

ON MONOTONICITY OF DIFFERENCE SCHEMES FOR COMPUTATIONAL PHYSICS*

V. S. BORISOV[†] AND S. SOREK[‡]

Abstract. Criteria are developed for monotonicity of linear as well as nonlinear difference schemes associated with the numerical analysis of systems of partial differential equations, integro-differential equations, etc. Difference schemes are converted into variational forms that satisfy the boundary maximum principle and also allow the investigation of monotonicity for nonlinear operators using linear patterns. Sufficient conditions are provided to review the monotonicity of single and coupled difference schemes. Necessary as well as necessary and sufficient conditions for monotonicity of explicit schemes are also developed. The notion of submonotone difference schemes is considered and the associated criteria are developed. We discuss the interrelationship between monotonicity, submonotonicity, and stability. Some known schemes serve as examples demonstrating the implementation of the developed approaches. Among these examples, we describe the possibility that stable schemes such as total variation diminishing (TVD) as well as monotonicity preserving can produce spurious oscillations.

Key words. differential equations, difference schemes, TVD schemes, monotonicity, grid connectedness, boundary maximum principle, stability

AMS subject classifications. 35B50, 39A11, 65L20, 65M12, 65N12

DOI. 10.1137/S1064827502406695

1. Introduction. In the numerical solution of a system of partial differential equations (PDEs), integrodifferential equations, etc., one is seeking to establish a scheme free from spurious oscillations (parasitic [11], physically unrealistic [23], [24], or wild [7]), namely a monotone scheme [11], [16], [24]. From a practical standpoint, a main drawback of nonmonotone schemes is a lack of accuracy in simulating complex physical processes with steep gradients (e.g., the case of shock waves propagation), and hence application of monotone schemes is fundamentally used for finding a physically realistic numerical solution to the corresponding PDEs (see, e.g., [2], [3], [7], [8], [10], [11], [12], [15], [16], [20], [24], [27]). Spurious oscillations could exist even if Lax's equivalence theorem [29] were valid, i.e., the scheme is stable and provides consistent approximation. Actually, this is concluded from Godunov's theorem [12, p. 277] stating that any monotone two-level linear scheme for a transport equation in the form

$$(1.1) \quad \frac{\partial y}{\partial t} + V \frac{\partial y}{\partial x} = 0$$

can be at most of first order accuracy. When investigating for the monotonic behavior of a numerical scheme, it is assumed that one can distinguish between spurious and natural oscillations. The notion of a “monotone difference scheme” originally appeared in Godunov [12]. His interpretation was that *a monotone scheme transforms a monotone increasing (or decreasing) 1-D scalar function $\check{u}(x)$, at a previous*

*Received by the editors May 1, 2002; accepted for publication (in revised form) August 11, 2003; published electronically March 3, 2004.

<http://www.siam.org/journals/sisc/25-5/40669.html>

[†]J. Blaustein Institute for Desert Research, Environmental Hydrology & Microbiology, Ben-Gurion University, Sede Boker Campus, 84990 Israel (borisov-liza@bezeqint.net).

[‡]J. Blaustein Institute for Desert Research, Environmental Hydrology & Microbiology, Ben-Gurion University, Sede Boker Campus, 84990 Israel, and Department of Mechanical Engineering, Perlstone Center for Aeronautical Studies, Ben-Gurion University, Beer Sheva, 84105 Israel (sorek@bgumail.bgu.ac.il).

time level, into a monotone increasing (or decreasing, respectively) function $u(x)$ at a prescribed time level. In view of this, the monotonicity represents an internal characteristic feature of a difference scheme. Specifically, this can refer to the equations of hydrodynamics, heat transfer problems, etc. (see, e.g., [16], [24]). Hereinafter such schemes will also be referred to as monotonicity preserving [14] or monotone in terms of Godunov (Godunov-monotone). Godunov [12, p. 275] proved that explicit schemes with constant coefficients in the form

$$(1.2) \quad u_i = \sum_{n=-k}^k a_n \check{u}_{i+n}$$

will be monotonicity preserving *iff* (i.e., if and only if) $a_n \geq 0$. Godunov's definition of monotonicity has been extended by Harten et al. [13] to the case of a nonlinear scheme in the following way. A difference scheme

$$(1.3) \quad u_i = H(\check{u}_{i-k}, \check{u}_{i-k+1}, \dots, \check{u}_{i+k})$$

is said to be monotone if H is a monotone increasing function of each of its arguments, namely

$$(1.4) \quad a_n^* \equiv \frac{\partial}{\partial w_n} H(w_{-k}, \dots, w_k) \geq 0, \quad -k \leq n \leq k.$$

Hereinafter such schemes will be referred to as Harten-monotone. The foregoing notion of monotonicity has been extended in many works (e.g., [1], [7], [10], [11], [27]).

Starting with the above-mentioned Godunov's theorem [12, p. 275] and the notion of monotonicity developed by Harten et al. [13], we suggest that a_n^* of (1.4) should be associated with the notion of a variational scheme for (1.3) in which these coefficients are functions of the $\check{u}$ values,

$$(1.5) \quad \delta u_i = \sum_{n=-k}^k a_n^* \delta \check{u}_{i+n}, \quad a_n^* = a_n^*(\check{u}_{i-k}, \check{u}_{i-k+1}, \dots, \check{u}_{i+k}),$$

and thus (1.5) represents a tangent space at a point of the manifold (1.3). We stress that (1.5) suggests that *any scheme in the form of (1.3) should be investigated for monotonicity in terms of its corresponding variational scheme, which will always be linear (although it may be emanating from a nonlinear operator) in terms of $\delta \check{u}_{i+n}$* . Moreover, we note that a variational representation implemented in a scheme (e.g., (1.5)) inherently addresses the notion of monotonicity given that the module of variation of the variable u_i should not exceed the module of variations of the variables $\check{u}_{i+n}$.

Godunov and Harten had a considerable impact on the evolution of methodologies for investigating the monotonicity of difference schemes, specifically with Harten's theorem [14, Theorem 2.1] concerning the interconnection between the two (Godunov [12], Harten et al. [13]) notions of monotonicity. The essence of this theorem states that

- $\langle H - 1 \rangle$ any monotone (i.e., Harten-monotone) scheme is total variation diminishing (TVD);
- $\langle H - 2 \rangle$ any TVD scheme is monotonicity preserving (i.e., monotone in terms of Godunov [12]).

Harten's theorem, considering $\langle H - 1 \rangle$ and $\langle H - 2 \rangle$, was proven in [14] for a specific class of conservative schemes approximating a 1-D initial value problem of a nonlinear hyperbolic PDE. Harten's theorem [14, p. 360] is also valid for a wider range of difference schemes as was proven, e.g., in [11] concerning scheme (1.3) for $H \equiv \check{u}_i - \check{h}(\check{u}_{i-k}, \check{u}_{i-k+1}, \dots, \check{u}_{i+k})$, where the difference operator $\check{h}$ satisfies the condition $\check{h}(u, u, \dots, u) = 0$. Let us also note that the $\langle H - 2 \rangle$ statement of Harten's theorem [14] is valid for the scheme

$$(1.6) \quad u_i = H_i(\check{u}_{i-k}, \check{u}_{i-k+1}, \dots, \check{u}_{i+k}),$$

where the nonlinear scheme operators H_i satisfy the condition

$$(1.7) \quad H_i(u, u, \dots, u) = u;$$

i.e., (1.6) transforms a constant into the same constant (see, e.g., [12], [11]). Thus we have the following proposition.

PROPOSITION 1.1. *Let a difference scheme be expressed in the form of (1.6). If (1.6) is TVD and (1.7) is valid, then the difference scheme is monotonicity preserving.*

Proposition 1.1 can be proven in exact analogy to the proof in [14, Theorem 2.1, statement (ii)] (see also [11]). Notice that the scheme operators H_i in (1.6) may depend on variable parameters.

Equation (1.7) is an important condition for a scheme to be monotonicity preserving. To demonstrate, we consider the scheme

$$(1.8) \quad u_i = \frac{1 - \alpha_i}{1 + \beta_i} \check{u}_i + \frac{\alpha_i}{1 + \beta_i} \check{u}_{i-1}, \quad 0 \leq \alpha_i \leq 1, \quad \beta_i \geq 0, \quad i = 0, \pm 1, \dots,$$

where $\beta_i = 0$ for all $i < 1$, $\beta_1 = (1 - \alpha_1)\varepsilon$, and $\beta_i = 2\varepsilon$ for all $i > 1$, $\varepsilon \geq 0$. We note that (1.7) is not valid for (1.8) if $\varepsilon > 0$. If we assign the grid function $\check{u}$ so that $\check{u}_i = 1$ for all points $i \leq 0$ and $\check{u}_i = 1 + \varepsilon$ for all points $i > 0$, then in view of (1.8) the grid function u becomes $u_i = 1$ for all $i \leq 1$, and $u_i = (1 + \varepsilon)/(1 + 2\varepsilon)$ for all $i > 1$. Obviously, the scheme (1.8) is Harten-monotone, as all its coefficients are nonnegative. Moreover, if $\alpha_{i+1} \leq \alpha_i$ for all $i = 0, \pm 1, \dots$, then the scheme operator is L_1 -contraction mapping (in the sense of [28]), and hence this scheme is TVD [14, p. 360]. Nevertheless, the monotone increasing function $\check{u}_i$ is transformed into a decreasing function u_i if $\varepsilon > 0$. However, in view of (1.7) we obtain from (1.8) that $\varepsilon = 0$, i.e., $\beta_i = 0$ for all $i = 0, \pm 1, \dots$. In such a case, scheme (1.8) with variable coefficients ($\alpha_i \neq \text{const}$) becomes

$$(1.9) \quad u_i = \check{u}_i - \alpha_i(\check{u}_i - \check{u}_{i-1}), \quad 0 \leq \alpha_i \leq 1, \quad i = 0, \pm 1, \dots$$

We may interpret (1.9) as an upwind scheme (with variable space increment) approximating (1.1). Notice that, even though the scheme operator of (1.9) would not be L_1 -contraction mapping, scheme (1.9) itself is TVD (the proof is analogous to the proof in [11, Theorem 2.3, item 2]), and hence, in view of Proposition 1.1, is monotonicity preserving.

We thus realize that nonnegativity of the coefficients is not, in general, a sufficient condition for linear schemes with variable coefficients to be monotonicity preserving. Moreover, the necessary condition associated with the nonnegativity was proven only for linear schemes with constant coefficients. Hence, Godunov's theorem [12, p. 277] (see also [14, p. 361]) concerning the order of accuracy of monotone schemes is proven

only for schemes with constant coefficients, as the proof is based on the assertion that a scheme is monotone only if its coefficients are nonnegative.

Let us formulate the necessary condition for a linear scheme to be monotonicity preserving. Consider the scheme written in the form

$$(1.10) \quad u_i = a_i^j \check{u}_j.$$

Here and in what follows, repeating superscripts together with subscript indexes denote summation. Assuming that (1.10) will be monotone in terms of Godunov, it will transform a constant ($\check{u}_j = A \neq 0$ for all j) into a constant ($u_i = B$ for all i), since such functions can simultaneously be considered as increasing and decreasing. Then, in view of (1.10), we have $B = A \sum_j a_i^j$. Consequently, the set of equalities

$$(1.11) \quad \sum_j a_i^j = C = \text{const} \quad \forall i$$

is the necessary condition for the monotonicity (in terms of Godunov) of scheme (1.10) with possible variable coefficients.

Let us note that the constant C in (1.11) is equal to a unit if (1.10) is an approximation to (1.1), since a constant being a solution to (1.1) must be a solution to (1.10), as it is often assumed (e.g., [12], [11]). In such a case we obtain instead of (1.11) that

$$(1.12) \quad \sum_j a_i^j = 1 \quad \forall i.$$

In a similar manner we can prove that (1.7) is a necessary condition for (1.6) to be monotonicity preserving. Thus (see also Proposition 1.1) we obtain the validity of the following.

PROPOSITION 1.2. *A TVD scheme written in the form of (1.6) will be monotonicity preserving iff condition (1.7) is valid.*

However, Harten's theorem [14, p. 360] does not guarantee absence of spurious oscillations, as can be demonstrated by the following scheme with nonnegative and constant coefficients:

$$(1.13) \quad u_i = \alpha \check{u}_{i-1} + \beta \check{u}_i + \alpha \check{u}_{i+1}, \quad \beta = C - 2\alpha, \quad i = 0, \pm 1, \dots,$$

where $C > 0$ and $0 \leq \alpha \leq C/2$. Scheme (1.13), being Harten-monotone ($\alpha, \beta \geq 0$), is also monotonicity preserving [12, p. 275] and TVD [11, Theorem 2.3]. Assigning a grid function $\check{u}_i$ such that it has a unique local maximum, namely

$$(1.14) \quad \check{u}_i = 1 \text{ if } i = 0, \quad \check{u}_i = 0 \text{ if } i = \pm 1, \pm 2, \dots,$$

we obtain $u_{-1} = u_1 = \alpha$, $u_0 = \beta$, and $u_i = 0$ at all other nodes i . We note that the grid function u_i can have two local maxima at $i = \pm 1$ if we take $\alpha = 5C/12$, for which $\beta = C/6$ by virtue of (1.13). If $C = 1$ and $\alpha = \tau/h^2$, where τ and h are time and space increments, respectively, then (1.13) will be an explicit scheme (e.g., [4], [29], [22], [16], [32]) approximating the heat equation $U_t = U_{xx}$. In view of analytical solutions (e.g., [4], [22], [16]) to the initial value problem for the above heat equation, we conclude that the oscillations exhibited by (1.13) are spurious.

As the Godunov–Harten approach was mainly developed for building free-of-spurious-oscillations (monotone) schemes associated with hyperbolic systems, let us

investigate the monotonicity in terms of the Godunov–Harten approach of a scheme approximating (1.1). Consider the following difference scheme (Lax scheme [26]; “triangle” [12]) approximating (1.1) ($V = \text{const}$) with first order accuracy:

$$(1.15) \quad u_i = \check{u}_i - \lambda(\bar{f}_{i+0.5} - \bar{f}_{i-0.5}),$$

where

$$(1.16) \quad \bar{f}_{i+0.5} = \frac{\check{u}_{i+1} + \check{u}_i}{2} - \frac{1}{2\lambda}(\check{u}_{i+1} - \check{u}_i),$$

$\lambda = V\tau/h$. If $|\lambda| \leq 1$, then (1.15) with (1.16) will be conservative and consistent with (1.1) (see [14, p. 358], [13, p. 299]), consistent with the entropy condition (see [14, p. 358], [13, p. 300]), Harten-monotone [13, p. 299], TVD [11, Theorem 2.3], and monotonicity preserving [12, p. 275]. Nevertheless, this scheme can still produce spurious oscillations. Actually, if we take $\lambda = 0.5$ and the grid function $\check{u}_i$ as in (1.14), then we obtain from (1.15) the grid function u_i as follows: $u_i = 0.25$ for $i = -1$, $u_i = 0.75$ for $i = 1$, and $u_i = 0$ for all $i \neq \pm 1$. Thus, $\check{u}_i$ having a unique local maximum (at $i = 0$) is transferred by (1.15) with (1.16) into u_i with two local maxima, namely at $i = \pm 1$ (cf. [7]). In view of the well-known analytical solution of (1.1), being a stationary wave $y = f(x - Vt)$ with $y(x, 0) = f(x)$, we note that (1.15) subject to (1.16) produces spurious oscillations. Hence, we conclude that the Godunov–Harten approach is not sufficient in guaranteeing the absence of spurious oscillations.

Following Godunov [12] and Ostapenko [27], the monotonicity of a scheme can be defined more restrictively. In Ostapenko [27] the notion of strong monotonicity for two-time-level schemes is introduced and the associated criteria for three-space-point schemes are developed to avoid spurious oscillations of difference derivatives.

In view of Ostapenko [27], let a grid function v_i be referred to as a “ $\wedge$ -function” (or a “ $\vee$ -function”) if there exist grid nodes m and n such that $m \leq n$; $v_m > v_{m-1}$ and $v_i \geq v_{i-1}$ ($v_m < v_{m-1}$ and $v_i \leq v_{i-1}$ for the $\vee$ -function) if $i < m$; $v_i = \text{const}$ if $m \leq i \leq n$; $v_n > v_{n+1}$ and $v_i \geq v_{i+1}$ ($v_n < v_{n+1}$ and $v_i \leq v_{i+1}$ for the $\vee$ -function) if $i > n$. Notice that the set of μ -functions (or η -functions) introduced in [27] is the union of the $\wedge$ -functions (or the $\vee$ -functions, respectively) and the set of monotone functions.

DEFINITION 1.3. *A scheme will be referred to as Godunov–Ostapenko-monotone (or GO-monotone) if it is monotonicity preserving and transforms a $\wedge$ -function (or $\vee$ -function) into a μ -function (or into an η -function, respectively). A nonlinear scheme, e.g., (1.3), will be referred to as linearly GO-monotone if its variational scheme, e.g., (1.5), is GO-monotone.*

PROPOSITION 1.4. *A linear scheme that can be expressed by (1.10) will be GO-monotone only if (1.11) is valid and the scheme coefficients of any column of the scheme matrix are the values of a μ -function, i.e., a_i^j is a μ -function of i for all j .*

Proof. The necessity of (1.11) was already proven above. Let k be the scheme matrix column number and δ_j^k denote the Kronecker delta. Assuming that (1.10) will be GO-monotone, it will transform $\check{u}_j = \delta_j^k$ into a μ -function as δ_j^k is a $\wedge$ -function of j . From (1.10) we obtain that $u_i = a_i^j \delta_j^k = a_i^k$. Hence, a_i^k is a μ -function of i . $\square$

An analogy to Proposition 1.4 for schemes with constant coefficients was proven in [26]. By virtue of Proposition 1.4 we note that the (1.13) scheme can be GO-monotone only if $0 \leq \varepsilon \leq C/3$, $C \geq 0$. Analogously, in view of Proposition 1.4 we obtain that the scheme (1.15) with (1.16) can be GO-monotone only if $|\lambda| = 1$.

DEFINITION 1.5. *A numerical scheme approximating a PDE will be referred to as shape preserving if the scheme possesses qualitative characteristics of the PDE, such as (i) an initially monotone function transformed into a monotone one; (ii) a function initially having one local extremum (maximum or minimum) transformed into a function with no more than one local extremum (maximum or minimum, respectively).*

Following Definition 1.5 we note that GO-monotone schemes are shape preserving. However, a shape preserving scheme (even being GO-monotone) can nevertheless produce spurious oscillations. Let us assume that scheme (1.10) is GO-monotone. Note that multiplication of a grid function by a positive constant is shape preserving mapping; thus, using a constant parameter $\beta > 1$, we convert (1.10) to read

$$(1.17) \quad v_i = b_i^j \check{u}_j, \quad b_i^j = \beta a_i^j, \quad \beta > 1.$$

Scheme (1.17) is GO-monotone, as it transforms a grid function $\check{u}_j$ into $v_i \equiv \beta u_i$, where u_i is defined by scheme (1.10) assumed as GO-monotone. However, with a sufficiently large value of β , scheme (1.17) will not be TVD, as $TV(v) = \beta TV(u) > TV(\check{u})$ if $\beta > TV(\check{u})/TV(u)$. Hence, (1.17) does not preserve the monotonicity property (see Harten [14]) and can produce spurious oscillations. As the GO-monotone scheme is associated with shape preservation, we also need to account for quantitative measures related to contraction [28]. A scheme operator will be an L_1 -contraction mapping only if the scheme is TVD [14, p. 360]. Moreover, one can assume that the set of grid functions $\{u, v, \dots\}$ is equipped with the semimetric $\rho(u, v) \equiv TV(u - v)$. However, as the TVD scheme operator will be a contraction mapping of the semimetric space of grid functions into itself, we conclude that contraction by itself is not always sufficient for a scheme to be free of spurious oscillations. We will emphasize this fact in section 5.

Recall that there exists another approach to investigating the monotonicity of a difference scheme, namely the maximum principle [30], [32]. In its simplest form it states [11], [17], [22], [23], [32] that in the absence of a source term in a PDE (parabolic, elliptic, or transport equation) the maximum positive as well as minimum negative values of the dependent variable occur at the boundary (referred to as a boundary maximum principle in [6], [36]). Hereinafter, in the case of time-dependent equations, initial conditions are also referred to as boundary conditions (see, e.g., [17]) if not stated otherwise. If the boundary maximum principle is preserved by a difference scheme approximating the PDE, then the scheme, in general, is asymptotically stable and does not produce spurious oscillations [3], [6], [20], [23], [22], [27], [31], [34]. A scheme is regarded as monotone if the boundary maximum principle is maintained [2], [30], [32]. Hereinafter a scheme will be referred to as monotone in terms of Samarskiy (Samarskiy-monotone) if the boundary maximum principle is valid.

The boundary maximum principle is valid [23], [32] for (1.10) if

$$(1.18) \quad a_i^j \geq 0, \quad \sum_j a_i^j \leq 1.$$

We note that in (1.18) there is no requirement for constant coefficients, and hence scheme (1.8) is Samarskiy-monotone since (1.18) is fulfilled for (1.8); however, in view of (1.11), scheme (1.8) is not Godunov-monotone if $\varepsilon > 0$. In a similar way we obtain that scheme (1.13) is monotone in terms of Samarskiy if $0 \leq C \leq 1$ and $0 \leq \alpha \leq C/2$. In the case where $C > 1$, scheme (1.13) being Harten-monotone is not monotone in terms of Samarskiy. Let us note that scheme (1.13) can be both GO-monotone and

Samarskiy-monotone only if $0 \leq \alpha \leq C/3$, $0 \leq C \leq 1$. Hence, a scheme can be monotone in one sense yet not so in another.

This paper is mainly devoted to elaboration of the methods for testing difference scheme monotonicity on the basis of such a quantitative characteristic as contraction [28]. On the basis of the notion of the boundary maximum principle and that of the variational scheme (1.5), we thus suggest the following definition.

DEFINITION 1.6. Consider $\mathbf{y}$ and $\mathbf{g}$ to denote the vectors belonging to the linear vector space X with $\delta\mathbf{y}$ and $\delta\mathbf{g}$ as their variations, and let $\mathbf{H}$ denote a difference operator (linear or nonlinear). The scheme $\mathbf{H}(\mathbf{y}) = \mathbf{g}$ will be referred to as Samarskiy-monotone if variations of its solution satisfy the boundary maximum principle. The scheme will be referred to as linearly monotone if solutions to its variational scheme,

$$(1.19) \quad \tilde{\mathbf{A}} \cdot \mathbf{x} = \tilde{\mathbf{f}}, \quad \tilde{\mathbf{A}} \equiv \frac{\partial \mathbf{H}}{\partial \mathbf{y}}, \quad \mathbf{x} \equiv \delta\mathbf{y}, \quad \tilde{\mathbf{f}} \equiv \delta\mathbf{g},$$

satisfy the maximum principle.

Hereinafter all functions are assumed to be differentiable in an appropriate sense. In view of Definition 1.6 and (1.19), we use linear schemes to investigate conditions for Samarskiy monotonicity, addressing linear as well as nonlinear operators. Justification of this approach is presented in section 6.

In what follows, scheme (1.19) will often be rewritten in the form

$$(1.20) \quad \mathbf{x} = \mathbf{A} \cdot \mathbf{x} + \tilde{\mathbf{f}}; \quad \mathbf{I} - \mathbf{A} = \tilde{\mathbf{A}} = \begin{Bmatrix} \tilde{\mathbf{A}}_{11} & \tilde{\mathbf{A}}_{12} \\ 0 & \tilde{\mathbf{I}} \end{Bmatrix},$$

where $\mathbf{I}$ and $\tilde{\mathbf{I}}$ denote the identity matrices; diagonal terms of $\tilde{\mathbf{A}}$ are equal to units.

If $\tilde{\mathbf{A}}$ in (1.19) is partitioned [18] into blocks of matrices and $\mathbf{x}$ and $\tilde{\mathbf{f}}$ into lower dimensional vectors, then (1.19) will be referred to as a vector difference scheme or as a scalar difference scheme, otherwise.

The maximum principle has been developed for scalar difference schemes to evaluate, in the Chebyshev norm, the stability of the schemes approximating elliptic, parabolic, and transport equations [2], [3], [16], [23], [22], [29], [32], [33], [35], [36]. Habitually, we find that vector schemes are converted into scalar ones for the investigation of their monotonicity via the maximum principle.

Investigations led to the conclusion that the validity of the maximum principle for difference schemes restricts the off-diagonal coefficients of $\mathbf{A}$ in (1.20) to be nonnegative or of $\tilde{\mathbf{A}}$ in (1.19) to be monotone [35]. These restrictions could be too demanding [20], and thus many attempts to circumvent them were reported [2], [3], [33], [35]. Stoyan [35] attempted to extend the validity of the maximum principle to $\tilde{\mathbf{A}}$ matrices that are not monotone. A theorem was suggested [35] claiming that the maximum principle is valid for the absolute values of the elements of $\mathbf{x}$ in (1.19) iff $\tilde{\mathbf{A}}$ is a strictly diagonally dominant Hermitian matrix. However, the claim in [35] about the sufficient condition is not proven in the case of a nonzero right-hand side (RHS) of (1.19), since it is assumed to be obvious with the understanding that $\tilde{\mathbf{A}}$ in (1.19) is nonsingular. Notice also that the diagonal dominance of $\tilde{\mathbf{A}}$ is not, in general, a necessary condition for Samarskiy monotonicity in the absolute values [35, p. 159] of a difference scheme approximating some boundary-value problem for a linear PDE.

This can be demonstrated following (1.19) in the form

$$(1.21) \quad \tilde{\mathbf{A}} \cdot \mathbf{x} = \begin{Bmatrix} 1 & -1 & 1 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{Bmatrix} \begin{Bmatrix} -\frac{2}{3}f \\ \frac{1}{3}f \\ f \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ f \end{Bmatrix} = \tilde{\mathbf{f}}.$$

Notice that (1.21) is Samarskiy-monotone in $|x_i|$, yet $\tilde{\mathbf{A}}$ is not diagonally dominant.

In section 2, we address the classic boundary maximum principle (see, e.g., [32], [23]) and suggest that it was proven only for the case of a specific grid connectedness, which could be associated with the numerical solution of an elliptic PDE. This feature of the classic maximum principle also appears in [36] stating that if $\tilde{\mathbf{A}}$ in (1.20) is a nonsingular M -matrix and weakly row diagonally dominant, then $\tilde{\mathbf{A}}_{11}$ is irreducible¹ to ensure that the boundary maximum principle [36] is valid. In [34] the maximum principle was proven assuming that $\tilde{\mathbf{A}}_{11}$ is reducible and nonsingular. We prove that the boundary maximum principle is valid in the case of a scalar scheme (1.19) such that the directed graph [18] for $\tilde{\mathbf{A}}$ need not be strongly connected. We prove that a weakly row diagonally dominant L -matrix [36] $\tilde{\mathbf{A}}$ satisfies the maximum principle [36] even if $\tilde{\mathbf{A}}_{11}$ could be reducible with no presumption of $\tilde{\mathbf{A}}_{11}$ being nonsingular. Furthermore, we later remove the need for $\tilde{\mathbf{A}}$ to be an L -matrix when proving the validity of the boundary maximum principle for the absolute values of the elements of $\mathbf{x}$ in (1.19). We develop also a criterion to investigate the monotonicity of vector difference schemes (without conforming them into scalar schemes) in terms of the boundary maximum principle. In section 3, we establish an additional notion for the monotonicity of vector difference schemes and its associated monotonicity criteria in terms of Definition 1.6. Furthermore, we provide the concept of submonotonicity and its criteria for vector difference schemes. In section 5 we discuss the interrelationship between different criteria of monotonicity applied to the vector schemes, as well as to the scalar form of these schemes, and discuss the interplay between monotonicity and stability, as well as exemplify (also in section 4) the monotonicity criteria.

2. Maximum principle. Let $\mathbb{R}$ and $\mathbb{C}$ denote the fields of real and complex numbers, respectively, and $\mathbb{K}$ denote either of these fields. Let L denote the finite dimensional vector space over the field $\mathbb{K}$ in which complex conjugates are defined; $L^n = L \times L \times \cdots \times L$ (n -times) denotes the Cartesian product; $\mathbf{y} = \{y_1, \dots, y_N\}^T$ denotes the column-vector which is an element of L ; $\mathbf{y}^* \equiv \bar{\mathbf{y}}^T$, for which $\bar{\mathbf{y}}$ is the member of L whose elements are the complex conjugates of $\mathbf{y}$; $\mathbf{A}$ denotes the matrix (operator) defined on a linear vector space; $(\mathbf{x}, \mathbf{y}) \equiv \mathbf{x}^* \cdot \mathbf{y}$ denotes the inner product of two $\mathbf{x}$ and $\mathbf{y}$ vectors of the space; $\|\mathbf{x}\|$ denotes any vector norm ($\|\mathbf{x}\| = |\mathbf{x}|$ if $N = 1$) and $\|\mathbf{A}\|$ denotes the matrix norm compatible with, or subordinate to (induced by) [18] a prescribed vector norm; $\|\mathbf{A}\|_p$ denotes the matrix norm induced by the vector norm $\|\mathbf{x}\|_p = (\sum_i |x_i|^p)^{1/p}$, $p \geq 1$.

Let Ω be a finite set of grid nodes of a finite dimensional Euclidean space; $i \in \Omega$ denotes a node of the grid Ω , and P_i denotes a prescribed subset (or stencil in [11]) surrounding the i node and including it.

Let $\mathbf{y}_i \in L$ denote vector-valued functions of grid (VFG) nodes. The equation associated with $\mathbf{y}_i$ will be written in the canonical form [32]

$$(2.1) \quad \Lambda \mathbf{y}_i \equiv \mathbf{A}_i \cdot \mathbf{y}_i - \mathbf{B}_i^j \cdot \mathbf{y}_j = \mathbf{f}_i, \quad j \in P_i^*, \quad i = 1, 2, \dots, I,$$

¹A square matrix is irreducible iff its directed graph is strongly connected [18].

where the regular matrix $\mathbf{A}_i$ and the matrix $\mathbf{B}_i^j$ are prescribed matrix-valued functions of grid nodes; $\mathbf{f}_i \in L$ denotes a prescribed VFG and $P_i^* (= P_i \setminus \{i\})$ denotes a neighborhood of the i node, which is the stencil P_i excluding the i node.

A grid node will be referred to as a boundary node if at this point the VFG is equal to a prescribed vector $\mathbf{g}_i$. In such a case we write

$$(2.2) \quad \mathbf{y}_i = \mathbf{g}_i.$$

Actually, (2.2) can be viewed as (2.1) for which $\mathbf{A}_i$ becomes the identity matrix, P_i^* becomes an empty set, and $\mathbf{f}_i = \mathbf{g}_i$. A node will be referred to as interior if the neighborhood P_i^* is not an empty subset. Denoting ω as the set of interior nodes and γ as the set of boundary nodes, we have $\Omega = \omega + \gamma$.

In [4] and [9], we find that a grid connectedness sequence is introduced in association with a scheme for solving a PDE at a region where it becomes hyperbolic, and a different grid connectedness sequence is introduced for a region where it becomes elliptic. Moreover, in [4] it is demonstrated that spurious oscillations will occur if the first grid connectedness sequence is implemented in the elliptic region or the second in the hyperbolic region.

Let the Ω grid be referred to as E-connected (EC) if there is a possibility of passing from any $i \in \omega$ node to an arbitrary $j \in \Omega$ node through the use of a sequence of P_i^* neighborhoods of nodes; i.e., there exist the nodes $i_1, i_2, \dots, i_k$ such that $i_1 \in P_i^*$, $i_2 \in P_{i_1}^*, \dots, i_k \in P_{i_{k-1}}^*, j \in P_{i_k}^*$. This actually reflects features associated with the solution of an elliptic PDE (as will be demonstrated in Examples 4.3 and 4.5).

The Ω grid will be referred to as H-connected (HC) if there exists a path from any interior $i \in \omega$ node to any that belongs to Ω_i^{HC} , a subset of Ω such that $\Omega_i^{HC} \subset \Omega$ but $\cup_i \Omega_i^{HC} \neq \Omega$ (see Example 4.1). We will consider the HC grids containing at least one boundary node in every Ω_i^{HC} . The HC notion relates to features associated with the solution of a hyperbolic PDE.

We will refer to a connected (C) grid as one that is either an EC or HC grid or a combination of these two grids with the possibility of passing from the EC grid to the HC .

2.1. Implementation of the classic maximum principle to C grids. The classic maximum principle was developed to investigate the monotonicity of scalar difference schemes associated with specific grid configuration [23], [32] which, according to our above definition, can be regarded as the EC grid. Therefore, we will consider the C grid containing the HC grid; hence, by default, there exists a path from every interior node to at least one boundary node.

THEOREM 2.1. *Consider a scalar grid function y_i defined on a C grid at $i \in \omega$ and a scalar function g_i prescribed on the $i \in \gamma$ boundary, and let*

$$(2.3) \quad \Lambda y_i \equiv a_i y_i - b_i^j y_j = f_i, \quad j \in P_i^*, \quad i \in \omega; \quad y_i = g_i, \quad i \in \gamma;$$

$$(2.4) \quad a_i > 0, \quad b_i^j > 0, \quad c_i \equiv a_i - \sum_{j \in P_i^*} b_i^j \geq 0 \quad \forall i \in \omega.$$

If $\Lambda y_i \leq 0$ (or $\Lambda y_i \geq 0$) for all i , then the maximum positive (or minimum negative, respectively) value of y_i at the interior i nodes cannot be greater (or less, respectively) than the value corresponding to y_i at the boundary nodes. Hence,

$$(2.5) \quad \max_{i \in \omega} y_i \leq \max_{i \in \gamma} y_j \quad \left(\min_{i \in \omega} y_i \geq \min_{i \in \gamma} y_j \right).$$

In what follows we will focus on the condition $\Lambda y_i \leq 0$, noting that it is actually also similar to $\Lambda y_i \geq 0$ when replacing y_i with $-y_i$.

Proof. Suppose that y_m will denote the maximal positive value at the interior $m \in \omega$ node so that

$$(2.6) \quad y_m = M = \max_{i \in \Omega} y_i.$$

Hence, in view of (2.6) we can argue that

$$(2.7) \quad M \geq \max_{j \in \gamma} y_j.$$

As $y_m \geq y_j$ for all $j \in P_m^*$, and in view of (2.3) and (2.4), we write

$$(2.8) \quad \Lambda y_m = c_m y_m + \sum_{j \in P_m^*} b_m^j (y_m - y_j) \geq c_m y_m \geq 0.$$

Considering the condition $\Lambda y_m \leq 0$, we conclude that

$$(2.9) \quad \Lambda y_m = 0, \quad c_m = 0, \quad y_j = y_m = M \quad \forall j \in P_m^*.$$

In the case when $\Lambda y_m < 0$ for at least one member of P_m^* , we obtain the contradiction to (2.8) that proves Theorem 2.1. By virtue of the C grid, there exist a boundary $l \in \gamma$ node and interior $i_1, i_2, \dots, i_k$ nodes such that the path from $m \in \omega$ to l can be given by

$$(2.10) \quad i_1 \in P_m^*, \quad i_2 \in P_{i_1}^*, \dots, i_k \in P_{i_{k-1}}^*, \quad l \in P_{i_k}^*.$$

In view of (2.9) and (2.10) we conclude that $y_{i_1} = M$. Using similar arguments it can be shown that $y_{i_2} = M, \dots, y_{i_k} = M, y_l = M$. However, since along the boundary we have

$$(2.11) \quad M = y_l \leq \max_{j \in \gamma} y_j,$$

hence in view of (2.6), (2.7), and (2.11), the only valid option becomes

$$(2.12) \quad M = \max_{i \in \Omega} y_i = \max_{j \in \gamma} y_j.$$

We note that (2.12) actually manifests Theorem 2.1; namely, it implies that the classic maximum principle is valid for the general C grid. $\square$

2.2. The maximum principle for vector difference equations. The maximum principle for (2.1) can be addressed by the following.

THEOREM 2.2. *Let a VFG $\mathbf{y}_i$ be defined on the C grid Ω ; in reference to (2.1), consider $\mathbf{A}_i^j \equiv \mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j \neq \mathbf{0}$ for all $i \in \omega$ and for all $j \in P_i^*$ together with $\mathbf{f}_i = \mathbf{0}$. If*

$$(2.13) \quad B_i \equiv \sum_{j \in P_i^*} \|\mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j\| \leq 1 \quad \forall i \in \omega,$$

then

$$(2.14) \quad \max_{i \in \omega} \|\mathbf{y}_i\| \leq \max_{i \in \gamma} \|\mathbf{y}_i\|.$$

Proof. As $\mathbf{f}_i = 0$, we can conclude from (2.1) that

$$(2.15) \quad \mathbf{y}_i = \mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j \cdot \mathbf{y}_j \equiv \mathbf{A}_i^j \cdot \mathbf{y}_j, \quad j \in P_i^*.$$

By virtue of (2.15) we obtain $\|\mathbf{y}_i\| \leq \|\mathbf{A}_i^j\| \|\mathbf{y}_j\|$ or

$$(2.16) \quad \Lambda \|\mathbf{y}_i\| \equiv \|\mathbf{y}_i\| - \|\mathbf{A}_i^j\| \|\mathbf{y}_j\| \leq 0, \quad j \in P_i^* \quad \forall i \in \omega.$$

Since $\mathbf{A}_i^j \neq \mathbf{0}$, then $\|\mathbf{A}_i^j\| > 0$, and by virtue of (2.13) and (2.16), we conclude that the difference operator $\Lambda \|\mathbf{y}_i\|$ fulfills the condition of Theorem 2.1. Thus (2.14) is valid. $\square$

Remark 2.3. Let us note that if in (2.15) we consider the case of $\mathbf{A}_i^j = \mathbf{0}$, then $\mathbf{B}_i^j = \mathbf{0}$, as the matrix $\mathbf{A}_i$ is assumed to be regular. In fact, $\mathbf{B}_i^j = \mathbf{0}$ means that the associated $\mathbf{y}_j$ vector does not belong to the neighborhood P_i^* . Consequently, the restriction $\mathbf{A}_i^j \neq \mathbf{0}$ is not significant for the existence of (2.13).

Hereinafter, in line with Definition 1.6, a scheme will be referred to as Samarskiy-monotone if the scheme complies with (2.13).

Remark 2.4. If (1.19) is a scalar scheme, then (2.13) when applied to scheme (1.19) signifies that $\tilde{\mathbf{A}}$ of (1.19) will be weakly row diagonally dominant to ensure the Samarskiy monotonicity of (1.19) (cf. [35]).

The use of (2.13) for scalar schemes will be demonstrated in Examples 4.1, 4.2, 4.3, and 4.4. The application of (2.13) to vector schemes will be described in Examples 4.5 and 4.6.

3. Monotonicity and submonotonicity. The demand that vector difference schemes comply with (2.4) or (2.13) can be too restrictive. Let us hence introduce for system (2.1) a new criterion of monotonicity based on a new notion.

3.1. Monotone schemes. Consider (2.1) in the form

$$(3.1) \quad \mathbf{Y} = \mathbf{A} \cdot \mathbf{Y} + \mathbf{F},$$

where $\mathbf{Y} = \{\mathbf{y}_1^T, \dots, \mathbf{y}_I^T\}^T$, $\mathbf{F} = \{(\mathbf{A}_1^{-1} \cdot \mathbf{f}_1)^T, \dots, (\mathbf{A}_I^{-1} \cdot \mathbf{f}_I)^T\}^T$,

$$(3.2) \quad \mathbf{A} = \{\mathbf{A}_i^j\}, \quad \mathbf{A}_i^j = \begin{cases} \mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j & \text{if } j \in P_i^*, \\ \mathbf{0} & \text{if } j \notin P_i^*, \end{cases} \quad i, j, = 1, \dots, I.$$

We adopt the Chebyshev norm $\|\cdot\|_C$ for the VFG $\mathbf{y}_i$ in the form

$$(3.3) \quad \|\mathbf{Y}\|_C = \max_{1 \leq i \leq I} \|\mathbf{y}_i\|,$$

for which the norm $\|\mathbf{y}_i\|$ on the RHS of (3.3) can be specified arbitrarily. Moreover, we can extend the notion of monotonicity for vector difference equations.

DEFINITION 3.1. Scheme (2.1) will be considered C-monotone if $\mathbf{A}$ in (3.1) is a contraction mapping operator [28], i.e., its matrix norm is subject to the Chebyshev vector norm and fulfills the condition

$$(3.4) \quad \|\mathbf{A}\|_C \leq 1.$$

On the basis of Definition 3.1, let us establish a new monotonicity criterion for a vector scheme.

Consider (2.1) as a transformation from the vector space L^{n_i} onto L , obtained by applying an operator $\mathbf{W}_i$ defined by

$$(3.5) \quad \mathbf{W}_i \equiv \{\mathbf{A}_i^1, \dots, \mathbf{A}_i^j, \dots, \mathbf{A}_i^{n_i}\}, \quad \mathbf{A}_i^j = \mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j.$$

Here and in what follows, n_i denotes the number of nodes in the neighborhood P_i^* . Using (3.5), we rewrite (2.1) to read

$$(3.6) \quad \mathbf{y}_i = \mathbf{W}_i \cdot \mathbf{Y}_i + \mathbf{A}_i^{-1} \cdot \mathbf{f}_i, \quad \mathbf{Y}_i = \{\mathbf{y}_{i_1}^T, \dots, \mathbf{y}_{i_{n_i}}^T\}^T \in L^{n_i}.$$

Let a vector norm in the vector space L^{n_i} be the cubic norm, i.e.,

$$(3.7) \quad \|\mathbf{Y}_i\|_\infty = \max_{j \in P_i^*} \|\mathbf{y}_j\|_\infty.$$

If the condition

$$(3.8) \quad B_i \equiv \|\mathbf{W}_i\|_\infty \leq 1 \quad \forall i \in \Omega$$

is valid, then the scheme can be regarded as *C-monotone*, i.e., a contraction mapping operator with a matrix norm subordinated to (3.3). This proposition can be verified when conforming (2.1) to the form of (3.1) and accounting for $\|\mathbf{Y}\|_\infty = \max_{1 \leq i \leq I} \|\mathbf{y}_i\|_\infty = \max_{1 \leq i \leq I} \|\mathbf{Y}_i\|_\infty$ and $\|\mathbf{A}\|_\infty = \max_{1 \leq i \leq I} \|\mathbf{W}_i\|_\infty$, from which we can conclude that $\mathbf{A}$ of (3.2) is the contraction mapping operator on the $(I \times N)$ -dimensional linear space with the cubic vector norm.

Remark 3.2. We note that (3.8) for a vector scheme is equivalent to (2.13) for the scalar form of the vector scheme. Thus, (3.8) is actually also a criterion of monotonicity in terms of Definition 1.6.

Consider the case when the operators $\mathbf{A}_i^j$ ($\equiv \mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j$) in (2.1) depend on several pairwise permutable normal operators, namely

$$(3.9) \quad \mathbf{A}_i^j = \varphi_i^j(\mathbf{D}^1, \mathbf{D}^2, \dots, \mathbf{D}^l), \quad l \geq 1.$$

Let $\mathbb{K}^l = \mathbb{K} \times \mathbb{K} \times \dots \times \mathbb{K}$ (l -times) denote the l th Cartesian power of the real line (under $\mathbb{K} = \mathbb{R}$) or the complex plane (under $\mathbb{K} = \mathbb{C}$) and let S denote a closed subset of $\mathbb{K}^l$. Let $s(\mathbf{D}^k)$ denote the spectrum of the operator $\mathbf{D}^k$ and let $s(\mathbf{D}^1) \times s(\mathbf{D}^2) \times \dots \times s(\mathbf{D}^l) \subseteq S$. We will assume that S is the Cartesian product of the closed subsets S^k ($k = 1, 2, \dots, l$) belonging to the field $\mathbb{K}$. In the case when the operators $\mathbf{D}^k$ are normal, we will assume that $S^k = \{\lambda \mid \nu^k \leq |\lambda| \leq \mu^k\}$, where $\lambda \in \mathbb{C}$, $\nu^k = \min_{1 \leq m \leq N} |\lambda_m^k|$, $\mu^k = \max_{1 \leq m \leq N} |\lambda_m^k|$, $\lambda_m^k \in s(\mathbf{D}^k)$. In the specific case when all matrices $\mathbf{D}^k$ are Hermitian, we will assume that $S^k = \{\lambda \mid \nu^k \leq \lambda \leq \mu^k\}$, where $\lambda \in \mathbb{R}$, and ν^k and μ^k denote, respectively, the left and right boundaries of the spectrum $s(\mathbf{D}^k)$. Let us also assume that the function $\varphi_i^j(\lambda_1, \lambda_2, \dots, \lambda_l)$ can be represented by a convergent Laurent series at each point $\Lambda \equiv \{\lambda_1, \dots, \lambda_l\} \in S$.

THEOREM 3.3. *Let scheme (2.1) be constructed of the operators $\mathbf{A}_i^j$ ($\equiv \mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j$) such that each obeys (3.9). If (2.1) fulfills*

$$(3.10) \quad B_i \equiv \max_{\Lambda \in S} \sum_{j \in P_i^*} |\varphi_i^j(\Lambda)| \leq 1, \quad \Lambda \equiv \{\lambda_1, \dots, \lambda_l\} \quad \forall i \in \omega,$$

then (2.1) is regarded as C-monotone.

Proof. As $\mathbf{D}^k$ of (3.9) belongs to the set of pairwise commutative normal operators, the matrices of the set are simultaneously unitarily similar to diagonal matrices [21]; i.e., there exists a unitary $\mathbf{U}$ matrix such that $\mathbf{U}^{-1} \cdot \mathbf{D}^k \cdot \mathbf{U}$ will be diagonal, namely

$$(3.11) \quad \mathbf{U}^{-1} \cdot \mathbf{D}^k \cdot \mathbf{U} = \{\lambda_m^k \delta_{mn}\}, \quad m, n = 1, \dots, N,$$

where λ_m^k denotes the m th eigenvalue of $\mathbf{D}^k$, and δ_{mn} denotes the Kronecker delta. Let us rewrite (2.1) in the form of (3.1) and let $\mathbf{q}_j \equiv \mathbf{U}^{-1} \cdot \mathbf{y}_j$, $j = 1, 2, \dots, I$. Then we obtain

$$(3.12) \quad \mathbf{Q} = \mathbf{W} \cdot \mathbf{Q} + \mathbf{F}, \quad \mathbf{Q} = \{\mathbf{q}_1^T, \dots, \mathbf{q}_I^T\}^T, \quad \mathbf{W} = \{\mathbf{W}_i^j\}, \quad i, j = 1, \dots, I,$$

where $\mathbf{F} \equiv \{\mathbf{F}_1^T, \dots, \mathbf{F}_I^T\}^T$, $\mathbf{F}_j \equiv \mathbf{U}^{-1} \cdot \mathbf{A}_j^{-1} \cdot \mathbf{f}_j$. As φ_i^j can be expanded into a Laurent series, every block $\mathbf{W}_i^j \equiv \mathbf{U}^{-1} \cdot \mathbf{A}_i^j \cdot \mathbf{U}$ in (3.12) can be written in the form

$$(3.13) \quad \mathbf{W}_i^j \equiv \mathbf{U}^{-1} \cdot \varphi_i^j(\mathbf{D}^1, \dots, \mathbf{D}^l) \cdot \mathbf{U} = \varphi_i^j(\mathbf{U}^{-1} \cdot \mathbf{D}^1 \cdot \mathbf{U}, \dots, \mathbf{U}^{-1} \cdot \mathbf{D}^l \cdot \mathbf{U}).$$

In view of (3.13) and (3.11) we write $\mathbf{W}_i^j$ of (3.13) in a diagonal form

$$(3.14) \quad \mathbf{W}_i^j = \{\varphi_i^j(\lambda_m^1, \lambda_m^2, \dots, \lambda_m^l) \delta_{mn}\}, \quad m, n = 1, \dots, N.$$

Applying the cubic norm $\|\mathbf{Q}\|_\infty$ on the vector space, we conclude from (3.14) that

$$(3.15) \quad \|\mathbf{W}\|_\infty = \max_i \left(\max_{1 \leq m \leq N} \sum_{j \in P_i^*} |\varphi_i^j(\lambda_m^1, \lambda_m^2, \dots, \lambda_m^l)| \right).$$

As λ_m^k is the m th eigenvalue of $\mathbf{D}^k$ (see (3.9)), we conclude that $\{\lambda_m^1, \lambda_m^2, \dots, \lambda_m^l\} \in S$ for all m , and hence we write

$$(3.16) \quad \max_{1 \leq m \leq N} \sum_{j \in P_i^*} |\varphi_i^j(\lambda_m^1, \lambda_m^2, \dots, \lambda_m^l)| \leq \max_{\Lambda \in S} \sum_{j \in P_i^*} |\varphi_i^j(\Lambda)|.$$

Let us note that the matrix norm in (3.15) is induced by the Chebyshev norm ($\max_j \|\mathbf{y}_j\|$) of VFG $\mathbf{y}_j$ with $\|\mathbf{y}_j\| \equiv \|\mathbf{U}^{-1} \cdot \mathbf{y}_j\|_\infty$. By virtue of (3.10) and (3.16) we conclude from (3.15) that the matrix $\mathbf{W}$ is a contraction mapping operator with a matrix norm subordinated to the Chebyshev norm; i.e., the scheme can be regarded as C-monotone. $\square$

COROLLARY 3.4. *If the grid associated with the scalar form of scheme (2.1) is connected and (3.10) is valid, then scheme (2.1) will be Samarskiy-monotone. Actually, the matrix norm in (3.15) fulfills the inequality $\|\mathbf{W}\|_\infty \leq 1$, which means that (3.8) is valid for scheme (3.12). Hence, in view of Remark 3.2, we conclude that (3.12) is Samarskiy-monotone. Since the cubic matrix norm in (3.15) is subordinated to the vector norm $\|\mathbf{Q}\|_\infty = \max_j \|\mathbf{q}_j\|_\infty = \max_j \|\mathbf{U}^{-1} \cdot \mathbf{y}_j\|_\infty = \max_j \|\mathbf{y}_j\|$, we conclude that (2.1) is also Samarskiy-monotone.*

The criterion of monotonicity, (3.10), will be addressed in Example 4.6.

Remark 3.5. In view of (3.15) we conclude that Theorem 3.3 will also be valid in the case when the left-hand side (LHS) of (3.10) will be superseded by the LHS of (3.16).

Remark 3.6. If the subset S in (3.10) is not closed, then the “maximum” in (3.10) must be superseded by the “supremum” so that Theorem 3.3 will be valid.

3.2. Submonotone schemes. Let us introduce for scheme (2.1) a less restrictive notion of monotonicity.

DEFINITION 3.7. *Scheme (2.1) will be considered submonotone if the matrix norm of $\mathbf{A}$ in (3.1) subject to a vector norm fulfills the condition*

$$(3.17) \quad \|\mathbf{A}\| \leq 1.$$

Following the conditions and notation of Theorem 3.3, we can formulate the following corollary.

COROLLARY 3.8. *If (2.1) fulfills*

$$(3.18) \quad B_j \equiv \max_{\Lambda \in S} \sum_{i \in \omega} |\varphi_i^j(\Lambda)| \leq 1, \quad \Lambda \equiv \{\lambda_1, \dots, \lambda_l\} \quad \forall j \in \omega,$$

then (2.1) is regarded as submonotone. The proof of Corollary 3.8 can be obtained by subordinating the matrix norm of $\mathbf{A}$ in (3.1) to $\|\mathbf{Y}\| = \sum_j \|\mathbf{U}^{-1} \cdot \mathbf{y}_j\|_\infty$, after which it becomes similar to the proof of Theorem 3.3.

An additional condition leading to submonotonicity can be formulated.

THEOREM 3.9. *Let scheme (2.1) be constructed of the operators $\mathbf{A}_i^j$ ($\equiv \mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j$) such that each obeys (3.9). If (2.1) fulfills*

$$(3.19) \quad B_i = \max_{\Lambda \in S} \sqrt{\sum_{j \in P_i^*} |\varphi_i^j(\Lambda)|^2} \leq 1, \quad \Lambda = \{\lambda_1, \dots, \lambda_l\} \quad \forall i \in \omega,$$

then (2.1) is regarded as submonotone.

Proof. Since $\mathbf{D}^k$ of (2.1) can be superseded by a diagonal matrix $\mathbf{U}^{-1} \cdot \mathbf{D}^k \cdot \mathbf{U}$ (see the proof of Theorem 3.3), we will consider $\mathbf{D}^k$ in a diagonal form without any loss of generality.

Consider scheme (2.1) in the form of (3.6). Let us assume that the vector norm in (3.1) is $\|\mathbf{Y}\| = \max_i \|\mathbf{Y}_i\|_2$. Then we obtain

$$(3.20) \quad \|\mathbf{A}\| = \max_{\mathbf{Y} \neq 0} \frac{\|\mathbf{A} \cdot \mathbf{Y}\|}{\|\mathbf{Y}\|} = \max_{\mathbf{Y} \neq 0} \frac{\max_i \|\mathbf{W}_i \cdot \mathbf{Y}_i\|_2}{\max_i \|\mathbf{Y}_i\|_2} \leq \max_i \|\mathbf{W}_i\|_2.$$

Let $\mathbf{W}_i^*$ be the adjoint of $\mathbf{W}_i$. Then $\|\mathbf{W}_i\|_2^2 (= \|\mathbf{W}_i^* \cdot \mathbf{W}_i\|_2 = \|\mathbf{W}_i \cdot \mathbf{W}_i^*\|_2)$, by virtue of the fact that every $\mathbf{D}_i^l$ is a diagonal matrix, will read

$$(3.21) \quad \|\mathbf{W}_i\|_2^2 = \max_{1 \leq m \leq N} \sum_{j \in P_i^*} |\varphi_i^j(\lambda_m^1, \dots, \lambda_m^l) \overline{\varphi_i^j(\lambda_m^1, \dots, \lambda_m^l)}|,$$

leading to

$$(3.22) \quad \|\mathbf{W}_i\|_2 = \sqrt{\max_{1 \leq m \leq N} \sum_{j \in P_i^*} |\varphi_i^j(\lambda_m^1, \dots, \lambda_m^l)|^2} \leq \max_{\Lambda \in S} \sqrt{\sum_{j \in P_i^*} |\varphi_i^j(\Lambda)|^2},$$

which in view of (3.20) manifests the proof of Theorem 3.9. $\square$

Submonotone criterion (3.19) will be demonstrated in Example 4.4 for a complex-valued scheme.

Corollary 3.8 and Theorem 3.9 can be followed by remarks that are similar to Remarks 3.5 and 3.6.

DEFINITION 3.10. *The scheme $\mathbf{H}(\mathbf{y}) = \mathbf{g}$ (see Definition 1.6) will be referred to as linearly submonotone if its variational scheme (1.19) is submonotone.*

3.3. Necessary conditions for monotonicity. Thus far we have dealt with the establishment of sufficient conditions to ensure the monotonicity of a scheme. In what follows we will consider necessary conditions for explicit schemes. The explicit schemes are defined as having neighborhoods of all interior points that contain the boundary nodes (2.2) only (the schemes will be defined as being implicit, otherwise). Consider system (2.1) in the form of (3.1). Let $\mathbf{Y} = \{\mathbf{Z}^T, \mathbf{X}^T\}^T$, $\mathbf{Z} = \{\mathbf{z}_1^T, \dots, \mathbf{z}_n^T\}^T$, $\mathbf{X} = \{\mathbf{x}_1^T, \dots, \mathbf{x}_{I-n}^T\}^T$, where the VFG $\mathbf{z}_i \equiv \mathbf{y}_i$ ($i = 1, \dots, n$) is defined at the interior nodes and the VFG $\mathbf{x}_{i-n} \equiv \mathbf{y}_i$ ($i = n+1, \dots, I$) is defined at the boundary nodes. Since (3.1) is explicit, for this scheme we write

$$(3.23) \quad \mathbf{Z} = \mathbf{A}' \cdot \mathbf{X}.$$

THEOREM 3.11. *Let R_i^* be a subset of a neighborhood P_i^* , $i \in \omega$. If scheme (2.1) is explicit, then*

$$(3.24) \quad \left\| \sum_{j \in R_i^*} \mathbf{A}_i^j \right\| \equiv \left\| \sum_{j \in R_i^*} \mathbf{A}_i^{-1} \cdot \mathbf{B}_i^j \right\| \leq 1 \quad \forall R_i^* \subseteq P_i^*, \forall i \in \omega$$

can be addressed as a necessary condition for C-monotonicity of scheme (2.1).

Proof. Let there exist the subset $R_i^* \subseteq P_i^*$ such that

$$(3.25) \quad \left\| \sum_{j \in R_i^*} \mathbf{A}_i^j \right\| > 1.$$

Assuming that (2.1) is C-monotone, we have that $\|\mathbf{A}\|_C$ fulfills condition (3.4), and in view of (3.2), we can write

$$(3.26) \quad \max_{\|\mathbf{x}\|_C=1} \max_{i \in \omega} \left\| \sum_{j=n+1}^I \mathbf{A}_i^j \cdot \mathbf{x}_{j-n} \right\| = \max_{\|\mathbf{x}\|_C=1} \max_{i \in \omega} \left\| \sum_{j \in P_i^*} \mathbf{A}_i^j \cdot \mathbf{x}_{j-n} \right\| \leq 1.$$

Considering that all nodes in the neighborhood P_i^* are boundary nodes, we see that $\mathbf{x}_j$ (for all $j \in P_i^*$) could be chosen as an arbitrary vector; therefore $\mathbf{x}_j = \mathbf{x}$ (for all $j \in P_i^*$) with $\|\mathbf{x}\| = 1$ for all $j \in R_i^*$ and $\|\mathbf{x}_j\| = 0$ for all $j \in P_i^* \setminus R_i^*$ will lead to

$$(3.27) \quad \left\| \sum_{j \in P_i^*} \mathbf{A}_i^j \cdot \mathbf{x}_{j-n} \right\| = \left\| \left(\sum_{j \in R_i^*} \mathbf{A}_i^j \right) \cdot \mathbf{x} \right\| \leq \left\| \sum_{j \in R_i^*} \mathbf{A}_i^j \right\| \quad \forall i \in \omega.$$

As the norm is a continuous function of $\mathbf{x}$ on the closed and bounded set $\|\mathbf{x}\| = 1$, there exists an $\mathbf{x}_0$ ($\|\mathbf{x}_0\| = 1$) for which the maximum of the LHS of (3.27) is attained, i.e.,

$$(3.28) \quad \max_{\|\mathbf{x}\|=1} \left\| \left(\sum_{j \in R_i^*} \mathbf{A}_i^j \right) \cdot \mathbf{x} \right\| = \left\| \left(\sum_{j \in R_i^*} \mathbf{A}_i^j \right) \cdot \mathbf{x}_0 \right\| = \left\| \sum_{j \in R_i^*} \mathbf{A}_i^j \right\| \quad \forall i \in \omega.$$

In view of (3.28) we conclude that

$$(3.29) \quad \left\| \sum_{j \in R_i^*} \mathbf{A}_i^j \right\| = \max_{\|\mathbf{x}\|=1} \left\| \left(\sum_{j \in R_i^*} \mathbf{A}_i^j \right) \cdot \mathbf{x} \right\| \leq \max_{\|\mathbf{x}\|_C=1} \left\| \sum_{j \in P_i^*} \mathbf{A}_i^j \cdot \mathbf{x}_{j-n} \right\| \quad \forall i \in \omega.$$

In view of (3.25), (3.29), and (3.26) we obtain the contradiction

$$(3.30) \quad 1 < \left\| \sum_{j \in R_i^*} \mathbf{A}_i^j \right\| \leq \max_{\|\mathbf{X}\|_C=1} \max_{i \in \omega} \left\| \sum_{j \in P_i^*} \mathbf{A}_i^j \cdot \mathbf{x}_{j-n} \right\| \leq 1,$$

which manifests the proof of Theorem 3.11. $\square$

THEOREM 3.12. *Let (2.1) be an explicit homogeneous scheme. Then condition (3.4) of C-monotonicity is necessary and sufficient for Samarskiy monotonicity.*

Proof. Let scheme (2.1) be C-monotone; then in view of (3.23) we obtain $\|\mathbf{Z}\|_C \leq \|\mathbf{A}'\|_C \|\mathbf{X}\|_C \leq \|\mathbf{X}\|_C$ manifesting that a C-monotone scheme will be Samarskiy-monotone. In proving the necessity, let there exist a Samarskiy-monotone scheme which does not comply with C-monotonicity, i.e.,

$$(3.31) \quad \|\mathbf{Z}\|_C \leq \|\mathbf{X}\|_C, \quad \|\mathbf{A}'\|_C > 1.$$

As scheme (3.23) is explicit, $\mathbf{X}$ could be chosen as an arbitrary vector. Therefore, $\mathbf{X} = \mathbf{X}_0$ such that $\|\mathbf{X}_0\|_C = 1$, leading to

$$(3.32) \quad \max_{\|\mathbf{X}\|_C=1} \|\mathbf{A}' \cdot \mathbf{X}\|_C = \|\mathbf{A}' \cdot \mathbf{X}_0\|_C = \|\mathbf{A}'\|_C.$$

We note that (3.32) is valid, since the norm is a continuous function of $\mathbf{X}$ on the closed and bounded set $\|\mathbf{X}\|_C = 1$. In view of (3.31) and (3.32), we thus obtain the contradiction

$$(3.33) \quad 1 = \|\mathbf{X}_0\|_C \geq \|\mathbf{Z}\|_C = \|\mathbf{A}' \cdot \mathbf{X}_0\|_C = \|\mathbf{A}'\|_C > 1,$$

which proves the necessity. $\square$

COROLLARY 3.13. *In view of Theorem 3.12 the necessary condition (3.24) for C-monotonicity is also the necessary condition for Samarskiy monotonicity.*

COROLLARY 3.14. *In the case of the scalar form of (2.1), the matrix norm in (3.4) is equivalent to the matrix norm in (3.8), and in view of Theorem 3.12, (3.8) becomes a necessary and sufficient condition for the Samarskiy monotonicity of an explicit scheme.*

COROLLARY 3.15. *If (2.1) is an explicit scheme and $\mathbf{A}_i^j = \varphi_i^j(\mathbf{D})$ in (3.9), then in view of Corollary 3.4 and Remark 3.5 we obtain that*

$$(3.34) \quad B_i \equiv \max_{\lambda \in s(\mathbf{D})} \sum_{j \in P_i^*} |\varphi_i^j(\lambda)| \leq 1 \quad \forall i \in \omega$$

will be necessary and sufficient conditions for the Samarskiy monotonicity of (2.1).

PROPOSITION 3.16. *If (2.1) is an explicit scheme and $\mathbf{A}_i^j = \varphi_i^j(\mathbf{D}^j)$ in (3.9), then*

$$(3.35) \quad B_j \equiv \max_{\lambda \in s(\mathbf{D}^j)} \sum_{i \in \omega} |\varphi_i^j(\lambda)| \leq 1 \quad \forall j \in \omega$$

will be necessary and sufficient conditions for the submonotonicity of (2.1). We consider $\mathbf{D}^j$ in a diagonal form without loss of generality (see the proof of Theorem 3.3). The sufficiency is evident in view of Corollary 3.8. To prove necessity, let us assume that (3.17) is valid; namely, $\|\mathbf{A}\|_1 \leq 1$, but $\exists j : B_j > 1$. Then we obtain the contradiction $1 < B_j \leq \|\mathbf{A}\|_1 \leq 1$, which manifests the proof of Proposition 3.16.

4. Examples. The developed monotonicity criteria are demonstrated throughout various examples of physical models expressed in finite-difference schemes. In doing so one can judge different grid connectedness sequences and the interrelationship between the monotonicity and the stability of the scheme (scalar or vector) in question.

Example 4.1. Consider (1.1) with the grids $\Omega_\tau = \{t_k, k = 0, 1, \dots\}$, $\Omega_x = \{x_i, i = 0, 1, \dots, I\}$, $\Omega = \Omega_\tau \times \Omega_x$. Let h ($= \text{const}$) denote the spatial intervals and τ ($= \text{const}$) the time increments, and let $x_i \equiv ih$, $t_k \equiv k\tau$; $u_i \equiv u(x_i, t_{k+1})$, $\check{u}_i \equiv u(x_i, t_k)$; $u_x \equiv u_{x,i} \equiv (u_{i+1} - u_i)/h$, $u_{\bar{x}} \equiv u_{\bar{x},i} \equiv (u_i - u_{i-1})/h$, $u_{\bar{t},i} \equiv (u_i - \check{u}_i)/\tau$. To solve (1.1), the two-parametric scheme [31] is proposed in the form

$$(4.1) \quad \alpha u_{\bar{t},i} + (1 - \alpha)u_{\bar{t},i-1} + \beta V u_x + (1 - \beta)V \check{u}_{\bar{x}} = 0,$$

which in view of the above notation is rewritten in the form of (2.3) to read

$$(4.2) \quad \left(\frac{1}{\tau} + \frac{V\sigma}{h} \right) u_i - \left[\frac{V\sigma}{h} u_{i-1} + \left(\frac{1}{\tau} - \frac{V(1-\sigma)}{h} \right) \check{u}_i + \frac{V(1-\sigma)}{h} \check{u}_{i-1} \right] = 0,$$

where $\sigma = \beta - h(1 - \alpha)/(V\tau)$. We note in view of (4.2) that there is a path from any interior i node at t_{k+1} only to other $j < i$ nodes at t_{k+1} as well as to $j \leq i$ at a previous time level. We thus conclude that (4.2) is associated with a sequence of P_i^* neighborhoods of nodes describing the HC grid. Scheme (4.2) will be Samarskiy-monotone, in view of (2.4), when subject to the conditions

$$(4.3) \quad \tau < \frac{\alpha}{V(1-\beta)} h, \quad 0 < \sigma < 1, \quad V > 0.$$

Upon substituting $\alpha = \beta = 0.5$ into (4.2) we obtain Wendroff's scheme [31]. Wendroff's scheme is unconditionally stable, yet we note that (4.3) is not fulfilled, and thus it is not Samarskiy-monotone in terms of (2.4). Condition (2.13) for Wendroff's scheme in view of (4.2) reads

$$(4.4) \quad |\theta - 1| + |1 - \theta| + |1 + \theta| \leq |1 + \theta|, \quad \theta \equiv \frac{V\tau}{h},$$

from which we note that only for the specific condition of $\theta = 1$ is Wendroff's scheme Samarskiy-monotone in contrast to the conditions for the scheme to be Godunov-monotone, namely, in view of Godunov's theorem [12, p. 277], Wendroff's scheme is monotone in the Godunov sense if $\theta = 0, 1, 2, \dots$.

Verifying the scheme on the basis of the various developed criteria is the procedure suggested hereinafter.

Example 4.2. The explicit scheme in [20, p. 186] is claimed to be nonmonotone (not Samarskiy-monotone) as some of its coefficients are negative and thus do not fulfill the requirements of the *maximum principle* (or in our case, (2.4)). Yet, based on practice [20], the developed scheme demonstrates monotonicity (free of spurious oscillations). However, we note that the obtained conditions ((22)–(24) in [20]) lead to the fulfillment of (2.13), which in turn proves the Samarskiy monotonicity of the scheme. Notice that the scheme in [20, p. 186] is not Harten-monotone.

Example 4.3. Let us consider the 1-D linear Schrödinger equation [32],

$$(4.5) \quad i \frac{\partial z}{\partial t} = \frac{\partial^2 z}{\partial x^2}, \quad 0 < x < 1, \quad t > 0; \quad z(x, 0) = z_0(x), \quad z(0, t) = z(1, t) = 0,$$

where $\mathbf{i}$ denotes the imaginary unit. Using the implicit scheme of [32] we write

$$(4.6) \quad \mathbf{i} \frac{u_i - \check{u}_i}{\tau} = \sigma u_{\bar{x}x} + (1 - \sigma) \check{u}_{\bar{x}x}, \quad \sigma = \alpha + \beta \mathbf{i}.$$

In [32] it is shown that (4.6) will be stable if

$$(4.7) \quad \tau \leq \begin{cases} +\infty, & \alpha \geq 0.5, \\ \frac{C_0 h^4}{(1-2\alpha)16}, & 0 \leq \alpha < 0.5, \quad C_0 = \text{const.} \end{cases}$$

Let us rewrite (4.6) in the canonical form (2.1), namely

$$\left(\frac{\mathbf{i}}{\tau} + \frac{2\sigma}{h^2} \right) u_i = \left[\frac{\mathbf{i}}{\tau} - \frac{2(1-\sigma)}{h^2} \right] \check{u}_i + \frac{\sigma(u_{i-1} + u_{i+1}) + (1-\sigma)(\check{u}_{i-1} + \check{u}_{i+1})}{h^2}.$$

Condition (2.13) in this case is written in the form

$$(4.8) \quad \theta(|\sigma| + |1 - \sigma|) + |\mathbf{i} - \theta(1 - \sigma)| \leq |\mathbf{i} + \theta\sigma|, \quad \frac{\theta = 2\tau}{h^2}.$$

We note that (4.8) cannot be fulfilled for any values of σ and θ . Hence, scheme (4.6) is not Samarskiy-monotone in terms of conditions (2.13).

Example 4.4. Consider again Example 4.3. On the basis of submonotone criterion (3.19) when applied to scheme (4.6), we obtain the condition

$$(4.9) \quad \theta^2(|\sigma|^2 + |1 - \sigma|^2) + |\mathbf{i} - \theta(1 - \sigma)|^2 \leq |\mathbf{i} + \theta\sigma|^2,$$

from which we note that the implicit scheme under $\sigma = 1$ is unconditionally submonotone while the explicit ($\sigma = 0$) one is not submonotone.

Example 4.5. Consider a 2-D diffusion problem for a vector-valued function $\mathbf{y} = \mathbf{y}(x_1, x_2)$ in the form

$$(4.10) \quad \frac{\partial}{\partial x_1} \left(\mathbf{D}_{11} \cdot \frac{\partial \mathbf{y}}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left(\mathbf{D}_{22} \cdot \frac{\partial \mathbf{y}}{\partial x_2} \right) = 0, \quad 0 \leq x_j \leq 1, \quad j = 1, 2,$$

where $\mathbf{D}_{11}$ and $\mathbf{D}_{22}$ are positive definite matrices with constant elements. Let us assign the grid $\Omega = \Omega_1 \times \Omega_2$, where $\Omega_j = \{x_{i_j}, i_j = 0_j, 1_j, \dots, I_j\}$, $x_{i_j} \equiv i_j h_j$, $h_j \equiv 1/I_j$, $j = 1, 2$. A possible difference scheme (see notation in Example 4.1) can be

$$(4.11) \quad (\mathbf{D}_{11} \cdot \mathbf{u}_{\bar{x}_1})_{x_1} + (\mathbf{D}_{22} \cdot \mathbf{u}_{\bar{x}_2})_{x_2} = 0.$$

We rewrite scheme (4.11) in the canonical form (2.1) to read

$$(4.12) \quad \mathbf{T} \cdot \mathbf{u}_{i_1, i_2} = \mathbf{T}_1 \cdot (\mathbf{u}_{i_1+1, i_2} + \mathbf{u}_{i_1-1, i_2}) + \mathbf{T}_2 \cdot (\mathbf{u}_{i_1, i_2+1} + \mathbf{u}_{i_1, i_2-1}),$$

where $\mathbf{T}_1 = \mathbf{D}_{11}/(h_1)^2$, $\mathbf{T}_2 = \mathbf{D}_{22}/(h_2)^2$, $\mathbf{T} = 2(\mathbf{T}_1 + \mathbf{T}_2)$. Condition (2.13) for (4.12) becomes

$$\left\| \left[\mathbf{D}_{11} + \left(\frac{h_1}{h_2} \right)^2 \mathbf{D}_{22} \right]^{-1} \cdot \mathbf{D}_{11} \right\| + \left\| \left[\left(\frac{h_2}{h_1} \right)^2 \mathbf{D}_{11} + \mathbf{D}_{22} \right]^{-1} \cdot \mathbf{D}_{22} \right\| \leq 1,$$

which is unconditionally fulfilled for the case of a scalar equation; however, it is fulfilled for specific cases of a vector equation, e.g., when $\mathbf{D}_{11} = \alpha \mathbf{D}_{22}$ for all $\alpha > 0$ (cf.

[17, p. 571]), for which scheme (4.11) will be Samarskiy-monotone in terms of criterion (2.13).

Example 4.6. Consider again Example 4.5. In addition to the conditions appearing in Example 4.5, it is assumed that the matrices $\mathbf{D}_{11}$ and $\mathbf{D}_{22}$ are symmetric and permutable.

When applying (3.10) to scheme (4.12), we obtain

$$(4.13) \quad \max_{\nu_i \leq \lambda_i \leq \mu_i, i=1,2} \left[(h_2)^2 \left| \frac{\lambda_1}{(h_2)^2 \lambda_1 + (h_1)^2 \lambda_2} \right| + (h_1)^2 \left| \frac{\lambda_2}{(h_2)^2 \lambda_1 + (h_1)^2 \lambda_2} \right| \right] \leq 1,$$

where ν_i and μ_i denote the left and right boundaries, respectively, of the spectrum $s(\mathbf{D}_{ii})$. For $\nu_i > 0$ we note that (4.13) is valid. Hence, for positive definite ($\nu_i > 0$) symmetric and permutable matrices, the scheme (4.12) unconditionally fulfills (3.10); i.e., this scheme will be regarded as unconditionally Samarskiy-monotone.

5. Remarks and discussion. Let us prove that an assertion similar to that of Godunov's [12, p. 277] is valid for Samarskiy-monotone schemes with, in general, variable coefficients. In view of Corollary 3.14, scheme (1.10) will be Samarskiy-monotone iff

$$(5.1) \quad \sum_j |a_i^j| \leq 1.$$

In view of (1.12) we obtain

$$(5.2) \quad 1 = \left| \sum_j a_i^j \right| \leq \sum_j |a_i^j| \quad \forall i.$$

By virtue of (5.1) and (5.2) we conclude that

$$(5.3) \quad \sum_j |a_i^j| = 1 \quad \forall i.$$

Thus, (1.12) and (5.3) must be valid simultaneously. It is possible iff all coefficients comply with $a_i^j \geq 0$. Hence, any Samarskiy-monotone explicit linear scheme approximating (1.1) has nonnegative coefficients. Godunov [12] has proven that an explicit linear scheme, be it monotonicity preserving or not, cannot be of second order accuracy if its coefficients are nonnegative. Forasmuch as any implicit linear scheme having a unique solution can be conformed into an explicit one [12], we have proven that any Samarskiy-monotone two-level linear scheme for (1.1) can be at most of first order accuracy excluding the case where each grid node (x_i, t_{k+1}) (see the notation of Example 4.1) belongs to the same characteristic curve as a node (x_j, t_k) , where j depends on i . Obviously, the above assertion (see also [12, p. 277]) is valid for Harten-monotone linear schemes with, in general, variable coefficients.

Let us consider a 1-D diffusion problem associated with a vector-valued function $\mathbf{y} = \mathbf{y}(x, t) \in L$

$$(5.4) \quad \frac{\partial \mathbf{y}}{\partial t} = \frac{\partial}{\partial x} \left(\mathbf{D} \cdot \frac{\partial \mathbf{y}}{\partial x} \right), \quad -\infty < x < \infty, \quad t > 0; \quad \mathbf{y}(x, 0) = \mathbf{y}_0(x),$$

where $\mathbf{D}$ is a symmetric and positive definite ($0 < d(\mathbf{y}, \mathbf{y}) \leq (\mathbf{D} \cdot \mathbf{y}, \mathbf{y})$ for all $\mathbf{y} \in L$, $d = \text{const}$) matrix with constant elements, L denotes a linear vector space, and $\mathbf{y}_0(x)$

denotes a prescribed vector-valued function. Using the notation of Example 4.1, we write

$$(5.5) \quad \frac{\mathbf{u}_i - \check{\mathbf{u}}_i}{\tau} = \sigma(\mathbf{D} \cdot \mathbf{u}_{\bar{x}})_x + (1 - \sigma)(\mathbf{D} \cdot \check{\mathbf{u}}_{\bar{x}})_x.$$

We rewrite (5.5) in the canonical form of (2.15) to read

$$(5.6) \quad \mathbf{u}_i = \mathbf{T} \cdot \left[\frac{\sigma \mathbf{D}}{h^2} \cdot (\mathbf{u}_{i-1} + \mathbf{u}_{i+1}) + \check{\mathbf{T}} \cdot \check{\mathbf{u}}_i + \frac{(1 - \sigma)}{h^2} \mathbf{D} \cdot (\check{\mathbf{u}}_{i-1} + \check{\mathbf{u}}_{i+1}) \right],$$

where

$$\mathbf{T} = \left(\frac{1}{\tau} \mathbf{I} + \frac{2\sigma}{h^2} \mathbf{D} \right)^{-1}, \quad \check{\mathbf{T}} = \frac{1}{\tau} \mathbf{I} - \frac{2(1 - \sigma)}{h^2} \mathbf{D}.$$

Using (3.8) for (5.6) yields that the explicit ($\sigma = 0$) scheme (5.6) will be Samarskiy-monotone iff

$$(5.7) \quad \frac{2\tau}{h^2} |D_i^i| + \left| 1 - \frac{2\tau}{h^2} D_i^i \right| + \frac{4\tau}{h^2} \sum_{j \neq i} |D_j^i| \leq 1, \quad i = 1, \dots, N.$$

In view of (5.7), the explicit ($\sigma = 0$) scheme (5.6) can be Samarskiy-monotone only if $\mathbf{D}$ becomes a diagonal matrix, which is the case for an uncoupled (5.4) system. As (3.8) addresses scalar schemes (see Remark 3.2), we therefore conclude that the scalar form of a vector scheme cannot always be suitable for investigation of monotonicity. However, in the case of a vector scheme, using (3.10) as well as (3.24), we note that the explicit scheme (5.6) will be Samarskiy-monotone iff

$$(5.8) \quad \tau \leq \frac{h^2}{2\|\mathbf{D}\|_2}.$$

Condition (3.10) applied to the implicit ($0 < \sigma \leq 1$) scheme (5.6) (or (5.5)) reads

$$(5.9) \quad \max_{d \leq \lambda \leq \|\mathbf{D}\|_2} \left| \frac{1}{\tau} + \frac{2\sigma}{h^2} \right|^{-1} \left(2 \left| \frac{\sigma}{h^2} \lambda \right| + \left| \frac{1}{\tau} - 2 \frac{1 - \sigma}{h^2} \lambda \right| + 2 \left| \frac{1 - \sigma}{h^2} \lambda \right| \right) \leq 1,$$

yielding the sufficient conditions of Samarskiy monotonicity in the form

$$(5.10) \quad \tau \leq \frac{h^2}{2\|\mathbf{D}\|_2(1 - \sigma)}, \quad 0 < \sigma \leq 1.$$

Let us also note that conforming the implicit scheme (5.6) into an explicit form, condition (3.10) can express necessary and sufficient conditions. Conformation of (5.5) into an explicit form, for a scalar equation, is cited in [16]. When (5.5) is a vector scheme, the explicit scheme form reads

$$(5.11) \quad \mathbf{u}_i = \mathbf{A}_0 \cdot \check{\mathbf{u}}_i + \sum_{j=1}^{\infty} \mathbf{A}_j \cdot (\check{\mathbf{u}}_{i-j} + \check{\mathbf{u}}_{i+j}),$$

where

$$(5.12) \quad \mathbf{A}_0 = \mathbf{I} - 4\tau \mathbf{G}^{-1} \cdot (h\mathbf{I} + \mathbf{G})^{-1} \cdot \mathbf{D}, \quad \mathbf{A}_1 = 4\tau h \mathbf{G}^{-1} \cdot (h\mathbf{I} + \mathbf{G})^{-2} \cdot \mathbf{D},$$

$$(5.13) \quad \mathbf{A}_j = \mathbf{q} \cdot \mathbf{A}_{j-1}, \quad j \geq 2,$$

$$(5.14) \quad \mathbf{q} = 4\sigma\tau(h\mathbf{I} + \mathbf{G})^{-2} \cdot \mathbf{D}, \quad \mathbf{G} \equiv (h^2\mathbf{I} + 4\sigma\tau\mathbf{D})^{0.5}.$$

Since all operators in (5.11) depend on the same symmetric matrix $\mathbf{D}$, in view of Corollary 3.15, scheme (5.11) as well as (5.6) will be Samarskiy-monotone iff

$$(5.15) \quad \tau \leq \frac{(2 - \sigma)h^2}{4\|\mathbf{D}\|_2(1 - \sigma)^2}.$$

When comparing (5.10) and (5.15), we note that (3.10) as well as (3.8) provides, in general, only sufficient conditions for Samarskiy monotonicity of a difference scheme. Forasmuch as the condition that is necessary and sufficient for the monotonicity of a scheme must not depend on the scheme formulation (explicit or implicit form), we conclude that the diagonal dominance of $\tilde{\mathbf{A}}$ in (1.19) (see Stoyan's theorem [35, p. 159]) is not, in general, a necessary condition for the Samarskiy monotonicity of a scheme approximating some boundary-value problem for a linear PDE.

Let us refer to the relation between monotonicity and stability concepts. In the case of (5.5), the necessary and sufficient condition for the so-called L_2 -stability of the scheme can be written [5, p. 1132] in the form

$$(5.16) \quad \tau \leq \begin{cases} +\infty & \text{if } 0.5 \leq \sigma \leq 1, \\ \frac{h^2}{2(1-2\sigma)\|\mathbf{D}\|_2} & \text{if } 0 \leq \sigma < 0.5. \end{cases}$$

In the case of implicit schemes, Anderson, Tannehill, and Pletcher [4] examine various finite-difference algorithms, including the scalar version of (5.5), applying the unconditionally stable (5.16) Crank–Nicholson scheme ($\sigma = 0.5$), which is recommended because of its second order temporal and spatial accuracy. However, by virtue of (5.15), scheme (5.5) will be Samarskiy-monotone (if $\sigma = 0.5$) when subject to the condition

$$(5.17) \quad \tau \leq \frac{3h^2}{2\|\mathbf{D}\|_2},$$

which is a forcing limitation on the time step, unlike condition (5.16). If (5.17) is violated, then the scheme can produce spurious oscillations. Hence, the recommendation of [4] is questionable.

Notice that the necessary and sufficient condition for the scalar version of the Crank–Nicholson scheme to be monotonicity preserving, originally established in Godunov [12], coincides with (5.17) applied to that scalar scheme. We note that actually (5.17) is the necessary and sufficient condition for monotonicity in Samarskiy's sense. The condition similar to (5.15) for a scalar version of scheme (5.5) to be monotonicity preserving can be found in [16]. It is worth noting that in [15], among other results, the necessary and sufficient condition was obtained for the two-parametric scheme approximating a 1-D heat equation to be monotonicity preserving.

Using Proposition 1.4 we obtain that the scalar version of scheme (5.11) can be GO-monotone only if

$$(5.18) \quad 0 \leq \gamma \leq \Gamma_\sigma, \quad \gamma = \frac{D\tau}{h^2}, \quad \Gamma_\sigma = \begin{cases} \frac{1}{3}, & \sigma = 0, \\ \frac{\sigma + 3 - \sqrt{(9 - \sigma)(1 - \sigma)}}{8\sigma(1 - \sigma)}, & 0 < \sigma \leq 1. \end{cases}$$

By virtue of (5.18) we obtain that the Crank–Nicholson scheme ($\sigma = 0.5$ in (5.11)) can be GO-monotone only if $0 \leq \gamma \leq \Gamma_{0.5}$, $\Gamma_{0.5} \approx 0.72$. Hence, in view of (5.17), we obtain that if $\Gamma_{0.5} < \gamma \leq 1.5$, then the Crank–Nicholson scheme will be Samarskiy-monotone but not GO-monotone. In [22, pp. 34–36] the model problem is solved by

the Crank–Nicholson scheme for which, when $\gamma = 1$, the scheme exhibits spurious oscillations. It can be easily verified that the scheme does not produce spurious oscillations if $0 \leq \gamma \leq \Gamma_{0.5}$. Thus, a Samarskiy-monotone scheme can nevertheless produce spurious oscillations if the scheme is not GO-monotone.

Let us consider the feasibility of monotonicity (Definition 1.6), submonotonicity (Definition 3.7), GO monotonicity (Definition 1.3), and the TVD [14, p. 360] notions, using the scalar conservation law ((1.1) in [27]),

$$(5.19) \quad \frac{\partial V}{\partial t} + \frac{\partial}{\partial x} F(V) = 0, \quad -\infty < x < \infty, \quad t > 0, \quad V(x, 0) = V_0(x),$$

and the explicit scheme approximating it ((1.2)–(1.11) in [27]),

$$(5.20) \quad U_i = \check{U}_i + \beta_{i+0.5}(\check{U}_{i+1} - \check{U}_i) - \alpha_{i-0.5}(\check{U}_i - \check{U}_{i-1}),$$

where α and β depend on $\check{U}$. In view of Theorem 6.1 (see the appendix in section 6) the Samarskiy monotonicity of the nonlinear scheme (5.20) can be tested on the basis of its variational scheme. Scheme (5.20) will be TVD (see, e.g., [14], [27]) if

$$(5.21) \quad \alpha_{i+0.5} \geq 0, \quad \beta_{i+0.5} \geq 0, \quad \alpha_{i+0.5} + \beta_{i+0.5} \leq 1 \quad \forall i.$$

Let us demonstrate that the TVD scheme (5.20) with (5.21) may produce spurious oscillations. Consider that $F(V) = V^2/2$ (inviscid Burgers equation) and assume that $\alpha_{i-0.5} = 0.5(1 + \gamma b_{i-0.5})$, $\beta_{i+0.5} = 0.5(1 - \gamma b_{i+0.5})$, where $b_{i-0.5} = (\check{U}_i + \check{U}_{i-1})/2$, $\gamma = \tau/h$; then we obtain the Lax scheme reading

$$(5.22) \quad U_i - \frac{\check{U}_{i+1} + \check{U}_{i-1}}{2} + \gamma \frac{\check{U}_{i+1}^2 - \check{U}_{i-1}^2}{4} = 0,$$

and the variational form of (5.22) becomes

$$(5.23) \quad \delta U_i = \frac{1 + \gamma \check{U}_{i-1}}{2} \delta \check{U}_{i-1} + \frac{1 - \gamma \check{U}_{i+1}}{2} \delta \check{U}_{i+1}.$$

In view of (5.21) scheme (5.22) will be TVD if

$$(5.24) \quad \gamma \equiv \frac{\tau}{h} \leq \frac{2}{\max_i |\check{U}_{i+1} + \check{U}_i|}.$$

The necessary conditions for linear monotonicity of (5.22), i.e., Samarskiy monotonicity of (5.23), in view of Corollary 3.13, can be written in the form

$$(5.25) \quad \gamma \equiv \frac{\tau}{h} \leq \frac{1}{\max_i |\check{U}_i|}.$$

The necessary and sufficient conditions for linear monotonicity of (5.22) in view of Corollary 3.14 can be written in the form

$$(5.26) \quad \gamma \equiv \frac{\tau}{h} \leq \frac{1}{\max_i |\check{U}_i|}, \quad \check{U}_{i-1} \leq \check{U}_{i+1} \quad \forall i.$$

When comparing (5.24), (5.25), and (5.26), we note that the Lax scheme, being TVD, need not be linearly monotone.

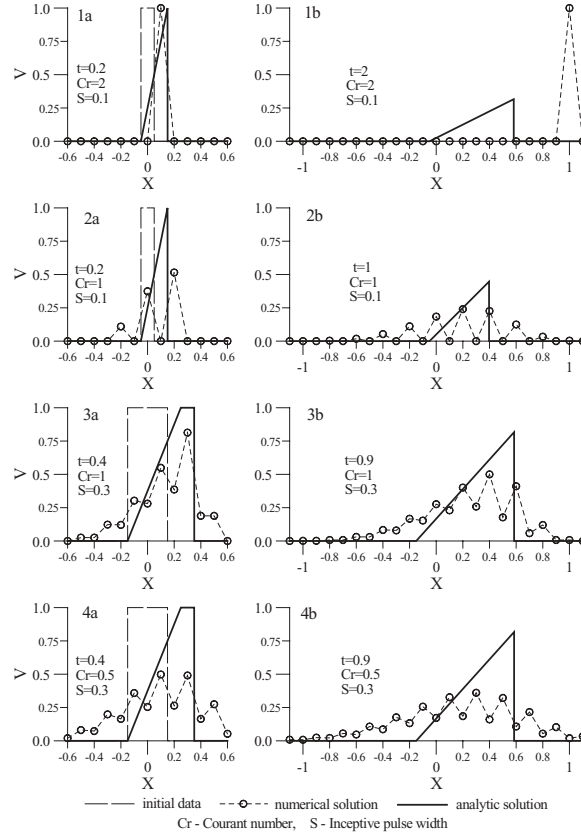


FIG. 5.1. Inviscid Burgers equation. Lax scheme versus the analytical solution.

Figure 5.1 shows the results calculated by scheme (5.22) for the Burgers equation with the initial conditions

$$(5.27) \quad V(x, 0) = \{0, x < -S/2; V_0, -S/2 \leq x \leq S/2; 0, x > S/2\},$$

where V_0 and S denote the inceptive pulse amplitude and width, respectively.

The solid curves in Figure 5.1 depict the exact solutions for different values of t : If $0 < t \leq T$, $T = 2S/V_0$, then

$$(5.28) \quad V(x, t) = \begin{cases} \frac{2x+S}{2b+S} & \text{if } -0.5S \leq x \leq b, b = V_0t - 0.5S, \\ V_0 & \text{if } b < x \leq 0.5(V_0t + S), \\ 0 & \text{if } x < -0.5S \text{ or } x > 0.5(V_0t + S). \end{cases}$$

If $t > T$, $T = 2S/V_0$, then

$$(5.29) \quad V(x, t) = \begin{cases} \frac{2V_0S(x+0.5S)}{(L+0.5S)^2} & \text{if } -0.5S \leq x \leq L, \\ 0 & \text{if } x < -0.5S \text{ or } x > L, \end{cases}$$

where $L = 2\sqrt{S^2 + 0.5V_0S(t - T)} - 0.5S$.

The numerical solutions were computed on a uniform grid with spatial increments of $h = 0.1$, i.e., $\Omega_x = \{0, \pm 0.1, \pm 0.2, \dots\}$, the velocity $V_0 = 1$, and the Courant number of $C_r \equiv V_0\tau/h = 2, 1, 0.5$.

Let us note that in the case when $S < 2h$ we obtain that (5.22), together with (5.27) and subject to (5.24), will be TVD if $0 < C_r \leq 2$. However, in view of (5.25) this scheme can be linearly monotone only if $0 < C_r \leq 1$. Hence we note that the TVD solution in Figures 5.1(1a) and 5.1(1b) is not asymptotically stable, and the numerical error is growing with time. However, when $C_r = 1$ the TVD scheme cannot be linearly monotone, as the second inequality in (5.26) is not valid for the grid function $\check{U}_i$. We thus note in Figures 5.1(2a) and 5.1(2b) that the TVD scheme produced spurious oscillations. In the case when $S > 2h$ we obtain by virtue of (5.24) that this scheme will be TVD if $0 < C_r \leq 1$. However, in view of (5.26) scheme (5.22) together with (5.27) cannot be linearly monotone, as the second inequality in (5.26) is not valid for the grid function $\check{U}_i$. Consequently, Figures 5.1(3a)–5.1(4b) demonstrate that the TVD scheme produced spurious oscillations for the Courant number of $C_r = 1$ (Figures 5.1(3a) and 5.1(3b)) and, for a refined time interval, $C_r = 0.5$ (Figures 5.1(4a) and 5.1(4b)). Thus, in contradiction to the statement appearing in [7] we note that a new extremum can be generated in spite of the fact that a scheme is both Harten-monotone and TVD. Hence we conclude that the Lax scheme (5.22), being stable and TVD, may still produce spurious oscillations; however, we note in [27, p. 1120] that this scheme becomes linearly monotone and does not produce spurious oscillations under a condition similar to (5.26). However, when we apply (5.22) to the case of shock waves (see, e.g., [4], [27]) we note that the Lax scheme is monotonicity preserving (see Proposition 1.2) even if (5.26) is not valid; namely, when the grid function $\check{U}_i$ is monotone decreasing. Hence, the necessary and sufficient conditions for linear monotonicity can be too restrictive for a scheme to be free of spurious oscillations.

Considering Lax scheme (5.22), we will, in view of Proposition 1.2, use (5.24) as a sufficient condition to prove that the scheme is monotonicity preserving. We had implemented the Lax scheme for the numerical solution of the Burgers equation (Figure 5.1), yet Figures 5.1(2a)–5.1(4b) demonstrate that a monotonicity preserving nonlinear scheme can still exhibit spurious oscillations.

In [27] the notion of strong monotonicity for Godunov-monotone schemes is introduced and the associated criteria are developed to avoid spurious oscillations of difference derivatives. For such monotonicity, let us consider the upwind scheme [27], which is strongly monotone in terms of Ostapenko [27], in contrast to the Lax scheme, which is not strongly monotone [27]. If we assume in (5.20) that $\alpha_{i-0.5} = \gamma(|b_{i-0.5}| + b_{i-0.5})$, $\beta_{i+0.5} = \gamma(|b_{i+0.5}| - b_{i+0.5})$, where $b_{i-0.5} = (\check{U}_i + \check{U}_{i-1})/2$, $\gamma = \tau/h$, then we obtain the upwind scheme approximating the Burgers equation. This (for $\check{U}_i \geq 0$ for all i) can be written in the form

$$(5.30) \quad \frac{U_i - \check{U}_i}{\tau} + \frac{\check{U}_i^2 - \check{U}_{i-1}^2}{2h} = 0.$$

In view of (5.21) scheme (5.30) will be TVD if condition (5.24) is fulfilled. The variational scheme of (5.30) reads

$$(5.31) \quad \delta U_i = \gamma \check{U}_{i-1} \delta \check{U}_{i-1} + (1 - \gamma \check{U}_i) \delta \check{U}_i.$$

In view of Corollary 3.13 we obtain (5.25) as the necessary conditions for linear monotonicity of (5.30), i.e., Samarskiy monotonicity of (5.31). In view of Corollary 3.14 scheme (5.30) will be linearly monotone iff

$$(5.32) \quad \gamma \equiv \frac{\tau}{h} \leq \frac{1}{\max_i |\check{U}_i|}, \quad \check{U}_{i-1} \leq \check{U}_i \quad \forall i.$$

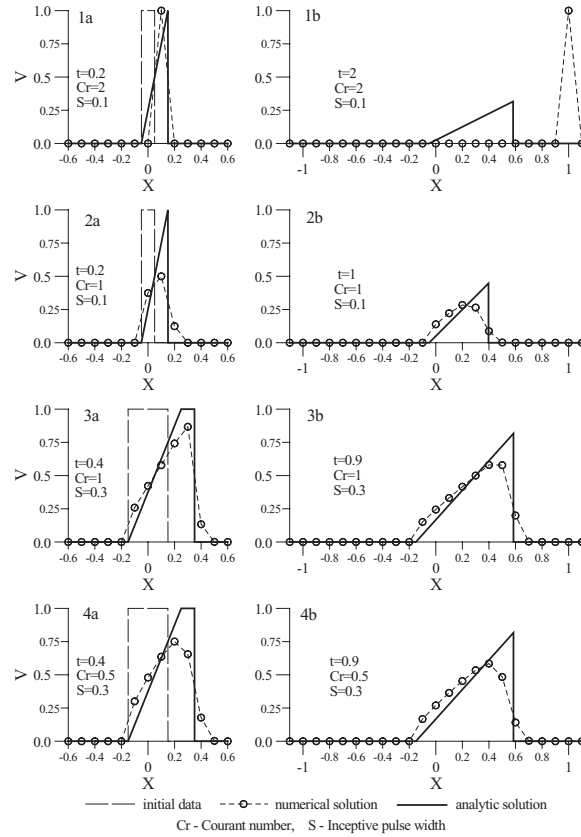


FIG. 5.2. *Inviscid Burgers equation. Upwind scheme versus the analytical solution.*

In Figure 5.2 we consider the upwind scheme (5.30) under the same conditions as were implemented for the Lax scheme demonstrated in Figure 5.1. Considering $S = 0.1$ and $C_r = 2$, Figures 5.2(1a) and 5.2(1b) demonstrate the same results depicted in Figures 5.1(1a) and 5.1(1b); namely, the scheme is strongly monotone in terms of Ostapenko [27] as well as TVD (as condition (5.24) is fulfilled) but is not asymptotically stable.

In Figures 5.2(2a)–5.2(4b) (for $0 < C_r \leq 1$) we note that, unlike in Figures 5.1(2a)–5.1(4b), spurious oscillations are not exhibited by the upwind scheme (5.30). In these figures we note that the second inequality of (5.32) is not valid; namely, the necessary and sufficient conditions for linear monotonicity may be too restrictive. However, the absence of spurious oscillations and asymptotic stability is maintained, as in view of Proposition 3.16 we obtain the necessary and sufficient conditions for linear submonotonicity of (5.30) in the form of (5.25), which is valid if $0 < C_r \leq 1$. As the TVD scheme was applied to all examples demonstrated in Figure 5.2, yet the notion of linear submonotonicity was demonstrated on a subset of these (Figures 5.2(2a)–5.2(4b)), we conclude that linear submonotonicity is more restrictive than TVD. Such a phenomenon is caused by the fact that TV is not a norm but a seminorm, and hence the TVD operator is L_1 -contractive, but in the quotient [28] space (see [5, p. 1133]).

In view of Proposition 3.16 and (5.25) (for $0 < C_r \leq 1$), the linear submonotonicity condition will be valid for the Lax scheme (5.22) which will, however, produce spurious oscillations (Figures 5.1(2a)–5.1(4b)).

In view of Proposition 1.4 and (5.23) we note that the Lax scheme (5.22) will not be linearly GO-monotone. The upwind scheme (5.30) subject to the conditions displayed in Figures 5.2(2a)–5.2(4b) and (5.31) can be linearly GO-monotone.

An implementation of the linear submonotonicity notion to a nonlinear vector scheme approximating a hyperbolic conservation law can be found in [5, p. 1132].

It is necessary to stress that all notions such as C-monotonicity, submonotonicity, and TVD can be viewed as concepts within Samarskiy monotonicity. Let us demonstrate this for a linear scheme, which can be written in the form of (1.10), conformed to (3.23),

$$(5.33) \quad \mathbf{u} = \mathbf{A} \cdot \check{\mathbf{u}}, \quad \mathbf{A} = \{a_i^j\}, \quad \mathbf{u}, \check{\mathbf{u}} \in X.$$

We note [5, p. 1126] that (5.33) can be considered as a two-node explicit scheme, which will be Samarskiy-monotone (see Definition 1.6) iff $\|\mathbf{u}\| \leq \|\check{\mathbf{u}}\|$, and hence [5, p. 1132] iff $\|\mathbf{A}\| \leq 1$. If the space X is equipped with the metric $\rho(\mathbf{u}, \mathbf{v}) \equiv \|\mathbf{u} - \mathbf{v}\|_1$ for all $\mathbf{u}, \mathbf{v} \in X$, i.e., $\|\mathbf{A}\| \equiv \|\mathbf{A}\|_1$, then, in view of Definition 3.7, scheme (1.10) will be submonotone. Analogously, for $\rho(\mathbf{u}, \mathbf{v}) \equiv \|\mathbf{u} - \mathbf{v}\|_\infty$, we note that (1.10) is C-monotone. It can be proven [5, p. 1133] that (1.10), being TVD, is submonotone. A scheme will be monotone, i.e., free of spurious oscillations and asymptotically stable, if it is both GO-monotone and Samarskiy-monotone. We realize that a scheme should (i) be associated with a C grid (see section 2), (ii) be shape preserving (Definition 1.5), and (iii) have an operator that is a contraction mapping [28]. As all of these are features embedded within the GO- and Samarskiy-monotone notions, we suggest referring to them as GOS schemes.

6. Appendix. In what follows we will establish the notion that the monotonicity of a variational scheme will guarantee that its original scheme also will be monotone. Let us consider, e.g., the explicit nonlinear scheme (1.6), which will be written in the vector form

$$(6.1) \quad \mathbf{u} = \mathbf{H}(\check{\mathbf{u}}), \quad \mathbf{u} \in X, \quad \check{\mathbf{u}} \in \Omega \subset X,$$

in which X denotes a linear vector space and Ω denotes a closed and bounded convex set.

THEOREM 6.1. *Let a nonlinear explicit scheme be written in the form (6.1). Then for any $\check{\mathbf{u}}, \check{\mathbf{u}} + \delta\check{\mathbf{u}} \in \Omega$ the scheme will be Samarskiy-monotone if its variational scheme is Samarskiy-monotone.*

Proof. Scheme (6.1) can be considered [5, p. 1126] as a two-node explicit scheme, and its variational scheme reads

$$(6.2) \quad \delta\mathbf{u} = \mathbf{H}'(\check{\mathbf{u}}) \cdot \delta\check{\mathbf{u}}, \quad \mathbf{H}'(\check{\mathbf{u}}) \equiv \frac{\partial \mathbf{H}(\check{\mathbf{u}})}{\partial \check{\mathbf{u}}}.$$

In view of [5, p. 1126] the variational scheme (6.2) is Samarskiy-monotone iff $\|\mathbf{H}'(\check{\mathbf{u}})\| \leq 1$ for all $\check{\mathbf{u}} \in \Omega$, and hence (6.2) will be Samarskiy-monotone iff

$$(6.3) \quad \sup_{\check{\mathbf{u}} \in \Omega} \|\mathbf{H}'(\check{\mathbf{u}})\| \leq 1.$$

Here $\|\cdot\|$ denotes the matrix norm induced by a prescribed vector norm defined on the vector space X . By virtue of the mean-value theorem [25, item 3.2.3] we obtain

from (6.1)

$$(6.4) \quad \begin{aligned} \|\delta \mathbf{u}\| &\equiv \|\mathbf{H}(\check{\mathbf{u}} + \delta \check{\mathbf{u}}) - \mathbf{H}(\check{\mathbf{u}})\| \\ &\leq \sup_{0 \leq t \leq 1} \|\mathbf{H}'(\check{\mathbf{u}} + t\delta \check{\mathbf{u}})\| \|\delta \check{\mathbf{u}}\| \leq \sup_{\check{\mathbf{u}} \in \Omega} \|\mathbf{H}'(\check{\mathbf{u}})\| \|\delta \check{\mathbf{u}}\|. \end{aligned}$$

In view of (6.3) we conclude from (6.4) that $\|\delta \mathbf{u}\| \leq \|\delta \check{\mathbf{u}}\|$; i.e., (6.1) is Samarskiy-monotone. $\square$

REFERENCES

- [1] L. ABRAHAMSSON AND S. OSHER, *Monotone difference schemes for singular perturbation problems*, SIAM J. Numer. Anal., 19 (1982), pp. 979–992.
- [2] V. V. AKIMENKO, *Monotone high-order accurate schemes for transport and parabolic equations*, Comput. Math. Math. Phys., 39 (1999), pp. 590–600.
- [3] V. V. AKIMENKO, *The maximum principle and nonlinear monotone schemes for parabolic equations*, Comput. Math. Math. Phys., 39 (1999), pp. 805–816.
- [4] D. A. ANDERSON, J. C. TANNEHILL, AND R. H. PLETCHER, *Computational Fluid Mechanics and Heat Transfer*, Hemisphere Publishing Corporation, New York, 1984.
- [5] V. S. BORISOV, *On discrete maximum principles for linear equation systems and monotonicity of difference schemes*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 1110–1135.
- [6] P. G. CIARLET, *Discrete maximum principle for finite-difference operators*, Aequationes Math., 4 (1970), pp. 338–352.
- [7] S. R. CHAKRAVARTHY AND S. OSHER, *High Resolution Applications of the Osher Upwind Scheme for the Euler Equations*, AIAA paper 83-1943, in 6th Computational Fluid Dynamics Conference (Danvers, MA) Collection of Technical Papers, 1983.
- [8] M. ELIASSI AND R. J. GLASS, *On the continuum-scale modeling of gravity-driven fingers in unsaturated porous media: The inadequacy of the Richards equation with standard monotonic constitutive relations and hysteretic equations of state*, Water Resources Research, 37 (2001), pp. 2019–2035.
- [9] B. ENGQUIST AND S. OSHER, *Stable and entropy satisfying approximations for transonic flow calculations*, Math. Comp., 34 (1980), pp. 45–75.
- [10] P. A. FORSYTH AND M. C. KROPINSKI, *Monotonicity considerations for saturated–unsaturated subsurface flow*, SIAM J. Sci. Comput., 18 (1997), pp. 1328–1354.
- [11] V. G. GANZHA AND E. V. VOROZHTSOV, *Numerical Solutions for Partial Differential Equations*, CRC Press, New York, 1996.
- [12] S. K. GODUNOV, *A finite difference method for the numerical computation and discontinuous solutions of fluid dynamics*, Mat. Sb., 47 (1959), pp. 271–306 (in Russian).
- [13] A. HARTEN, J. M. HYMAN, AND P. D. LAX, WITH APPENDIX BY B. KEYFIZ, *On finite-difference approximations and entropy conditions for shocks*, Comm. Pure Appl. Math., 29 (1976), pp. 297–322.
- [14] A. HARTEN, *High resolution schemes for hyperbolic conservation laws*, J. Comput. Phys., 49 (1983), pp. 357–393.
- [15] R. HORVATH, *On the monotonicity conservation in numerical solutions of the heat equation*, Appl. Numer. Math., 42 (2002), pp. 189–199.
- [16] N. N. KALITKIN, *Numerical Methods*, Nauka, Moscow, 1978 (in Russian).
- [17] O. A. LADYŽENSKAJA, V. A. SOLONIKOV, AND N. N. URALCEVA, *Linear and Quasi-Linear Equations of Parabolic Type*, AMS, Providence, RI, 1968.
- [18] P. LANCASTER, *Theory of Matrices*, Academic Press, New York, 1969.
- [19] P. D. LAX AND R. D. RICHTMYER, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math., 9 (1956), pp. 267–293.
- [20] L. LEBOUCHER, *Monotone scheme and boundary conditions for finite volume simulation of magnetohydrodynamic internal flows at high Hartman number*, J. Comput. Phys., 150 (1999), pp. 181–198.
- [21] L. MIRSKY, *An Introduction to Linear Algebra*, Dover, New York, 1990.
- [22] K. W. MORTON AND D. F. MAYERS, *Numerical Solution of Partial Differential Equations*, Cambridge University Press, Cambridge, UK, 1994.
- [23] K. W. MORTON, *Numerical Solution of Convection-Diffusion Problems*, Chapman & Hall, London, 1996.
- [24] E. S. ORAN AND J. P. BORIS, *Numerical Simulation of Reactive Flow*, Elsevier, New York, 1987.

- [25] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press, New York, London, 1970.
- [26] V. V. OSTAPENKO, *On monotonicity of difference schemes*, Siberian Math. J., 39 (1998), pp. 959–972.
- [27] V. V. OSTAPENKO, *On the strong monotonicity of nonlinear difference schemes*, Comput. Math. Math. Phys., 38 (1998), pp. 1119–1133.
- [28] M. REED AND B. SIMON, *Methods of Modern Mathematical Physics*, Academic Press, New York, 1972.
- [29] R. D. RICHTMYER AND K. W. MORTON, *Difference Methods for Initial-Value Problems*, 2nd ed., Wiley-Interscience, New York, 1967.
- [30] A. A. SAMARSKIY, *On the monotone difference schemes for elliptical and parabolic equations in the case of not self-conjugate elliptical operator*, Zh. Vychisl. Mat. Mat. Fiz., 5 (1965), pp. 548–551 (in Russian).
- [31] A. A. SAMARSKIY AND A. V. GULIN, *Stability of Finite Difference Schemes*, Nauka, Moscow, 1973 (in Russian).
- [32] A. A. SAMARSKIY, *Theory of Finite Difference Schemes*, Nauka, Moscow, 1977 (in Russian).
- [33] A. H. SCHATZ, *A weak discrete maximum principle and stability of the finite element method in L_∞ on plane polygonal domains. I*, Math. Comp., 34 (1980), pp. 77–91.
- [34] V. V. SMELOV, *Algebraic aspect of the discrete maximum principle*, Russian J. Numer. Anal. Math. Modelling, 16 (2001), pp. 175–190.
- [35] G. STOYAN, *On maximum principle for monotone matrices*, Linear Algebra Appl., 78 (1986), pp. 147–161.
- [36] G. WINDISH, *M-Matrices in Numerical Analysis*, BSB Teubner, Leipzig, 1989.