ON SPACE-TIME INVERSION INVARIANCE AND ITS RELATION TO NON-DISSIPATEDNESS OF A CESE CORE SCHEME

Sin-Chung Chang
NASA Glenn Research Center, Cleveland, OH 44135
E-mail: sin-chung.chang-1@nasa.gov

The core motivating ideas of the space-time CESE method are clearly presented and critically analyzed. It is explained why these ideas result in all the simplifying and enabling features of the CESE method. A thorough discussion of the \( a \) scheme, a two-level non-dissipative CESE solver of the PDE \( \frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0 \) with two independent mesh variables and two equations per mesh point is also presented. It is shown that the scheme possesses some rather intriguing properties such as: (i) its two independent mesh variables \( u_n^j \) and \( (u_x)_n^j \) separately satisfy two decoupled three-level leapfrog schemes and (ii) it shares with the leapfrog scheme the same amplification factors, even though the \( a \) scheme and the leapfrog scheme have completely different origins and structures. It is also explained why the leapfrog scheme is not as robust as the \( a \) scheme. The amplification factors/matrices of several non-dissipative schemes are carefully studied and the key properties that contribute to their non-dissipatedness are clearly spelled out. Finally we define and establish space-time inversion (STI) invariance for several non-dissipative schemes and show that their non-dissipatedness is a result of their STI invariance.

1. Introduction

The space-time conservation element and solution element (CESE) method is a high-resolution and genuinely multidimensional method for solving conservation laws \([1–71]\). Its nontraditional features include: (i) a unified treatment of space and time; (ii) the introduction of conservation elements (CEs) and solution elements (SEs) as the vehicles for enforcing space-time flux conservation; (iii) a novel time marching strategy that has a space-time staggered stencil at its core and, as such, fluxes at an interface can be evaluated without using any interpolation or extrapolation procedure (which, in turn, leads to the method’s ability to capture shocks without using Riemann solvers); (iv) the requirement that each scheme be built from a non-dissipative core scheme and, as a result, the numerical dissipation can be controlled effectively; and (v) the fact that mesh values of the physical dependent variables and their spatial derivatives are considered as independent marching variables to be solve for simultaneously. Note that CEs are non-overlapping space-time subdomains introduced such that (i) the computational domain can be filled by these subdomains; and (ii) flux conservation can be enforced over each of them and also over the union of any combination of them. On the other hand, SEs are space-time subdomains introduced such that (i) the boundary of each CE can be divided into several component parts with each of them belonging to a unique SE; and (ii) within a SE, any physical flux vector is approximated using simple smooth functions. In general, a CE does not coincide with a SE.

Without using flux-splitting or other special techniques, since its inception in 1991 [1] the unstructured-mesh compatible CESE method has been used to obtain numerous accurate 1D, 2D and 3D steady and unsteady flow solutions with Mach numbers ranging from 0.0028 to 10 [51]. The physical phenomena modeled include traveling and interacting shocks, acoustic waves, vortex shedding, viscous flows, detonation waves, cavitation, flows in fluid film bearings, heat conduction with melting and/or freezing, electrodynamics, MHD vortex, hydraulic jump, crystal growth, and chromatographic problems [3–71]. In particular, the rather unique capability of the CESE method to resolve both strong shocks and small disturbances (e.g., acoustic waves) simultaneously [13,15,16] makes it an effective tool for attacking computational aeroacoustics (CAA) problems. Note that the fact that second-order CESE schemes can solve CAA problems accurately is an
exception to the commonly-held belief that a second-order scheme is not adequate for solving CAA problems. Also note that, while numerical dissipation is needed for shock capturing, it may also result in annihilation of small disturbances. Thus a solver that can handle both strong shocks and small disturbances simultaneously must be able to overcome this difficulty.

In spite of its nontraditional features and potent capabilities, the core ideas of the CESE method are simple. In fact, all of its key features are the inescapable results of an honest pursuit driven by these simple ideas. The first and foremost is the belief that the method must be solid in physics. As such, in the CESE development, conservation laws are enforced locally and globally in their natural space-time unity forms for 1D, 2D and 3D cases. Moreover, because direct physical interaction generally occurs only among the immediate neighbors, use of the simplest stencil also becomes a CESE requirement. Obviously, this requirement is also very helpful in simplifying boundary-condition implementation.

The second idea is derived from the realization that stability and accuracy are two competing issues in time-accurate computations, i.e., too much numerical dissipation would degrade accuracy while too little of it will cause instability. In other words, to meet both accuracy and stability requirements, computation must be performed away from the edge (“cliff”) of instability but not too far from it. This represents a real dilemma in numerical method development. As an example, schemes with high-order accuracy generally has high accuracy and low numerical dissipation. However, it is susceptible to instability. In fact, in dealing with complicated real-world problems, stability of these schemes often is difficult to maintain without resorting to ad hoc treatments. To confront this issue head-on, in CESE development, it is required that a solver be built from a non-dissipative (i.e., neutrally stable) core scheme. By definition, computations involving a neutrally stable scheme are performed right on the edge of instability and therefore the numerical results generated are non-dissipative. As such numerical dissipation can be controlled effectively if the deviation of a solver from its non-dissipative core scheme can be adjusted using some built-in parameters. Note that the above idea also plays an essential role in the recent successful development of a family of Courant number insensitive schemes [59,61,64,65,67].

Other CESE ideas are: (i) the flux at an interface be evaluated in a simple and consistent manner; (ii) genuinely multidimensional schemes be built as simple, consistent and straightforward extensions of 1D schemes; (iii) triangular and tetrahedral meshes be used in 2D and 3D cases, respectively, so that the method is compatible to the simplest unstructured meshes and thus can be used to solve problems with complex geometries; and (iv) logical structures and approximation techniques used be as simple as possible, and special techniques that has only limited applicability and may cause undesirable side effects be avoided. Fortunately for the CESE development, as it turns out, the realization of the above lesser ideas (i)–(iv) follows effortlessly from that of the first two core ideas.

From the above review, it is seen that the most critical task in the CESE development is to develop conservation-law enforcing and neutrally stable schemes. To give a preview over how this was accomplished, consider the simple advection equation

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0$$

(1.1)

where $a \neq 0$ is a constant. Let $x_1 = x$, and $x_2 = t$ be considered as the coordinates of a two-dimensional Euclidean space $E_2$. Then, because Eq. (1.1) can be expressed as $\nabla \cdot \vec{h} = 0$ with $\vec{h} \overset{\text{def}}{=} (au, u)$, Gauss’ divergence theorem in the space-time $E_2$ implies that Eq. (1.1) is the differential form of the integral conservation law

$$\oint_{S(V)} \vec{h} \cdot d\vec{s} = 0$$

(1.2)

As depicted in Fig. 1, here (i) $S(V)$ is the boundary of an arbitrary space-time region $V$ in $E_2$, and (ii) $d\vec{s} = d\sigma \vec{n}$ with $d\sigma$ and $\vec{n}$, respectively, being the area and the unit outward normal of a surface element on $S(V)$. Note that: (i) because $\vec{h} \cdot d\vec{s}$ is the space-time flux of $\vec{h}$ leaving the region $V$ through the surface element $d\vec{s}$, Eq. (1.2) simply states that the total space-time flux of $\vec{h}$ leaving $V$ through $S(V)$ vanishes; (ii) in $E_2$, $d\sigma$ is the length of a line segment on the simple closed curve $S(V)$; and (iii) all mathematical operations can be carried out as though $E_2$ were an ordinary two-dimensional Euclidean space.
It is well known that a solution to Eq. (1.1) (or Eq. (1.2)) represents non-dissipative data propagation along its characteristic lines defined by $dx/dt = a$. Moreover, as will be further discussed in Sec. 4, Eq. (1.1) is invariant under space-time inversion (STI), i.e., it transforms back to itself if $x$ and $t$ are replaced by $-x$ and $-t$, respectively. Note that, in physics, STI invariance generally is referred to as $PT$ invariance where $P$ denotes parity (i.e., a mirror-image or spatial-reflection operation) while $T$ denotes a time-reversal operation.

From the above observations, it is obvious to the current author that non-dissipative behavior of the physical solutions to Eq. (1.1) will be preserved by the numerical solutions generated by a solver of Eq. (1.1) only if the solver also possesses the same STI invariance property. In other words, a non-dissipative solver of Eq. (1.1) must be STI invariant in some yet-to-be-defined sense. This conviction was further enhanced by the observation that the configurations of the stencils of all classical non-dissipative solvers of Eq. (1.1) such as the Wendroff [p.503, 72], the Crank-Nicolson [p.504, 72] and the leapfrog [p.100, 73] schemes do not change under space-time inversion while those of dissipative solvers of Eq. (1.1) such as the Lax scheme [p.97, 73] do change under space-time inversion. The preliminary results of a study along this line of thoughts were presented in a conference that took place in 1992 [2].

In a 1993 AIAA paper [74], the concept of STI invariance was also used by Thomas and Roe* to construct non-dissipative numerical schemes. (note: for the case where space-time is a 2D Euclidean space $E_2$, STI invariance is equivalent to the rotational symmetry referred to in [74]. However, in a Euclidean space of odd dimension, STI invariance and rotational symmetry are different because, in $E_N$, the determinant of the transformation matrix for STI is $(-1)^N$ while that for a proper rotation is always 1). In a 1994 ICASE report [75], the concept of STI invariance is again used by Roe to construct non-dissipative linear Bicharacteristic schemes.

In addition to the non-dissipative property referred to above, a solution to Eq. (1.1) also possesses the following properties: (i) it is completely determined by the data specified at the initial time level; (ii) its value at a spatial point at a later time has a finite domain of dependence (a point) at an earlier time. As such, in the initial CESE development, the focus is on the construction of an ideal core solver of Eq. (1.1) that is flux-conserving and also possesses the properties similar to those listed above, i.e., a flux-conserving two-level, explicit and non-dissipative solver. For a reason to be presented in Sec. 4, none of classical solvers meets the above requirement. Fortunately, by utilizing the concept of STI invariance, we are able to construct such an ideal scheme which is referred to as the $a$ scheme. It is a two level scheme with two independent mesh variables and two equations per mesh point.

The rest of the paper is organized as follows: In sec. 2, a thorough discussion of the $a$ scheme and its many equivalent forms is given. In addition, we establish a rather surprising result, i.e., the two independent mesh variables $u^n_{a}$ and $(u_{a})^n_{x}$ of the $a$ scheme separately satisfy two decoupled leapfrog schemes Eqs. (2.23) and (2.24) even though the $a$ scheme and the leapfrog scheme have completely different origins and structures. We also explain why the leapfrog scheme is not as robust as the $a$ scheme. In Sec. 3, we study the amplification factors/matrices of several non-dissipative schemes and point out the key properties that contribute to their non-dissipatedness. In Sec. 4, we define and establish STI invariance for each of the schemes discussed in Sec. 3 and show that their non-dissipatedness is a result of their STI invariance. Conclusions and discussions are given in Sec. 5. Finally, a computer code which can be used to verify Eqs. (2.23) and (2.24) numerically is listed in Appendix.

2. The $a$ scheme and the leapfrog scheme

To proceed, consider the set $\Omega_1$ of space-time staggered mesh points $(j, n)$ (dots in Fig. 2(a)), where
\[
\Omega_1 \overset{\text{def}}{=} \{(j, n)| n = 0, \pm 1, \pm 2, \pm 3, \ldots, (j + n) \text{ is an odd integer}\} 
\] (2.1)
Each $(j, n) \in \Omega_1$ is associated with a solution element, i.e., $SE(j, n)$. By definition, $SE(j, n)$ is the interior of the space-time region bounded by a dashed curve depicted in Fig. 2(b). It includes a horizontal line segment, a vertical line segment, and their immediate neighborhood.

* A participant of the 23rd Conference on Modeling and Simulation, April 30-May 1, 1992, Pittsburgh, PA. He was in the audience when the paper [2] was presented by this author, and participated in the discussion afterward.
At this juncture, the reader is warned that the notation used here may differ from that used in the previous CESE papers [1-71]. In particular, (i) the mesh indices \( j \) and \( n \) are only allowed to be whole integers here; and (ii) the spatial and temporal intervals that are denoted by \( \Delta x / 2 \) and \( \Delta t / 2 \), respectively, in [1-71] are represented by \( \Delta x \) and \( \Delta t \), respectively, here. These changes are introduced so that the CESE schemes and traditional schemes to be discussed here can be compared on the same footing.

Let \( (x, t) \in SE(j, n) \). Then Eq. (1.2) will be simulated numerically assuming that \( \bar{u}(x, t) \) and \( \bar{h}(x, t) \), respectively, are approximated by

\[
\bar{u}^*(x, t; j, n) \triangleq u^n_j + (u_x^n)_j (x - x_j) + (u_t^n)_j (t - t^n) \tag{2.2}
\]

and

\[
\bar{h}^*(x, t; j, n) \triangleq \left( au^*(x, t; j, n), \ u^*(x, t; j, n) \right) \tag{2.3}
\]

Note that: (i) \( u^n_j, (u_x)_j^n \), and \( (u_t)_j^n \) are constants in \( SE(j, n) \), and they can be considered as the numerical analogues of the values of \( u, \partial u / \partial x \), and \( \partial u / \partial t \) at the mesh point \( (j, n) \), respectively, (ii) \( (x_j, t^n) \) are the coordinates of the mesh point \( (j, n) \) with \( x_j = j \Delta x \) and \( t^n = n \Delta t \), and (iii) Eq. (2.3) is the numerical analogy of the definition \( \bar{h} = (au, u) \).

Let \( u = u^*(x, t; j, n) \) satisfy Eq. (1.1) within \( SE(j, n) \). Then one has \( (u_t)_j^n = -a (u_x)^n_j \). As a result, Eq. (2.2) reduces to

\[
u\bar{u}^*(x, t; j, n) = u^n_j + (u_x^n)_j [(x - x_j) - a (t - t^n)] \tag{2.4}
\]

i.e., \( \bar{u}_j^n \) and \( (u_x)^n_j \) are the only independent mesh variables associated with \( (j, n) \).

Let \( E_2 \) be divided into non-overlapping rectangular regions (see Fig. 2(a)) referred to as basic conservation elements (BCEs). As depicted in Figs. 2(c)–2(e), (i) each \( BCE \) has one and only one pair of diagonally opposite vertices which belong to \( \Omega_1 \); (ii) the space-time \( E_2 \) can be filled by \( CE_+ (j, n) \) and \( CE_-(j, n) \), \( (j, n) \in \Omega_1 \); and (iv) \( CE(j, n) \), which is the union of \( CE_-(j, n) \) and \( CE_+(j, n) \), is referred to as a compounded conservation element (CCE). Let the flux of \( \bar{h}^* \) conserve over all BCEs, i.e.,

\[
\int_{S(CE_+(j, n))} \bar{h}^* \cdot d\vec{s} = 0, \quad (j, n) \in \Omega_1 \tag{2.5}
\]

and

\[
\int_{S(CE_-(j, n))} \bar{h}^* \cdot d\vec{s} = 0, \quad (j, n) \in \Omega_1 \tag{2.6}
\]

Note that, among the line segments forming the boundary of \( CE_-(j, n) \), \( \overline{AB} \) and \( \overline{AD} \) belong to \( SE(j, n) \), while \( \overline{CB} \) and \( \overline{CD} \) belong to \( SE(j - 1, n - 1) \). Similarly, the boundary of \( CE_+(j, n) \) belongs to either \( SE(j, n) \) or \( SE(j + 1, n - 1) \). With the aid of the above observations and Eqs. (2.3) and (2.4), it can be shown that Eqs. (2.5) and (2.6) are equivalent to

\[
(1 - \nu) [u + (1 + \nu) u_x]_j^n = (1 - \nu) [u - (1 + \nu) u_x]_j^{n-1}, \quad (j, n) \in \Omega_1 \tag{2.7}
\]

and

\[
(1 + \nu) [u - (1 - \nu) u_x]_j^n = (1 + \nu) [u + (1 - \nu) u_x]_j^{n-1}, \quad (j, n) \in \Omega_1 \tag{2.8}
\]

respectively. Here: (i) \( \nu \triangleq a \Delta t / \Delta x \) is the Courant number; (ii)

\[
(u_x)_j^n \triangleq \frac{\Delta x}{2} (u_x)_j^n \tag{2.9}
\]

and (iii) to simplify notation, in the above and hereafter we adopt a convention that can be explained using an expression on the left side of Eq. (2.7) as an example, i.e.,

\[
[u + (1 + \nu) u_x]_j^n = u_j^n + (1 + \nu)(u_x)_j^n
\]
At this juncture, note that:

(a) Because

\[ \frac{\partial u}{\partial \bar{x}} = \frac{\Delta x}{2} \frac{\partial u}{\partial x} \]

if \( \bar{x} \equiv x/(\Delta x/2) \), the normalized parameter \((u_\bar{x})^n_j\) may be interpreted as the numerical analogy of the value at \((j, n)\) of the derivative of \(u\) with respect to the normalized coordinate \(\bar{x}\).

(b) Derivation of Eqs. (2.7) and (2.8) can be facilitated by the following observations: because \(u^*(x, t; j, n)\) is linear in \(x\) and \(t\), it can be shown that the total flux of \(\vec{h}^*\) leaving \(C_{E-}(j, n)\) or \(C_{E+}(j, n)\) through any of the four line segments that form its boundary is equal to the scalar product of the vector \(\vec{h}^*\) evaluated at the midpoint of the line segment and the “surface” vector (i.e., the unit outward normal multiplied by the length) of the line segment.

(c) By definition, point \(B\) depicted in Fig. 2(c) does not belong to either \(SE(j, n)\) or \(SE(j - 1, n - 1)\), i.e., \(\vec{h}^*\) is not defined there. However, this is not a problem for flux evaluation over \(AB\) or \(CB\) because the value of \(\vec{h}^*\) at an isolated point does not contribute to the flux of \(\vec{h}^*\) over a finite line segment.

Let \(1 - \nu^2 \neq 0\), i.e. \(1 - \nu \neq 0\) and \(1 + \nu \neq 0\). Then Eqs. (2.7) and (2.8) reduce to

\[ [u + (1 + \nu)u_{\bar{x}}]^n_j = [u - (1 + \nu)u_{\bar{x}}]^n+1_{j+1}, \quad (j, n) \in \Omega_1 \quad (2.10) \]

and

\[ [u - (1 - \nu)u_{\bar{x}}]^n_j = [u + (1 - \nu)u_{\bar{x}}]^n+1_{j-1}, \quad (j, n) \in \Omega_1 \quad (2.11) \]

respectively. Note that, for each \((j, n) \in \Omega_1\), Eq. (2.10) is a stronger condition than Eq. (2.7) (and therefore its equivalent Eq. (2.5)) in the sense that the former implies the latter for any given \(\nu\) while the latter implies the former only if an extra condition (i.e., \(\nu \neq 1\) for this case) is imposed. Similarly, for each \((j, n) \in \Omega_1\), Eq. (2.11) is also a stronger condition than Eq. (2.8) and its equivalent Eq. (2.6). Hereafter, for any given \(\nu\), the system of equations defined by Eqs. (2.10) and (2.11) for all \((j, n) \in \Omega_1\) is referred to as the first basic form of the a scheme. According to the above observation, any solution to this form automatically satisfies the conservation conditions represented by Eqs. (2.5) and (2.6).

For any given \((j, n) \in \Omega_1\), Eqs. (2.10) and (2.11) form a pair of linear equations for the unknowns \(u^n_j\) and \((u_{\bar{x}})^n_j\) if the expressions on the right sides are treated as known source terms. It can be shown that, for any given \(\nu\), the unique solution to this system is given by

\[ u^n_j = \frac{1}{2}\{[u - (1 + \nu)u_{\bar{x}}]_{j+1}^{n-1} + (1 + \nu)[u + (1 - \nu)u_{\bar{x}}]_{j-1}^{n-1}\}, \quad (j, n) \in \Omega_1 \quad (2.12) \]

and

\[ (u_{\bar{x}})^n_j = \frac{1}{2}\{[u + (1 + \nu)u_{\bar{x}}]_{j+1}^{n-1} - [u + (1 - \nu)u_{\bar{x}}]_{j-1}^{n-1}\}, \quad (j, n) \in \Omega_1 \quad (2.13) \]

In other words, for any given \(\nu\) and any \((j, n) \in \Omega_1\), Eqs. (2.10) and (2.11) imply Eqs. (2.12) and (2.13), and vice versa. Because \(u^n_j\) and \((u_{\bar{x}})^n_j\) can be explicitly evaluated in terms of \(u_{\bar{x}}^{n-1}_{j\pm 1}\) and \((u_{\bar{x}})^{n-1}_{j\pm 1}\) using Eqs. (2.12) and (2.13), hereafter, for any given \(\nu\), the system of equations defined by these equations for all \((j, n) \in \Omega_1\) is referred to as the forward marching form of the a scheme. Obviously a solution to the forward marching form must also be a solution to the first basic form and vice versa.

Next, by (i) replacing the indices \(j\) and \(n\) in Eq. (2.10) with \(j - 1\) and \(n + 1\), respectively, and (ii) using the fact that \((j, n) \in \Omega_1 \Leftrightarrow (j + 1, n - 1) \in \Omega_1 \Leftrightarrow (j - 1, n + 1) \in \Omega_1\) (hereafter the symbol \(\Leftrightarrow\) is used as the shorthand for “if and only if”), one can show that, for any given \(\nu\), the system of equations defined by Eq. (2.10) for all \((j, n) \in \Omega_1\) is identical to that defined by

\[ [u - (1 + \nu)u_{\bar{x}}]^n_j = [u + (1 + \nu)u_{\bar{x}}]_{j-1}^{n+1}, \quad (j, n) \in \Omega_1 \quad (2.14) \]
for all \((j, n) \in \Omega_1\). Similarly by (i) replacing the indices \(j\) and \(n\) in Eq. (2.11) with \(j+1\) and \(n+1\), respectively, and (ii) using the fact that \((j, n) \in \Omega_1 \iff (j-1, n-1) \in \Omega_1 \iff (j+1, n+1) \in \Omega_1\), one can show that, for any given \(\nu\), the system of equations defined by Eq. (2.11) for all \((j, n) \in \Omega_1\) is identical to that defined by

\[
[u + (1 - \nu)u_{\bar{x}}]^n_j = [u - (1 - \nu)u_{\bar{x}}]^{n+1}_{j+1}, \quad (j, n) \in \Omega_1
\]

for all \((j, n) \in \Omega_1\). Hereafter, for any given \(\nu\), the system of equations defined by Eqs. (2.14) and (2.15) for all \((j, n) \in \Omega_1\) is referred to as the second basic form of the scheme. Obviously the first and the second basic forms are simply different representations of the same system of equations.

For any given \((j, n) \in \Omega_1\), Eqs. (2.14) and (2.15) form a pair of linear equations for the unknowns \(u^n_j\) and \((u_{\bar{x}})^n_j\) if the expressions on the right sides are treated as known source terms. It can be shown that, for any given \(\nu\), the unique solution to this system is given by

\[
u^n_j = \frac{1}{2} \left\{ (1 + \nu) \left[ u - (1 - \nu)u_{\bar{x}} \right]^{n+1}_{j+1} + (1 - \nu) \left[ u + (1 + \nu)u_{\bar{x}} \right]^{n+1}_{j-1} \right\}, \quad (j, n) \in \Omega_1
\]

and

\[
(u_{\bar{x}})^n_j = \frac{1}{2} \left\{ [u - (1 - \nu)u_{\bar{x}}]^{n+1}_{j+1} - [u + (1 + \nu)u_{\bar{x}}]^{n+1}_{j-1} \right\}, \quad (j, n) \in \Omega_1
\]

In other words, for any given \(\nu\) and any \((j, n) \in \Omega_1\), Eqs. (2.14) and (2.15) imply Eqs. (2.16) and (2.17), and vice versa. Because \(u^n_j\) and \((u_{\bar{x}})^n_j\) can be explicitly evaluated in terms of \(u^{n+1}_{j+1}\) and \((u_{\bar{x}})^{n+1}_{j+1}\) using Eqs. (2.16) and (2.17), hereafter, for any given \(\nu\), the system of equations defined by these equations for all \((j, n) \in \Omega_1\) will be referred to as the backward marching form of the scheme. Obviously a solution to the backward marching form must also be a solution to the second basic form and vice versa.

The above developments are summarized and further elucidated in the following remarks:

(a) According to the above discussions, among the four forms of the scheme defined above, any one implies each of the other three and vice versa. Because of this equivalency, hereafter the same essential conditions represented by any one of the above four forms will be referred to simply as the scheme.

(b) Because any solution to the first basic form of the scheme automatically satisfies the conservation conditions Eqs. (2.5) and (2.6), it follows from the remarks made in item (a) that any solution to any other form also satisfies Eqs. (2.5) and (2.6).

(c) Each component equation of the first or second basic form represents an implicit relation involving the mesh variables at two diagonally opposite neighboring mesh points, i.e., each “basic” stencil of the scheme is formed by two neighboring mesh points. On the other hand, each component equation of the forward or backward marching form represents an explicit relation involving the mesh variables at three neighboring mesh points, i.e., each “marching” stencil of the scheme is formed by three neighboring mesh points.

(d) Because (i) the vector \(\vec{h}^*\) at any surface element lying on any interface separating two neighboring BCEs is evaluated using the information from a single SE, and (ii) the unit outward normal vector on the surface element pointing outward from one of these two neighboring BCEs is exactly the negative of that pointing outward from another BCE, one concludes that the flux leaving one of these BCEs through the interface is the negative of that leaving another BCE through the same interface. Due to this interface flux cancelation, the local conservation relations Eqs. (2.5) and (2.6) lead to a global flux conservation relation, i.e., the total flux of \(\vec{h}^*\) leaving the boundary of any space-time region that is the union of any combination of BCEs vanishes. In particular, because CE\((j, n)\) is the union of CE\(-j, n\) and CE\(+j, n\),

\[
\int_{\partial(CE(j, n))} \vec{h}^* \cdot d\vec{s} = 0, \quad (j, n) \in \Omega_1
\]

must follow from Eqs. (2.5) and (2.6). In fact, it can be shown that Eq. (2.18) is equivalent to Eq. (2.12). Also it can be shown that the conservation condition over the union of CE\(+j-1, n\) and CE\(-j+1, n\) is equivalent to Eq. (2.16) (see Figs. 2(f)-(h)).
(e) As shown in [5], the two amplification factors of the $a$ scheme are identical to those of the leapfrog scheme. As a result, the $a$ scheme is non-dissipative and it is stable if $|\nu| < 1$. (see the additional discussions given in Sec. 3).

(f) For the dissipative extensions [5] of the non-dissipative forward marching form of the $a$ scheme, only the less stringent local flux conservation condition Eq. (2.18) is imposed. Because CE($j,n$), by definition, is a CCE, it follows that the solution to such an extension satisfies the following less stringent global flux conservation relation: the total flux of $h^a$ leaving the boundary of any space-time region that is the union of any combination of CCEs vanishes. Moreover, because Eq. (2.18) is equivalent to Eq. (2.12), for each of these extensions, $u_j^n$ is still evaluated using Eq. (2.12) while $(u_\bar{x})_j^n$ is evaluated using an equation different from Eq. (2.13).

Next we will show that, in a sense to be defined shortly, any solution of the $a$ scheme automatically satisfies the classical leapfrog scheme. Consider any $(j,n+1) \in \Omega_1$. Then because $(j,n+1) \in \Omega_1 \Leftrightarrow (j+1,n) \in \Omega_1 \Leftrightarrow (j-1,n) \in \Omega_1 \Leftrightarrow (j,n+1) \in \Omega_1$, Eqs. (2.12), (2.13), (2.16), and (2.17) imply that

$$u_j^{n+1} = \frac{1}{2} \left( (1-\nu) [u - (1+\nu)u_{x,j+1}^n ] + (1+\nu) [u + (1-\nu)u_{x,j-1}^n ] \right), \quad (j,n+1) \in \Omega_1 \tag{2.19}$$

$$\left( u_{\bar{x}}^\nu \right)^{n+1}_j = \frac{1}{2} \left( [u - (1+\nu)u_{x,j+1}^n ] - [u + (1-\nu)u_{x,j-1}^n ] \right), \quad (j,n+1) \in \Omega_1 \tag{2.20}$$

$$u_j^{-n} = \frac{1}{2} \left( (1+\nu) [u - (1-\nu)u_{x,j+1}^n ] + (1-\nu) [u + (1+\nu)u_{x,j-1}^n ] \right), \quad (j,n-1) \in \Omega_1 \tag{2.21}$$

$$\left( u_{\bar{x}}^\nu \right)^{-n}_j = \frac{1}{2} \left( [u - (1-\nu)u_{x,j+1}^n ] - [u + (1+\nu)u_{x,j-1}^n ] \right), \quad (j,n-1) \in \Omega_1 \tag{2.22}$$

respectively. By subtracting Eqs. (2.21) and (2.22) from Eqs. (2.19) and (2.20), respectively, and using the definition $\nu = a \Delta t / \Delta x$, one has two leapfrog schemes in which the mesh variables $u_j^n$ and $(u_\bar{x})_j^n$ are decoupled, i.e.,

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} + a \frac{u_{x,j+1}^{n} - u_{x,j-1}^{n}}{2\Delta x} = 0, \quad (j,n+1) \in \Omega_1 \tag{2.23}$$

and

$$\frac{(u_\bar{x})^{n+1}_j - (u_\bar{x})^{n-1}_j}{2\Delta t} + a \frac{(u_\bar{x})_{x,j+1}^{n} - (u_\bar{x})_{x,j-1}^{n}}{2\Delta x} = 0, \quad (j,n+1) \in \Omega_1 \tag{2.24}$$

Hereafter, the system of equations defined by Eqs. (2.23) and (2.24) over all $(j,n+1) \in \Omega_1$ will be referred to as the leapfrog form of the $a$ scheme. Note that: (i) the $a$ scheme implies the leapfrog form but not vice versa; and (ii) given the initial data at two consecutive time levels, the leapfrog form can be used for both forward and backward time marching.

Let $u_j^0$ and $(u_\bar{x})_j^0$, $j = \pm 1, \pm 3, \ldots$, be given. Then Eqs. (2.12) and (2.13) can be used to define uniquely a set of $u_j^n$ and $(u_\bar{x})_j^n$ where $(j,n) \in \Omega_1$ and $n = 1, 2, 3, \ldots$. On the other hand, with the aid of (i) the given initial data $u_j^0$ and $(u_\bar{x})_j^0$, $j = \pm 1, \pm 3, \ldots$, and (ii) the values of $u_j^1$ and $(u_\bar{x})_j^1$, $j = 0, \pm 2, \pm 4, \ldots$, determined above, the leapfrog schemes Eqs. (2.23) and (2.24) can be used to define uniquely a new set of $u_j^n$ and $(u_\bar{x})_j^n$ where $(j,n) \in \Omega_1$ and $n = 2, 3, 4, \ldots$. According to the result established in the last paragraph, for each $(j,n) \in \Omega_1$ where $n = 2, 3, 4, \ldots$, the values of $u_j^n$ and $(u_\bar{x})_j^n$ in the new set, respectively, should be identical to those in the set defined using the forward marching form of the $a$ scheme. This conclusion has been verified in numerical experiments using periodic initial data (refer to the code listed in the appendix).

This section is concluded with the following remarks:

(a) Because, at each $(j,n) \in \Omega_1$, the $a$ scheme is formed by two coupled equations involving two independent mesh variables $u_j^n$ and $(u_\bar{x})_j^n$, it is consistent with a pair of partial differential equations (PDEs) with one of them being Eq. (1.1) [1,11].

(b) The leapfrog scheme is a two-way marching scheme, i.e., given the data at any two consecutive time levels, Eq. (2.23) can be used for both forward and backward time marching.
(c) The leapfrog scheme Eq. (2.23) can be derived using a finite-difference [p.100, 73] or a finite-volume approach. In fact, for any \((j, n) \in \Omega_1\), Eq. (2.23) can be cast into the local conservation form

\[
\int_{S(FV(j+1,n))} \vec{h}^* \cdot d\vec{s} = 0, \quad (j + 1, n) \in \Omega_1
\]  

(2.25)

where (i) \(FV(j+1,n)\) is the region \(ABCD\) depicted in Fig. 3; and (ii) the average flux vectors \(\vec{h}^*\) at \(AB, BC, CD,\) and \(DA\) are taking to be \((au_j^{n+1}, u_j^{n+1}), (au_j^n, u_j^n), (au_{j-1}^{n-1}, u_{j-1}^{n-1})\), and \((au_{j+1}, u_{j+1}^{n+1})\), respectively. Thus, for the leapfrog scheme, hereafter \(FV(j+1,n)\) is referred to as the conservation element (CE) assigned to the mesh point \((j + 1, n)\).

(d) Let (i)

\[
\Omega_{11} \stackrel{\text{def}}{=} \{(j, n)|(j, n) \in \Omega_1 \text{ and } n = \pm 1, \pm 3, \ldots\}
\]

(2.26)

(ii)

\[
\Omega_{12} \stackrel{\text{def}}{=} \{(j, n)|(j, n) \in \Omega_1 \text{ and } n = 0, \pm 2, \pm 4, \ldots\}
\]

(2.27)

and (iii) \(\Sigma_{11} (\Sigma_{12}) \stackrel{\text{def}}{=} \) the set of the CEs of the leapfrog scheme assigned to all \((j, n) \in \Omega_{11} (\Omega_{12})\). Then (i) \(\Omega_1 = \Omega_{11} \cup \Omega_{12}\), and (ii) the CEs in \(\Sigma_{11}\) and \(\Sigma_{12}\), separately, are non-overlapping and fill the entire space-time \(E_2\). Thus, by using Eq. (2.25) and the same argument used to establish the global flux conservation relation for the \(a\) scheme, one can establish two separate global flux conservation relations for the leapfrog scheme, i.e., the total flux of \(\vec{h}^*\) leaving the boundary of any space-time region that is the union of any combination of CEs \(\in \Omega_{11} (\Omega_{12})\) vanishes.

(e) The leapfrog scheme shares with the \(a\) scheme three key nontraditional features, i.e., they both are explicit two-way marching schemes, have space-time staggered stencils and their interface fluxes are evaluated without using any interpolation or extrapolation technique. However, due to its less sophisticated structure, it is much more difficult (if it is possible at all) to construct space-time flux-conserving dissipative extensions of the leapfrog scheme than those of the \(a\) scheme. Because (i) stable nonlinear simulations can only be achieved using dissipative extensions and (ii) a two level scheme is easier to handle than a three-level scheme, CESE development (which is based on the \(a\) scheme) is much simpler and more robust than that based on the leapfrog scheme.

(f) By using Eq. (2.9), one can show that Eq. (2.24) is still valid if each \((u_x)_j^n\), \((j, n) \in \Omega_1\), is replaced by its mesh-size independent version \((u_x)_j^n\). Because the leapfrog scheme is second-order in accuracy for both space and time, one can infer from Eq. (2.23) and the mesh-size independent version of Eq. (2.24) that the \(a\) scheme is second-order in accuracy for both \(u_j^n\) and \((u_x)_j^n\). In fact, this conclusion was proved rigorously as Eqs. (6.32) and (6.33) in [1], and can also be verified numerically using the code listed in Appendix.

(g) In the above, the \(a\) and leapfrog schemes are defined using only the mesh points \(\in \Omega_1\). Independently, they can also be defined using only the mesh points \(\in \Omega_2\) where

\[
\Omega_2 \stackrel{\text{def}}{=} \{(j, n)|j, n = 0, \pm 1, \pm 2, \pm 3, \ldots, \text{ and } (j + n) \text{ is an even integer}\}
\]

(2.28)

For the current 1D case where a structured mesh is used, both the \(a\) scheme and the leapfrog scheme defined over \(\Omega_1\) are completely decoupled from those defined over \(\Omega_2\). Thus, there is no practical reason to carry out computations using the two decoupled schemes simultaneously. Nevertheless, the leapfrog scheme is traditionally referred but unjustifiably to that formed by the two decoupled leapfrog schemes. Moreover, through a twisted logic, the leapfrog scheme is falsely blamed for its solution becoming decoupled after a long time marching.

To simplify mathematical manipulation, hereafter we will exclusively deal with the "dual" \(a\) scheme and the "dual" leapfrog scheme which, respectively, are formed from two decoupled \(a\) schemes and two decoupled leapfrog schemes. As such these dual schemes are defined over

\[
\Omega \stackrel{\text{def}}{=} \Omega_1 \cup \Omega_2 = \{(j, n)|j, n = 0, \pm 1, \pm 2, \ldots\}
\]

(2.29)
Moreover, from now on it is to be understood that (i) the a scheme and the leapfrog scheme refer to the dual schemes which are defined over $\Omega$; and (ii) the “dual” version of a previously defined system of equations will be identified using the same equation number of the original version with an extra symbol “d” being attached at the end.

3. von Neumann Analysis

As a preliminary to later developments, the results of the von Neumann analysis for several non-dissipative (i.e., neutrally stable) solvers for Eq. (1.1) will be studied in this section. In particular, we will point out key properties shared by the amplification matrices (factors) of these schemes.

The Wendroff scheme \[p.503, 72\], i.e.,

\[(1 + \nu)(u_{j+1}^n - u_j^n) + (1 - \nu)(u_j^{n+1} - u_{j+1}^n) = 0, \quad (j, n) \in \Omega \tag{3.1}\]

is a two-level and two-way marching implicit non-dissipative solver of Eq. (1.1). Let

\[u_j^n = [g(\nu, \theta)]^n e^{ij\theta}, \quad (j, n) \in \Omega \tag{3.2}\]

be a solution to Eq. (3.1) where (i) $i \equiv \sqrt{-1}$, (ii) $\theta (-\infty < \theta < \infty)$ is a phase angle, and (iii) $g(\nu, \theta) (\neq 0)$ is a complex function of $\nu$ and $\theta$ (note: $[g(\nu, \theta)]^n$ is undefined for any $n < 0$ if $g(\nu, \theta) = 0$). By substituting Eq. (3.2) into Eq. (3.1), one has

\[\left[(1 + \nu)(ge^{i\theta} - 1) + (1 - \nu)(g - e^{i\theta})\right]g^n = 0, \quad n = 0, \pm 1, \pm 2, \ldots, \tag{3.3}\]

Because $g^0 = 1$, Eq. (3.3) $\iff$

\[\left[(1 + \nu)e^{i\theta} + (1 - \nu)\right]g = (1 - \nu)e^{i\theta} + (1 + \nu) \tag{3.4}\]

Dividing the expressions on both sides of Eq. (3.4) by $e^{i\theta/2}$ and using basic trigonometry, one has

\[g = g_w(\nu, \theta) \overset{\text{def}}{=} \frac{\cos(\theta/2) - i\nu\sin(\theta/2)}{\cos(\theta/2) + i\nu\sin(\theta/2)} \quad (\cos(\theta/2) \pm i\nu\sin(\theta/2) \neq 0) \tag{3.5}\]

Note that: (i) $\cos(\theta/2) + i\nu\sin(\theta/2) = 0 \iff \cos(\theta/2) - i\nu\sin(\theta/2) = 0 \iff \nu = 0$ and $\theta = \pm \pi, \pm 3\pi, \pm 5\pi, \ldots$, i.e., the domain of $g_w(\nu, \theta)$ is

\[D(g_w) \overset{\text{def}}{=} \{(\nu, \theta) | \nu \neq 0 \text{ with any real } \theta, \text{ or } \nu = 0 \text{ with any real } \theta \neq \pm \pi, \pm 3\pi, \pm 5\pi, \ldots\} \tag{3.5a}\]

(ii) by definition, $g_w(\nu, \theta)$ is the amplification factor of the Wendroff scheme; and (iii) over $D(g_w)$, we have

\[g(\nu, -\theta) \cdot g(\nu, \theta) = 1 \tag{3.6}\]

and

\[g(\nu, \theta) = g(\nu, -\theta) \tag{3.7}\]

if $g(\nu, \theta) = g_w(\nu, \theta)$. Here, for any complex number $c$, its complex conjugate is denoted by $\overline{c}$. Combining Eqs. (3.6) and (3.7), one concludes that, for all real $\theta$, we have

\[|g(\nu, \theta)| = 1 \tag{3.8}\]

if $g(\nu, \theta) = g_w(\nu, \theta)$ and $\nu \neq 0$. Thus, practically, the Wendroff scheme is unconditionally neutrally stable.

At this juncture, note that the property Eq. (3.7) is a result of the fact that the Wendroff scheme is a linear solver of Eq. (1.1) which has real constant coefficients and a single mesh variable per mesh point. This can be seen by the following arguments: By taking the complex conjugate of Eq. (3.1), it is seen that
the complex conjugate of any solution to the solver must also be a solution. Thus, if Eq. (3.2) is a solution, so is
\[ u^n_j = \left[ g(\nu, \theta) \right]^n e^{-ij\theta}, \quad (j, n) \in \Omega \] (3.9)

Because (i) for any \((\nu, \theta) \in D(g_w), g(\nu, \theta)\) is uniquely determined by the requirement that Eq. (3.2) is a solution to Eq. (3.1), and (ii) \((\nu, \theta) \in D(g_w) \leftrightarrow (\nu, -\theta) \in D(g_w),\) a comparison between Eq. (3.2) and the equation resulting from replacing \(\theta\) with \(-\theta\) in Eq. (3.9) reveals that, for any \((\nu, \theta) \in D(g_w),\)
\[ g(\nu, -\theta) = g(\nu, \theta) \] (3.7a)

Again because \((\nu, \theta) \in D(g_w) \leftrightarrow (\nu, -\theta) \in D(g_w),\) Eqs. (3.7) and (3.7a) are equivalent. QED. As such, the property Eq. (3.7) is shared by any solver of Eq. (1.1) which shares the same characteristics of the Wendroff scheme defined above. Because this solver can even be a dissipative solvers such as the Lax scheme [p.97, 73] the key factor contributing to the neutral stability of the Wendroff scheme is Eq. (3.6) rather than Eq. (3.7).

Another two-level and two-way marching implicit non-dissipative solver of Eq. (1.1) is the Crank-Nicolson scheme [p.504, 72], i.e.,
\[ u^{n+1}_j - u^n_j + \frac{\nu}{4} (u^{n+1}_{j+1} + u^n_{j+1} - u^{n+1}_{j-1} - u^n_{j-1}) = 0, \quad (j, n) \in \Omega \] (3.10)

Through a process similar to that presented above, it can be shown that the amplification factor for the Crank-Nicolson scheme is
\[ g_{cn}(\nu, \theta) \equiv \frac{1 - (i\nu/2) \sin \theta}{1 + (i\nu/2) \sin \theta}, \quad -\infty < \nu, \theta < +\infty \] (3.11)

It can be shown easily that, for any real \(\nu\) and \(\theta,\) Eqs. (3.6)–(3.8) are satisfied if \(g(\nu, \theta) = g_{cn}(\nu, \theta).\) Thus the Crank-Nicolson scheme is unconditionally neutrally stable.

There is no classical two-level neutrally stable explicit solver of Eq. (1.1). In fact, among the classical schemes, the three-level leapfrog scheme is the only well-known neutrally stable explicit scheme. To study its amplification factors, first we will cast it into an equivalent two-level forward marching matrix form. As a preliminary, let
\[ P \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad H(\nu) \equiv \begin{pmatrix} \nu & 0 \\ 0 & 0 \end{pmatrix} \] (3.12)

and
\[ \tilde{\phi}(j, n) \equiv \begin{pmatrix} \phi^n_j \\ \phi^n_{j+1} \end{pmatrix}, \quad (j, n) \in \Omega \] (3.13)

where \(\phi^n_j\) and \(\phi^n_{j+1}, (j, n) \in \Omega\) are arbitrary real constants. Then (i)
\[ P = P^{-1} \] (3.14)

where \(P^{-1}\) is the inverse of \(P;\) and (ii) the system of matrix equations
\[ \tilde{\phi}(j, n + 1) = P \tilde{\phi}(j, n) + H(\nu) \left[ \tilde{\phi}(j - 1, n) - \tilde{\phi}(j + 1, n) \right], \quad (j, n) \in \Omega \] (3.15)

is equivalent to
\[ \phi^{n+1}_j = \phi^n_j + \nu \left[ \phi^n_{j-1} - \phi^n_{j+1} \right], \quad (j, n) \in \Omega \] (3.16)

and
\[ \tilde{\phi}^{n+1}_j = \tilde{\phi}^n_j, \quad (j, n) \in \Omega \] (3.17)
From the above observation (ii), one concludes that Eqs. (3.13) and (3.15) are equivalent to

\[ \tilde{\phi}(j, n) = \left( \begin{array}{c} \phi^n_j \\ \phi^{n-1}_j \end{array} \right), \quad (j, n) \in \Omega \]  

(3.18)

and

\[ \phi^{n+1}_j = \phi^{n-1}_j + \nu [\phi^n_{j-1} - \phi^n_{j+1}], \quad (j, n) \in \Omega \]  

(3.19)

By comparing Eq. (3.19) with Eq. (2.23d) and using Eq. (3.18), it is seen that Eq. (2.23d) is equivalent to Eq. (3.15) if \( \phi^n_{j} \) is taken to be \( u^n_j \), i.e., if

\[ \tilde{\phi}(j, n) \overset{\text{def}}{=} \left( \begin{array}{c} u^n_j \\ u^{n-1}_j \end{array} \right), \quad (j, n) \in \Omega \]  

(3.20)

Hereafter, assuming Eq. (3.20), Eq. (3.15) is referred to as the forward marching form of the leapfrog scheme.

For any real \( \nu \) and \( \theta \), let \( A(\nu, \theta) \) be a \( 2 \times 2 \) nonsingular complex matrix function of \( \nu \) and \( \theta \) such that

\[ \tilde{\phi}(j, n) = e^{i\theta} [A(\nu, \theta)]^n \tilde{a}, \quad (j, n) \in \Omega \]  

(3.21)

is a solution to Eq. (3.15) for all possible complex constant \( 2 \times 1 \) column matrices \( \tilde{a} \) (note: \( [A(\nu, \theta)]^n \) is undefined for any \( n < 0 \) if \( A(\nu, \theta) \) is singular). By substituting Eq. (3.21) into Eq. (3.15), one has

\[ [A(\nu, \theta) - P - H(\nu)(e^{-i\theta} - e^{i\theta})] [A(\nu, \theta)]^n \tilde{a} = 0, \quad n = 0, \pm 1, \pm 2, \ldots \]  

(3.22)

Because (i) \( [A(\nu, \theta)]^0 = I \) where \( I \) is the \( 2 \times 2 \) identity matrix, and (ii) \( \tilde{a} \) can be any complex constant \( 2 \times 1 \) column matrix, with the aid of Eq. (3.12), one concludes that Eq. (3.22) \( \iff \)

\[ A(\nu, \theta) = P + H(\nu)(e^{-i\theta} - e^{i\theta}) = \begin{pmatrix} -2i\nu \sin \theta & 1 \\ 1 & 0 \end{pmatrix}, \quad -\infty < \nu, \theta < +\infty \]  

(3.23)

According to Eq. (3.23), \( \det[A(\nu, \theta)] = -1 \). Thus \( A(\nu, \theta) \) is indeed nonsingular over its domain. By definition, \( A(\nu, \theta) \) is referred to as the amplification matrix of the leapfrog scheme with respect to the form Eq. (3.15). It can be shown easily that

\[ A(\nu, -\theta)PA(\nu, \theta)P^{-1} = I, \quad -\infty < \nu, \theta < +\infty \]  

(3.24)

and

\[ \overline{A(\nu, \theta)} = A(\nu, -\theta), \quad -\infty < \nu, \theta < +\infty \]  

(3.25)

Here, for any matrix \( M \), its complex conjugate is denoted by \( \overline{M} \). Combining Eqs. (3.24) and (3.25), one has

\[ \overline{A(\nu, \theta)} PA(\nu, \theta)P^{-1} = I, \quad -\infty < \nu, \theta < +\infty \]  

(3.26)

Eqs. (3.24)—(3.26) are the current counterparts to Eqs. (3.6)—(3.8), respectively.

By definition, the two eigenvalues of \( A(\nu, \theta) \) are referred to as the amplification factors of the leapfrog scheme Eq. (2.23d). Let these factors be denoted by \( \lambda_+(\nu, \theta) \) and \( \lambda_-(\nu, \theta) \), respectively. Then the eigenvalues of \( A(\nu, \theta) \) are \( \lambda_+(\nu, \theta) \) and \( \lambda_-(\nu, \theta) \), respectively. On the other hand, because \( PA(\nu, \theta)P^{-1} \) is similar to \( A(\nu, \theta) \) (i.e., these two matrices are related by a similarity transformation), the eigenvalues of \( PA(\nu, \theta)P^{-1} \) are also \( \lambda_+(\nu, \theta) \) and \( \lambda_-(\nu, \theta) \), respectively [76]. Moreover, we have

\[ \det[A(\nu, \theta)] \cdot \det[PA(\nu, \theta)P^{-1}] = \det(I) = 1, \quad -\infty < \nu, \theta < +\infty \]  

(3.27)
which is a result of Eq. (3.26) and the matrix theorem: \( \det(AB) = \det(A) \cdot \det(B) \) for any \( N \times N \) matrices \( A \) and \( B \). By using the above observations and another matrix theorem, i.e., \textit{the determinant of a square matrix is equal to the product of the eigenvalues of the same matrix}, one now arrives at the conclusion that

\[
\lambda_+(\nu, \theta) \cdot \lambda_-(\nu, \theta) \cdot \lambda_+(\nu, \theta) \cdot \lambda_-(\nu, \theta) = 1, \quad -\infty < \nu, \theta < +\infty
\]
i.e.,

\[
|\lambda_+(\nu, \theta)| \cdot |\lambda_-(\nu, \theta)| = 1, \quad -\infty < \nu, \theta < +\infty \tag{3.28}
\]

For any given \( \nu \), stability of the leapfrog scheme requires that

\[
|\lambda_+(\nu, \theta)| \leq 1 \quad \text{and} \quad |\lambda_-(\nu, \theta)| \leq 1, \quad -\infty < \theta < +\infty \tag{3.29}
\]

Thus Eq. (3.28) implies that, for any given \( \nu \), the leapfrog scheme must be neutrally stable, i.e.,

\[
|\lambda_+(\nu, \theta)| = |\lambda_-(\nu, \theta)| = 1, \quad -\infty < \theta < +\infty \tag{3.30}
\]

if it is stable. As such, Eq. (3.26) does not imply neutral stability of the leapfrog scheme. However, it does imply that the scheme can only be neutrally stable (i.e., non-dissipative) if it is stable. Here we have reached this conclusion without using the explicit form of \( \lambda_+(\nu, \theta) \) and \( \lambda_-(\nu, \theta) \), i.e.,

\[
\lambda_{\pm}(\nu, \theta) = -i\nu \sin \theta \pm \sqrt{1 - \nu^2 \sin^2 \theta} \tag{3.31}
\]

In fact, by using Eqs. (3.23) and (3.31) along with a rigorous definition of stability given in [64], it can be shown that the leapfrog scheme is (i) neutrally stable if \( \nu^2 < 1 \); (ii) linearly unstable when \( \nu^2 = 1 \); and (iii) exponentially unstable if \( \nu^2 > 1 \). Note that the linear instability at \( \nu^2 = 1 \) is due to the fact that \( A(\nu, \theta) \) becomes defective when \( \nu^2 = \sin^2 \theta = 1 \) (see pp. 6-9 in [64]).

Note that, except for some simple 1D cases, explicit forms of amplification factors generally are unavailable. However, as explain earlier, the matrix relation such as Eq. (3.26) may be used to draw important conclusions about possible neutral stability of a scheme even if no explicit amplification factors are available.

Next we will show that: (i) the leapfrog scheme Eq. (2.23d) can be cast into different equivalent two-level matrix forms; and (ii) the amplification matrices associated with different matrix forms are related by similarity transformations and therefore they share the same eigenvalues, i.e., the same amplification factors.

Let

\[
\tilde{\phi}_{\pm}(j, n) \overset{\text{def}}{=} \left( \begin{array}{c} u^n_j \\ u^{n+1}_{j+1} \end{array} \right), \quad (j, n) \in \Omega \tag{3.32}
\]

\[
I_+ \overset{\text{def}}{=} \left( \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right), \quad I_- \overset{\text{def}}{=} \left( \begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \right) \tag{3.33}
\]

\[
H_+(\nu) \overset{\text{def}}{=} \left( \begin{array}{cc} \nu & 1 \\ 0 & 0 \end{array} \right), \quad H_-(\nu) \overset{\text{def}}{=} \left( \begin{array}{cc} -\nu & 0 \\ 1 & 0 \end{array} \right) \tag{3.34}
\]

and

\[
R(\theta) \overset{\text{def}}{=} I_+ + I_- e^{i\theta}, \quad -\infty < \theta < +\infty \tag{3.35}
\]

Eqs. (3.33) and (3.35) imply that

\[
(I_+)^2 = I_+, \quad (I_-)^2 = I_-, \quad I_- I_+ = I_+ I_- = 0, \quad I_- + I_+ = I \tag{3.36}
\]

and

\[
(R(\theta))^{-1} = R(-\theta), \quad -\infty < \theta < +\infty \tag{3.37}
\]

where \( 0 \) is the \( 2 \times 2 \) null matrix. Moreover, by using Eqs. (3.20) and (3.32)–(3.34), one can shown that: (i)

\[
\tilde{\phi}_{\pm}(j, n) = I_+ \tilde{\phi}(j, n) + I_- \tilde{\phi}(j \pm 1, n), \quad (j, n) \in \Omega \tag{3.38}
\]
and (ii) Eq. (2.23d) is equivalent to each of the following two forms:

\[
\begin{align*}
\vec{\phi}_+(j, n + 1) &= H_-(\nu)\vec{\phi}_+(j + 1, n) + H_+(\nu)\vec{\phi}_+(j - 1, n), \quad (j, n) \in \Omega \\
\vec{\phi}_-(j, n + 1) &= [H_-(\nu)]^t\vec{\phi}_-(j + 1, n) + [H_+(\nu)]^t\vec{\phi}_-(j - 1, n), \quad (j, n) \in \Omega
\end{align*}
\]

and

where \([H_\pm(\nu)]^t\) is the transpose of \(H_\pm(\nu)\).

Next, by substituting Eq. (3.21) into Eq. (3.38) and using Eq. (3.35), one concludes that Eq. (3.21) can be cast into the following equivalent forms:

\[
\vec{\phi}_\pm(j, n) = e^{i\theta} R(\pm \theta) [A(\nu, \theta)]^n \vec{a}, \quad (j, n) \in \Omega
\]

According to Eqs. (3.37), \(R(\mp \theta)R(\pm \theta) = I\). Thus Eq. (3.41) can be recast as

\[
\vec{\phi}_\pm(j, n) = e^{i\theta} [A_\pm(\nu, \theta)]^n \vec{a}_\pm, \quad (j, n) \in \Omega
\]

where

\[
A_\pm(\nu, \theta) \overset{\text{def}}{=} R(\pm \theta)A(\nu, \theta)R(\mp \theta) = R(\pm \theta)A(\nu, \theta) [R(\pm \theta)]^{-1}
\]

and

\[
\vec{a}_\pm \overset{\text{def}}{=} R(\pm \theta)\vec{a}
\]

Note that: (i) Eq. (3.21) is a solution to Eq. (3.15) for all possible complex constant \(2 \times 1\) column matrices \(\vec{a}\) \(\ni\) Eq. (3.23); (ii) Eqs. (3.15), (3.39), and (3.40) are equivalent forms of the same leapfrog scheme Eq. (2.23d), and (iii) non-singularity of \(R(\pm \theta)\) implies that \(\vec{a}_\pm\) can be any complex constant \(2 \times 1\) column matrix just like \(\vec{a}\). As such, a comparison between Eqs. (3.21) and (3.42) reveals that \(A_\pm(\nu, \theta)\) defined by Eqs. (3.23), (3.35), and (3.43), i.e.,

\[
A_+(\nu, \theta) = \begin{pmatrix}
-2i\nu \sin \theta & e^{-i\theta} \\
e^{i\theta} & 0
\end{pmatrix} \quad \text{and} \quad A_-(\nu, \theta) = \begin{pmatrix}
-2i\nu \sin \theta & e^{i\theta} \\
e^{-i\theta} & 0
\end{pmatrix}, \quad -\infty < \nu, \theta < +\infty
\]

must be the unique solutions to Eqs. (3.39) and (3.40), respectively. This fact can also be verified directly by substituting Eqs. (3.42) into Eqs. (3.39) and (3.40), respectively. Thus \(A_+(\nu, \theta)\) and \(A_-(\nu, \theta)\) defined by Eq. (3.45) are the amplification matrices of the leapfrog scheme with respect to the forms Eqs. (3.39) and (3.40), respectively. Moreover, (i) because \(A(\nu, \theta), A_+(\nu, \theta), \text{ and } A_-(\nu, \theta)\) are related by similarity transformations Eq. (3.44), they have the same eigenvalues (i.e., amplification factors) \(\lambda_+(\nu, \theta)\) and \(\lambda_-(\nu, \theta)\); and (ii) by using Eqs. (3.45), one can show that Eqs. (3.24)–(3.26) are still valid if \(A(\nu, \pm \theta)\) is replaced by \(A_+(\nu, \pm \theta)\) or \(A_-(\nu, \pm \theta)\).

Next we consider the \(a\) scheme. Let

\[
\vec{q}(j, n) \overset{\text{def}}{=} \begin{pmatrix}
\vec{u}_n^j \\
\vec{u}_{n+1}^j
\end{pmatrix}, \quad (j, n) \in \Omega
\]

and

\[
Q_+(\nu) = \frac{1}{2} \begin{pmatrix}
1 + \nu & 1 - \nu^2 \\
-1 & -(1 - \nu)
\end{pmatrix}, \quad Q_-(\nu) = \frac{1}{2} \begin{pmatrix}
1 - \nu & -(1 - \nu^2) \\
1 & -(1 + \nu)
\end{pmatrix}
\]

Then the forward marching form Eqs. (2.12d) and (2.13d) of the \(a\) scheme can be cast into the matrix form:

\[
\vec{q}(j, n) = Q_-(\nu)\vec{q}(j + 1, n - 1) + Q_+(\nu)\vec{q}(j - 1, n - 1), \quad (j, n) \in \Omega
\]

Let \(G(\nu, \theta)\) be a \(2 \times 2\) nonsingular complex matrix function of \(\nu\) and \(\theta\) such that

\[
\vec{q}(j, n) = e^{i\theta} [G(\nu, \theta)]^n \vec{b}, \quad (j, n) \in \Omega
\]
is a solution to Eq. (3.48) for all possible complex constant $2 \times 1$ column matrices $\bar{b}$. By substituting Eq. (3.49) into Eq. (3.48), one has

$$
[G(\nu, \theta) - e^{i\theta} Q_-(\nu) - e^{-i\theta} Q_+(\nu)] [G(\nu, \theta)]^n \bar{b} = 0, \quad n = 0, \pm 1, \pm 2, \ldots (3.50)
$$

Because (i) $[G(\nu, \theta)]^0 = I$, and (ii) $\bar{b}$ can be any complex constant $2 \times 1$ column matrix, with the aid of Eq. (3.47), one concludes that Eq. (3.50) holds.

$$
G(\nu, \theta) = e^{i\theta} Q_-(\nu) + e^{-i\theta} Q_+(\nu) = \begin{pmatrix} \cos \theta - i\nu \sin \theta & -i(1 - \nu^2) \sin \theta \\ i \sin \theta & -(\cos \theta + i\nu \sin \theta) \end{pmatrix}, \quad -\infty < \nu, \theta < +\infty (3.51)
$$

According to Eq. (3.51), $\det[G(\nu, \theta)] \equiv -1$. Thus $G(\nu, \theta)$ is indeed nonsingular over its domain. By definition, $G(\nu, \theta)$ is referred to as the amplification matrix of the forward marching form of the $a$ scheme. It can be shown that the eigenvalues of $G(\nu, \theta)$ are $\lambda_+(\nu, \theta)$ and $\lambda_-(\nu, \theta)$ defined in Eq. (3.31), i.e., the amplification factors of the forward marching form of the $a$ scheme are identical to those of the leapfrog scheme.

Next let

$$
U \overset{\text{def}}{=} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} (3.52)
$$

and observe that

$$
U = U^{-1} (3.53)
$$

Thus Eq. (3.51) implies that

$$
G(\nu, -\theta)UG(\nu, \theta)U^{-1} = I, \quad -\infty < \nu, \theta < +\infty (3.54)
$$

and

$$
\overline{G(\nu, \theta)} = G(\nu, -\theta), \quad -\infty < \nu, \theta < +\infty (3.55)
$$

Combining Eqs. (3.54) and (3.55), one has

$$
\overline{G(\nu, \theta)} U G(\nu, \theta) U^{-1} = I, \quad -\infty < \nu, \theta < +\infty (3.56)
$$

Eqs. (3.54)–(3.56) are the current counterparts to Eqs. (3.24)–(3.26), respectively.

Eqs. (3.26) and (3.56) are identical in their forms. Thus, without using the explicit forms of $\lambda_+(\nu, \theta)$ and $\lambda_-(\nu, \theta)$, the same argument that was given following Eq. (3.26) can be invoked to show that, for any given $\nu$, the forward marching form of the $a$ scheme must be neutrally stable if it is stable. In fact, by using Eqs. (3.51) and (3.31) along with a rigorous definition of stability given in [64], it can be shown that the forward marching form of the $a$ scheme is (i) neutrally stable if $\nu^2 < 1$; (ii) linearly unstable when $\nu^2 = 1$; and (iii) exponentially unstable if $\nu^2 > 1$. Note that the linear instability at $\nu^2 = 1$ is due to the fact that $G(\nu, \theta)$ becomes defective when $\nu^2 = \sin^2 \theta = 1$.

At this juncture, note that Eqs. (3.25) and (3.55), like Eq. (3.7), follow directly from a rather common property that Eqs. (3.15) and (3.48) are linear homogeneous systems of equations with real constant coefficients. Thus, one concludes that Eqs. (3.6), (3.24), and (3.54) are the deciding contributing factors to the non-dissipativeness of the Wendroff, Crank-Nicolson, leapfrog and $a$ schemes. In Sec. 4, we will show that these latter equations are the natural results of a unique space-time invariance property shared by the above non-dissipative schemes. To pave the way, this section is ended with the following remarks:

(a) The leapfrog scheme Eq. (2.23d) can also be cast into a two-level backward marching matrix form, i.e.,

$$
\bar{\psi}(j, n) = P \bar{\psi}(j, n + 1) + H(\nu) \left[ \bar{\psi}(j + 1, n + 1) - \bar{\psi}(j - 1, n + 1) \right], \quad (j, n) \in \Omega (3.57)
$$

where

$$
\bar{\psi}(j, n) \overset{\text{def}}{=} \begin{pmatrix} u_{j+1}^{n-1} \\ u_j^n \end{pmatrix} = P \bar{\psi}(j, n), \quad (j, n) \in \Omega (3.58)
$$
(b) The backward marching form Eqs. (2.16d) and (2.17d) of the $a$ scheme can be cast into the matrix form:

$$\vec{q}(j, n) = \hat{Q}_+(\nu)\vec{q}(j + 1, n + 1) + \hat{Q}_-(\nu)\vec{q}(j - 1, n + 1), \quad (j, n) \in \Omega$$

(3.59)

where

$$\hat{Q}_+(\nu) \overset{\text{def}}{=} \frac{1}{2} \begin{pmatrix} 1 + \nu & -(1 - \nu^2) \\ 1 & -(1 + \nu) \end{pmatrix} = UQ_+(\nu)U^{-1}$$

(3.60)

and

$$\hat{Q}_-(\nu) \overset{\text{def}}{=} \frac{1}{2} \begin{pmatrix} 1 - \nu & 1 - \nu^2 \\ -1 & -(1 + \nu) \end{pmatrix} = UQ_-(\nu)U^{-1}$$

(3.61)

4. Space-time inversion invariance and non-dissipatedness

Note that Eq. (1.1) is equivalent to

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0$$

(4.1)

where

$$x' \overset{\text{def}}{=} 2x_c - x \quad \text{and} \quad t' \overset{\text{def}}{=} 2t_c - t$$

(4.2)

with $x_c$ and $t_c$ being any pair of given real constants. Thus Eq. (1.1) is mapped into the same equation under the mapping

$$(x, t) \leftrightarrow (x', t')$$

(4.3)

Because Eq. (4.2) $\leftrightarrow x + x' = 2x_c$ and $t + t' = 2t_c$, the transformation defined in Eq. (4.2) represents a space-time inversion (i.e., a spatial reflection followed by a time reversal) with respect to the point $(x_c, t_c)$. As such Eq. (1.1) is said to possess the space-time inversion (STI) invariance.

As a result of the above invariance property, solution to Eq. (1.1) also possesses a similar invariant property. Let

$$u = u_o(x, t)$$

(4.4)

be a solution to Eq. (1.1). Then because (i) $u = u_o(x, t)$ is a solution to Eq. (1.1) $\leftrightarrow u = u_o(x', t')$ is a solution to Eq. (4.1), and (ii) Eqs. (1.1) and (4.1) are equivalent if $x'$ and $t'$ are related to $x$ and $t$ by Eq. (4.2), one concludes that

$$u = u_o(2x_c - x, 2t_c - t)$$

(4.5)

must also be a solution of Eq. (1.1). In fact, the general solution to Eq. (1.1) is given by

$$u = F(x - at)$$

(4.6)

where $F(s)$ is an arbitrary differentiable function of $s$. It is easy to see that $u = F((2x_c - x) - a(2t_c - t))$ is also a solution to Eq. (1.1)

Numerical schemes generally are constructed without considering the invariant properties of the physical equations they model. Thus numerical solutions generally do not possess the same invariant properties of physical solutions. As an example, the general solution Eq. (4.6) represents non-dissipative data propagation along the characteristic lines defined by $dx/dt = a$. However, numerical solutions generated by dissipative analogies of Eq. (1.1) are dissipative. Using the numerical analogies of Eq. (1.1) defined earlier as examples, it will be shown in this section that non-dissipatedness of a scheme is deeply related to its STI invariance property.

Let $j_c$ and $n_c$ denote any pair of given fixed integers. Then the mapping

$$(j, n) \leftrightarrow (2j_c - j, 2n_c - n), \quad (j, n) \in \Omega$$

(4.7)

represents a space-time inversion mapping of mesh points $(j, n) \in \Omega$ with respect to $(j_c, n_c)$. As an example, under the mapping Eq. (4.7) with $j_c = n_c = 0$, points $A$, $B$, $C$, and $D$ depicted in Fig. 4 map onto points
$A'$, $B'$, $C'$, and $D'$, respectively, and vice versa. Thus a system of equations, such as Eq. (3.1), is said to be STI invariant if the system maps onto itself under the one-to-one mapping

$$u_j^n \mapsto u_{2j-c-n}^{2n_c-n}, \quad (j, n) \in \Omega$$

(4.8)

In other words, a STI invariant system is a system which will transform back to the original system if, for each $(j, n) \in \Omega$, $u_j^n$ is replaced by $u_{2j-c-n}$ where $j_c$ and $n_c$ are any pair of given fixed integers.

Let $j' \overset{\text{def}}{=} j - j_c$ and $n' \overset{\text{def}}{=} n - n_c$. Then (i) $j = j_c$ and $n = n_c \Leftrightarrow j' = n' = 0$ and (ii) $(j, n) \in \Omega \Leftrightarrow (j', n') \in \Omega$. As a result, without any loss of generality, in the following developments we can and will consider only the special case of Eq. (4.8) with $j_c = n_c = 0$, i.e.,

$$u_j^n \mapsto u_{-j}^{-n}, \quad (j, n) \in \Omega$$

(4.9)

First we shall establish the STI invariance of the Wendroff scheme. Under the mapping Eq. (4.9), Eq. (3.1) is mapped onto

$$(1 + \nu)\left(u_{-(j+1)}^{-(n+1)} - u_{-j}^{-n}\right) + (1 - \nu)\left(u_{-j}^{-n} - u_{-(j+1)}^{-(n+1)}\right) = 0, \quad (j, n) \in \Omega$$

(4.10)

(note: Eq. (4.10) is not the equation that we would obtain from Eq. (3.1) by simply replacing the symbols “$j$” and “$n$” in it with $-j$ and $-n$, respectively). Let

$$j^* \overset{\text{def}}{=} -(j + 1) \quad \text{and} \quad n^* \overset{\text{def}}{=} -(n + 1), \quad (j, n) \in \Omega$$

(4.11)

Then, with the aid of Eq. (4.11) and the fact that $(j^*, n^*) \in \Omega \Leftrightarrow (j, n) \in \Omega$, Eq. (4.10) can be rearranged as

$$(1 + \nu)\left(u_{j^*+1}^{n^*+1} - u_{j^*}^{n^*}\right) + (1 - \nu)\left(u_{j^*}^{n^*} - u_{j^*+1}^{n^*+1}\right) = 0, \quad (j^*, n^*) \in \Omega$$

(4.12)

By comparing Eq. (4.12) with Eq. (3.1), one can see that the system of equations defined by Eq. (4.12) (and therefore that defined by Eq. (4.10)) is identical to that defined by Eq. (3.1). QED.

To gain a deeper insight about STI invariance, consider the two component equations of Eq. (3.1) with $(j, n) = (1, 1)$ and $(j, n) = (-2, -2)$. They are

$$(1 + \nu)(u_2^1 - u_1^1) + (1 - \nu)(u_1^2 - u_2^1) = 0$$

(4.13)

and

$$(1 + \nu)(u_{-2}^{-1} - u_{-1}^{-1}) + (1 - \nu)(u_{-1}^{-2} - u_{-2}^{-1}) = 0$$

(4.14)

respectively. Obviously Eq. (4.14) is the image of Eq. (4.13) under the mapping Eq. (4.9) and vice versa. This represents a concrete example of the fact that the image of a component equation of Eq. (3.1) under the mapping Eq. (4.9) is also a component equation of Eq. (3.1), and vice versa.

Moreover, (i) the mesh points involved in Eq. (4.13) are points $A$, $B$, $C$, and $D$ depicted in Fig. 4 while those involved in Eq. (4.14) are points $A'$, $B'$, $C'$, and $D'$; and (ii) under the mapping Eq. (4.7) with $j_c = n_c = 0$, i.e.,

$$(j, n) \leftrightarrow (-j, -n), \quad (j, n) \in \Omega$$

(4.15)

points $A'$, $B'$, $C'$, and $D'$ are the images of $A$, $B$, $C$, and $D$, respectively, and vice versa. Because, for a linear constant-coefficient solver of Eq. (1.1) with one mesh variable and one equation per mesh point, the stencil of all its component equations have the same configuration, the above observation implies that such a scheme cannot be STI invariant unless it meets the requirement that its stencil configuration is unchanged under the space-time inversion mapping Eq. (4.9). Obviously this requirement is met by the non-dissipative Wendroff, Crank-Nicolson, and leapfrog schemes.
On the other hand, any two-level explicit linear constant-coefficient (forward marching) solver of Eq. (1.1) with one mesh variable and one equation per mesh point can never meet the above requirement. By its explicit nature, the stencil of such a scheme can contain only one point at the upper time level. Whereas stability and other considerations demand that the stencil contains at least two points at the lower time level. As such, the configuration of its stencil will change under the mapping Eq. (4.15). In turn, this implies that such a scheme cannot be STI invariant. In view of the strong tie to be established between STI invariance and non-dissipatedness, this explains why none of such two-level explicit schemes is non-dissipative.

Next we will establish a relation between STI invariance and non-dissipatedness. Let \( s(j, n), (j, n) \in \Omega \), be a function of \((j, n)\) such that

\[
u^n = s(j, n), \quad (j, n) \in \Omega \quad (4.16)
\]

is a solution to Eq. (3.1). Then, by substituting Eq. (4.16) into Eq. (3.1), one obtains the system of identities:

\[
(1 + \nu) (s(j + 1, n + 1) - s(j, n)) + (1 - \nu) (s(j, n + 1) - s(j + 1, n)) \equiv 0, \quad (j, n) \in \Omega \quad (4.17)
\]

On the other hand, because Eq. (4.10) \(\iff\) Eq. (3.1), by substituting Eq. (4.16) into Eq. (4.10), one concludes that the system of identities

\[
(1 + \nu) (s(-j + 1, -(n + 1)) - s(-j, -n)) + (1 - \nu) (s(-j, -(n + 1)) - s((-j + 1), -n)) \equiv 0, \quad (j, n) \in \Omega \quad (4.18)
\]

is identical to that defined in Eq. (4.17). As a concrete example, the component identities in Eq. (4.17) with \((j, n) = (1, 1)\) and \((j, n) = (-2, -2)\) are

\[
(1 + \nu) (s(2, 2) - s(1, 1)) + (1 - \nu) (s(1, 2) - s(2, 1)) \equiv 0 \quad (4.19)
\]

and

\[
(1 + \nu) (s(-2, -2) - s(-1, -1)) + (1 - \nu) (s(-1, -2) - s(-2, -1)) \equiv 0 \quad (4.20)
\]

respectively. Whereas the component identities in Eq. (4.18) with \((j, n) = (1, 1)\) and \((j, n) = (-2, -2)\) are Eqs. (4.20) and (4.19), respectively.

Moreover, a direct comparison of Eqs. (3.1) and (4.18) reveals that

\[
u_j^n = s(-j, -n), \quad (j, n) \in \Omega \quad (4.21)
\]

represents another solution to the system of identities Eq. (3.1), i.e., Eq. (4.16) is a solution to Eq. (3.1) \(\iff\) Eq. (4.21) is a solution to Eq. (3.1). As an example, by substituting Eq. (4.16) into Eq. (4.13) and (4.14), one has Eqs. (4.19) and (4.20), respectively. Whereas, by substituting Eq. (4.21) into Eqs. (4.13) and (4.14), one has Eqs. (4.20) and (4.19), respectively. In the following, the key property Eq. (3.6) of the Wendroff scheme will be derived using the fact just established above without using the explicit form of \(g_w(\nu, \theta)\) given in Eq. (3.5).

Let Eq. (3.2) be a solution to Eq. (3.1) for all \((\nu, \theta) \in D(g_w)\). Then \(g(\nu, \theta)\) must take the unique form defined in Eq. (3.5). Moreover, the fact established in the last paragraph implies that, for all \((\nu, \theta) \in D(g_w)\),

\[
u_j^n = [g(\nu, \theta)]^{-n} e^{-ij\theta} \equiv [(g(\nu, \theta))^{-1}]^n e^{-ij\theta}, \quad (j, n) \in \Omega \quad (4.22)
\]

must also be a solution to Eq. (3.1). Because \((\nu, \theta) \in D(g_w) \iff (\nu, -\theta) \in D(g_w)\), by replacing \(\theta\) in Eq. (4.22) with \(-\theta\), one concludes that, for all \((\nu, \theta) \in D(g_w)\),

\[
u_j^n = [(g(\nu, -\theta))^{-1}]^n e^{ij\theta}, \quad (j, n) \in \Omega \quad (4.23)
\]

must also be a solution of Eq. (3.1). By comparing Eq. (3.2) and (4.23) and recalling the uniqueness of \(g(\nu, \theta)\), one concludes that \(g(\nu, \theta) = (g(\nu, -\theta))^{-1}\), i.e., Eq. (3.6). QED.
Using arguments similar to those just presented, one can easily show that the Crank-Nicolson scheme is also STI invariant and possesses all other properties established above for the Wendroff scheme. In particular, it can be shown that the fact that \( g(\nu, \theta) = g_{\text{STI}}(\nu, \theta) \) satisfies Eq. (3.6) is a result of its STI invariance.

Similarly, one can easily establish STI invariance for the leapfrog scheme Eq. (2.23d) and show that Eq. (4.16) is a solution to Eq. (2.23d) \( \iff \) Eq. (4.21) is a solution to the same equation. However, as will be shown later, it is trickier to establish the key relation Eq. (3.24).

As a preliminary, first we offer the following remarks:

(a) Let Eq. (4.16) be a solution to Eq. (2.23d). Then Eq. (3.20) implies that

\[
\tilde{\phi}(j, n) = \begin{pmatrix} s(j, n) \\ s(j, n-1) \end{pmatrix}, \quad (j, n) \in \Omega
\]

is a solution to Eq. (3.15), an equivalent form of Eq. (2.23d). Moreover, because Eq. (4.16) is a solution to Eq. (2.23d) \( \iff \) Eq. (4.21) is a solution to Eq. (2.23d), one concludes that

\[
\tilde{\phi}(j, n) = \begin{pmatrix} s(-j, n) \\ s(-j, 1-n) \end{pmatrix} \equiv P \begin{pmatrix} s(-j, 1-n) \\ s(-j, (1-n) - 1) \end{pmatrix}, \quad (j, n) \in \Omega
\]

must also be a solution to Eq. (3.15).

(b) Let \( j_o \) and \( n_o \) be arbitrary given fixed integers. Then \((j + j_o, n + n_o) \in \Omega \iff (j, n) \in \Omega\). Thus, by direct substitution, one can show that, for any real \( \nu \) and \( \theta \), Eq. (3.21) is a solution to Eq. (3.15) \( \iff \)

\[
\tilde{\phi}(j, n) = e^{ij + j_o} [A(\nu, \theta)]^{n + n_o} \tilde{a}, \quad (j, n) \in \Omega
\]

is a solution to the same equation.

For any real \( \nu \) and \( \theta \), let Eq. (3.21) be a solution to Eq. (3.15). Then, as was shown earlier, \( A(\nu, \theta) \) must be in the unique form defined in Eq. (3.23). Also we can assume that:

\[
\begin{pmatrix} s(j, n) \\ s(j, n-1) \end{pmatrix} \equiv e^{ij\theta}[A(\nu, \theta)]^n \tilde{a}, \quad (j, n) \in \Omega
\]

By replacing the indices \( j \) and \( n \) in Eq. (4.27) with \(-j\) and \( 1-n \), respectively and using the fact that \((j, n) \in \Omega \iff (-j, 1-n) \in \Omega\), one concludes that Eq. (4.27) \( \iff \)

\[
\begin{pmatrix} s(-j, 1-n) \\ s(-j, (1-n) - 1) \end{pmatrix} \equiv e^{-ij\theta}[A(\nu, \theta)]^{-n} \tilde{a}, \quad (j, n) \in \Omega
\]

By using Eq. (4.28) and the above remark (a), one concludes that, for any real \( \nu \) and \( \theta \),

\[
\tilde{\phi}(j, n) = Pe^{-ij\theta}[A(\nu, \theta)]^{-1} \tilde{a} \equiv e^{-ij\theta}P[A(\nu, \theta)]^{-1} \tilde{a}, \quad (j, n) \in \Omega
\]

must also be a solution to Eq. (3.15).

To cast Eq. (4.29) into another form, let

\[
\hat{A}(\nu, \theta) \overset{\text{def}}{=} PA(\nu, \theta)P^{-1}, \quad -\infty < \nu, \theta < +\infty
\]

Then, with the aid of the relation

\[
[PA(\nu, \theta)P^{-1}]^{-1} \equiv P[A(\nu, \theta)]^{-1}P^{-1}, \quad -\infty < \nu, \theta < +\infty
\]

one can show that

\[
P[A(\nu, \theta)]^{-1-n} = \left[ \hat{A}(\nu, \theta) \right]^{-1-n} P, \quad n = 0, \pm 1, \pm 2, \ldots; -\infty < \nu, \theta < +\infty
\]
By using Eq. (4.32), Eq. (4.29) can be expressed as

$$\phi(j, n) = e^{-ij\theta} \left[ \hat{A}(\nu, \theta) \right]^{1-n} P\hat{a}, \quad (j, n) \in \Omega$$  \hspace{1cm} (4.33)

Then, by replacing $\theta$ in Eq. (4.33) with $-\theta$, it is seen that, for any real $\nu$ and $\theta$,

$$\phi(j, n) = e^{ij\theta} \left[ \hat{A}(\nu, -\theta) \right]^{1-n} P\hat{a}, \quad (j, n) \in \Omega$$  \hspace{1cm} (4.34)

must be a solution to Eq. (3.15). In turn, by using the above remark (b), One concludes that

$$\phi(j, n) = e^{ij\theta} \left[ \hat{A}(\nu, -\theta) \right]^{-n} P\hat{a} \equiv e^{ij\theta} \left[ (\hat{A}(\nu, -\theta))^{-1} \right]^{n} P\hat{a}, \quad (j, n) \in \Omega$$  \hspace{1cm} (4.35)

must be a solution to Eq. (3.15). Because $P$ is nonsingular, $P\hat{a}$ can be any complex constant $2 \times 1$ column matrix just like $\hat{a}$. Thus, by (i) comparing Eq. (4.35) with Eq. (3.21), (ii) using Eq. (4.30), and (iii) recalling the uniqueness of $A(\nu, \theta)$, one has

$$A(\nu, \theta) = [PA(\nu, -\theta)P^{-1}]^{-1} \equiv P[A(\nu, -\theta)]^{-1}P^{-1}, \quad -\infty < \nu, \theta < +\infty$$  \hspace{1cm} (4.36)

which, by using Eq. (3.14), can be shown to be equivalent to Eq. (3.24). QED.

Next, we shall define and establish STI invariance for the $a$ scheme. Note that: (i) $\partial u/\partial x \leftrightarrow -\partial u/\partial x$ under the mapping $x \leftrightarrow -x$; and (ii) $(u_x)_n$ is the numerical analogy of the value of $\partial u/\partial x$ at the mesh point $(j, n)$. Because of the above observations and the definition Eq. (2.9), the a scheme is said to be STI invariant if and only if the associated system of equations maps onto itself under the one-to-one mapping

$$\left( \begin{array}{c} u^n_j \\ (u_x)_j^n \end{array} \right) \leftrightarrow \left( \begin{array}{c} u^{-n}_j \\ -(u_x)^{-n}_j \end{array} \right), \quad (j, n) \in \Omega$$  \hspace{1cm} (4.37)

By using Eqs. (3.46) and (3.52), Eq. (4.37) can be expressed as

$$\hat{q}(j, n) \leftrightarrow U\hat{q}(-j, -n), \quad (j, n) \in \Omega$$  \hspace{1cm} (4.38)

Under the mapping Eq. (4.38), the forward marching matrix form Eq. (3.48) is mapped onto

$$U\hat{q}(-j, -n) = Q_- (\nu)U\hat{q}(-j - 1, -n + 1) + Q_+ (\nu)U\hat{q}(-j + 1, -n + 1), \quad (j, n) \in \Omega$$  \hspace{1cm} (4.39)

Let Eq. (4.39) be multiplied from left by the nonsingular matrix $U$. Then, with the aid of Eqs. (3.53), (3.60) and (3.61), one concludes that Eq. (4.39) is equivalent to

$$\hat{q}(-j, -n) = \hat{Q}_+(\nu)\hat{q}(-j + 1, -n + 1) + \hat{Q}_-(\nu)\hat{q}(-j - 1, -n + 1), \quad (j, n) \in \Omega$$  \hspace{1cm} (4.40)

Moreover, because (i) $(j, n) \in \Omega \leftrightarrow (-j, -n) \in \Omega$, and (ii) by replacing the indices $j$ and $n$ with $-j$ and $-n$, respectively, Eq. (3.59) will turn into Eq. (4.40), one concludes that the system of equations defined by Eq. (4.40) is identical to that defined by Eq. (3.59), another equivalent form of the a scheme. Thus the a scheme is STI invariant. QED.

Next, we will establish Eq. (3.54) from the STI invariant property of the a scheme. For any real $\nu$ and $\theta$, let Eq. (3.49) be a solution to Eq. (3.48). Then the STI invariance of the a scheme and Eq. (4.38) imply that, for any real $\nu$ and $\theta$,

$$\hat{q}(j, n) = U e^{-ij\theta} [G(\nu, \theta)]^{-n} \hat{b} \equiv e^{-ij\theta} U [G(\nu, \theta)]^{-n} \hat{b}, \quad (j, n) \in \Omega$$  \hspace{1cm} (4.41)
must also be a solution to Eq. (3.48). By using a line of arguments similar to that used to obtain Eq. (4.36) from Eqs. (3.21) and (4.29), Eq. (3.49) and (4.41) can be used to show that

\[ G(\nu, \theta) = \frac{UG(\nu, -\theta)U^{-1}}{U - 1} \equiv U[G(\nu, -\theta)]^{-1}U^{-1}, \quad -\infty < \nu, \theta < +\infty \]

(4.42)

By using Eq. (3.53), it can be shown that Eq. (4.42) is equivalent to Eq. (3.54). QED.

This section is concluded with the following remarks (see Fig. 5):

(a) Let \((j, n) = (2, 2)\). Then \((j, n)\) is point \(A\). According to Eqs. (2.10d), (2.11d), (2.14d), and (2.15d), the basic stencils of the \(a\) scheme which involve point \(A\) are formed by the following four pairs of mesh points:

(i) \(A\) and \(B\),
(ii) \(A\) and \(C\),
(iii) \(A\) and \(D\), and
(iv) \(A\) and \(E\).

The forward marching relations Eqs. (2.12d) and (2.13d) are derived using the basic relations involving the mesh variables of pairs (iii) and (iv), while the backward marching relations Eqs. (2.16d) and (2.17d) are derived using those involving the mesh variables of pairs (i) and (ii).

(b) The images of points \(A\), \(B\), \(C\), and \(D\) under the mapping Eq. (4.15) are points \(A'\), \(B'\), \(C'\), and \(D'\), respectively. Obviously, the configuration of the image of any stencil referred to in the above item (a) is identical to that of the original stencil.

(c) From the observations given above, it becomes clear that a two-level explicit solver of Eq. (1.1) cannot be STI invariant unless it has at least two independent mesh variables assigned to each mesh point.

5. Conclusions and Discussions

A thorough and rigorous discussion of the core ideas and the core \(a\) scheme of the CESE method has been presented. These simple core ideas provide the CESE method a foundation which is solid in physics and yet simple in mathematics. In fact, it was clearly demonstrated that all the enabling and simplifying features of the CESE method are rooted in three key ideas, i.e., (i) conservation laws be enforced in their natural space-time unity forms; (ii) the simplest stencil be used; and (iii) a solver be built from a non-dissipative core scheme.

Moreover, it has been shown that the \(a\) scheme possesses some rather intriguing properties such as:

(i) its two independent mesh variables \(u_n^j\) and \((u_x)_n^j\) separately satisfy two decoupled three-level leapfrog schemes Eqs. (2.23d) and (2.24d), and
(ii) it shares with the leapfrog scheme the same amplification factors, even though the \(a\) scheme and the leapfrog scheme have completely different origins and structures. These properties are tributes to the power of beautiful and deep physical ideas underlying its development.

In addition, we also establish rigorously the deep relation between STI (i.e., PI) invariance and non-dissipatedness. It represents a good example of universal applicability of a fundamental physics idea.

Perhaps the biggest surprise that occurred during the CESE development is the realization that a two-level explicit solver cannot be STI invariant if, corresponding to one physical conservation law, it has only one mesh variable and one equation per mesh point. It implies strongly that a discrete system may behave fundamentally different from a continuous system.

References


71. Other CE/SE references are posted on: http://www.grc.nasa.gov/www/microbus.


APPENDIX

implicit real*8(a-h,o-z)
parameter (nxd=2000)
dimension u(nxd), un(nxd), ux(nxd), uxn(nxd), vo(nxd), v(nxd),*
v(n(nxd)), vxo(nxd), vx(nxd), vxn(nxd)

Program `lfa.for'--a combined dual `a' and leapfrog solver
for the pure advection equation. The computation is carried out
assuming a sine-wave initial data and periodic boundary
conditions. The sine-wave is of unit wavelength. It can be shown
that both exact and numerical solutions must be spatially
periodic with unit wavelength too.

u, un, ux, uxn---array associated with the $a$ scheme.
vo, v, vn, vxo, vx, vxn---array associated with the leapfrog
scheme.

ux, uxn, vxo, vx, and vxn represent the numerical analogues of
normalized spatial derivatives during time marching. However,
the output contains the non-normalized values.

it = no. of time-marching steps.
k = no. of spatial intervals per unit length.
cn = Courant number.
a = the advection speed.

it = 1000
k = 101
cn = 0.5d0
a = 1.d0

dx = 1.d0/dfloat(k)
dt = cn*dx/a
hdx = dx/2.d0
t = dt*dfloat(it)
k1 = k + 1
k2 = k + 2
pi = 3.1415926535897932d0
tp = pi*2.d0
pdx = pi*dx
cns = (1.d0 - cn**2)/2.d0
cnp = (1.d0 + cn)/2.d0
cnm = (1.d0 - cn)/2.d0

c
open (unit=8,file='for008')
write (8,10) it,k,dt
write (8,20) a,cn
write (8,30) t
do 100 j = 1,k2
  tpx = tp*dfloat(j-1)*dx
  u(j) = dsin(tpx)
  ux(j) = pdx*dcos(tpx)
  v(j) = u(j)
  vx(j) = ux(j)
100  continue
do 400 i = 1,it
  do 200 j = 2,k1
    un(j) = cnm*u(j+1) + cnp*u(j-1) + cns*(ux(j-1) - ux(j+1))
    uxn(j) = (u(j+1) - u(j-1))/2.d0 - cnp*ux(j+1) - cnm*ux(j-1)
    if (i.eq.1) goto 150
    vn(j) = vo(j) - cn*(v(j+1) - v(j-1))
    vxn(j) = vxo(j) - cn*(vx(j+1) - vx(j-1))
  goto 200

150  continue
do 200 j = 2,k1
  un(j) = cnm*u(j+1) + cnp*u(j-1) + cns*(ux(j-1) - ux(j+1))
  uxn(j) = (u(j+1) - u(j-1))/2.d0 - cnp*ux(j+1) - cnm*ux(j-1)
if (i.eq.1) goto 150
  vn(j) = vo(j) - cn*(v(j+1) - v(j-1))
  vxn(j) = vxo(j) - cn*(vx(j+1) - vx(j-1))

150  continue

goto 200
150 vn(j) = un(j)
    vxn(j) = uxn(j)
200 continue
    do 250 j = 2,k1
    vo(j) = v(j)
    vxo(j) = vx(j)
250 continue
    do 300 j = 2,k1
    u(j) = un(j)
    ux(j) = uxn(j)
    v(j) = vn(j)
    vx(j) = vxn(j)
300 continue
    u(k2) = u(2)
    ux(k2) = ux(2)
    v(k2) = v(2)
    vx(k2) = vx(2)
    u(1) = u(k1)
    ux(1) = ux(k1)
    v(1) = v(k1)
    vx(1) = vx(k1)
400 continue
    tpat = tp*a*t
    do 500 j = 1,k1
    x = dfloat(j-1)*dx
    tpx = tp*x
    ue = dsin(tpx-tpat)
    uxe = tp*dcos(tpx-tpat)
    eu = u(j) - ue
    ux(j) = ux(j)/hdx
    eu = u(j) - uxe
    ev = v(j) - ue
    vx(j) = vx(j)/hdx
    evx = vx(j) - uxe
    write (8,40)
    write (8,50) x,ue,ux
    write (8,60) u(j), eu, ux(j), eux
    write (8,70) v(j), ev, vx(j), evx
500 continue
    close (unit=8)
10 format (’it =’,i8,’ k =’,i8,’ dt =’,g14.7)
20 format (’a =’,g14.7,’ CFL =’,g14.7)
30 format (’t =’,g14.7)
40 format (’***************************************************’)
50 format (’x =’,g14.7,’ ue =’,g14.7,’ uxe =’,g14.7)
60 format (’u =’,g14.7,’ eu =’,g14.7,’ ux =’,g14.7,’ eux =’,g14.7)
70 format (’v =’,g14.7,’ ev =’,g14.7,’ vx =’,g14.7,’ evx =’,g14.7)
stop
eend
Figure 1.—A surface element on the boundary $S(V)$ of an arbitrary space-time volume $V$.

Figure 2.—The SEs, BCEs, and CCEs.

2(a).—The space-time mesh.

2(b).—$SE_{(j,n)}$.

2(c).—$CE_{(j,n)}$.

2(d).—$CE_{(j,n)}$.

2(e).—$CE_{(j,n)}$.

2(f).—$CE_{(j,n)}$.

2(g).—$CE_{(j,n)}$.

2(h).—The union of $CE_{(j-1,n+1)}$ and $CE_{(j+1,n+1)}$.

Figure 2.—The SEs, BCEs, and CCEs.
Figure 3.—FV(j,n+1).

Figure 4.—The stencil of Wendroff scheme and its STI image.

Figure 5.—The basic stencils of the $a$ scheme and their STI images.