered.

[1] Our definition of ω differs from the typical approach in the literature on VSS multistep methods, where only ratios of adjacent step sizes are used. The present definition is more convenient in what follows.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

We will develop LMMs (2.3) that satisfy the discrete monotonicity property

$$\|u_n\| \leq \max(\|u_{n-k}\|, \|u_{n-k+1}\|, \dots, \|u_{n-1}\|). \tag{2.10}$$

The class of methods that satisfy (2.10) whenever f satisfies (2.7) are known as SPP methods. The most widely used SSP methods are one-step (Runge–Kutta) methods. When using an SSP multistep method, an SSP Runge–Kutta method can be used to ensure monotonicity of the starting values. In the remainder of this work, we focus on conditions for monotonicity of subsequent steps (for any given starting values).

The following theorem refines a well-known result in the literature, by taking into account the dependence of h_{FE} on u.

THEOREM 1. *Suppose that f satisfies the forward Euler condition* (2.7) *and that the method* (2.3) *has nonnegative coefficients $\alpha_{j,n}, \beta_{j,n} \geq 0$. Furthermore, suppose that the time step is chosen so that*

$$0 \leq h_n \leq \min_{0\leq j\leq k-1} \left(\frac{\alpha_{j,n}}{\beta_{j,n}} \, h_{\mathrm{FE}}(u_{n-k+j}) \right) \tag{2.11}$$

for each n, where the ratio $\alpha_{j,n}/\beta_{j,n}$ is understood as $+\infty$ if $\beta_{j,n} = 0$. Then the solution of the initial value problem (2.1) *given by the LMM* (2.3) *satisfies the monotonicity condition* (2.10)*.*

Remark 1. The step-size restriction (2.11) is also *necessary* for monotonicity in the sense that, for any method (2.3), there exist some f and starting values such that the monotonicity condition (2.10) will be violated if the step size (2.11) is exceeded.

Remark 2. Even in the case of the fixed-step-size method (2.2), the theorem above generalizes results in the literature that are based on the assumption of a constant h_{FE}. It is natural to implement a step-size strategy that uses simply $h_{\mathrm{FE}}(u_{n-1})$ in (2.11), but this will not give the correct step size in general. On the other hand, since h_{FE} usually varies slowly from one step to the next, is often nondecreasing, and since the restriction (2.11) is often pessimistic, such a strategy will usually work well.

Since u_n (and hence h_{FE}) varies slowly from one step to the next, it seems convenient to separate the factors in the upper bound in (2.11) and consider the sufficient condition

$$0 \leq h_n \leq \mathcal{C}_n \mu_n, \tag{2.12}$$

where the SSP coefficient $\mathcal{C}_n$ is

$$\mathcal{C}_n = \begin{cases} \max\{r \in \mathbb{R}^+ : \alpha_{j,n} - r\beta_{j,n} \geq 0\} & \text{if } \alpha_{j,n} \geq 0,\ \beta_{j,n} \geq 0 \text{ for all } j, \\ 0 & \text{otherwise,} \end{cases} \tag{2.13}$$

and

$$\mu_n := \min_{0\leq j\leq k-1} h_{\mathrm{FE}}(u_{n-k+j}) \qquad (n \geq k). \tag{2.14}$$

Note that in general the SSP coefficient varies from step to step, since it depends on the method coefficients.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

2.2. Optimal SSP formulas. For a given order p, number of steps k, and previous step-size sequence $h_{n-1}, h_{n-2}, \ldots$, we say that a multistep formula is optimal if it gives the largest possible SSP coefficient $\mathcal{C}_n$ in (2.12) and satisfies the order conditions. In this section we formulate this optimization problem algebraically. The linear multistep formula takes the form

$$u_n = \sum_{j=0}^{k-1} \left(\alpha_j u_{n-k+j} + h_n \beta_j f(u_{n-k+j})\right). \tag{2.15}$$

Here we have omitted the subscript n on the method coefficients to simplify the notation. The conditions for formula (2.15) to be consistent of order p are

$$\sum_{j=0}^{k-1} \alpha_j = 1, \tag{2.16a}$$

$$\sum_{j=0}^{k-1} \left(\Omega_j^m \alpha_j + m\Omega_j^{m-1}\beta_j\right) = \Omega_k^m, \qquad 1 \leq m \leq p. \tag{2.16b}$$

Let us change variables by introducing

$$\delta_j := \alpha_j - r\beta_j.$$

Then the SSP coefficient of the formula is just the largest r such that all δ_j, β_j are nonnegative [9, 5]. In these variables, the order conditions (2.16) become

$$\sum_{j=0}^{k-1} (\delta_j + r\beta_j) = 1, \tag{2.17a}$$

$$\sum_{j=0}^{k-1} \left(\Omega_j^m(\delta_j + r\beta_j) + m\Omega_j^{m-1}\beta_j\right) = \Omega_k^m, \qquad 1 \leq m \leq p. \tag{2.17b}$$

We will refer to a formula by the triplet (ω, δ, β). Given p, k, and a set of step-size ratios ω_j, the formula with the largest SSP coefficient for the next step can be obtained by finding the largest r such that (2.17) has a nonnegative solution $\delta, \beta \geq 0$. This could be done following the approach of [5], by bisecting in r and solving a sequence of linear programming feasibility problems. The solution of this optimization problem would be the formula for use in the next step. We do not pursue that approach here. Instead, we derive families of formulas that can be applied based on the sequence of previous step sizes.

2.3. Two optimal methods. For convenience, we list the methods most likely to be of interest for practical application. These (and other methods) are derived and analyzed in the rest of the paper. Recall that μ_n has been defined in (2.14), and we assume $n \geq k$. The definition of the Ω quantities (with dependence on n suppressed) is given in (2.4)–(2.5).

2.3.1. SSPMSV32. Our three-step, second-order method is

$$u_n = \frac{\Omega_2^2 - 1}{\Omega_2^2}\left(u_{n-1} + \frac{\Omega_2}{\Omega_2 - 1} h_n f(u_{n-1})\right) + \frac{1}{\Omega_2^2} u_{n-3}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with SSP step-size restriction

$$h_n \leq \frac{h_{n-2} + h_{n-1}}{h_{n-2} + h_{n-1} + \mu_n} \cdot \mu_n.$$

If the step size is constant, this is equivalent to the known optimal second-order three-step SSP method.

2.3.2. SSPMSV43. Our four-step, third-order method is

$$\begin{aligned} u_n = & \frac{(\Omega_3 + 1)^2(\Omega_3 - 2)}{\Omega_3^3} \left(u_{n-1} + \frac{\Omega_3}{\Omega_3 - 2} h_n f(u_{n-1}) \right) \\ & + \frac{3\Omega_3 + 2}{\Omega_3^3} \left(u_{n-4} + \frac{\Omega_3(\Omega_3 + 1)}{3\Omega_3 + 2} h_n f(u_{n-4}) \right) \end{aligned}$$

with SSP step-size restriction

$$h_n \leq \frac{\sum_{j=1}^{3} h_{n-j}}{\left(\sum_{j=1}^{3} h_{n-j}\right) + 2\mu_n} \cdot \mu_n.$$

If the step size is constant, this is equivalent to the known optimal third-order four-step SSP method.

3. Existence and construction of optimal SSP formulas. In this section we consider the set of formulas satisfying (2.17) for fixed order p, number of steps k, and some step-size sequence Ω_j. It is natural to ask whether any such formula exists, what the supremum of achievable r values is (i.e., the optimal SSP coefficient $\mathcal{C}$), and whether that supremum is attained by some formula. Here we give answers for certain classes.

In section 3.1 we discuss how large an SSP coefficient can be, and prove the existence of a formula with the maximum SSP coefficient. In sections 3.2 and 3.3 we construct some practical optimal formulas of orders 2 and 3, while the existence of higher-order formulas is established in section 3.4. The theorems of the present section are proved in section 8. Our theorems are based on [11] by extending the corresponding results of that paper to the VSS case. The basic tools in [11] include Farkas' lemma, the duality principle, and the strong duality theorem of linear programming [13].

3.1. Upper bound on the SSP coefficient and existence of an optimal formula. In the fixed-step-size case, the classical upper bound $\mathcal{C} \leq \frac{k-p}{k-1}$ on the SSP coefficient $\mathcal{C}$ for a k-step explicit linear multistep formula of order p with $k \geq p$ was proved in [9] together with the existence of optimal methods.

THEOREM 2. *Suppose that some time-step ratios ω_j are given. Then the SSP coefficient for a k-step explicit linear multistep formula with order of accuracy $p \geq 2$ is bounded by*

$$\mathcal{C}(\omega, \delta, \beta) \leq \begin{cases} 0 & \text{if } \Omega_k \leq p, \\ \frac{\Omega_k - p}{\Omega_k - 1} & \text{if } \Omega_k > p. \end{cases} \tag{3.1}$$

Moreover, suppose there exists a k-step explicit linear multistep formula of order $p \geq 2$ with positive SSP coefficient. Then there is a k-step formula of order p whose SSP coefficient is equal to the optimal one.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3.2. Second-order formulas. The bound in Theorem 2 is sharp for $p = 2$, as the following result shows.

THEOREM 3 (optimal second-order formulas with $k \geq 2$ steps). *Suppose that some time-step ratios ω_j are given. Then there exists a second-order linear multistep formula with k steps and with positive SSP coefficient if and only if $\Omega_k > 2$. In this case, the optimal formula is*

$$u_n = \frac{\Omega_{k-1}^2 - 1}{\Omega_{k-1}^2}\left(u_{n-1} + \frac{\Omega_{k-1}}{\Omega_{k-1}-1}\, h_n\, f(u_{n-1})\right) + \frac{1}{\Omega_{k-1}^2}\, u_{n-k}, \tag{3.2}$$

and has SSP coefficient

$$\mathcal{C}(\omega, \delta, \beta) = \frac{\Omega_k - 2}{\Omega_k - 1}. \tag{3.3}$$

Remark 3. If the step size is fixed, method (3.2) with $k \geq 3$ is equivalent to the optimal k-step, second-order SSP method given in [1, section 8.2.1]:

$$u_n = \frac{(k-1)^2 - 1}{(k-1)^2}\left(u_{n-1} + \frac{k-1}{k-2}\, h f(u_{n-1})\right) + \frac{1}{(k-1)^2}\, u_{n-k} \tag{3.4}$$

with SSP coefficient $\mathcal{C} = (k-2)/(k-1)$.

3.3. Third-order formulas. Compared to the family of second-order formulas above, the optimal third-order formulas have a more complicated structure. Although we will eventually focus on two relatively simple third-order formulas (corresponding to $k = 4$ and $k = 5$), we present complete results in order to give a flavor of what may happen in the search for optimal formulas. The following Theorem 4 characterizes optimal third-order linear multistep formulas and their SSP coefficients, again for arbitrary step-size ratios ω_j. The theorem also provides an efficient way to find these optimal k-step formulas, since the sets of nonzero formula coefficients, denoted by $\mathcal{N} \equiv \mathcal{N}(\omega, \delta, \beta)$, are explicitly described. First we define $2k - 2$ quantities and sets that will appear in Theorem 4.

- For $j = 0$:
$$r_0 := \frac{\Omega_k - 3}{\Omega_k - 1}, \quad \mathcal{S}_0 := \{\delta_0, \beta_0, \beta_{k-1}\}.$$
- For $1 \leq j \leq k-2$:
$$r_j := \max\left(\frac{\Omega_k - \Omega_j - 3}{\Omega_k - \Omega_j - 1}, \frac{2}{\omega_j} + \frac{1}{\Omega_k - \Omega_{j-1}}\right),$$
and we have either $\mathcal{S}_j := \{\delta_j, \beta_j, \beta_{k-1}\}$ or $\mathcal{S}_j := \{\delta_j, \beta_{j-1}, \beta_j\}$ or $\mathcal{S}_j := \{\delta_j, \beta_{j-1}, \beta_j, \beta_{k-1}\}$, depending on whether the first expression is greater, or the second expression is greater, or the two expressions are equal in the above $\max(\ldots, \ldots)$.
- For $j = k-1$:
$$r_{k-1} := \frac{2}{\omega_{k-1}} + \frac{1}{\Omega_k - \Omega_{k-2}}, \quad \mathcal{S}_{k-1} := \{\delta_{k-1}, \beta_{k-2}, \beta_{k-1}\}.$$
- For $0 \leq j \leq k-3$: the quantities r_{k+j} are defined below, and $\mathcal{S}_{k+j} := \{\beta_j, \beta_{j+1}, \beta_{k-1}\}$. For any $0 \leq j \leq k-3$ we set
(3.5)
$$P_{k+j}(x) := \Delta_j \Delta_{j+1} x^3 - (\Delta_j \Delta_{j+1} + \Delta_j + \Delta_{j+1})\, x^2 + 2\,(\Delta_j + \Delta_{j+1} + 1)\, x - 6,$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where

$$\Delta_m := \Omega_k - \Omega_m. \tag{3.6}$$

If

$$\Delta_{j+1}^2 - (\Delta_j + 1)\,\Delta_{j+1} + 3\Delta_j > 0 \quad \text{or} \quad \Delta_j < 5 + 2\sqrt{6}, \tag{3.7}$$

then $P_{k+j}(\cdot)$ has a unique real root. We define

$$r_{k+j} := \begin{cases} \text{the real root of } P_{k+j} & \text{if (3.7) holds,} \\ +\infty & \text{if (3.7) does not hold.} \end{cases} \tag{3.8}$$

THEOREM 4 (optimal third-order formulas with $k \geq 2$ steps). *Let time-step ratios ω_j be given. Then the inequality $\Omega_k > 3$ is necessary and sufficient for the existence of a third-order, k-step explicit linear multistep formula with positive SSP coefficient. For $\Omega_k > 3$, the optimal SSP coefficient is*

$$\mathcal{C}(\omega, \delta, \beta) = \min_{0 \leq j \leq 2k-3} r_j,$$

and the set of nonzero coefficients of an optimal SSP formula satisfies $\mathcal{N} \subseteq \mathcal{S}_\ell$, where the index $\ell \in \{0, 1, \ldots, 2k-3\}$ is determined by the relation $r_\ell = \min_{0 \leq j \leq 2k-3} r_j$. If the index ℓ where the minimum is attained is not unique and we have $\min_{0 \leq j \leq 2k-3} r_j = r_{\ell_1} = \cdots = r_{\ell_m}$, then $\mathcal{N} \subseteq \mathcal{S}_{\ell_1} \cap \cdots \cap \mathcal{S}_{\ell_m}$.

Remark 4. Let us highlight some differences between the fixed-step-size and the VSS cases concerning the sets of nonzero formula coefficients.

1. The pattern of nonzero coefficients for optimal third- or higher-order formulas can be different from that of their fixed-step-size counterparts (this phenomenon does not occur in the class of optimal second-order formulas). A simple example is provided by the optimal third-order, five-step formula with

$$(\Omega_0, \ldots, \Omega_5) := \left(0, 1, \frac{7}{3}, \frac{11}{3}, 5, 6\right),$$

where $\mathcal{N} = \{\beta_0, \beta_1, \beta_4\}$ (the coefficient pattern being similar to the case of the optimal fixed-step-size third-order, six-step method).

2. If the index ℓ with $\mathcal{C}(\omega, \delta, \beta) = r_\ell$ is not unique, then the optimal formula has less than 3 nonzero coefficients in the general case.

3. If the index ℓ with $\mathcal{C}(\omega, \delta, \beta) = r_\ell$ satisfies $1 \leq \ell \leq k-2$ and the expressions in the $\max(\ldots)$ are equal, then the optimal formula is generally not unique and has more than 3 nonzero coefficients. For example, with $\omega_j := 5$ for $1 \leq j \leq k-2$, $\omega_{k-1} := 4$, and $\omega_k := 1$, we have a one-parameter family of optimal methods, and $\mathcal{N} \subseteq \{\delta_{k-2}, \beta_{k-3}, \beta_{k-2}, \beta_{k-1}\}$ with

$$\delta_{k-2} := \frac{1}{7}(2 - 5\beta_{k-2}), \; \beta_{k-3} := \frac{2}{63}(16\beta_{k-2} - 5), \quad \text{and } \beta_{k-1} := \frac{5}{63}(20 - \beta_{k-2})$$

for any $5/16 \leq \beta_{k-2} \leq 2/5$. However, for any fixed ω it can be shown that there is an optimal formula that has at most p nonzero coefficients just as in the fixed-step-size case [9].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The optimal fixed-step-size SSP method of order 3 and $k = 4$ or $k = 5$ steps has nonzero coefficients $\{\delta_0, \beta_0, \beta_{k-1}\}$ [1, section 8.2.2]. In the rest of this section we consider formulas with 4 or 5 steps that generalize the corresponding fixed-step-size methods. A continuity argument shows that the set of nonzero coefficients is preserved if the step sizes are perturbed by a small enough amount. Hence we solve the VSS order conditions (2.17) with $p = 3$ for r, δ_0, β_0, and β_{k-1}; by using $\Omega_{k-1} > 0$, we obtain that the unique solution is

$$r = \frac{\Omega_{k-1} - 2}{\Omega_{k-1}}, \tag{3.9a}$$

$$\delta_0 = \frac{4(\Omega_{k-1}+1) - \Omega_{k-1}^2}{\Omega_{k-1}^3}, \qquad \beta_0 = \frac{\Omega_{k-1}+1}{\Omega_{k-1}^2}, \qquad \beta_{k-1} = \frac{(\Omega_{k-1}+1)^2}{\Omega_{k-1}^2}. \tag{3.9b}$$

The resulting VSS formula reads

$$\begin{aligned} u_n = {} & \frac{(\Omega_{k-1}+1)^2(\Omega_{k-1}-2)}{\Omega_{k-1}^3}\, u_{n-1} + \frac{(\Omega_{k-1}+1)^2}{\Omega_{k-1}^2}\, h_n f(u_{n-1}) \\ & + \frac{3\Omega_{k-1}+2}{\Omega_{k-1}^3}\, u_{n-k} + \frac{\Omega_{k-1}+1}{\Omega_{k-1}^2}\, h_n f(u_{n-k}). \end{aligned} \tag{3.10}$$

Proposition 1. *For*

$$2 < \Omega_{k-1} \le 2(1+\sqrt{2}) \approx 4.828, \tag{3.11}$$

the SSP coefficient of (3.10) *is optimal, and is equal to* (3.9a).

Proof. By using (3.9), formula (3.10) takes the form

$$u_n = r\beta_{k-1}u_{n-1} + \beta_{k-1}h_n\, f(u_{n-1}) + (r\beta_0 + \delta_0)u_{n-k} + \beta_0 h_n\, f(u_{n-k}).$$

By definition (see [1, Chapter 8]), its SSP coefficient is given by

$$\mathcal{C}(\omega,\delta,\beta) = \min\left(\frac{r\beta_{k-1}}{\beta_{k-1}}, \frac{r\beta_0+\delta_0}{\beta_0}\right) \equiv \min\left(\frac{\Omega_{k-1}-2}{\Omega_{k-1}}, \frac{3\Omega_{k-1}+2}{\Omega_{k-1}(\Omega_{k-1}+1)}\right)$$

from which we see that

$$\mathcal{C}(\omega,\delta,\beta) = \begin{cases} \frac{\Omega_{k-1}-2}{\Omega_{k-1}} & \text{for } \ 2 \le \Omega_{k-1} \le 2(1+\sqrt{2}), \\ \frac{3\Omega_{k-1}+2}{\Omega_{k-1}(\Omega_{k-1}+1)} & \text{for } \ \Omega_{k-1} > 2(1+\sqrt{2}). \end{cases} \tag{3.12}$$

But Theorem 2 says that the SSP coefficient of any multistep formula with $p = 3$ can be at most $\frac{\Omega_k - 3}{\Omega_k - 1} \equiv \frac{\Omega_{k-1}-2}{\Omega_{k-1}}$, so for $2 < \Omega_{k-1} \le 2(1+\sqrt{2})$ the SSP coefficient of (3.10) is optimal. ∎

The natural requirement (3.11) also justifies our choice for k: in the fixed-step-size case we have $\Omega_{k-1} = k-1$, and $2 < k-1 \le 2(1+\sqrt{2})$ holds if and only if $k = 4$ or $k = 5$.

Remark 5. For $k \in \{4, 5\}$, we have the following strengthening of Proposition 1: the VSS formula (3.10) is optimal if and only if (3.11) holds. To see this, it is enough to show that (3.10) is not optimal for

$$\Omega_{k-1} > 2(1+\sqrt{2}). \tag{3.13}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Indeed, by fixing any $\Omega_{k-1} > 2(1+\sqrt{2})$, one checks by direct computation that

$$\frac{3\Omega_{k-1}+2}{\Omega_{k-1}(\Omega_{k-1}+1)} < r_j \quad (j = 0, 1, \ldots, 2k-3). \tag{3.14}$$

But the SSP coefficient of (3.10) is given by the left-hand side of (3.14) according to (3.12), and the optimal SSP coefficient for third-order formulas is $\min\limits_{0\le j\le 2k-3} r_j$ according to Theorem 4. Hence the SSP coefficient of (3.10) is not optimal when (3.13) holds.

Remark 6. One could develop optimal third-order explicit SSP formulas for $k > 5$ as well. However, their structure, as indicated by Theorem 4, would be more complicated, and the analysis performed in section 4.2 would become increasingly involved.

3.4. Higher-order formulas. The last theorem in this section reveals that arbitrarily high-order VSS SSP explicit linear multistep formulas exist, though they may require a large number of steps.

THEOREM 5. *Let $K_1, K_2 \ge 1$ be arbitrary and let p and $k > p^3K_1K_2/2$ be arbitrary positive integers. Suppose that ω_j are given and that*

$$1/K_1 \le \omega_j \le K_2 \quad \textit{for all } 1 \le j \le k. \tag{3.15}$$

Then there exists a k-step formula of order p with $\mathcal{C}(\omega, \delta, \beta) > 0$.

4. Step-size selection and asymptotic behavior of the step sizes.

To fully specify a method, we need not only a set of multistep formulas but also a prescription for the step size. When using a one-step SSP method to integrate a hyperbolic PDE, usually one chooses the step size $h_n := \gamma\,\mathcal{C}\,h_{\mathrm{FE}}(u_{n-1})$, where γ is a safety factor slightly less than unity. For SSP multistep methods, the choice of step size is more complicated. First, multiple previous steps must be taken into account when determining an appropriate h_{FE}, as already noted. But more significantly, the SSP coefficient $\mathcal{C}_n$ depends on the method coefficients, while the method coefficients depend on the choice of h_n. These coupled relations result in a step-size restriction that is a nonlinear function of recent step sizes. In this section we propose a greedy step-size selection algorithm and investigate the dynamics of the resulting step-size recursion for the formulas derived in the previous section.

Besides the step-size algorithms themselves, our main result will be that the step size remains bounded away from zero, so the computation is guaranteed to terminate. Because the step-size sequence is given by a recursion involving h_{FE}, we will at times require assumptions on h_{FE}:

$$\text{For all } n \text{ we have } \mu^- \le h_{\mathrm{FE}}(u_n) \le \mu^+ \text{ for some } \mu^\pm \in (0,\infty). \tag{4.1}$$

(4.2)

$$\text{For all } n \text{ we have } \varrho_{\mathrm{FE}} \le \frac{h_{\mathrm{FE}}(u_n)}{h_{\mathrm{FE}}(u_{n+1})} \le \frac{1}{\varrho_{\mathrm{FE}}} \text{ for some prescribed value } \varrho_{\mathrm{FE}} \in (0,1].$$

Assumption (4.1) states that the forward Euler permissible step size remains bounded and is also bounded away from zero. For stable hyperbolic PDE discretizations, this is very reasonable since it means that the maximum wave speed remains finite and nonzero. Assumption (4.2) states that the forward Euler step size changes little over a single numerical time step. Typically, this is reasonable since it is a necessary condition for the numerical solution to be accurate. It can easily be checked a posteriori.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

4.1. Second-order methods. Let us first analyze the three-step, second-order method in detail.

Set $k = 3$ in the second-order formula (3.2), and suppose that $h_j > 0$ has already been defined for $1 \leq j \leq n-1$ with some $n \geq 3$. The SSP step-size restriction (2.12) is implicit, since $\mathcal{C}_n$ depends on h_n. By (3.3) we have

$$\mathcal{C}_n = \frac{\omega_{1,n} + \omega_{2,n} - 1}{\omega_{1,n} + \omega_{2,n}} \equiv \frac{h_{n-2} + h_{n-1} - h_n}{h_{n-2} + h_{n-1}}.$$

Solving for h_n in (2.12) gives

$$h_n \leq \frac{h_{n-2} + h_{n-1}}{h_{n-2} + h_{n-1} + \mu_n} \cdot \mu_n.$$

It is natural to take the largest allowed step size, i.e., to define

$$h_n := \frac{h_{n-2} + h_{n-1}}{h_{n-2} + h_{n-1} + \mu_n} \cdot \mu_n. \tag{4.3}$$

For the general k-step, second-order formula (3.2), the same analysis leads to the following choice of step size, which guarantees monotonicity:

$$h_n := \frac{\sum_{j=1}^{k-1} h_{n-j}}{\left(\sum_{j=1}^{k-1} h_{n-j}\right) + \mu_n} \cdot \mu_n. \tag{4.4}$$

Note that this definition automatically ensures $\Omega_{k-1,n} > 1$, and hence $\mathcal{C}_n > 0$ for any $\mu_n > 0$.

4.1.1. Asymptotic behavior of the step size. Since (4.4) is a nonlinear recurrence, one might wonder if the step size could be driven to zero, preventing termination of the integration process. The following theorem shows that, under some natural assumptions, this cannot happen.

THEOREM 6. *Consider the solution of* (2.1) *by the second-order formula* (3.2) *with some $k \geq 3$. Let the initial $k-1$ step sizes be positive and let the subsequent step sizes h_n be chosen according to* (4.4). *Assume that* (4.1) *holds with some constants $\mu^{\pm}$. Then the step-size sequence h_n satisfies*

$$\frac{k-2}{k-1}\,\mu^- \leq \liminf_{n\to+\infty} h_n \leq \limsup_{n\to+\infty} h_n \leq \frac{k-2}{k-1}\,\mu^+. \tag{4.5}$$

As a special case, if $h_{\mathrm{FE}}(u_n)$ is constant, then

$$h_n \to \frac{k-2}{k-1}\,h_{\mathrm{FE}} \quad (n \to +\infty).$$

Remark 7. The asymptotic step size $\frac{k-2}{k-1}\,h_{\mathrm{FE}}$ given above is precisely the allowable step size for the fixed-step-size SSP method of k steps.

The proof of Theorem 6 is given in section 9.

Remark 8. Our greedy step-size selection (4.4) for second-order methods is optimal in the following sense. Let us assume that there is another step-size sequence, say h_n^-, with the following properties:

- The corresponding starting values are equal, that is, $h_j^- = h_j > 0$ for $j = 1, 2, \ldots, k-1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

- The μ_n quantities in (2.14) corresponding to the sequences h_n and h_n^- are all equal to a fixed common constant $\mu > 0$.
- The sequence h_n^- satisfies (4.4) with inequality, that is,

$$h_n^- \le \frac{\sum_{j=1}^{k-1} h_{n-j}^-}{\left(\sum_{j=1}^{k-1} h_{n-j}^-\right) + \mu} \cdot \mu.$$

Then—as a straightforward modification of the proof of (9.8) in section 9.2 shows—we have $h_n^- \le h_n$ for all $n \ge 1$.

4.2. Third-order methods. In this section we give our step-size selection algorithm for the third-order four-step and five-step SSP formulas (3.10) by following the same approach as in section 4.1.

By using the optimal SSP coefficient $\mathcal{C}_n = \frac{\Omega_{k-1,n}-2}{\Omega_{k-1,n}}$ given in Proposition 1 we see that $h_n = \mathcal{C}_n \mu_n$ in (2.12) if and only if

$$h_n = \frac{\sum_{j=1}^{k-1} h_{n-j}}{\left(\sum_{j=1}^{k-1} h_{n-j}\right) + 2\mu_n} \cdot \mu_n.$$

This relation also yields that

$$\Omega_{k-1,n} = 2 + \frac{1}{\mu_n} \sum_{j=1}^{k-1} h_{n-j},$$

so $\Omega_{k-1,n} > 2$ is guaranteed by $\mu_n > 0$. Therefore, (3.11) is equivalent to

$$\sum_{j=1}^{k-1} h_{n-j} \le \sqrt{8}\, \mu_n. \tag{4.6}$$

The definition of h_n in Theorem 7 below is based on these considerations. The assumptions of the theorem on the starting values and on the problem (involving the boundedness of the h_{FE} quantities and that of their ratios) are constructed to ensure (4.6). As a conclusion, Theorem 7 uses the maximum allowable SSP step size (2.12) together with the optimal SSP coefficient $\mathcal{C}_n > 0$.

THEOREM 7. *Consider the solution of* (2.1) *by the third-order formula* (3.10) *with $k = 4$ or $k = 5$. Let the initial $k - 1$ step sizes be positive and let the subsequent step sizes h_n be chosen according to*

$$h_n := \frac{\sum_{j=1}^{k-1} h_{n-j}}{\left(\sum_{j=1}^{k-1} h_{n-j}\right) + 2\mu_n} \cdot \mu_n. \tag{4.7}$$

Assume that (4.1)–(4.2) *and the condition*

$$0 < h_j \le \varrho \cdot h_{\mathrm{FE}}(u_j) \quad \text{for } j = 1, \ldots, k-1 \tag{4.8}$$

hold with

$$(\varrho, \varrho_{\mathrm{FE}}) = \begin{cases} \left(\frac{6}{10}, \frac{9}{10}\right) & \text{for } k = 4, \\ \left(\frac{57}{100}, \frac{962}{1000}\right) & \text{for } k = 5. \end{cases} \tag{4.9}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Then the step-size sequence h_n satisfies

$$\frac{k-3}{k-1}\,\mu^- \le \liminf_{n\to+\infty} h_n \le \limsup_{n\to+\infty} h_n \le \frac{k-3}{k-1}\,\mu^+. \tag{4.10}$$

As a special case, if $h_{\mathrm{FE}}(u_n)$ is constant, then

$$h_n \to \frac{k-3}{k-1}\,h_{\mathrm{FE}} \quad (n\to+\infty).$$

The proof of this theorem is given in section 9. From the proof we will see that it is possible to slightly adjust the simple $(\varrho, \varrho_{\mathrm{FE}})$ values given in (4.9) based on properties of the problem to be solved; see Figure 5.

Remark 9. Our greedy step-size selection (4.7) for the third-order methods in the above theorem is optimal in the same sense as it is described in Remark 8.

5. Stability and convergence. The conditions for a method to have a positive SSP coefficient are closely related to sufficient conditions to ensure stability and convergence.

Recall that an LMM is said to be zero stable if it produces a bounded sequence of values when applied to the initial value problem

$$u'(t) = 0, \qquad u(t_0) = u_0, \qquad t \in [t_0, t_0+T] \tag{5.1}$$

(see, e.g., [2, section III.5]). The following result shows that VSS SSP methods are zero stable as long as the step-size restriction for monotonicity is observed.

THEOREM 8. *Let $\alpha_{j,n}$ be the coefficients of a VSS LMM* (2.3) *and suppose that $\alpha_{j,n} \ge 0$ for all n, j. Then the method is zero stable.*

Proof. Application of (2.3) to (5.1) yields the recursion

$$u_n = \sum_{j=0}^{k-1} \alpha_{j,n} u_{n-k+j}.$$

Since the coefficients $\alpha_{j,n}$ are nonnegative and sum to one, each solution value u_n is a convex combination of previous solution values. Hence, the sequence is bounded. ∎

COROLLARY 1. *Let an SSP VSS LMM be given and suppose that the step sizes are chosen so that $\mathcal{C}_n > 0$ for each n. Then the method is zero stable.*

In order to prove convergence, we must also bound the local error by bounding the ratio of successive step sizes and bounding the method coefficients.

LEMMA 1. *For method* (3.2) *with $k \ge 3$ or method* (3.10) *with $k = 4$ or 5, let the step sizes be chosen so that $\mathcal{C}_n > 0$. Then there exists a constant Λ such that*

$$0 \le \alpha_{j,n}, \beta_{j,n} < \Lambda \quad \textit{for all } n.$$

Proof. For any method, $\mathcal{C}_n > 0$ implies that $\beta_{j,n} \ge 0$ and $0 \le \alpha_{j,n} \le 1$. For the second-order methods, $\mathcal{C}_n > 0$ implies $\Omega_{k-1,n} > 1$, which implies $\beta_{k-1,n} < 2$. For the third-order methods, $\mathcal{C}_n > 0$ implies $\Omega_{k-1,n} > 2$, which implies $\beta_{k-1,n} < 9/4$ and $\beta_{0,n} < 3/4$. ∎

We recall the following result from [2, section III.5, Theorem 5.8].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THEOREM 9. *Let a VSS LMM be applied to an initial value problem with a given f and on a time interval $[t_0, t_0 + T]$. Let $h = \max_n h_n$ be the largest step in the step sequence. Assume that*

1. *the method is stable, of order p, and the coefficients $\alpha_{j,n}, \beta_{j,n}$ are bounded uniformly as $h \to 0$;*
2. *the starting values satisfy $\|u(t_j) - u_j\| = \mathcal{O}(h_0^p)$, where h_0 is a bound on the starting step sizes;*
3. *$h_n/h_{n-1} \leq \eta$, where η is independent of n and h.*

Then the method is convergent of order p; i.e., $\|u(t_n) - u_n\| = \mathcal{O}(h^p)$ for all $t_n \in [t_0, t_0 + T]$.

The methods described in section 4 satisfy the conditions of Theorem 9, and are thus convergent, as shown in the following theorem.

THEOREM 10. *Under the assumptions of Theorem* 6*, our methods defined in section* 4.1 *and with $\|u(t_j) - u_j\| = \mathcal{O}(h_0^2)$ for $j \leq k$ are convergent of order* 2*. Similarly, under the assumptions of Theorem* 7*, our methods defined in section* 4.2 *and with $\|u(t_j) - u_j\| = \mathcal{O}(h_0^3)$ for $j \leq k$ are convergent of order* 3.

Proof. It is sufficient to show that assumptions 1 and 3 of Theorem 9 are fulfilled. Our construction in sections 4.1 and 4.2 guarantees $\mathcal{C}_n > 0$. Therefore Lemma 1 applies, so the coefficients $\alpha_{j,n}, \beta_{j,n}$ are uniformly bounded and nonnegative. Thus Theorem 8 applies, so the methods are stable. Furthermore, the condition $h_n/h_{n-1} \leq \eta < +\infty$ is implied by (4.5) or (4.10), hence Theorem 9 is applicable. ∎

6. Numerical examples. In this section we investigate the performance of the proposed methods by performing numerical tests. In section 6.4, the accuracy of our methods is verified by a convergence test on a linear equation with time-varying advection velocity. In sections 6.5–6.8 we apply the methods SSPMSV32 and SSPMSV43 to nonlinear hyperbolic conservation laws in one and two dimensions.

All code used to generate the numerical results in this work is available at https://github.com/numerical-mathematics/ssp-lmm-vss_RR.

6.1. Efficiency. Let N denote the number of steps (excluding the starting steps) used to march with a k-step VSS method from time t_0 to a final time $t_0 + T$. If the same method were used with a fixed step size, the number of steps required would be at least $N' = (T - \sum_{j=1}^{k-1} h_j)/h_{\min}$, where $h_{\min}$ is the smallest step size used by the VSS method. Thus the reduction in computational cost by allowing a VSS is given by

$$s := \frac{N}{N'} = \frac{h_{\min}}{h_{\text{avg}}},$$

where $h_{\text{avg}} := (T - \sum_{j=1}^{k-1} h_j)/N$ is the mean step size used by the VSS method.

6.2. Spatial discretization. In space, we use the wave-propagation semidiscretizations described in [7]. For the second-order temporal schemes, we use a spatial reconstruction based on the total-variation-diminishing (TVD) monotonized-central-difference limiter [14]; for the third-order temporal schemes, we use a fifth-order WENO reconstruction. The only exception is the Woodward–Colella problem, where we use the TVD spatial discretization for both second- and third-order methods. The second-order spatial semidiscretization is provably TVD (for scalar, one-dimensional problems) under forward Euler integration with a CFL number of $\nu_{\text{FE}} = 1/2$. We use this value in the step-size selection algorithm for all methods and all problems.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Algorithm 1. Second-order time-stepping algorithm.

Initialization: Choose an initial step size h_1 and compute the forward Euler step size $h_{\mathrm{FE}}(u_0)$.

Starting procedure:
for $n = 1, 2, \ldots, k-1$ **do**
  Compute u_n using the two-stage, second-order SSP Runge–Kutta method with step size h_n.
  Find the CFL number ν_n and compute $h_{\mathrm{FE}}(u_n)$.
  Set $h \leftarrow \gamma\, \mathcal{C}_0\, h_{\mathrm{FE}}(u_n)$.
  if $\nu_n > \mathcal{C}_0\, \nu_{\mathrm{FE}}$ **then**
    Set $h_n \leftarrow h$ and repeat the step.
  else if $n < k-1$ **then**
    Set $h_{n+1} \leftarrow h$.
  else
    Set h_k based on (4.4).
  end if
end for

Main method:
for $n = k, \ldots, N$ **do**
  Compute u_n using the method (3.2) with step size h_n.
  Compute h_{n+1} from (4.4).
end for

6.3. Time-stepping algorithm. The complete time-stepping algorithm is given in Algorithms 1 and 2. We denote by ν_n the CFL number at step n. By default we take the initial step size $h_1 = 0.1$. In order to ensure monotonicity of the starting steps, we use the two-stage, second-order SSP Runge–Kutta method to compute them, with step size

$$h_n := \gamma\, \mathcal{C}_0\, h_{\mathrm{FE}}(u_{n-1}), \quad n = 1, \ldots, k-1,$$

where $\mathcal{C}_0 = 1$ is the SSP coefficient of the Runge–Kutta method and $\gamma = 0.9$ is a safety factor (cf. section 4).

For our third-order methods we can check conditions (4.2) and (4.8) only a posteriori. If condition (4.2) or (4.8) is violated, then the computed solution is discarded and the current step is repeated with a smaller step size as described in Algorithm 2. These step-size reductions can be repeated if necessary, but in our computations the first reduction was always sufficient so that the new step was accepted. For the numerical examples of this section, the step size and CFL number corresponding to the starting methods are shown on the right of Figures 1 and 2.

Remark 10. There is no need to check the CFL condition $\nu_n \leq \mathcal{C}_0\, \nu_{\mathrm{FE}}$ for $n \geq k$ (i.e., when the LMM method is used) since it is automatically satisfied by the step-size selection (2.12).

Condition (4.2) is violated only when the maximum wave speed changes dramatically between consecutive steps, which may suggest insufficient temporal resolution of the problem. We have found that, for the problems considered herein, omitting the enforcement of condition (4.8) never seems to change the computed solution in a

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Algorithm 2. Third-order time-stepping algorithm.

Initialization: Choose an initial step size h_1 and compute the forward Euler step size $h_{\mathrm{FE}}(u_0)$.

Starting procedure:
 for $n = 1, 2, \ldots, k-1$ **do**
  Compute u_n using the two-stage, second-order SSP Runge–Kutta method with step size h_n.
  Find the CFL number ν_n and compute $h_{\mathrm{FE}}(u_n)$.
  if condition (4.2) or (4.8) does not hold **then** set

$$h_n \leftarrow \begin{cases} h_n/2 & \text{if at least (4.2) is violated,} \\ \gamma\, \mathcal{C}_0\, \varrho\, h_{\mathrm{FE}}(u_n) & \text{if (4.8) is violated,} \end{cases}$$

   and repeat the step.
  else
   Set $h \leftarrow \gamma\, \mathcal{C}_0\, h_{\mathrm{FE}}(u_n)$.
  end if
  if $\nu_n > \mathcal{C}_0\, \nu_{\mathrm{FE}}$ **then**
   Set $h_n \leftarrow h$ and repeat the step.
  else if $n < k-1$ **then**
   Set $h_{n+1} \leftarrow h$.
  else
   Set h_k based on (4.7).
  end if
 end for

Main method:
 for $n = k, \ldots, N$ **do**
  Compute u_n using the method (3.10) with step size h_n.
  if condition (4.2) is violated **then**
   Set $h_n \leftarrow h_n/2$ and repeat step.
  else
   Compute h_{n+1} from (4.7).
  end if
 end for

significant way. Since these two conditions were introduced only as technical assumptions for some of our theoretical results, they could perhaps be omitted in a practical implementation.

6.4. Convergence test. We consider the linear advection problem

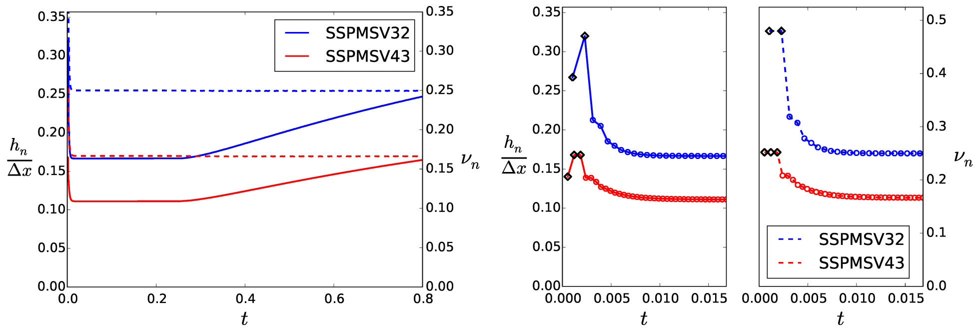
$$u_t + \left(2 + \frac{3}{2}\sin(2\pi t)\right) u_x = 0, \qquad (6.1)$$
$$u(x, 0) = \sin(2\pi x), \quad x \in [0, 1],$$

with periodic boundary conditions. We use a spatial step size $\Delta x = 2^{-d}$, $d = 6, \ldots, 11$, and compute the solution at a final time $t = 5$ (i.e., after 10 cycles). Table 1 shows

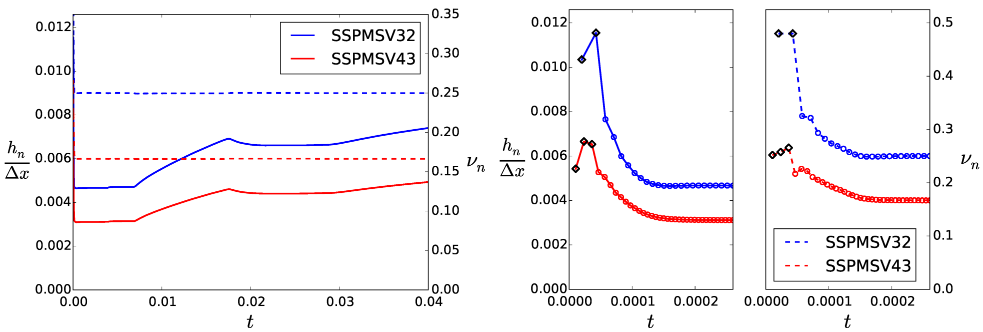
Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php



(a) Burgers' equation



(b) Woodward–Colella blast-wave problem

FIG. 1. *Evolution of step size for one-dimensional Burgers' and Euler equations with second- and third-order LMMs (solid lines). The dashed curves indicate the CFL number for each method. Figures on the right show a close-up view of the step size and CFL number at early times. The black marks indicate the starting Runge–Kutta method's step sizes.*

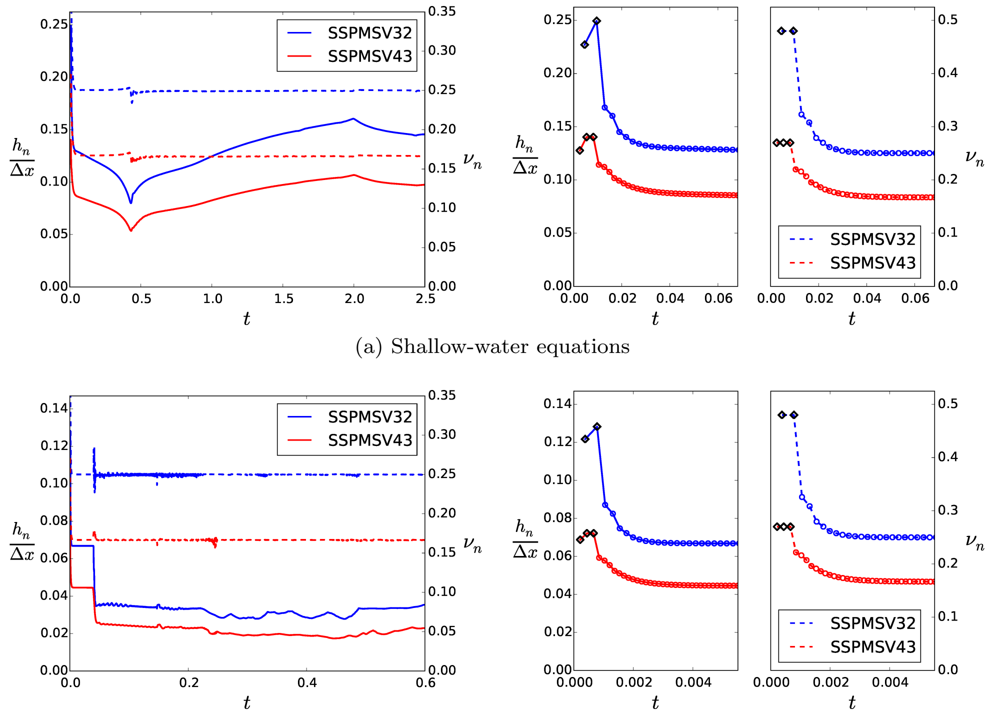
the L_1-norm of the error at the final time. The solution is computed by using second-order methods with $k = 3, 4$ steps and third-order methods with $k = 4, 5$ steps. All methods attain the expected order.

6.5. Burgers' equation. We consider the inviscid Burgers' initial value problem

$$u_t + uu_x = 0,$$
$$u(x, 0) = \frac{1}{2} + \sin(2\pi x), \quad x \in [0, 1],$$

with periodic boundary conditions and 256 grid cells. Figure 1(a) shows the evolution of the step size, up to a final time $t = 0.8$. The step size is constant until the shock forms, and then increases. The dashed curves in Figure 1(a) indicate how the CFL number ν_n varies with time for each method. Since $\nu_n = \mathcal{C}_n \nu_{\mathrm{FE}}$ and $\nu_{\mathrm{FE}} = 1/2$, the CFL number after the first few steps is $1/4$ and $1/6$, for the second- and third-order LMMs, respectively. Note that these are the theoretical maximum values for which the solution remains TVD. For the third-order method coupled with WENO discretization the TVD-norm of the solution does not increase more than 10^{-4} over a single time step. The sudden decrease of the step size after the first few steps of the simulation is due to the switch from the starting (Runge–Kutta) method to the LMM. For both methods we have $s \approx 0.88$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.



(a) Shallow-water equations

(b) Shock-bubble interaction problem

FIG. 2. *Evolution of step size for two-dimensional shallow-water and Euler equations with second- and third-order LMMs (solid lines). The dashed curves indicate the CFL number for each method. Figures on the right show a close-up view of the step size and CFL number at early times. The black marks indicate the starting Runge–Kutta method's step sizes.*

TABLE 1

L_1-norms of the error at final time $t = 5$ for the variable-coefficient advection problem (6.1) *for second- and third-order LMMs. For each method, the second column denotes the convergence order. The number of spatial points is indicated by N.*

N	SSPMSV32		SSPMSV42		SSPMSV43		SSPMSV53	
128	1.50×10^{-2}		1.83×10^{-2}		9.20×10^{-6}		6.08×10^{-5}	
256	4.30×10^{-3}	1.80	5.34×10^{-3}	1.78	1.30×10^{-6}	2.82	8.10×10^{-6}	2.91
512	1.15×10^{-3}	1.90	1.44×10^{-3}	1.89	1.68×10^{-7}	2.95	1.04×10^{-6}	2.96
1024	3.01×10^{-4}	1.93	3.81×10^{-4}	1.92	2.13×10^{-8}	2.98	1.32×10^{-7}	2.98
2048	7.74×10^{-5}	1.96	9.84×10^{-5}	1.95	2.67×10^{-9}	2.99	1.66×10^{-8}	2.99

6.6. Woodward–Colella blast-wave problem. Next we consider the one-dimensional Euler equations for inviscid, compressible flow:

$$\begin{aligned} \rho_t + (\rho u)_x &= 0, \\ (\rho u)_t + (\rho u^2 + p)_x &= 0, \\ E_t + (u(E+p))_x &= 0. \end{aligned}$$

Here ρ, u, E, and p denote the density, velocity, total energy, and pressure, respectively. The fluid is an ideal gas, thus the pressure is given by $p = \rho(\gamma - 1)e$, where e is internal energy and $\gamma = 1.4$. The domain is the unit interval with $\Delta x = 1/512$,

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and we solve the Woodward–Colella blast-wave problem [15] with initial conditions

$$\rho(x,0) = 1, \quad v(x,0) = 0, \quad \text{and} \quad p(x,0) = \begin{cases} 1000 & \text{if } 0 \le x < 0.1, \\ 0.01 & \text{if } 0.1 \le x < 0.9, \\ 100 & \text{if } 0.9 \le x \le 1. \end{cases}$$

The initial conditions consist of two discontinuities. Reflecting boundary conditions are used and initial conditions lead to strong shock waves, and rarefactions that eventually interact with each other. The second-order TVD semidiscretization is used with both the second- and third-order methods and the solution is computed at time $t = 0.04$, after the two shock waves have collided.

Figure 1(b) shows how the step size evolves over time. The maximum wave speed decreases as the shock waves approach each other, then increases when they collide, and finally decreases when the shocks move apart. The step size exhibits exactly the opposite behavior since it is inversely proportional to the maximum wave speed. As before, the CFL number ν_n remains close to 1/4 and 1/6 for the second- and third-order LMMs, respectively. Note that these are the theoretical maximum values for which the characteristic variables remain TVD. For this problem $s \approx 0.76$.

6.7. Two-dimensional shallow-water flow. Consider the two-dimensional shallow-water equations

$$\begin{aligned} h_t + (hu)_x + (hv)_y &= 0, \\ (hu)_t + (hu^2 + \frac{1}{2}gh^2)_x + (huv)_y &= 0, \\ (hv)_t + (huv)_x + (hv^2 + \frac{1}{2}gh^2)_y &= 0, \end{aligned}$$

where h is the depth, (u, v) the velocity vector, and hu, hv the momenta in each direction. The gravitational constant g is equal to unity. The domain is a square $[-2.5, 2.5] \times [-2.5, 2.5]$ with 250 grid cells in each direction. The initial condition consists of an uplifted, inward-flowing cylindrically symmetric perturbation given by

$$h(x, y, 0) = 1 + g(x, y), \quad u(x, y, 0) = -xg(x, y), \quad \text{and} \quad v(x, y, 0) = -yg(x, y),$$

where $g(x, y) := e^{-10\left(\sqrt{x^2+y^2}-1\right)^2}$. We apply reflecting boundary conditions at the top and right boundary, whereas the bottom and left boundaries are outflow. The wave is initially directed towards the center of the domain, and a large peak forms as it converges there, before subsequently expanding outward. We compute the solution up to time $t = 2.5$, after the solution has been reflected from the top and right boundaries. Figure 2(a) shows how the step size varies in time. It decreases as the initial profile propagates towards the center of the domain since the maximum wave speed increases. As the solution expands outwards, higher step sizes are allowed. Note that at time $t \approx 2$ the wave hits the boundaries, so a small decrease in the allowed step size is observed. Again, the dashed curves indicate the variation of the CFL number. Now we have the ratio $s \approx 0.62$ for both second- and third-order methods.

6.8. Shock-bubble interaction. Finally, we consider the Euler equations of compressible fluid dynamics in cylindrical coordinates. Originally, the problem is three dimensional but can be reduced to two dimensions by using cylindrical symmetry. The resulting system of equations is

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$
\begin{aligned}
\rho_t + (\rho u)_z + (\rho v)_r &= -\frac{\rho v}{r}, \\
(\rho u)_t + (\rho u^2 + p)_z + (\rho uv)_r &= -\frac{\rho uv}{r}, \\
(\rho v)_t + (\rho uv)_z + (\rho v^2 + p)_r &= -\frac{\rho v^2}{r}, \\
(\rho E)_t + (u(\rho E + p))_z + (v(\rho E + p))_r &= -\frac{(\rho E + p)v}{r}.
\end{aligned}
$$

Here, ρ is the density, p is the pressure, E is the total energy, while u and v are the z- and r-components of the velocity. The z- and r-axes are parallel and perpendicular, respectively, to the axis of symmetry. The problem involves a planar shock wave traveling in the z-direction that impacts a spherical low-density bubble. The initial conditions are taken from [6]. We consider a cylindrical domain $[0, 2] \times [0, 0.5]$ using a 640×160 grid and impose reflecting boundary conditions at the bottom of the domain and outflow conditions at the top and right boundaries.

Since the fluid is a polytropic ideal gas, the internal energy e is proportional to the temperature. Initially the bubble is at rest, hence, there is no difference in pressure between the bubble and the surrounding air. Therefore, the temperature inside the bubble is high resulting in large sound speed $c = \sqrt{\gamma(\gamma - 1)e}$, where $\gamma = 1.4$. Consequently, this results in reduced step sizes right after the shock wave hits the bubble at $t \approx 0.05$, as shown in Figure 2(b). The efficiency ratio s is about 0.82.

7. Conclusions. The methods presented in this paper are the first adaptive multistep methods that provably preserve any convex monotonicity property satisfied under explicit Euler integration. We have provided a step-size selection algorithm—yielding an optimal step-size sequence h_n—that strictly enforces this property while ensuring that the step size is bounded away from zero. The methods are proved to converge at the expected rate for any ODE system satisfying the forward Euler condition.

As suggested by Theorems 6 and 7, for all tests the CFL number remains close to $\frac{k-p}{k-1}\nu_{\mathrm{FE}}$, where k and p are the steps and order of the method, respectively. The numerical results verify that the step size is approximately inversely proportional to the maximum wave speed and the proposed step-size strategy successfully chooses step sizes closely related to the maximum allowed CFL number. In practice, we expect that conditions (4.2) and (4.8), which were introduced only as technical assumptions for some of our theoretical results, could usually be omitted without impacting the solution.

The methods presented herein are of orders two and three. We have proved the existence of methods of arbitrary order. However, the optimal methods for orders at least 3 seem to have a complicated structure. It would be useful to develop adaptive (and possibly suboptimal) multistep SSP methods of order higher than three having a simple structure.

8. The proofs of the theorems in section 3. We present three lemmas, Lemma 2 below, Lemma 3 in section 8.1, and Lemma 4 in section 8.2, that are used in the proofs of Theorems 2–5 in sections 8.1–8.4. Lemmas 2–4 are straightforward generalizations of the corresponding results of [11] for VSS formulas, hence their proofs are omitted here.

LEMMA 2. *Let $r \geq 0$ be arbitrary and let p and $n \geq k$ be arbitrary positive integers. Suppose that some time-step ratios $\omega_{j,n} > 0$ are given. Then the following two statements are equivalent.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(i) *For all formulas with k steps and order of accuracy p we have $\mathcal{C}_n(\omega,\delta,\beta)\le r$.*
(ii) *There exists a nonzero real univariate polynomial q_n of degree at most p, that satisfies the conditions*

$$q_n(\Omega_{j,n})\ge 0 \qquad \textit{for all } 0\le j\le k-1, \tag{8.1a}$$
$$q_n'(\Omega_{j,n})+rq_n(\Omega_{j,n})\ge 0 \qquad \textit{for all } 0\le j\le k-1, \tag{8.1b}$$
$$q_n(\Omega_{k,n})=0. \tag{8.1c}$$

Furthermore, if statement (i) *holds, then the polynomial q_n can be chosen to satisfy the following conditions.*
1. *The degree of q_n is exactly p.*
2. *The real parts of all roots of q_n except for $\Omega_{k,n}$ lie in the interval $[\Omega_{0,n},\Omega_{k-1,n}]$.*
3. *The set of all real roots of q_n is a subset of $\{\Omega_{0,n},\Omega_{1,n},\dots,\Omega_{k,n}\}$.*
4. *For even p, $\Omega_{0,n}$ is a root of q_n with odd multiplicity. For odd p, if $\Omega_{0,n}$ is a root of q_n, then its multiplicity is even.*
5. *The multiplicity of the root $\Omega_{k,n}$ is one. For $1\le j\le k-1$, if $\Omega_{j,n}$ is a root of q_n, then its multiplicity is* 2.
6. *The polynomial q_n is nonnegative on the interval $[\Omega_{0,n},\Omega_{k,n}]$.*

The following observation will be useful in the proofs. If $q_n(\Omega_{j,n})=0$, then the inequality (8.1b) for this index j simplifies to $q_n'(\Omega_{j,n})\ge 0$. Otherwise, if $q_n(\Omega_{j,n})\ne 0$, then the inequality (8.1b) for this index j can be written by using the logarithmic derivative of q_n as

$$\frac{q_n'(\Omega_{j,n})}{q_n(\Omega_{j,n})}=\sum_{\ell=1}^{p}\frac{1}{\Omega_{j,n}-\lambda_{\ell,n}}\ge -r,$$

where the $\lambda_{\ell,n}$ numbers are the complex (including real) roots of q_n, and q_n has been chosen such that its degree is exactly p.

8.1. The proof of Theorem 2. The following lemma will be used in the second step of the proof of Theorem 2.

LEMMA 3. *Let $r>0$ be arbitrary and let p and $n\ge k$ be arbitrary positive integers. Suppose that some time-step ratios $\omega_{j,n}>0$ are given. Then exactly one of the following statements is true.*
(i) *There is a formula with k steps, order of accuracy p and $\mathcal{C}_n(\omega,\delta,\beta)\ge r$.*
(ii) *There is a real univariate polynomial q_n of degree at most p, that satisfies the conditions* (8.1a), (8.1b), *and*

$$q_n(\Omega_{k,n})<0. \tag{8.2}$$

The proof of Theorem 2. *Step* 1. To prove the bound (3.1), we first consider the $k\ge 2$ case, and let q_n denote the polynomial

$$q_n(x):=x^{p-1}(\Omega_{k,n}-x). \tag{8.3}$$

• If $\Omega_{k,n}>p$, this q_n satisfies conditions (8.1) with $r:=(\Omega_{k,n}-p)/(\Omega_{k,n}-1)$, since for each $1\le j\le k-1$ we have

$$\frac{q_n'(\Omega_{j,n})}{q_n(\Omega_{j,n})}=\frac{p-1}{\Omega_{j,n}}-\frac{1}{\Omega_{k,n}-\Omega_{j,n}}\ge\frac{p-1}{\Omega_{k-1,n}}-\frac{1}{\Omega_{k,n}-\Omega_{k-1,n}}=-\frac{\Omega_{k,n}-p}{\Omega_{k,n}-1},$$

and $q_n'(\Omega_{0,n})\equiv q_n'(0)\ge 0$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

• If $\Omega_{k,n} \le p$, the conditions (8.1) hold with $r := 0$, because for $x \in [\Omega_{0,n}, \Omega_{k-1,n}] \equiv [0, \Omega_{k,n} - 1]$ we have

$$q_n'(x) \ge x^{p-2}\left((p-1)\Omega_{k,n} - p\left(\Omega_{k,n} - 1\right)\right) = x^{p-2}\left(p - \Omega_{k,n}\right) \ge 0.$$

For $k = 1$, the same q_n as in (8.3) satisfies (8.1) with $r := 0$ and $\Omega_{k,n} = 1$. Then, in each case, (3.1) is an immediate consequence of Lemma 2.

Step 2. Suppose now

$$\exists \text{ a } k\text{-step explicit linear multistep formula of order } p \ge 2 \text{ with } \mathcal{C}_n(\omega, \delta, \beta) > 0. \tag{8.4}$$

We prove that there is a k-step formula of order p whose SSP coefficient is equal to the optimal one.

Let us set $\mathcal{H} := \{r > 0 : \text{the statement (ii) of Lemma 3 holds}\}$. Step 1 implies that the largest possible $\mathcal{C}_n(\omega, \delta, \beta)$ for all k-step formulas of order p and having the given time-step ratios ω_j is finite. So $\mathcal{H} \ne \varnothing$ by Lemma 3. We also see that $\mathcal{H}$ is an infinite interval, because if a polynomial q_n satisfies the conditions (8.1a), (8.1b) with some $r := \rho > 0$, and (8.2), then it satisfies the same conditions with any $r \in (\rho, +\infty)$. Thus, with a suitable $r^* \ge 0$, we have $(r^*, +\infty) \subseteq \mathcal{H} \subseteq [r^*, +\infty)$. Clearly, $r^* > 0$ due to assumption (8.4) and Lemma 3. We claim that $r^* \notin \mathcal{H}$.

Suppose to the contrary that $r^* \in \mathcal{H}$. Then there is a polynomial q_n satisfying the conditions (8.1a), (8.1b) with $r := r^*$, and (8.2). Now we define $\widetilde{q}_n := q_n + |q_n(\Omega_{k,n})|/2$. One easily checks that this $\widetilde{q}_n$ polynomial satisfies the same set of conditions with

$$r := \left(\frac{1 + \max\limits_{0 \le j \le k-1} q_n(\Omega_{j,n})}{1 + |q_n(\Omega_{k,n})|/2 + \max\limits_{0 \le j \le k-1} q_n(\Omega_{j,n})} \right) \cdot r^* \in (0, r^*),$$

so we would get $(0, r^*) \cap \mathcal{H} \ne \varnothing$, a contradiction.

Hence $\mathcal{H} = (r^*, +\infty)$. Therefore, in view of Lemma 3, r^* is the optimal SSP coefficient and there exists an optimal method with $\mathcal{C}_n(\omega, \delta, \beta) = r^*$. □

8.2. The proof of Theorem 3. In the following we apply the usual terminology and say that an inequality constraint is *binding* if the inequality holds with equality. The lemma below is used in the proofs of Theorems 3 and 4.

LEMMA 4. *Let $n \ge k$ be arbitrary positive integers and $p \ge 2$. Suppose that the time-step ratios $\omega_{j,n} > 0$ are given, and that there exists an explicit linear multistep formula with k steps, order of accuracy p, and $\mathcal{C}_n > 0$. Let $\delta_{j,n}$ and $\beta_{j,n}$ denote the coefficients of a formula with the largest SSP coefficient $\mathcal{C}_n(\omega_n, \delta_n, \beta_n) \in (0, +\infty)$. Let q_n be a polynomial that satisfies all the conditions of Lemma 2 with $r := \mathcal{C}_n(\omega_n, \delta_n, \beta_n)$. Then the following statements hold.*

(i) *If $\delta_{j,n} \ne 0$ for some $0 \le j \le k-1$, then $q_n(\Omega_{j,n}) = 0$.*

(ii) *If $\beta_{j,n} \ne 0$ for some $0 \le j \le k-1$, then $q_n'(\Omega_{j,n}) + r q_n(\Omega_{j,n}) = 0$.*

(iii) *This q_n can be chosen so that the total number of binding inequalities in* (8.1a)–(8.1b) *is at least p.*

The proof of Theorem 3. The necessity of the condition $\Omega_{k,n} > 2$ for the existence of a second-order formula with positive SSP coefficient is an immediate consequence of Theorem 2. On the other hand, we easily see that, up to a positive multiplicative constant, $q_n(x) = x(\Omega_{k,n} - x)$ is the unique polynomial satisfying (8.1a), (8.1c), and Properties 1 and 4 of Lemma 2. If $\Omega_{k,n} > 2$, then q_n does not satisfy (8.1b) with

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$r := 0$ for $j = k-1$; therefore there exists a formula with $\mathcal{C}_n > 0$ due to Lemma 2. Thus the condition $\Omega_{k,n} > 2$ is also sufficient.

From now on we assume $\Omega_{k,n} > 2$, and that a formula with the optimal SSP coefficient $\mathcal{C}_n(\omega_n, \delta_n, \beta_n)$ is considered. From the above form of q_n we see that $q_n(\Omega_{j,n}) = 0$ $(0 \le j \le k-1)$ holds only for $j = 0$, so the statement (iii) of Lemma 4 guarantees that at least one of the inequalities in (8.1b) is binding with $r := \mathcal{C}_n(\omega_n, \delta_n, \beta_n)$. Since $q_n'(\Omega_{0,n}) + rq_n(\Omega_{0,n}) = \Omega_{k,n} > 0$ and for all $1 \le j \le k-2$ we have

$$\frac{q_n'(\Omega_{j,n})}{q_n(\Omega_{j,n})} = \frac{1}{\Omega_{j,n}} - \frac{1}{\Omega_{k,n} - \Omega_{j,n}} > \frac{1}{\Omega_{k-1,n}} - \frac{1}{\Omega_{k,n} - \Omega_{k-1,n}} = -\frac{\Omega_{k-1,n} - 1}{\Omega_{k-1,n}},$$

the unique binding inequality in (8.1b) must be the one corresponding to the index $j = k-1$. But $q_n'(\Omega_{k-1,n}) + rq_n(\Omega_{k-1,n}) = 0$ implies $r = (\Omega_{k,n} - 2) / (\Omega_{k,n} - 1)$.

Finally, the statements (i) and (ii) of Lemma 4 guarantee that only $\delta_{0,n}$ and $\beta_{k-1,n}$ can differ from zero. Solving the conditions for order two (see (2.17)) then yields the unique formula stated in (3.2). ∎

8.3. The proof of Theorem 4. First we present a lemma that will be used in the proof of Theorem 4. Moreover, statement 2 of the lemma proves our claim about the unique real root of the polynomial P_{k+j} in (3.5) under assumption (3.7), hence guaranteeing that the quantity r_{k+j} in (3.8) has a proper definition. The Δ_j quantities defined in (3.6) clearly satisfy the condition $1 < \Delta_{j+1} < \Delta_j$ below.

LEMMA 5. *Let some $1 < \Delta_{j+1} < \Delta_j$ numbers be given, and for any $\rho \in \mathbb{R}$ let*

$$A_j(\rho) := \begin{pmatrix} \rho - 2 & \rho - 1 & \rho - 3 \\ \rho\Delta_j^2 - 2\Delta_j & \rho\Delta_j - 1 & \rho\Delta_j^3 - 3\Delta_j^2 \\ \rho\Delta_{j+1}^2 - 2\Delta_{j+1} & \rho\Delta_{j+1} - 1 & \rho\Delta_{j+1}^3 - 3\Delta_{j+1}^2 \end{pmatrix}. \tag{8.5}$$

Then the following statements hold.

1. *Every real root of the polynomial P_{k+j} (appearing in* (3.5)*) is positive.*
2. *Under condition* (3.7)*, P_{k+j} has a unique real root.*
3. *If*

$$\exists (r, a, b) \in \mathbb{R}^3 \text{ satisfying } A_j(r) \cdot (a, b, 1)^\top = 0 \text{ and } a^2 - 4b < 0, \tag{8.6}$$

then

$$(3.7) \quad \text{and} \quad P_{k+j}(r) = 0 \tag{8.7}$$

hold.

4. *If* (8.7) *holds with some $r \in \mathbb{R}$, then* (8.6) *is true, and the triplet $(r, a, b) \in \mathbb{R}^3$ is unique.*

Proof. Throughout the proof we always assume $1 < \Delta_{j+1} < \Delta_j$ and $r \in \mathbb{R}$. First we define some auxiliary polynomials—their dependence on Δ_{j+1} is suppressed for brevity.

$\widetilde{P}_1(r) := (r^2 - r)\Delta_{j+1} - r + 2,$
$\widetilde{P}_2(r) := (r^2 - 2r)\Delta_{j+1} - 2r + 6,$
$\widetilde{P}_3(r) := (r^2 - r^3)\Delta_{j+1}^2 + 2(r^2 - 2r)\Delta_{j+1} - 2r + 6,$
$\widetilde{P}_4(r) := r^2\left(r^2 - 4r + 2\right)\Delta_{j+1}^2 - 4r\left(r^2 - 5r + 3\right)\Delta_{j+1} + 2\left(r^2 - 6r + 6\right),$
$\widetilde{P}_5(r) := \left(-\left(5 + 2\sqrt{6}\right)r^3 + 2\left(3 + \sqrt{6}\right)r^2 - 2r\right)\Delta_{j+1} + \left(5 + 2\sqrt{6}\right)r^2 - 4\left(3 + \sqrt{6}\right)r + 6,$
$\widetilde{P}_6(r) := (r - 1)\Delta_{j+1} - r + 3,$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$\widetilde{P}_7(r) := r^2\Delta_{j+1}^2 - 4r\Delta_{j+1} + 6,$
$\widetilde{P}_8(x,a,r) := (r^2 - r)x^2 + (r-1)(ar + r - 3)x + a(2-r) - r + 3,$
$\widetilde{P}_9(r) := (r - r^2)\Delta_{j+1}^2 - \left(r^2 - 4r + 3\right)\Delta_{j+1} + r - 3,$
$\widetilde{P}_{10}(r) := (r^2 - r)\Delta_{j+1} - r + 3,$
$\widetilde{P}_{11}(r) := r\Delta_{j+1}^2 - (r+3)\Delta_{j+1} + 1.$

Step 1. Statement 1 follows from the fact that the coefficient of x^m in $P_{k+j}(x)$ is positive for odd m and negative for even m, so P_{k+j} cannot have a nonpositive root.

Step 2. We now prove statement 2. Since P_{k+j} is cubic, it has a real root. By taking into account Statement 1, we will show that

$$\exists r > 0 \ : \ (3.7), \ P_{k+j}(r) = 0, \quad \text{and} \quad P'_{k+j}(r) \leq 0 \tag{8.8}$$

is *false*, implying that $P'_{k+j} > 0$ at every real root of P_{k+j}, which is possible only if P_{k+j} has a unique real root.

If $\widetilde{P}_1(r) = 0$, we obtain

$$(r, \Delta_{j+1}) = (2 - \sqrt{2}, 3 + 2\sqrt{2}) \tag{8.9}$$

by using $P_{k+j}(r) = 0$ also. But then one easily shows that (8.8) is impossible.

So we can suppose in the rest of Step 2 that $\widetilde{P}_1(r) \neq 0$. Then from $P_{k+j}(r) = 0$ we get

$$\Delta_j = \widetilde{P}_2(r)/(r\widetilde{P}_1(r)). \tag{8.10}$$

This form of Δ_j is substituted into (8.8) and we obtain

$$r > 0 \quad \text{and} \quad \widetilde{P}_3(r)/\widetilde{P}_1(r) > 0 \quad \text{and} \quad \widetilde{P}_4(r)/\widetilde{P}_1(r) \leq 0 \quad \text{and} \tag{8.11}$$

$$\left(\widetilde{P}_5(r)/\widetilde{P}_1(r) < 0 \quad \text{or} \quad \widetilde{P}_6(r)\,\widetilde{P}_7(r)/\widetilde{P}_1(r) > 0\right). \tag{8.12}$$

Because all $\widetilde{P}_j$ polynomials $(1 \leq j \leq 7)$ are at most quadratic in Δ_{j+1}, one can systematically reduce (8.11)–(8.12) to a system of univariate polynomial inequalities in r, and verify that (8.11)–(8.12) has no solution. This finishes the proof of Step 2.

Step 3. To prove statement 3, first notice that

$$A_j(r) \cdot (a, b, 1)^\top = 0 \tag{8.13}$$

implies

$$0 = \det A_j(r) \equiv (\Delta_j - 1)\,(\Delta_{j+1} - 1)\,(\Delta_j - \Delta_{j+1})\,P_{k+j}(r), \tag{8.14}$$

so we have $P_{k+j}(r) = 0$. To show (3.7), depicted in Figure 3, we separate two cases.

If $r = 1$, then we obtain $a = -2$ from the first component of (8.13), and $b = 4\Delta_j - \Delta_j^2$ with

$$\Delta_{j+1} = 4 - \Delta_j \quad \text{and} \quad 2 < \Delta_j < 3 \tag{8.15}$$

from the other two components and from $1 < \Delta_{j+1} < \Delta_j$. So (3.7) holds.

We can thus suppose in the rest of Step 3 that $r \neq 1$. Then from (8.13) we get

$$b = \frac{(2-r)a - r + 3}{r - 1} \tag{8.16}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

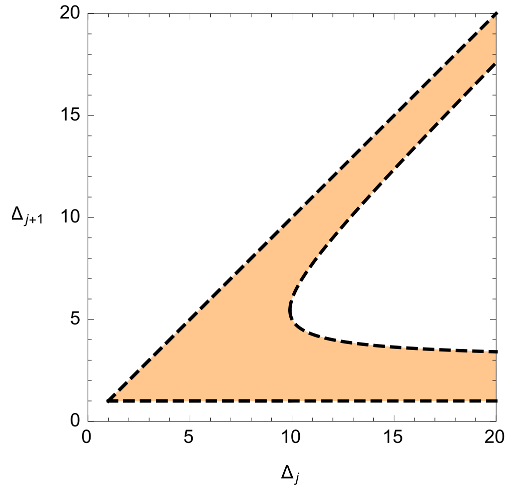


FIG. 3. *The set defined by* (3.7).

and

$$\widetilde{P}_8(\Delta_j, a, r) = 0 = \widetilde{P}_8(\Delta_{j+1}, a, r). \tag{8.17}$$

We again separate two cases.

If $\widetilde{P}_1(r) = 0$, then by using (8.17) and $1 < \Delta_{j+1} < \Delta_j$ we obtain

$$r = 2 - \sqrt{2},\ \Delta_j > \Delta_{j+1} = 3 + 2\sqrt{2}, \quad \text{and} \quad a = -\left(\Delta_j + 1 + 1/\sqrt{2}\right). \tag{8.18}$$

But it is easy to check that (8.18), the inequality $a^2 - 4b < 0$ from (8.6), and the *negation* of (3.7) cannot hold simultaneously. So (3.7) is proved in this case.

Therefore, in the rest of Step 3, we can suppose that $\widetilde{P}_1(r) \neq 0$. Then a is expressed from $\widetilde{P}_8(\Delta_{j+1}, a, r) = 0$ and we get

$$a = \widetilde{P}_9(r)/\widetilde{P}_1(r). \tag{8.19}$$

Moreover, we express Δ_j from $P_{k+j}(r) = 0$ (shown to hold in (8.14)) as in (8.10). To finish Step 3, we will verify that

$$a^2 - 4b < 0 \text{ from (8.6)}, \text{ and the negation of (3.7)} \tag{8.20}$$

cannot hold. The substitution of (8.19) and (8.10) into (8.20) yields

$$r > 0, \quad \widetilde{P}_6(r)\widetilde{P}_{10}(r)\widetilde{P}_{11}(r) < 0, \quad \widetilde{P}_5(r)/\widetilde{P}_1(r) \geq 0, \quad \text{and} \quad \widetilde{P}_6(r)\widetilde{P}_7(r)/\widetilde{P}_1(r) \leq 0. \tag{8.21}$$

Again, these $\widetilde{P}_j$ polynomials are at most quadratic in Δ_{j+1}, so one can check that (8.21) has no solution.

Step 4. To prove statement 4, we first suppose that

$$r \neq 1 \quad \text{and} \quad \widetilde{P}_1(r) \neq 0. \tag{8.22}$$

Then we have seen in Step 3 that if $(a, b) \in \mathbb{R}^2$ solves (8.13), then a and b necessarily have the form (8.19) and (8.16), respectively, and the uniqueness of r is guaranteed

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

by Step 2. So, under condition (8.22), any triplet solving (8.13) is unique. On the other hand, direct substitution into (8.13) shows that the above (r, a, b) is indeed a solution—the nontrivial component to check in (8.13) is the second one, which is just $(\Delta_j - 1)\,(\Delta_j - \Delta_{j+1})\,P_{k+j}(r)/\widetilde{P}_1(r) = 0$ due to $P_{k+j}(r) = 0$ from assumption (8.7). Finally, to finish Step 4 in the case when (8.22) holds, we notice that for this unique triplet (r, a, b) we have $a^2 - 4b = \widetilde{P}_6(r)\widetilde{P}_{10}(r)\widetilde{P}_{11}(r)/(\widetilde{P}_1(r))^2$. But, with Δ_j expressed as in (8.10), one can again show that

$$\widetilde{P}_6(r)\widetilde{P}_{10}(r)\widetilde{P}_{11}(r) \geq 0, \quad 1 < \Delta_{j+1} < \Delta_j, \quad \text{and} \quad (3.7) \text{ from assumption } (8.7)$$

cannot hold. This means that $a^2 - 4b < 0$ from (8.6) also holds.

To prove Statement 4 when $r = 1$, we recall from Step 3 that in this case the only possible triplet is $(r, a, b) = (1, -2, 4\Delta_j - \Delta_j^2)$ with (8.15), which indeed satisfy (8.6).

Last, we consider the case when $\widetilde{P}_1(r) = 0$. We have seen in Step 2 that this time (8.9) holds. Since $\Delta_{j+1} < \Delta_j$, we also have $3 + 2\sqrt{2} < \Delta_j$. These yield that the unique solution to (8.13) is

$$(r, a, b) = (2 - \sqrt{2},\ -(\Delta_j + 1 + 1/\sqrt{2}),\ (2 + \sqrt{2})\Delta_j).$$

Now (3.7) implies $\Delta_j < 5 + 7/\sqrt{2}$, so $a^2 - 4b = \Delta_j^2 - 3(2 + \sqrt{2})\Delta_j + \sqrt{2} + 3/2 < 0$, and the proof is complete. ∎

The proof of Theorem 4. *Step* 1. The necessity of the condition $\Omega_{k,n} > 3$ for the existence of a third-order formula with positive SSP coefficient, is again an immediate consequence of Theorem 2. To see that this condition is also sufficient, suppose to the contrary that there is no third-order formula with k steps, positive SSP coefficient, and $\Omega_{k,n} > 3$, and let us apply Lemma 2 with $r := 0$. First we show that, up to a positive multiplicative constant, $q_n(x) = x^2(\Omega_{k,n} - x)$ is the unique polynomial satisfying (8.1) and properties 1–2 and 4–5 of Lemma 2. Indeed, (8.1a)–(8.1b) with $j = 0$ and $r = 0$, together with properties 2 and 5 imply that $\Omega_{0,n} = 0$ is a root of q_n (otherwise the logarithmic derivative of q_n at 0 would be negative), and then the multiplicity of 0 is precisely two due to properties 1 and 4, so the desired form of q_n follows. Now, applying (8.1b) with $j = k - 1$ and $r = 0$ to this q_n gives $(3 - \Omega_{k,n})\,(\Omega_{k,n} - 1) \geq 0$, contradicting $\Omega_{k,n} > 3$. This contradiction proves the sufficiency of the condition $\Omega_{k,n} > 3$.

Step 2. From now on we assume $\Omega_{k,n} > 3$, and that a formula with the optimal SSP coefficient $r := \mathcal{C}_n(\omega_n, \delta_n, \beta_n) \in (0, +\infty)$ is considered. Let q_n be a polynomial satisfying conditions (8.1), Properties 1–6 of Lemma 2, and condition (iii) of Lemma 4.

Case I. In the case when $\Omega_{k,n}$ is not the only real root of q_n, then, due to properties 1, 3, and 5 of Lemma 2, q_n is unique (up to a positive multiplicative constant) and takes the form $q_n(x) = (x - \Omega_{j_0,n})^2(\Omega_{k,n} - x)$ with an appropriate index $0 \leq j_0 \leq k-1$. We notice that the only binding inequality in (8.1a) is the one corresponding to $j = j_0$, and the $j = j_0$ inequality in (8.1b) is also binding (independently of the value of r). On the other hand, now all roots of q_n are real, so the logarithmic derivative of q_n is strictly decreasing on the intervals $[\Omega_{0,n}, \Omega_{j_0,n})$ and $(\Omega_{j_0,n}, \Omega_{k-1,n}]$. Thus, by (iii) of Lemma 4, (8.1b) contains either two or three binding inequalities, and they correspond to

A. $j \in \{0, k-1\}$ for $j_0 = 0$,
B. $j \in \{j_0 - 1, j_0\}$ or $j \in \{j_0, k-1\}$ for $1 \leq j_0 \leq k-2$,
C. $j \in \{k-2, k-1\}$ for $j_0 = k-1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In case A, (8.1b) is solved with $j = k-1$, and we obtain $r = (\Omega_{k,n} - 3)/(\Omega_{k,n} - 1)$. Due to (i) and (ii) of Lemma 4, all of the formula coefficients except for possibly $\{\delta_{0,n}, \beta_{0,n}, \beta_{k-1,n}\}$ vanish.

In case B, (8.1b) is solved with $j = j_0 - 1$ and $j = k-1$ to get

$$r = 2/(\Omega_{j_0,n} - \Omega_{j_0-1,n}) + 1/(\Omega_{k,n} - \Omega_{j_0-1,n})$$

and

$$r = (\Omega_{k,n} - \Omega_{j_0,n} - 3)/(\Omega_{k,n} - \Omega_{j_0,n} - 1),$$

respectively, then the maximum is chosen. If these two expressions for r are not equal, we get from (i) and (ii) of Lemma 4 that the number of nonzero formula coefficients is at most three, whereas, if the two expressions for r are equal, the number of nonzero formula coefficients is at most four.

In case C, (8.1b) is solved with $j = k-2$ to yield $r = 2/(\Omega_{k-1,n} - \Omega_{k-2,n}) + 1/(\Omega_{k,n} - \Omega_{k-2,n})$. All formula coefficients except for possibly $\{\delta_{k-1,n}, \beta_{k-2,n}, \beta_{k-1,n}\}$ vanish.

Therefore, the optimal SSP coefficient is equal to one of the r_j quantities ($0 \le j \le k-1$) defined in section 3.3, and the nonzero formula coefficients also have the form stated there.

Case II. In the case when $\Omega_{k,n}$ is the only real root of q_n, then, by taking into account (8.1a), and properties 1 and 5 of Lemma 2, q_n (up to a positive multiplicative constant) has the form

$$q_n(x) = (\Omega_{k,n} - x)\left((\Omega_{k,n} - x)^2 + a(\Omega_{k,n} - x) + b\right)$$

with some coefficients $a, b \in \mathbb{R}$ satisfying the discriminant condition $a^2 - 4b < 0$. Let us introduce the abbreviation $Q_3 := q_n' + rq_n$.

This time there are no binding inequalities in (8.1a), so, due to (iii) of Lemma 4, precisely three inequalities are binding in (8.1b)—there can be no more, since the polynomial Q_3 is cubic. Let $0 \le j_0 < j_1 < j_2 \le k-1$ denote the three indices corresponding to these binding inequalities, then $Q_3(\Omega_{j_0,n}) = Q_3(\Omega_{j_1,n}) = Q_3(\Omega_{j_2,n}) = 0$. We observe that the leading coefficient of Q_3 is $-r < 0$, and $Q_3(\Omega_{k,n}) = -b < 0$ because of $a^2 - 4b < 0$. As $\Omega_{j,n}$ is strictly increasing in j, we get that Q_3 is positive on $(-\infty, \Omega_{j_0,n}) \cup (\Omega_{j_1,n}, \Omega_{j_2,n})$, and negative on $(\Omega_{j_0,n}, \Omega_{j_1,n}) \cup (\Omega_{j_2,n}, +\infty)$. Hence neither $j_1 \ge j_0+2$ nor $j_2 \le k-2$ can occur (otherwise $Q_3(\Omega_{j_0+1,n}) < 0$ or $Q_3(\Omega_{k-1,n}) < 0$ would contradict (8.1b)). This shows that the binding inequalities in (8.1b) are the ones corresponding to the index set $j \in \{j_0, j_0+1, k-1\}$ with an appropriate index $0 \le j_0 \le k-3$. The previous sentence can be rewritten as $A_{j_0}(r) \cdot (a, b, 1)^\top = 0$, with the matrix A_{j_0} defined in (8.5). Since now $a^2 - 4b < 0$ as well, (8.6) is also satisfied. Hence (3.7) and $P_{k+j_0}(r) = 0$ hold by Lemma 5. Due to uniqueness we see that $r = r_{k+j_0}$, with r_{k+j_0} defined in (3.8).

Thus, the optimal SSP coefficient is equal to one of the r_{k+j_0} quantities ($0 \le j_0 \le k-3$), and we get from (ii) of Lemma 4 that all formula coefficients except for possibly $\{\beta_{j_0,n}, \beta_{j_0+1,n}, \beta_{k-1,n}\}$ vanish.

Step 3. Finally, in order to conclude that the optimal SSP coefficient $\mathcal{C}_n(\omega_n, \delta_n, \beta_n)$ is in fact the minimum of the r_j expressions defined in section 3.3, we show that $\mathcal{C}_n(\omega_n, \delta_n, \beta_n) \le r_j$ for any $0 \le j \le 2k-3$.

Indeed, fix any $0 \le j_0 \le k-1$ and define $q_n(x) := (x - \Omega_{j_0,n})^2(\Omega_{k,n} - x)$. The inequalities (8.1a) and (8.1c) are trivially satisfied. We verify (8.1b) with $r := r_{j_0}$ by setting $\widetilde{Q}_3 := q_n' + rq_n$ and distinguishing three cases.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

A. If $j_0 = 0$, then one checks that the three roots of $\widetilde{Q}_3$ are found at $\{x_1, 0, \Omega_{k-1,n}\}$ with $x_1 = -2\Omega_{k,n}/(\Omega_{k,n} - 3) < 0$.

B. If $1 \le j_0 \le k-2$ and

$$\frac{\Omega_{k,n} - \Omega_{j_0,n} - 3}{\Omega_{k,n} - \Omega_{j_0,n} - 1} < \frac{2}{\omega_{j_0,n}} + \frac{1}{\Omega_{k,n} - \Omega_{j_0-1,n}}, \tag{8.23}$$

then the three roots of $\widetilde{Q}_3$ are found at $\{\Omega_{j_0-1,n}, \Omega_{j_0,n}, x_3\}$ with $\Omega_{k-1,n} < x_3 < \Omega_{k,n}$. If we have ">" in (8.23), then the three roots of $\widetilde{Q}_3$ are found at $\{x_1, \Omega_{j_0,n}, \Omega_{k-1,n}\}$ with $\Omega_{j_0-1,n} < x_1 < \Omega_{j_0,n}$. Finally, if there is "=" in (8.23), then the roots of $\widetilde{Q}_3$ are found at $\{\Omega_{j_0-1,n}, \Omega_{j_0,n}, \Omega_{k-1,n}\}$.

C. If $j_0 = k-1$, then the three roots of $\widetilde{Q}_3$ are found at $\{\Omega_{k-2,n}, \Omega_{k-1,n}, x_3\}$ with $\Omega_{k-1,n} < x_3 < \Omega_{k,n}$.

These, combined with the fact that the leading coefficient of $\widetilde{Q}_3$ is negative, mean that all inequalities in (8.1) are satisfied. Thus, in view of Lemma 2, $\mathcal{C}_n \le r_{j_0}$.

Now fix an arbitrary index $0 \le j_0 \le k-3$. The inequality $\mathcal{C}_n \le r_{k+j_0}$ is trivial if $r_{k+j_0} = +\infty$. Otherwise, if $r_{k+j_0} < +\infty$ in (3.8), then, by definition, (3.7) and $P_{k+j_0}(r_{k+j_0}) = 0$ hold. So due to Lemma 5, (8.6) holds with $r := r_{k+j_0}$ and $j := j_0$. By defining

$$q_n(x) := (\Omega_{k,n} - x)\left((\Omega_{k,n} - x)^2 + a(\Omega_{k,n} - x) + b\right)$$

and with a and b given by (8.6), we see that (8.1a) and (8.1c) are true because $a^2 - 4b < 0$. On the other hand, the $A_{j_0}(r) \cdot (a, b, 1)^\top = 0$ relation in (8.6) expresses the fact that the inequalities corresponding to $j \in \{j_0, j_0+1, k-1\}$ are binding in (8.1b). So, just as in Case II in Step 2, we see that $q_n' + rq_n$ is positive on $(-\infty, \Omega_{j_0,n}) \cup (\Omega_{j_0+1,n}, \Omega_{k-1,n})$, and negative on $(\Omega_{j_0,n}, \Omega_{j_0+1,n}) \cup (\Omega_{k-1,n}, +\infty)$. Thus all inequalities in (8.1b) hold. This means that q_n satisfies conditions (8.1) with this r value, so, in view of Lemma 2, $\mathcal{C}_n \le r_{k+j_0}$. The proof of the theorem is complete. ∎

8.4. The proof of Theorem 5.

Proof. We follow the ideas of the proof given in [11] for the fixed-step-size case.

Suppose to the contrary that p, k, and $\omega_{j,n}$ ($1 \le j \le k$) satisfy the conditions of the theorem for some $n \ge k$ and $K_1, K_2 \ge 1$, but for all formulas with k steps and order of accuracy p, we have $\mathcal{C}_n(\omega, \delta, \beta) = 0$. Then by Lemma 2, there exists a nonzero real polynomial q_n that satisfies the conditions (8.1a)–(8.1c) with $r = 0$, moreover, $q_n \ge 0$ on $[0, \Omega_{k,n}]$ and $\deg q_n = p$ (properties 1 and 6 of Lemma 2).

First we define $A := \max_{x \in [0, \Omega_{k,n}]} q_n(x)$ and $b := \max_{x \in [0, \Omega_{k,n}]} |q_n'(x)|$ and we introduce the polynomial $P(x) := q_n\left(\frac{x+1}{2} \cdot \Omega_{k,n}\right) - \frac{A}{2}$. Then the Markov brothers' inequality (see, e.g., [11]) for the first derivative implies that

$$\max_{x \in [-1,1]} |P'(x)| \le p^2 \cdot \max_{x \in [-1,1]} |P(x)|,$$

that is, $b\frac{\Omega_{k,n}}{2} \le p^2 \frac{A}{2}$. On the other hand, summing the lower estimates in (3.15) we get $\frac{k}{K_1} \le \Omega_{k,n}$, implying $\frac{b}{2} \cdot \frac{k}{K_1} \le \frac{b}{2}\Omega_{k,n}$. Thus

$$bk/K_1 \le p^2 A. \tag{8.24}$$

Now, because of $q_n(\Omega_{k,n}) = 0$, the Newton–Leibniz formula and elementary estimates yield that

$$A \le \int_0^{\Omega_{k,n}} \max(0, -q_n'(t))dt. \tag{8.25}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Here we notice that the polynomial q_n' is of degree at most $p-1$, and $q_n'(\Omega_{j,n}) \geq 0$ for all $0 \leq j \leq k-1$, so the set

$$\{j \in \mathbb{Z} \cap [0, k-1] : \exists x \in [\Omega_{j,n}, \Omega_{j+1,n}] \text{ with } q_n'(x) < 0\}$$

has at most $p/2$ elements. Therefore—by decomposing the interval $[0, \Omega_{k,n}]$ as the union of the appropriate subintervals of length $\omega_{j,n}$, and then applying the upper estimate in (3.15) and the estimate $-q_n' \leq b$ for at most $p/2$ times—we will get that $\int_0^{\Omega_{k,n}} \max(0, -q_n'(t))dt \leq \frac{p}{2}K_2 b$, and hence

$$A \leq bpK_2/2. \tag{8.26}$$

Inequalities (8.24) and (8.26) imply that $k \leq p^3 K_1 K_2/2$, which contradicts the assumption of Theorem 5. Hence there is a formula with k steps, order of accuracy p, and $\mathcal{C}_n(\omega, \delta, \beta) > 0$. ∎

9. The proofs of the theorems in section 4. In section 9.1 we first prove a theorem about the convergence of some rational recursions. This Theorem 11 will then be used in sections 9.2 and 9.3 to prove Theorems 6 and 7, respectively.

9.1. Global attractivity in a class of higher-order rational recursions.

THEOREM 11. *Let us fix an integer $k \geq 3$ and a real number $A > 0$. Suppose that for $1 \leq j \leq k-1$ the initial values $\tau_j \geq 0$ are given such that $\sum_{j=1}^{k-1} \tau_j > 0$. For any $n \geq k$ we define*

$$\tau_n := \frac{\sum_{j=1}^{k-1} \tau_{n-j}}{A + \sum_{j=1}^{k-1} \tau_{n-j}}. \tag{9.1}$$

Then

$$\lim_{n\to+\infty} \tau_n = \begin{cases} 0 & \text{if} \quad k-1 \leq A, \\ \dfrac{k-1-A}{k-1} & \text{if} \quad 0 < A < k-1. \end{cases}$$

To prove Theorem 11, we will apply the following lemma.

LEMMA 6 (Theorem A.0.1 in [8]). *Suppose that $a \leq b$ are given real numbers, $k \geq 3$ is a fixed integer, and the numbers x_j are chosen such that $x_j \in [a,b]$ for $1 \leq j \leq k-1$. Assume further that*

1. *$f\colon [a,b]^{k-1} \to [a,b]$ is continuous,*
2. *f is nondecreasing in each of its arguments,*
3. *there is a unique $\overline{x} \in [a,b]$ such that $f(\overline{x}, \overline{x}, \ldots, \overline{x}) = \overline{x}$,*
4. *and the sequence x_n is defined for $n \geq k$ as*

$$x_n := f(x_{n-1}, x_{n-2}, \ldots, x_{n-(k-1)}).$$

Then $\lim_{n\to+\infty} x_n = \overline{x}$.

The straightforward proof of Lemma 6 is found in [8] (see their Theorem 1.4.8 (for $k=3$), Theorem A.0.1 (for $k=4$), or Theorem A.0.9 (for general k))—the idea of the proof is the same in the easiest case when f is nondecreasing in each of its arguments. Notice that in [8, Theorem 1.4.8] one should have "The equation $f(x,x) = x$ has a unique solution in $[a,b]$" instead of "... a unique positive solution." Now we give the proof of Theorem 11.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The proof of Theorem 11. For some $a \le b$ (to be specified soon) we set

$$f(z_1, z_2, \ldots, z_{k-1}) := \frac{\sum_{j=1}^{k-1} z_j}{A + \sum_{j=1}^{k-1} z_j} \tag{9.2}$$

with $z_j \in [a, b]$ $(j = 1, 2, \ldots, k-1)$. Then

$$(\partial_j f)(z_1, z_2, \ldots, z_{k-1}) = \frac{A}{\left(A + \sum_{j=1}^{k-1} z_j\right)^2} > 0, \tag{9.3}$$

hence f is nondecreasing in each of its arguments (and trivially continuous). Notice that due to (9.1) we have $\tau_n \in (0, 1)$ for any $n \ge k$, so by shifting the indices we can assume that $\tau_j \in (0, 1)$ for $1 \le j \le k-1$ and $\tau_n = f(\tau_{n-1}, \tau_{n-2}, \ldots, \tau_{n-(k-1)})$ for $n \ge k$. We distinguish two cases.

1. *The case* $k - 1 \le A$. Then f maps $[a, b]^{k-1}$ to $[a, b]$ with $a := 0$ and $b := 1$. Now for any $\overline{x} \in [0, 1]$ we have

$$\overline{x} - f(\overline{x}, \overline{x}, \ldots, \overline{x}) \equiv \frac{\overline{x}[A - (k-1) + (k-1)\overline{x}]}{A + (k-1)\overline{x}} = 0 \tag{9.4}$$

   precisely if $\overline{x} = 0$, so Lemma 6 yields $\lim_{n\to+\infty} \tau_n = 0$.
2. *The case* $0 < A < k - 1$. We set $\tau^* := \min_{1\le j\le k-1} \tau_j$. Then $\tau^* > 0$ and $\tau_j \in [\tau^*, 1]$ for $1 \le j \le k-1$.
   (a) *The case* $\tau^* \ge \frac{k-1-A}{k-1}$. We choose $a := \frac{k-1-A}{k-1} < 1 =: b$, and notice that $f(a, a, \ldots, a) = a$ and $f(b, b, \ldots, b) < b$. By also using the nondecreasing property of f in each of its arguments we obtain that f maps $[a, b]^{k-1}$ to $[a, b]$. Now for any $\overline{x} \in [a, b]$ the equality (9.4) holds if and only if $\overline{x} = a$, so Lemma 6 yields $\lim_{n\to+\infty} \tau_n = \frac{k-1-A}{k-1}$.
   (b) *The case* $0 < \tau^* < \frac{k-1-A}{k-1}$. This time we choose $a := \tau^* < 1 =: b$. Since now

$$a < f(a, a, \ldots, a) \Longleftrightarrow \tau^* < f(\tau^*, \tau^*, \ldots, \tau^*) \Longleftrightarrow \tau^* < \frac{k-1-A}{k-1},$$

   we have just as before that f maps $[a, b]^{k-1}$ to $[a, b]$. For any $\overline{x} \in [a, b]$ we have (9.4) precisely if $\overline{x} = \frac{k-1-A}{k-1}$; therefore, we can use Lemma 6 again to get $\lim_{n\to+\infty} \tau_n = \frac{k-1-A}{k-1}$. ∎

EXAMPLE 1. *The sequence* $(\tau_n)_{n\ge1}$ *defined by* (9.1) *can have long, nonmonotonic starting slices. Consider, for example, the case* $k = 4$ *with*

$$\tau_1 := 1, \quad \tau_2 := \frac{1}{200}, \quad \tau_3 := \frac{95638788642}{100000000000}$$

and

$$\tau_n := \frac{\tau_{n-1} + \tau_{n-2} + \tau_{n-3}}{1 + \tau_{n-1} + \tau_{n-2} + \tau_{n-3}} \qquad \text{for } n \ge 4.$$

Then the consecutive monotone nonincreasing subsequences of τ_n *for* $1 \le n \le 1000$ *has lengths*

$$(2, 3, 2, 1, 2, 1, 2, 1, 2, 1, 2, 3, 3, 3, 3, 2, 1, 2, 1, 2, 1, \\ 2, 1, 2, 3, 3, 3, 3, 2, 1, 2, 1, 2, 1, 2, 3, 3, 3, 3, 3, 917).$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

9.2. The proof of Theorem 6.

Proof. We prove the theorem for $k=3$ first. We define two sequences

$$h_n^- := \frac{h_{n-2}^- + h_{n-1}^-}{h_{n-2}^- + h_{n-1}^- + \mu^-} \cdot \mu^-, \quad h_n^+ := \frac{h_{n-2}^+ + h_{n-1}^+}{h_{n-2}^+ + h_{n-1}^+ + \mu^+} \cdot \mu^+, \quad h_1^\pm := h_1,\ h_2^\pm := h_2, \tag{9.5}$$

and their scaled counterparts $\tau_n^- := h_n^-/\mu^-$, $\tau_n^+ := h_n^+/\mu^+$ $(n \geq 1)$. Then τ_n^- and τ_n^+ satisfy

$$\tau_n^- = \frac{\tau_{n-2}^- + \tau_{n-1}^-}{\tau_{n-2}^- + \tau_{n-1}^- + 1}, \quad \tau_n^+ = \frac{\tau_{n-2}^+ + \tau_{n-1}^+}{\tau_{n-2}^+ + \tau_{n-1}^+ + 1}, \quad \tau_1^\pm > 0,\ \tau_2^\pm > 0.$$

By applying Theorem 11 with $k=3$ and $A=1$ we see that $\tau_n^- \to 1/2$, hence $h_n^- \to \mu^-/2$ as $n \to +\infty$. Similarly, we get $h_n^+ \to \mu^+/2$. We now define

$$(0,+\infty)^3 \ni (a,x,y) \mapsto \widetilde{f}(a,x,y) := a \cdot \frac{x+y}{a+x+y} \tag{9.6}$$

(cf. (9.2)). It is elementary to see that for any $(a,x,y) \in (0,+\infty)^3$ we have

$$\partial_1 \widetilde{f}(a,x,y) = \frac{(x+y)^2}{(a+x+y)^2} > 0, \quad \partial_2 \widetilde{f}(a,x,y) = \partial_3 \widetilde{f}(a,x,y) = \frac{a^2}{(a+x+y)^2} > 0 \tag{9.7}$$

(cf. (9.3), and notice that the function $a \mapsto \widetilde{f}(a,x,y)/a$, for example, would be monotone *decreasing*). Clearly, for $n = 1,2$ we have

$$h_n^- \leq h_n \leq h_n^+, \tag{9.8}$$

so we can suppose that (9.8) has already been proved up to some $n \geq 2$. Then by repeatedly using the inequality $\mu^- \leq \mu_n \leq \mu^+$ (implied by the assumption (4.1)), (9.7), and (9.8), we obtain

$$\begin{aligned} h_{n+1}^- &\equiv \frac{h_{n-1}^- + h_n^-}{h_{n-1}^- + h_n^- + \mu^-} \cdot \mu^- \leq \frac{h_{n-1} + h_n^-}{h_{n-1} + h_n^- + \mu^-} \cdot \mu^- \leq \frac{h_{n-1} + h_n}{h_{n-1} + h_n + \mu^-} \cdot \mu^- \\ &\leq \frac{h_{n-1} + h_n}{h_{n-1} + h_n + \mu_{n+1}} \cdot \mu_{n+1} \equiv h_{n+1} \leq \frac{h_{n-1} + h_n}{h_{n-1} + h_n + \mu^+} \cdot \mu^+ \\ &\leq \frac{h_{n-1} + h_n^+}{h_{n-1} + h_n^+ + \mu^+} \cdot \mu^+ \leq \frac{h_{n-1}^+ + h_n^+}{h_{n-1}^+ + h_n^+ + \mu^+} \cdot \mu^+ \equiv h_{n+1}^+. \end{aligned}$$

This shows the validity of (9.8) for all $n \geq 1$ by induction. By taking lim inf and lim sup in (9.8), Theorem 6 for $k=3$ is proved.

The proof of Theorem 6 in the general case requires only formal modifications of the argument given above: Theorem 11 with a general $k \geq 3$ and with $A=1$ implies that for the corresponding sequences we have $\tau_n^\pm \to \frac{k-2}{k-1}$ as $n \to +\infty$, and the corresponding function $\widetilde{f}\colon (0,+\infty)^k \to (0,+\infty)$ increases in each of its arguments. ☐

Figure 4 gives a graphical illustration of Theorem 6 for $k=3$, using a hypothetical sequence of values for μ_n.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

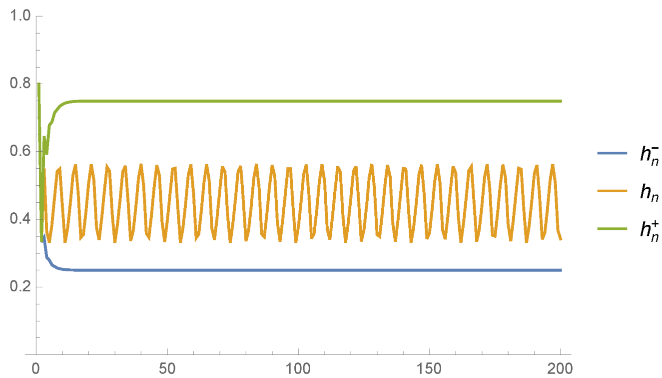


FIG. 4. *The first* 200 *terms of the sequences* h_n^-, h_n, *and* h_n^+ *with* $\mu_n := 1+\sin(n)/2$, $\mu^- := 1/2$, *and* $\mu^+ := 3/2$*; see* (4.3) *and* (9.5).

9.3. The proof of Theorem 7.

Proof. Step 1. Initially we suppose that some values of $\varrho > 0$ and $0 < \varrho_{\mathrm{FE}} \leq 1$ have already been chosen; we will make an actual choice for them in Step 5 so that the inductive argument given below becomes valid. First let us set $n = k$.

Step 2. We know from (4.8) that

$$h_m \leq \varrho \cdot h_{\mathrm{FE}}(u_m) \quad \text{for} \quad m = 1, 2, \ldots, n-1. \tag{9.9}$$

We will prove that

$$\sum_{j=1}^{k-1} h_{n-j} \leq \sqrt{8}\,\mu_n \tag{9.10}$$

and

$$h_n \leq \varrho \cdot h_{\mathrm{FE}}(u_n). \tag{9.11}$$

Step 3. Applying (9.9) and (4.2) repeatedly we have

$$\sum_{j=1}^{k-1} h_{n-j} \leq \varrho \cdot \sum_{j=1}^{k-1} h_{\mathrm{FE}}(u_{n-j}) \leq \varrho \cdot \sum_{j=1}^{k-1} \frac{h_{\mathrm{FE}}(u_{n-1})}{(\varrho_{\mathrm{FE}})^{j-1}} = h_{\mathrm{FE}}(u_{n-1}) \cdot \sum_{j=1}^{k-1} \frac{\varrho}{(\varrho_{\mathrm{FE}})^{j-1}}. \tag{9.12}$$

On the other hand, the definition of μ_n in (2.14), the repeated application of (4.2), and $0 < \varrho_{\mathrm{FE}} \leq 1$ imply that

$$\sqrt{8}\,\mu_n \geq \sqrt{8} \min_{0 \leq j \leq k-1} \left((\varrho_{\mathrm{FE}})^j \cdot h_{\mathrm{FE}}(u_{n-1}) \right) = h_{\mathrm{FE}}(u_{n-1}) \cdot \sqrt{8}\,(\varrho_{\mathrm{FE}})^{k-1}. \tag{9.13}$$

By comparing the right-hand sides of (9.12) and (9.13) after a division by $h_{\mathrm{FE}}(u_{n-1}) \geq \mu^- > 0$, we get that if ϱ and ϱ_{FE} are chosen such that

$$\sum_{j=1}^{k-1} \frac{\varrho}{(\varrho_{\mathrm{FE}})^{j-1}} \leq \sqrt{8}\,(\varrho_{\mathrm{FE}})^{k-1}, \qquad \varrho > 0, \quad 0 < \varrho_{\mathrm{FE}} \leq 1, \tag{9.14}$$

then (9.10) holds for this particular n value.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

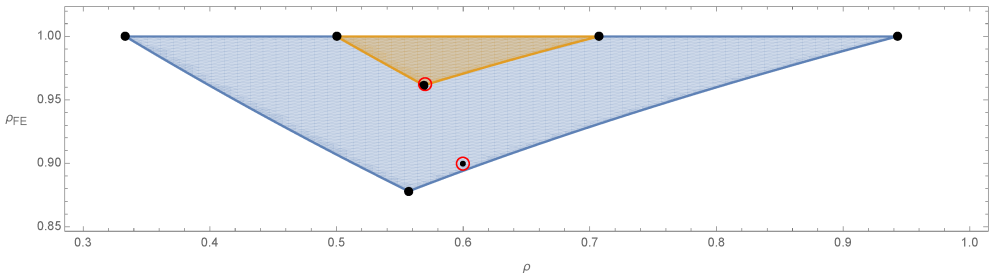


FIG. 5. *The solution set of the inequalities* (9.14)–(9.15). *In the* $k = 4$ *case, the blue region has vertices at* $(1/3, 1)$, $(\sqrt{8}/3, 1)$, $\approx (0.557, 0.878)$. *In the* $k = 5$ *case, the orange region has vertices at* $(1/2, 1)$, $(1/\sqrt{2}, 1)$, $\approx (0.569, 0.9615)$. *The points corresponding to* (4.9) *have been circled.*

Step 4. In order to show (9.11), we first notice that the function

$$(0, +\infty)^2 \ni (z, a) \mapsto \widetilde{f}(z, a) := \frac{z}{z + 2a} \cdot a$$

is increasing in each of its arguments (cf. (9.6)–(9.7)). This monotonicity property, the definition of h_n in (4.7), (9.12), and the inequality $\mu_n \le h_{\mathrm{FE}}(u_{n-1})$ yield that

$$\begin{aligned} h_n &= \frac{\sum_{j=1}^{k-1} h_{n-j}}{\left(\sum_{j=1}^{k-1} h_{n-j}\right) + 2\mu_n} \cdot \mu_n \\ &\le \frac{h_{\mathrm{FE}}(u_{n-1}) \cdot \sum_{j=1}^{k-1} \frac{\varrho}{(\varrho_{\mathrm{FE}})^{j-1}}}{\left(h_{\mathrm{FE}}(u_{n-1}) \cdot \sum_{j=1}^{k-1} \frac{\varrho}{(\varrho_{\mathrm{FE}})^{j-1}}\right) + 2h_{\mathrm{FE}}(u_{n-1})} \cdot h_{\mathrm{FE}}(u_{n-1}) \\ &= \frac{\sum_{j=1}^{k-1} \frac{\varrho}{(\varrho_{\mathrm{FE}})^{j-1}}}{2 + \sum_{j=1}^{k-1} \frac{\varrho}{(\varrho_{\mathrm{FE}})^{j-1}}} \cdot h_{\mathrm{FE}}(u_{n-1}). \end{aligned}$$

On the other hand, from (4.2) we see that $\varrho \cdot h_{\mathrm{FE}}(u_n) \ge \varrho \cdot \varrho_{\mathrm{FE}} \cdot h_{\mathrm{FE}}(u_{n-1})$. Therefore if ϱ and ϱ_{FE} are chosen such that

$$\frac{\sum\limits_{j=1}^{k-1} \frac{\varrho}{(\varrho_{\mathrm{FE}})^{j-1}}}{2 + \sum\limits_{j=1}^{k-1} \frac{\varrho}{(\varrho_{\mathrm{FE}})^{j-1}}} \le \varrho \cdot \varrho_{\mathrm{FE}}, \tag{9.15}$$

then (9.11) holds for the actual n value.

Step 5. So (9.10)–(9.11) will be proved as soon as we have found some ϱ and ϱ_{FE} satisfying (9.14) and (9.15). Figure 5 depicts the solution set of this system of inequalities (9.14)–(9.15) in the variables $(\varrho, \varrho_{\mathrm{FE}})$. The choice made in (4.9) is a simple rational pair; clearly, one could, for example, relax the assumption on ϱ (by choosing it larger), but then condition (4.2) would, in general, become more stringent.

Step 6. According to the discussion preceding Theorem 7, on the one hand, we have $\Omega_{k-1,n} > 2$, hence $\mathcal{C}_n > 0$. On the other hand, (9.10) guarantees $h_n = \mathcal{C}_n \mu_n > 0$ and this is the maximum value of h_n preserving the SSP property.

Step 7. Now we repeat Steps 2–6 inductively for each $n \ge k + 1$ (Step 5 is no longer needed since $(\varrho, \varrho_{\mathrm{FE}})$ have already been given some particular values). The range of m in the induction hypothesis (9.9) is extended step-by-step by (9.11).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Step 8. Finally, to prove (4.10), we make use of the fact that

$$(0,+\infty)^k \ni (a, z_1, z_2, \ldots, z_{k-1}) \mapsto \bar{f}(a, z_1, z_2, \ldots, z_{k-1}) := a \cdot \frac{\sum_{j=1}^{k-1} z_j}{\left(\sum_{j=1}^{k-1} z_j\right) + 2a}$$

is increasing in each of its arguments (cf. (9.6)–(9.7)), and repeat the steps presented in section 9.2: we define the corresponding auxiliary sequences $h_n^{\pm}$ and $\tau_n^{\pm}$ and apply Theorem 11 with $k \in \{4,5\}$ and $A = 2$ to show that $\tau_n^{\pm} \to \frac{k-3}{k-1}$ as $n \to +\infty$. ∎

Acknowledgment. The authors would like to thank the anonymous referees for their suggestions, which have improved the presentation of the material.

REFERENCES

[1] S. Gottlieb, D. I. Ketcheson, and C.-W. Shu, *Strong Stability Preserving Runge–Kutta And Multistep Time Discretizations*, World Scientific, Haekensack, NJ, 2011.
[2] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations* I: *Nonstiff Problems*, 2nd ed., Springer Ser. Comput. Math., Springer, Berlin, 1993.
[3] W. Hundsdorfer and S. J. Ruuth, *On monotonicity and boundedness properties of linear multistep methods*, Math. Comp., 75 (2005), pp. 655–672.
[4] W. Hundsdorfer, S. J. Ruuth, and R. J. Spiteri, *Monotonicity-preserving linear multistep methods*, SIAM J. Numer. Anal., 41 (2003), pp. 605–623.
[5] David I. Ketcheson, *Computation of optimal monotonicity preserving general linear methods*, Math. Comp., 78 (2009), pp. 1497–1513.
[6] D. I. Ketcheson, K. Mandli, A. J. Ahmadia, A. Alghamdi, M. Quezada de Luna, M. Parsani, M. G. Knepley, and M. Emmett, *PyClaw: Accessible, extensible, scalable tools for wave propagation problems*, SIAM J. Sci. Comput., 34 (2012), pp. C210–C231.
[7] D. I. Ketcheson, M. Parsani, and R. J. LeVeque, *High-order wave propagation algorithms for hyperbolic systems*, SIAM J. Sci. Comput., 35 (2013), pp. A351–A377.
[8] M. R. S. Kulenović and G. Ladas, *Dynamics of Second Order Rational Difference Equations with Open Problems and Conjectures*, Chapman & Hall/CRC, Boca Raton, FL, 2002.
[9] H. W. J. Lenferink, *Contractivity-preserving explicit linear multistep methods*, Numer. Math., 55 (1989), pp. 213–223.
[10] H. W. J. Lenferink, *Contractivity-preserving implicit linear multistep methods*, Math. Comp., 56 (1991), pp. 177–199.
[11] A. Németh and D. I. Ketcheson, *Existence and Optimality of Strong Stability Preserving Linear Multistep Methods: A Duality-Based Approach*, preprint, arXiv:1504.03930, 2015.
[12] S. J. Ruuth and W. Hundsdorfer, *High-order linear multistep methods with general monotonicity and boundedness properties*, J. Comput. Phys., 209 (2005), pp. 226–248.
[13] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley Chichester, England, 1998.
[14] B. van Leer, *Towards the ultimate conservative difference scheme.* IV. *A new approach to numerical convection*, J. Comput. Phys., 23 (1977), pp. 276–299.
[15] P. Woodward and P. Colella, *The numerical simulation of two-dimensional fluid flow with strong shocks*, J. Comput. Phys., 54 (1984), pp. 115–173.
[16] X. Zhang and C.-W. Shu, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, J. Comput. Phys., 229 (2010), pp. 3091–3120.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 05/22/18 to 109.171.137.212. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php