SIAM J. NUMER. ANAL.
Vol. 32, No. 6, pp. 1839–1875, December 1995

007

FINITE DIFFERENCE CALCULUS INVARIANT STRUCTURE OF A CLASS OF ALGORITHMS FOR THE NONLINEAR KLEIN–GORDON EQUATION*

S. LI† AND L. VU-QUOC†

"Not all integrals of the motion, however, are of equal importance in mechanics. There are some whose constancy is of profound significance, deriving from the fundamental homogeneity and isotropy of space and time."
L. D. Landau and E. M. Lifshitz, *Mechanics*

Dedicated to Professor Karl. S. Pister on the occasion of his 70th birthday.

Abstract. In a previous work, the authors have presented a formalism for deriving systematically invariant, symmetric finite difference algorithms for nonlinear evolution differential equations that admit conserved quantities. This formalism is herein cast in the context of *exact* finite difference calculus. The algorithms obtained from the proposed formalism are shown to derive exactly from discrete scalar potential functions using finite difference calculus, in the same sense as that of the corresponding differential equation being derivable from its associated energy function (a conserved quantity). A clear ramification of this result is that the derived algorithms preserve certain discrete invariant quantities, which are the consistent counterpart of the invariant quantities in the continuous case. Results on the nonlinear stability of a class of algorithms that are derived using the proposed formalism, and that preserve energy or linear momentum, are discussed in the context of finite difference calculus. Some numerical experiments are presented to illustrate the conservation property of the proposed algorithms.

Key words. conservation laws, nonlinear Klein–Gordon equation, solitons, finite difference methods, nonlinear stability, finite difference calculus

AMS subject classifications. 35L65, 35L70, 39A11, 39A12, 65M06, 65M12

1. Introduction.

The practice of designing finite difference algorithms that possess a conserved quantity dates back to the late 1920s with the work of Courant, Friedrichs, and Lewy [1928]. The basic idea of earlier attempts is to construct special algorithms that are not only consistent with the differential equations, but also yield a *global* quadratic invariant form referred to as "energy." Although not all of these conserved quantities are truly energy in the physical sense, they are usually just some positive definite quantities. The primary motivation there is to devise a norm that can guarantee the global stability. This so-called "energy-conserving method" was mainly applied to certain linear mixed initial/boundary-value problems, and the results are well documented (see, e.g., Richtmyer and Morton [1967]).

There is a recent surge of interest in methods that preserve certain invariant quantities. A few examples are listed here: Hughes, Caughey, and Liu [1978] for nonlinear elastodynamics, Strauss and Vazquez [1978] for solving the nonlinear Klein–Gordon equation (NLKGE), Greenspan [1984] for general second-order ordinary differential equations, Feng [1984] and Feng and Qin [1986] for general linear Hamiltonian sys-

*Received by the editors July 13, 1992; accepted for publication (in revised form) March 28, 1994. This research was supported by a grant from the National Science Foundation and the Space Research Assistantship Program of the Florida Space Grant Consortium.

†Aerospace Engineering, Mechanics and Engineering Science, University of Florida, Gainesville, Florida 32611 (vql@cauchy.aero.ufl.edu). The first author was supported by a graduate research assistantship.

tems, Chin and Qin [1989] for solving the three-body problem, Huang [1991] for solving KdV equations, Fei and Vazquez [1991a] for the sine-Gordon equation, Simo and Wong [1991] for solving rigid body dynamics, and Simo, Rifai, and Fox [1992] for nonlinear dynamics of shell structures.

However, this resurgence of the "energy-conserving methods" is not simply a reminiscence of the classical energy method, but does carry a new trend of designing accurate numerical algorithms that inherit in the discrete sense certain conservation properties of the original differential equations. In other words, stability is not the primary concern here; instead, the main interest shifts to the development of algorithms that can preserve characteristic invariants of the original differential equations, such as energy, momenta, or some other conserved quantities, and even the Hamiltonian structure (Feng [1984]). This line of work is basically motivated by the fact that *in some areas, the ability to preserve some invariant properties of the original differential equation is a criterion to judge the success of a numerical simulation.*

Conservation laws play an important role in the theory of solitons (e.g., Whitham [1974], Lamb [1980], Dodd et al. [1982]). Among the important equations that possess an infinite number of conservation laws is the sine-Gordon equation, which is also found in many applications (e.g., Steudel [1975] and Barone, Esposito, and Magee [1971]). The conservation of energy is an essential property in an elastic collision of two solitons. In general, most numerical algorithms fail, however, to preserve the system energy. For instance, the energy potential is reported to fluctuate between 2% and 6% for a spatial increment of $h = 0.05$ and a grid ratio[1] of $\lambda = 1$ in a problem solved using the algorithm proposed by Ablowitz, Kruskal, and Ladik [1979], a very efficient and popular algorithm for solving the NLKGE. The necessity of employing algorithms that conserve a discrete energy is further elaborated in Fei and Vazquez [1991b], in which a nonconserving algorithm is shown to yield incorrect results, while conserving algorithms yield correct results.

Strauss and Vazquez [1978] proposed an "energy-conserving" algorithm for the NLKGE. Still, the problem remains incompletely solved, since the general formulational procedure and the relationship between the algorithm and the boundary conditions are not satisfactorily clear. Besides, the expression for the discrete energy given in Strauss and Vazquez [1978] is not absolutely positive definite, which implies that it lacks physical sense. These issues will be discussed thoroughly in this paper.[2]

In this formulation, we consider the problem in a more comprehensive manner as compared with the classical energy methods. Due to the theory of infinitesimal symmetry (Lie group) of differential operators, some evolution equations possess several, or even an infinity of, local conservation laws (see, e.g., Olver [1986]). By combining these evolution equations with certain boundary conditions, these local conservation laws can be converted into global conserved quantities (first integrals). The issue now is whether one can design *finite difference equations that are not only consistent with the original differential equations, but also themselves possess inherent algebraic invariant properties—i.e., discrete conservation laws.* Here, the finite difference equations are shown to derive from the discrete conservation law in the context of exact finite difference calculus, which essentially replicates the equivalence between the original differential equations and the local conservation laws in the continuous context.

[1] I.e., the ratio of time increment k over space increment h.

[2] Additional analysis of the Strauss and Vazquez [1978] algorithm with respect to long-time behavior of the solution and comparisons with three other algorithms, which include the algorithm proposed by Ablowitz, Kruskal, and Ladik [1979], are established in Jimenez and Vazquez [1990].

Unlike the classical energy methods, the discrete conserved quantity is not the only issue here; the invariant algebraic structure in the proposed approach is believed to carry richer information on the discrete evolution system. In the present paper, we are considering explicit and implicit algorithms for the general NLKGE; we refer to Fei and Vazquez [1991a] for explicit algorithms for the particular sine-Gordon equation.

Various invariant properties of the NLKGE in the continuous context are summarized in §2. In §3, we will review a general formalism to construct invariant-conserving algorithms, which hinges essentially on a particular symmetry property of finite difference calculus (Vu-Quoc and Li [1993]). Here, the novelty of the proposed formalism is shown to lie in the systematic use of exact finite difference calculus as a constructive tool to design algorithms that preserve certain algebraic invariants. In §4, three algorithms for the NLKGE are proposed following the derivation formalism of §3. These algorithms possess either a companion local discrete energy expression or a companion local discrete linear momentum expression; these expressions do have their counterparts in the continuous context. Using symmetry conditions, we derive the global discrete energy expression and global discrete momentum expression, respectively, and therefore the name "invariant-conserving algorithms." Nonlinear stability theorems are established as direct by-products of the algebraic conservation properties of the proposed algorithms. Moreover, we reexamine the Strauss and Vazquez [1978] algorithm to derive its algebraic invariant and its new global conserved quantity, which is, however, not strongly positive definite. Thus, no strong nonlinear stability statement can be made for the Strauss–Vazquez algorithm, which reconfirms the results of the stability analysis for this algorithm on the linear Klein–Gordon equation in Vu-Quoc and Li [1993]. Numerical results testifying to the workability of the proposed algorithms are presented in §5. Not only is the system's discrete energy preserved exactly, the system's linear and angular momenta are accurate up to 10 significant digits on a very coarse grid.

2. Conservation laws and invariants for NLKGE. Consider the function $U\colon \mathbb{R}\times\mathbb{R}_+ \rightarrow \mathbb{R}$. The Cauchy problem for the one-dimensional (1-D) NLKGE is given by[3]

$$\boxed{\begin{aligned} \frac{\partial U}{\partial t} &= V, \\ \frac{\partial V}{\partial t} &= U_{xx} - G'(U), \end{aligned}} \tag{2.1ab}$$

with the following initial and boundary conditions

$$U(x,0) = f(x) \quad \text{and} \quad V(x,0) = g(x), \tag{2.1c}$$

$$|V|, \quad |U_x| \longrightarrow 0 \quad \text{as} \quad |x| \longrightarrow \infty. \tag{2.1d}$$

2.1. Energy conservation law. Integrating a *local conservation energy law* for the NLKGE given by

$$\boxed{\begin{aligned} \frac{\partial U}{\partial t} &= V, \\ \frac{\partial}{\partial t}\left[\frac{1}{2}(V)^2 + \frac{1}{2}(U_x)^2 + G(U)\right] &= \frac{\partial}{\partial x}(U_x V), \end{aligned}} \tag{2.2ab}$$

[3]The subscript x (or t) designates partial differentiation with respect to x (or t).

and using the boundary conditions (2.1d), we obtain the conservation of the total energy $\mathbf{E}$

$$\boxed{\mathbf{E}(t) = \mathbf{E}(0) \quad \forall t \in \mathbb{R}_+,} \tag{2.3}$$

$$\mathbf{E}(t) := \int_{-\infty}^{+\infty} \mathcal{E}(x,t)dx, \quad \mathcal{E}(x,t) := \left[\frac{1}{2}(U_t)^2 + \frac{1}{2}(U_x)^2 + G(U)\right], \tag{2.4}$$

where $\mathcal{E}$ is called the energy density, and $\mathbf{E}(0)$ is evaluated using the initial conditions (2.1c).

Remark 2.1. The functions G, f, and g in (2.2b) and (2.1c) are assumed to be differentiable as much as required and well behaved such that $\mathbf{E}(0)$ is finite.

Remark 2.2. The Neumann-type boundary condition (2.1d) is more than enough to obtain (2.3); a less restrictive boundary condition, which is suitable for numerical computation, is[4]

$$\boxed{[U_x V]_{-\infty}^{+\infty} = 0.} \tag{2.5}$$

We refer to (2.5) as the symmetry condition.

2.2. Linear-momentum conservation law. Multiplying the NLKGE (2.1b) by U_x, one obtains another local conservation law (for linear momentum)

$$\boxed{\begin{aligned} &\frac{\partial U}{\partial t} = V, \\ &\frac{\partial}{\partial t}(U_x V) = \frac{\partial}{\partial x}\left[\frac{1}{2}(V)^2 + \frac{1}{2}(U_x)^2 - G(U)\right], \end{aligned}} \tag{2.6ab}$$

which then leads to the conservation of the global linear momentum $\mathbf{M}$

$$\boxed{\mathbf{M}(t) = \mathbf{M}(0) \quad \forall t \in \mathbb{R}_+,} \tag{2.7a}$$

$$\mathbf{M}(t) := \int_{-\infty}^{+\infty} \mathcal{M}(x,t)dx, \quad \mathcal{M} := U_t U_x, \tag{2.7b}$$

provided that the following boundary condition is satisfied:

$$\boxed{\mathcal{J}(x,t)\Big|_{x=-\infty}^{x=+\infty} = 0,} \tag{2.8}$$

where

$$\mathcal{J}(x,t) := \left[\frac{1}{2}(V)^2 + \frac{1}{2}(U_x)^2 - G(U)\right].$$

[4]See also Remark 4.2.

The initial linear momentum $\mathbf{M}(0)$ can be computed using the initial conditions (2.1c).

2.3. Angular-momentum conservation law. Using the above definition of $\mathcal{E}$, $\mathcal{M}$, and $\mathcal{J}$, one can verify that (2.1b) is equivalent to

$$\boxed{\frac{\partial}{\partial t}\{x\mathcal{E}(x,t) + t\mathcal{M}(x,t)\} - \frac{\partial}{\partial x}\{x\mathcal{M}(x,t) + t\mathcal{J}(x,t)\} = 0,} \tag{2.9}$$

which leads to the conservation of global angular momentum $\mathbf{A}$:

$$\boxed{\mathbf{A}(t) = \mathbf{A}(0) \quad \forall t \in \mathbb{R}_+,} \tag{2.10a}$$

$$\mathbf{A}(t) := \int_{-\infty}^{+\infty} \mathcal{A}(x,t)dx, \quad \mathcal{A}(x,t) := x\mathcal{E}(x,t) + t\mathcal{M}(x,t), \tag{2.10b}$$

provided that the following boundary condition is satisfied[5]

$$\boxed{[x\mathcal{M}(x,t) + t\mathcal{J}(x,t)]_{-\infty}^{+\infty} = 0.} \tag{2.11}$$

3. Algorithm design formalism. In this section, we present a refinement of the formalism for the derivation of invariant-conserving algorithms proposed in Vu-Quoc and Li [1993] in the context of exact finite difference calculus.

3.1. Notations and definitions. Let u designate the finite difference solution.[6] Let L and T be positive numbers. We consider the rectangle $\overline{\Omega} := \{(x,t) \mid -L \leq x \leq L,\ 0 \leq t \leq T\} \subset \mathbb{R}^2$ within which the computation is performed. Consider a grid of points in $\overline{\Omega}$ with spatial increment h and time increment k. Let N and M be integers such that $h = L/N$ and $k = T/M$. Then the coordinates (x,y) of the grid points in $\overline{\Omega}$ can be described by

$$\begin{aligned} x &= ih, \qquad i \in \underline{N}^{\pm} := \{-N,\ldots,-1,0,1,\ldots,N\}, \\ t &= jk, \qquad j \in \underline{M}^{+} := \{0,1,\ldots,M\}. \end{aligned}$$

The exact solution U at a grid point $(x = ph, y = qk)$ is written as $U_{p,q}$, i.e.,

$$U_{p,q} := U(ph, qk), \tag{3.1}$$

whereas the finite difference solution at the same point is denoted as $u_{p,q}$. The short-hand notation (p,q) designates the point with coordinates $(x = ph, y = qk)$ in $\overline{\Omega}$, whereas (i,j)—with $i \in \underline{N}^{\pm}$ and $j \in \underline{M}^{+}$—designates a finite difference grid point with coordinates $(x = ih, y = jk)$ in $\overline{\Omega}$.[7] In what follows, we will let $p = i$ or $i \pm \frac{1}{2}$, and $q = j$ or $j \pm \frac{1}{2}$.

[5] For a general discussion on invariant properties of NLKGE, see Strauss [1978].

[6] Recall that U designates the exact solution.

[7] In other words, p and q are not necessarily integers.

DEFINITION 3.1 (shift operator). *We define the shift operators* $\mathfrak{S}_x^r$ *in space and* $\mathfrak{S}_t^s$ *in time as*

$$\mathfrak{S}_x^r u_{p,q} := u_{p+r,q}, \quad \mathfrak{S}_t^s u_{p,q} := u_{p,q+s}. \tag{3.2}$$

DEFINITION 3.2 (arithmetic averaging operator). *At the point* (p,q), *the arithmetic averaging operators* $\mathfrak{A}_x$ *in space and* $\mathfrak{A}_t$ *in time are defined as follows:*

$$\boxed{\mathfrak{A}_x u_{p,q} := \frac{1}{2}\left(\mathfrak{S}_x^{\frac{1}{2}} + \mathfrak{S}_x^{-\frac{1}{2}}\right) u_{p,q} = \frac{1}{2}\left(u_{p+\frac{1}{2},q} + u_{p-\frac{1}{2},q}\right),} \tag{3.3a}$$

$$\boxed{\mathfrak{A}_t u_{p,q} := \frac{1}{2}\left(\mathfrak{S}_t^{\frac{1}{2}} + \mathfrak{S}_t^{-\frac{1}{2}}\right) u_{p,q} = \frac{1}{2}\left(u_{p,q+\frac{1}{2}} + u_{p,q-\frac{1}{2}}\right).} \tag{3.3b}$$

DEFINITION 3.3a (central difference operator). *The central difference operators* $\mathfrak{C}_x$ *and* $\mathfrak{C}_t$ *at point* (p,q) *are defined to be*

$$\boxed{\mathfrak{C}_x^{n+1} u_{p,q} := \left(\mathfrak{S}_x^{\frac{1}{2}} - \mathfrak{S}_x^{-\frac{1}{2}}\right) \mathfrak{C}_x^n u_{p,q} = \left[\mathfrak{C}_x^n u_{p+\frac{1}{2},q} - \mathfrak{C}_x^n u_{p-\frac{1}{2},q}\right] \quad \forall n \in \mathbb{N},} \tag{3.4a}$$

$$\boxed{\mathfrak{C}_t^{n+1} u_{p,q} := \left(\mathfrak{S}_t^{\frac{1}{2}} - \mathfrak{S}_t^{-\frac{1}{2}}\right) \mathfrak{C}_t^n u_{p,q} = \left[\mathfrak{C}_t^n u_{p,q+\frac{1}{2}} - \mathfrak{C}_t^n u_{p,q-\frac{1}{2}}\right] \quad \forall n \in \mathbb{N},} \tag{3.4b}$$

with $\mathfrak{C}_x^0$ *and* $\mathfrak{C}_t^0$ *being the identity, i.e.,*

$$\mathfrak{C}_x^0 u_{p,q} = \mathfrak{C}_t^0 u_{p,q} = u_{p,q}. \tag{3.4c}$$

DEFINITION 3.3b ($\mathfrak{H}$ operator). *At point* (p,q), *we define the finite difference operators* $\mathfrak{H}_x$ *and* $\mathfrak{H}_t$ *as*

$$\mathfrak{H}_x u_{p,q} := \left(\mathfrak{S}_x^1 - \mathfrak{S}_x^{-1}\right) u_{p,q} = u_{p+1,q} - u_{p-1,q}\ , \tag{3.5a}$$

$$\mathfrak{H}_t u_{p,q} := \left(\mathfrak{S}_t^1 - \mathfrak{S}_t^{-1}\right) u_{p,q} = u_{p,q+1} - u_{p,q-1}. \tag{3.5b}$$

DEFINITION 3.4 (geometric averaging operator). *Let the geometric averaging operators* $\mathfrak{G}_x$ *and* $\mathfrak{G}_t$ *be defined as*

$$\boxed{\mathfrak{G}_x u_{p,q} := \sqrt{u_{p+\frac{1}{2},q} u_{p-\frac{1}{2},q}}, \quad \mathfrak{G}_t u_{p,q} := \sqrt{u_{p,q+\frac{1}{2}} u_{p,q-\frac{1}{2}}}.} \tag{3.6}$$

Remark 3.1. Without the subscripts x or t, the notations $\mathfrak{S}$, $\mathfrak{A}$, $\mathfrak{C}$, and $\mathfrak{H}$ are used to indicate that these operators are applicable in both x and t.

DEFINITION 3.5 (bilinear operator $\langle\cdot,\cdot\rangle$). *Consider two operators* α *and* β, *which can be any of the operators*[8] $\{\mathfrak{S}, \mathfrak{A}, \mathfrak{C}, \mathfrak{H}, \mathfrak{G}\}$, *and two functions* $f, g\colon \overline{\Omega} \to \mathbb{R}$. *We define the operator* $\langle\cdot,\cdot\rangle$ *as follows:*

$$\langle \alpha, \beta\rangle (f,g)_{p,q} = \alpha f_{p,q}\ \beta g_{p,q}, \tag{3.7}$$

[8]To help the reading, it may be useful to note that the German letters $\mathfrak{S}$, $\mathfrak{A}$, $\mathfrak{C}$, $\mathfrak{H}$, $\mathfrak{G}$ correspond to the Roman letters S, A, C, H, G, respectively.

at any point $(p, q) \in \overline{\Omega}$.

Remark 3.2. In particular, for the shift operator $\mathfrak{S}$, we have

$$\mathfrak{S}^r (fg)_{p,q} = \mathfrak{S}^r f_{p,q}\, \mathfrak{S}^r g_{p,q} = \langle \mathfrak{S}^r, \mathfrak{S}^r \rangle (f, g)_{p,q} \Longrightarrow \mathfrak{S}^r \equiv \langle \mathfrak{S}^r, \mathfrak{S}^r \rangle. \tag{3.8a}$$

Note that this property is not true for the other operators. For example,

$$\mathfrak{H} = \mathfrak{S}^1 - \mathfrak{S}^{-1} = \langle \mathfrak{S}^1, \mathfrak{S}^1 \rangle - \langle \mathfrak{S}^{-1}, \mathfrak{S}^{-1} \rangle \neq \langle \mathfrak{H}, \mathfrak{H} \rangle. \tag{3.8b}$$

For the velocity $V = \frac{\partial U}{\partial t}$ in (2.1a), we introduce the central difference approximation at point (p, q)

$$v_{p,q} = \frac{1}{k} \mathfrak{C}_t u_{p,q}. \tag{3.9}$$

Likewise, the spatial derivative

$$W := \frac{\partial U}{\partial x} \tag{3.10a}$$

has its central difference approximation as

$$w_{p,q} = \frac{1}{h} \mathfrak{C}_x u_{p,q}. \tag{3.10b}$$

3.2. The basic lemmata. Some of the results in finite difference calculus that will be used in §4 are presented here. We refer to Milne-Thomson [1933], Jordan [1950], and Vu-Quoc and Li [1993] for further details.

LEMMA 3.1 (commutativity of $\mathfrak{C}_x^m$ and $\mathfrak{C}_t^n$).[9] *At the point* (p, q), *we have*

$$\mathfrak{C}_t^m \mathfrak{C}_x^n u_{p,q} = \mathfrak{C}_x^n \mathfrak{C}_t^m u_{p,q} \quad \forall m, n \in \mathbb{N}. \tag{3.11}$$

Remark 3.3. Since $\mathfrak{C}_x^m$ and $\mathfrak{C}_x^n$, as well as $\mathfrak{C}_t^m$ and $\mathfrak{C}_t^n$, commute with each other, we can write unambiguously that (with (3.11) in mind)

$$\mathfrak{C}^m \mathfrak{C}^n u_{p,q} = \mathfrak{C}^n \mathfrak{C}^m u_{p,q}, \tag{3.12}$$

at any point (p, q), and in which any combination of subscripts x and t for $\mathfrak{C}$ is permissible.

LEMMA 3.2 (commutativity of $\mathfrak{A}$ and $\mathfrak{C}^n$).[10] *The following relations hold:*

$$\mathfrak{A}_x \mathfrak{C}_x^n u_{p,q} = \mathfrak{C}_x^n \mathfrak{A}_x u_{p,q}, \quad \mathfrak{A}_t \mathfrak{C}_t^n u_{p,q} = \mathfrak{C}_t^n \mathfrak{A}_t u_{p,q}, \tag{3.13a}$$

$$\mathfrak{A}_x \mathfrak{C}_t^n u_{p,q} = \mathfrak{C}_t^n \mathfrak{A}_x u_{p,q}, \quad \mathfrak{A}_t \mathfrak{C}_x^n u_{p,q} = \mathfrak{C}_x^n \mathfrak{A}_t u_{p,q}, \tag{3.13b}$$

at any point (p, q), *or unambiguously*

$$\boxed{\mathfrak{A} \mathfrak{C}^n u_{p,q} = \mathfrak{C}^n \mathfrak{A} u_{p,q}.} \tag{3.14}$$

[9] See Vu-Quoc and Li [1993] for a proof.

[10] See Vu-Quoc and Li [1993] for a proof.

LEMMA 3.3 (discrete Leibniz rule). *Consider two functions $f, g: \overline{\Omega} \rightarrow \mathbb{R}$. The central difference operator $\mathfrak{C}$ when applied to the product function (fg) yields*

$$\boxed{\mathfrak{C}(fg)_{p,q} = (\mathfrak{C}f_{p,q})(\mathfrak{A}g_{p,q}) + (\mathfrak{A}f_{p,q})(\mathfrak{C}g_{p,q}),} \tag{3.15a}$$

with $p = i$ or $i \pm \frac{1}{2}$, and $q = j$ or $j \pm \frac{1}{2}$ $\forall(i,j)$. In other words, when applied to a product of functions,

$$\boxed{\mathfrak{C} \equiv \langle \mathfrak{C}, \mathfrak{A} \rangle + \langle \mathfrak{A}, \mathfrak{C} \rangle.} \tag{3.15b}$$

Proof. Since $\mathfrak{C} = \mathfrak{S}^{1/2} - \mathfrak{S}^{-1/2}$, it follows from Definition 3.5 and Remark 3.2 that

$$\begin{aligned}
\mathfrak{C}(fg)_{p,q} &= \left(\langle \mathfrak{S}^{\frac{1}{2}}, \mathfrak{S}^{\frac{1}{2}} \rangle - \langle \mathfrak{S}^{-\frac{1}{2}}, \mathfrak{S}^{-\frac{1}{2}} \rangle \right) (f, g)_{p,q} \\
&= \frac{1}{2} \Big(\langle (\mathfrak{S}^{\frac{1}{2}} - \mathfrak{S}^{-\frac{1}{2}}), (\mathfrak{S}^{\frac{1}{2}} + \mathfrak{S}^{-\frac{1}{2}}) \rangle \\
&\qquad + \langle (\mathfrak{S}^{\frac{1}{2}} + \mathfrak{S}^{-\frac{1}{2}}), (\mathfrak{S}^{\frac{1}{2}} - \mathfrak{S}^{-\frac{1}{2}}) \rangle \Big) (f, g)_{p,q} \\
&= \left(\langle \mathfrak{C}, \mathfrak{A} \rangle + \langle \mathfrak{A}, \mathfrak{C} \rangle \right) (f, g)_{p,q} = \mathfrak{C}f_{p,q}\ \mathfrak{A}g_{p,q} + \mathfrak{A}f_{p,q}\ \mathfrak{C}g_{p,q}.
\end{aligned} \tag{3.16}$$

Remark 3.4. Relations (3.15a) and (3.15b) are basically the discrete counterparts of the derivation property in the continuous case.[11] Thus the linearity of $\mathfrak{C}$ and the discrete Leibniz rule (3.15a) effectively makes $\mathfrak{C}$ into a *discrete tangent operator*, which plays a crucial role in the development of the invariant-conserving algorithms that follow.

COROLLARY 3.1 (general discrete Leibniz rule). *In general, the nth-order central difference operator when applied on a product of functions yields*

$$\boxed{\begin{aligned}
\mathfrak{C}^n(fg)_{p,q} &= \sum_{m=0}^{n+1} \binom{n}{m} \left(\mathfrak{C}^{n-m} \mathfrak{A}^m f_{p,q} \right) \left(\mathfrak{A}^{n-m} \mathfrak{C}^m g_{p,q} \right) \\
&= \left[\sum_{m=0}^{n+1} \binom{n}{m} \langle \mathfrak{C}^{n-m} \mathfrak{A}^m, \mathfrak{A}^{n-m} \mathfrak{C}^m \rangle \right] (f, g)_{p,q}.
\end{aligned}} \tag{3.17}$$

Proof.

$$\begin{aligned}
\mathfrak{C}^n(fg)_{p,q} &= \left(\langle \mathfrak{C}, \mathfrak{A} \rangle + \langle \mathfrak{A}, \mathfrak{C} \rangle \right)^n (f, g)_{p,q} \\
&= \left[\sum_{m=0}^{n+1} \binom{n}{m} \langle \mathfrak{C}, \mathfrak{A} \rangle^{(n-m)} \langle \mathfrak{A}, \mathfrak{C} \rangle^m \right] (f, g)_{p,q} \\
&= \sum_{m=0}^{n+1} \binom{n}{m} \left(\mathfrak{C}^{n-m} \mathfrak{A}^m f_{p,q} \right) \left(\mathfrak{A}^{n-m} \mathfrak{C}^m g_{p,q} \right).
\end{aligned} \tag{3.18}$$

[11] If D represents a differential operator, then the counterpart of (3.15b) is $D = \langle D, \mathbf{1} \rangle + \langle \mathbf{1}, D \rangle$, where $\mathbf{1}$ is the identity operator.

LEMMA 3.4 (averaging rule for function products). *Consider two functions f, $g: \overline{\Omega} \to \mathbb{R}$. The averaging operator when applied to the product function (fg) yields*

$$\boxed{\mathfrak{A}(fg)_{p,q} = \mathfrak{A}f_{p,q}\ \mathfrak{A}g_{p,q} + \frac{1}{4}\mathfrak{C}f_{p,q}\ \mathfrak{C}g_{p,q},} \tag{3.19a}$$

or

$$\boxed{\mathfrak{A} \equiv \langle \mathfrak{A}, \mathfrak{A}\rangle + \frac{1}{4}\langle \mathfrak{C}, \mathfrak{C}\rangle,} \tag{3.19b}$$

with $p = i$ or $i \pm \frac{1}{2}$, and $q = j$ or $j \pm \frac{1}{2}$ $\forall (i, j)$.

Proof. Since $\mathfrak{A} = \frac{1}{2}(\mathfrak{S}^{1/2} + \mathfrak{S}^{-1/2})$, by (3.10a) we have

$$\begin{aligned}
\mathfrak{A}(fg)_{p,q} &= \frac{1}{2}\left(\langle \mathfrak{S}^{\frac{1}{2}}, \mathfrak{S}^{\frac{1}{2}}\rangle + \langle \mathfrak{S}^{-\frac{1}{2}}, \mathfrak{S}^{-\frac{1}{2}}\rangle\right)(f,g)_{p,q} \\
&= \left[\left\langle \frac{1}{2}(\mathfrak{S}^{\frac{1}{2}} + \mathfrak{S}^{-\frac{1}{2}}), \frac{1}{2}(\mathfrak{S}^{\frac{1}{2}} + \mathfrak{S}^{-\frac{1}{2}})\right\rangle\right. \\
&\qquad \left. + \left\langle \frac{1}{2}(\mathfrak{S}^{\frac{1}{2}} - \mathfrak{S}^{-\frac{1}{2}}), \frac{1}{2}(\mathfrak{S}^{\frac{1}{2}} - \mathfrak{S}^{-\frac{1}{2}})\right\rangle\right](f,g)_{p,q} \\
&= \left(\langle \mathfrak{A}, \mathfrak{A}\rangle + \frac{1}{4}\langle \mathfrak{C}, \mathfrak{C}\rangle\right)(f,g)_{p,q} = \mathfrak{A}f_{p,q}\ \mathfrak{A}g_{p,q} + \frac{1}{4}\mathfrak{C}f_{p,q}\ \mathfrak{C}g_{p,q}.
\end{aligned} \tag{3.20}$$

COROLLARY 3.2. *Generally, the nth-order averaging operator when applied to a product of functions yields*

$$\boxed{\begin{aligned}
\mathfrak{A}^n(fg)_{p,q} &= \sum_{m=0}^{n+1}\binom{n}{m}\left(\frac{1}{4}\right)^m (\mathfrak{A}^{n-m}\mathfrak{C}^m f_{p,q})(\mathfrak{A}^{n-m}\mathfrak{C}^m g_{p,q}) \\
&= \left\{\sum_{m=0}^{n+1}\binom{n}{m}\left(\frac{1}{4}\right)^m \langle \mathfrak{A}^{n-m}\mathfrak{C}^m, \mathfrak{A}^{n-m}\mathfrak{C}^m\rangle\right\}(f,g)_{p,q}.
\end{aligned}} \tag{3.21}$$

Proof. Similar to the proof of Corollary 3.1.

LEMMA 3.5. *The following identities hold for any single function $f: \overline{\Omega} \to \mathbb{R}$:*

$$\boxed{\begin{aligned}
\mathfrak{A}^2 f_{p,q} &= \frac{1}{4}\mathfrak{C}^2 f_{p,q} + f_{p,q}, \\
(\mathfrak{S} f_{p,q})^2 &= 2\,(\mathfrak{A}f_{p,q})^2 - \mathfrak{A}\,(f_{p,q})^2, \\
\mathfrak{H} f_{p,q} &= 2\mathfrak{A}\mathfrak{C} f_{p,q}.
\end{aligned}} \tag{3.22abc}$$

Proof. (1) (3.22a) is obtained from the following identities:

$$\begin{aligned}
\mathfrak{A}^2 &= \left[\frac{1}{2}\left(\mathfrak{S}^{\frac{1}{2}} + \mathfrak{S}^{-\frac{1}{2}}\right)\right]^2 = \frac{1}{4}\left(\mathfrak{S}^1 + 2 + \mathfrak{S}^{-1}\right), \\
\mathfrak{C}^2 &= \left(\mathfrak{S}^{\frac{1}{2}} - \mathfrak{S}^{-\frac{1}{2}}\right)^2 = \left(\mathfrak{S}^1 - 2 + \mathfrak{S}^{-1}\right) = 4\mathfrak{A}^2 - 4.
\end{aligned} \tag{3.23a}$$

(2) (3.22b) can be obtained first by considering $\mathfrak{G}_x$ applied to f as follows:

$$(3.23b)\quad \begin{aligned}(\mathfrak{G}_x f_{p,q})^2 &:= f_{p+\frac{1}{2},q} f_{p-\frac{1}{2},q} \\ &= 2\left(\frac{1}{2}\left(f_{p+\frac{1}{2},q}+f_{p-\frac{1}{2},q}\right)\right)^2 - \frac{1}{2}\left(\left(f_{p+\frac{1}{2},q}\right)^2+\left(f_{p-\frac{1}{2},q}\right)^2\right) \\ &= 2\left(\mathfrak{A}_x f_{p,q}\right)^2 - \mathfrak{A}_x\left(f_{p,q}\right)^2.\end{aligned}$$

That a similar argument can be made for $\mathfrak{G}_t$ leads to (3.22b).

(3) (3.22c) holds true since the identity

$$\mathfrak{H} = \mathfrak{S}^1 - \mathfrak{S}^{-1} = (\mathfrak{S}^{\frac{1}{2}} + \mathfrak{S}^{-\frac{1}{2}})(\mathfrak{S}^{\frac{1}{2}} - \mathfrak{S}^{-\frac{1}{2}}) = 2\mathfrak{A}\,\mathfrak{C}$$

is valid for both x and t.

4. Invariant-conserving algorithms for NLKGE. In this section, we will examine a class of invariant-conserving algorithms proposed in Vu-Quoc and Li [1993] for the NLKGE with particular reference to their algebraic invariants in the context of finite difference calculus and to their nonlinear stability property.

4.1. Energy-conserving Algorithm I.

4.1.1. Algorithm I (ALGO_E1) and its algebraic invariant. Consider the evolution equation (2.2) at a point $(i, j+\frac{1}{2})$. By applying the central difference operators $\mathfrak{C}$ in (3.4) to approximate the differential operators $\partial/\partial t$ and $\partial/\partial x$, and the averaging operator $\mathfrak{A}$ in (3.3b) to approximate any functions at $(i, j+\frac{1}{2})$, we obtain the following algorithm

$$(4.1ab)\quad \boxed{\begin{aligned}&\frac{1}{k}\mathfrak{C}_t u_{i,j+\frac{1}{2}} = \mathfrak{A}_t v_{i,j+\frac{1}{2}},\\ &\frac{1}{k}\mathfrak{C}_t v_{i,j+\frac{1}{2}} + \frac{\mathfrak{C}_t G(u_{i,j+\frac{1}{2}})}{\mathfrak{C}_t u_{i,j+\frac{1}{2}}} = \frac{1}{h^2}\mathfrak{A}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}},\end{aligned}}$$

which involves only the values of u and v at the grid points (i,j) or $(i\pm 1, j\pm 1)$. More conventionally, (4.1ab) takes the form

(4.1cd)

$$\boxed{\begin{aligned}&\frac{1}{k}\left(u_{i,j+1}-u_{i,j}\right) = \frac{1}{2}\left(v_{i,j+1}+v_{i,j}\right),\\ &\frac{1}{k}\left(v_{i,j+1}-v_{i,j}\right) + \frac{G(u_{i,j+1})-G(u_{i,j})}{u_{i,j+1}-u_{i,j}}\\ &\quad = \frac{1}{2h^2}\left[\left(u_{i+1,j+1}-2u_{i,j+1}+u_{i-1,j+1}\right)+\left(u_{i+1,j}-2u_{i,j}+u_{i-1,j}\right)\right].\end{aligned}}$$

Remark 4.1. In (4.1a), we made the following assumption:

$$(4.2)\quad (\mathfrak{A}_t - \mathbf{1})\, f_{i,j+\frac{1}{2}} = 0.$$

In other words, we identify the averaging operator $\mathfrak{A}_t$ with the identity operator $\mathbf{1}$.[12] This assumption turned out to be valid for second-order algorithms as shown in Vu-Quoc and Li [1993]. The reader is also referred to that reference for a geometric

[12] In fact, the identification of the averaging operator $\mathfrak{A}_t$ and the identity operator $\mathbf{1}$ is also consistent with (3.15ab); see also Remark 3.4 and its footnote.

interpretation and the linear stability aspects of **ALGO_E1**. We are concerned here with the exact algebraic invariant structure of **ALGO_E1** and its nonlinear stability aspects. We emphasize that the identification (4.2) is only applied in the *design* of **ALGO_E1**—i.e., to obtain expressions (4.1ab)—but *not* in the subsequent analysis of **ALGO_E1** in the context of finite difference calculus.

THEOREM 4.1 (algebraic invariant). **ALGO_E1** *is derivable from its local algebraic invariant* (*a discrete conservation law*) *expressed as follows*:

$$\boxed{\begin{aligned} &\frac{1}{2k}\mathfrak{C}_t\left[\left(v_{i,j+\frac{1}{2}}\right)^2+\mathfrak{A}_x\left(w_{i,j+\frac{1}{2}}\right)^2+2G(u_{i,j+\frac{1}{2}})\right]\\ &\quad-\frac{1}{h}\mathfrak{C}_x\left[\left(\mathfrak{A}_t w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x v_{i,j+\frac{1}{2}}\right)\right]=0.\end{aligned}} \tag{4.3}$$

Proof. We will show that (4.1ab) can be derived from (4.3) using finite difference calculus. Application of the commutativity between $\mathfrak{C}$ and $\mathfrak{A}$ (Lemma 3.2), and the discrete Leibniz rule (3.15), to (4.3) (multiplied by k) leads to

(4.4a)

$$\begin{aligned}\frac{1}{2}\mathfrak{C}_t&\left[\left(v_{i,j+\frac{1}{2}}\right)^2+\mathfrak{A}_x\left(w_{i,j+\frac{1}{2}}\right)^2+G(u_{i,j+\frac{1}{2}})\right]-\frac{k}{h}\mathfrak{C}_x\left[\left(\mathfrak{A}_t w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x v_{i,j+\frac{1}{2}}\right)\right]\\ &=\left(\mathfrak{C}_t v_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_t v_{i,j+\frac{1}{2}}\right)+\mathfrak{A}_x\left[\left(\mathfrak{C}_t w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_t w_{i,j+\frac{1}{2}}\right)\right]\\ &\quad+\mathfrak{C}_t G(u_{i,j+\frac{1}{2}})-\frac{k}{h}\left[\left(\mathfrak{A}_t\mathfrak{C}_x w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x^2 v_{i,j+\frac{1}{2}}\right)\right.\\ &\quad\left.+\left(\mathfrak{A}_t\mathfrak{A}_x w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x\mathfrak{C}_x v_{i,j+\frac{1}{2}}\right)\right]=0.\end{aligned}$$

Then by Lemma 3.4,

$$\begin{aligned}\mathfrak{A}_x\left[\left(\mathfrak{C}_t w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_t w_{i,j+\frac{1}{2}}\right)\right]&=\left(\mathfrak{A}_x\mathfrak{C}_t w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x\mathfrak{A}_t w_{i,j+\frac{1}{2}}\right)\\ &\quad+\frac{1}{4}\left(\mathfrak{C}_x\mathfrak{C}_t w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{C}_x\mathfrak{A}_t w_{i,j+\frac{1}{2}}\right),\end{aligned}$$

which, together with (3.10b), when substituted into (4.4a) yields

(4.4b)

$$\begin{aligned}&\mathfrak{C}_t v_{i,j+\frac{1}{2}}\ \mathfrak{A}_t v_{i,j+\frac{1}{2}}+\frac{1}{h^2}\mathfrak{A}_x\left[\left(\mathfrak{C}_t\mathfrak{C}_x u_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_t\mathfrak{C}_x u_{i,j+\frac{1}{2}}\right)\right]+\mathfrak{C}_t G(u_{i,j+\frac{1}{2}})\\ &\quad-\frac{k}{h^2}\left(\mathfrak{A}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x^2 v_{i,j+\frac{1}{2}}\right)-\frac{k}{h^2}\left(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x u_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x\mathfrak{C}_x v_{i,j+\frac{1}{2}}\right)\\ &=\mathfrak{C}_t v_{i,j+\frac{1}{2}}\ \mathfrak{A}_t v_{i,j+\frac{1}{2}}+\frac{1}{h^2}\left(\mathfrak{A}_x\mathfrak{C}_t\mathfrak{C}_x u_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x\mathfrak{A}_t\mathfrak{C}_x u_{i,j+\frac{1}{2}}\right)\\ &\quad+\frac{1}{4h^2}\left(\mathfrak{C}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}\right)+\mathfrak{C}_t G(u_{i,j+\frac{1}{2}})\\ &\quad-\frac{k}{h^2}\left(\mathfrak{A}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x^2 v_{i,j+\frac{1}{2}}\right)-\frac{k}{h^2}\left(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x u_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x\mathfrak{C}_x v_{i,j+\frac{1}{2}}\right)\\ &=\mathfrak{C}_t v_{i,j+\frac{1}{2}}\ \mathfrak{A}_t v_{i,j+\frac{1}{2}}+\frac{1}{h^2}\left(\mathfrak{A}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}\right)\left(\frac{1}{4}\mathfrak{C}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}-\mathfrak{A}_x^2\mathfrak{C}_t u_{i,j+\frac{1}{2}}\right)+\mathfrak{C}_t G(u_{i,j+\frac{1}{2}})\\ &=0.\end{aligned}$$

Applying (3.22a) of Lemma 3.5 to the second term of (4.4b), we obtain

$$\frac{1}{h^2}\left(\mathfrak{A}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}\right)\left(\frac{1}{4}\mathfrak{C}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}-\mathfrak{A}_x^2\mathfrak{C}_t u_{i,j+\frac{1}{2}}\right)=-\frac{1}{h^2}\mathfrak{A}_t\mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}\left(\mathfrak{C}_t u_{i,j+\frac{1}{2}}\right).$$

Thus, (4.4b) becomes

$$\mathfrak{C}_t v_{i,j+\frac{1}{2}} \mathfrak{A}_t v_{i,j+\frac{1}{2}} - \frac{1}{h^2}\left(\mathfrak{A}_t \mathfrak{C}_x^2 u_{i,j+\frac{1}{2}}\right)\left(\mathfrak{C}_t u_{i,j+\frac{1}{2}}\right) + \mathfrak{C}_t G(u_{i,j+\frac{1}{2}}) = 0. \tag{4.4c}$$

For the factor $\mathfrak{A}_t v_{i,j+1/2}$ in the first term of (4.4c), we invoke relation (4.1a) (the trapezoidal rule)

$$\mathfrak{A}_t v_{i,j+\frac{1}{2}} = \frac{1}{k}\mathfrak{C}_t u_{i,j+\frac{1}{2}}. \tag{4.4d}$$

A division of (4.4c) by $\mathfrak{C}_t u_{i,j+1/2}$ yields

$$\frac{1}{k}\mathfrak{C}_t v_{i,j+\frac{1}{2}} - \frac{1}{h^2}\mathfrak{A}_t \mathfrak{C}_x^2 u_{i,j+\frac{1}{2}} + \frac{\mathfrak{C}_t G(u_{i,j+\frac{1}{2}})}{\mathfrak{C}_t u_{i,j+\frac{1}{2}}} = 0, \tag{4.4e}$$

i.e., (4.1a). In fact, (4.3) is the discrete local energy conservation law, a counterpart of (2.2b).

4.1.2. Energy conservation and nonlinear stability. We show that the conservation of the total system energy, with appropriate boundary conditions, is a property of **ALGO_E1**. We define the discrete energy $\mathbf{E}_{N,j}^{(E1)}$ for **ALGO_E1** at time $t = jk$ as

$$\boxed{\mathbf{E}_{N,j}^{(E1)} := h \sum_{i=-N}^{N} \mathcal{E}_{i,j}^{(1)},} \tag{4.5a}$$

where the superscript $(E1)$ refers to **ALGO_E1**, and where the discrete energy density $\mathcal{E}_{i,j}^{(E1)}$ is given by

$$\boxed{\mathcal{E}_{i,j}^{(E1)} = \frac{1}{2}(v_{i,j})^2 + \frac{1}{2h^2}\left(\mathfrak{C}_x^{(1)} \mathfrak{A}_x u_{i,j}\right)^2 + G(u_{i,j}).} \tag{4.5b}$$

Therefore, the local algebraic invariant (4.3) can be rewritten as

$$\mathfrak{C}_t \mathcal{E}_{i,j+\frac{1}{2}}^{(E1)} - \frac{k}{h}\mathfrak{C}_x\left(\mathfrak{A}_t w_{i,j+\frac{1}{2}}\right)\left(\mathfrak{A}_x v_{i,j+\frac{1}{2}}\right) = 0, \tag{4.6a}$$

or

$$\left[\mathcal{E}_{i,j+1}^{(E1)} - \mathcal{E}_{i,j}^{(E1)}\right] - \frac{k}{h}\left[(\mathfrak{A}_t w\ \mathfrak{A}_x v)_{i+\frac{1}{2},j+\frac{1}{2}} - (\mathfrak{A}_t w\ \mathfrak{A}_x v)_{i-\frac{1}{2},j+\frac{1}{2}}\right] = 0. \tag{4.6b}$$

Multiplying (4.6b) by h then summing from $i = -N$ to $i = N$, and in view of the definition of the discrete energy in (4.5a), we have

$$\begin{aligned} &\frac{1}{h}\left[\mathbf{E}_{N,j+1}^{(E1)} - \mathbf{E}_{N,j}^{(E1)}\right] \\ &= -k\left[\left(\mathfrak{A}_t w_{N+\frac{1}{2},j+\frac{1}{2}}\right)\left(\mathfrak{A}_x v_{N+\frac{1}{2},j+\frac{1}{2}}\right) - \left(\mathfrak{A}_t w_{-N-\frac{1}{2},j+\frac{1}{2}}\right)\left(\mathfrak{A}_x v_{-N-\frac{1}{2},j+\frac{1}{2}}\right)\right]. \end{aligned} \tag{4.7}$$

A choice of symmetric boundary conditions such that the last bracketed term in (4.7) vanishes yields the conservation of the discrete energy $\mathbf{E}_{N,\cdot}^{(E1)}$, thus Theorem 4.2 follows.[13]

THEOREM 4.2 (discrete energy conservation). **ALGO_E1** *described in* (4.1) *conserves the discrete energy* $\mathbf{E}_{N,\cdot}^{(E1)}$ *defined in* (4.5a) *and* (4.5b) *in the sense that*

$$\mathbf{E}_{N,j+1}^{(E1)} = \mathbf{E}_{N,j}^{(E1)} \quad \forall j \in \underline{M}^+, \tag{4.8a}$$

provided that the boundary conditions are chosen to be symmetric, such that

$$\boxed{\mathfrak{A}_t w_{N+\frac{1}{2},j+\frac{1}{2}} \, \mathfrak{A}_x v_{N+\frac{1}{2},j+\frac{1}{2}} = \mathfrak{A}_t w_{-N-\frac{1}{2},j+\frac{1}{2}} \, \mathfrak{A}_x v_{-N-\frac{1}{2},j+\frac{1}{2}}.} \tag{4.8b}$$

By induction, it follows that

$$\boxed{\mathbf{E}_{N,j}^{(E1)} = \mathbf{E}_{N,0}^{(E1)} \quad \forall j \in \underline{M}^+.} \tag{4.8c}$$

Remark 4.2. (1) The symmetry boundary condition (4.8b) is essentially the discrete counterpart of (2.5) keeping in mind the identification of the averaging operator $\mathfrak{A}$ and the identity $\mathbf{1}$. This boundary condition ensures the discrete energy conservation (4.8c), in the same manner as (2.5) ensures the conservation law (2.3). Note, however, that (4.8b) leads a *nonlinear* constraint equation to be imposed on the unknowns. The following simpler set of linear constraint equations, which ensures that (4.8b) is satisfied, has been proposed in Vu-Quoc and Li [1993]:

$$\boxed{\begin{aligned} \left(u_{N+1,j+1} - u_{N,j+1}\right) &- \left(u_{-N,j+1} - u_{-N-1,j+1}\right) \\ &= -\left(u_{N+1,j} - u_{N,j}\right) + \left(u_{-N,j} - u_{-N-1,j}\right), \\ \left(u_{N+1,j+1} + u_{N,j+1}\right) &- \left(u_{-N,j+1} + u_{-N-1,j+1}\right) \\ &= \left(u_{N+1,j} + u_{N,j}\right) - \left(u_{-N,j} + u_{-N-1,j}\right). \end{aligned}} \tag{4.9ab}$$

(2) In practical calculations, boundary conditions have to be imposed at finite range; thus (2.1d) is not used in practice. Instead, the boundary condition (4.8b)—or rather (4.9ab)—is used. To simulate a condition as close to (2.1d) as possible, for a given number ϵ arbitrarily small, the spatial dimension number L is estimated through numerical experiments to ensure that

$$\left|\frac{1}{h}\left(u((N+1)h, jk) - u(Nh, jk)\right)\right| < \epsilon \qquad \forall j \in \underline{M}^+, \tag{4.10a}$$

$$\left|\frac{1}{k}\left(u(Nh, (j+1)k) - u(Nh, jk)\right)\right| < \epsilon \qquad \forall j \in \underline{M}^+. \tag{4.10b}$$

[13]The first subscript N in $\mathbf{E}_{N,\cdot}^{(E1)}$ is fixed; the second subscript, not fixed, is represented by a "dot." The "dot" is used when we do not want to give any particular attention to a subscript.

Thus, the essence of the *Neumann* boundary condition (2.1d) is still preserved.

COROLLARY 4.1. *Consider the potential function $G(U)$ of the form*

$$G(U) = \frac{m^2}{2}U^2 + G_1(U), \tag{4.11}$$

where $G_1(U) \geq 0$, $m \neq 0$, and assume that the discrete energy (4.5a) *is finite. Then, the numerical solution of* **ALGO_E1** *is uniformly bounded.*

Proof. Define the collection of all numerical values $u_{i,j}$ as

$$\underline{u}_j := \left\{u_{i,j} \mid i \in \underline{N}^{\pm}\right\}, \quad \underline{u} := \left\{\underline{u}_j \mid j \in \underline{M}^{+}\right\}, \tag{4.12a}$$

and a norm of $\underline{u}$ as

$$\mathfrak{N}(\underline{u}) := \max_{[j \in \underline{M}^+]} \left\{ \sqrt{\frac{h}{2} \sum_{i=-N}^{N} (u_{i,j})^2} \right\}, \tag{4.12b}$$

with $\mathfrak{N}^{(E1)}(\underline{u})$ indicating that the solution $\underline{u}$ comes from **ALGO_E1**. Then since (4.12c)

$$\mathbf{E}_{N,j}^{(E1)} = h \sum_{i=-N}^{N} \left[\frac{1}{2}(v_{i,j})^2 + \frac{1}{2h^2}(\mathfrak{A}_x \mathfrak{C}_x u_{i,j})^2 + \frac{m^2}{2}(u_{i,j})^2 + G_1(u_{i,j})\right] = \mathbf{E}_{N,0}^{(E1)},$$

we have the bound on $\mathfrak{N}^{(E1)}(\underline{u})$ as

$$\mathfrak{N}^{(E1)}(\underline{u}) \leq \max_{[j \in \underline{M}^+]} \left\{ \frac{1}{m}\sqrt{\mathbf{E}_{N,j}^{(E1)}(\underline{u})} \right\} = \frac{1}{m}\sqrt{\mathbf{E}_{N,0}^{(E1)}(\underline{u})}. \tag{4.12d}$$

This bound indicates that for the type of potential function (4.11), the difference operator defined in (4.1ab) is uniformly bounded by a constant dependent only on the initial conditions.

Remark 4.3. Denote

$$|u|_{\max} := \max_{\substack{i \in \underline{N}^{\pm} \\ j \in \underline{M}^{+}}} |u_{i,j}|; \tag{4.13a}$$

then it is obvious that $\exists c_0 \in \mathbb{R}$ is constant such that

$$|u|_{\max} \leq c_0 \mathfrak{N}(u) \leq \frac{c_0}{m}\sqrt{\mathbf{E}_{N,0}^{(E1)}}. \tag{4.13b}$$

In the continuous case, for the potential function of the NLKGE of the form (4.11), the boundedness of the energy $\mathbf{E}^{(E1)}(t)$ or $\mathfrak{N}(\underline{u})$ does not automatically imply that the quantity $|U|_{\max} = \max_x |U(x,t)|$ will be bounded. Therefore, the above result is valid for the discrete case only.

The norm $\mathfrak{N}$ in (4.12b) is defined based on the first term of the particular potential function $G(\cdot)$ in (4.11). In general, $G(u)$ can be split into two parts: $G(u) = G_0(u) + G_1(u)$, where $G_0(u)$ is a nonnegative, convex function and $G_1(u)$ a nonnegative function. An example is the sine-Gordon equation, where $G(U) \equiv G_1(U) = 1 - \cos(U)$ is a nonnegative, but nonconvex, function. Even though definitions of discrete energy other than (4.12c) are possible, these definitions do not in general induce a norm, as in the case of (4.12c). For this reason, a stability analysis for the general case must be attacked from a different angle; the result is a nonlinear stability theorem for **ALGO_E1** that will be presented shortly.

Let the quantities $\underline{v}_j$ and $\underline{w}_j$ be defined in the same fashion as in $(4.12a)_1$. We introduce the following definitions:

$$\text{(4.14a)} \qquad ||\underline{u}_l||^2 := h \sum_{i=-N}^{N} (u_{i,l})^2, \quad ||\underline{v}_l||^2 := h \sum_{i=-N}^{N} (v_{i,l})^2,$$

$$\text{(4.14b)} \qquad \mathfrak{A}_x||\underline{w}_l||^2 := h \sum_{i=-N}^{N} \mathfrak{A}_x (w_{i,l})^2, \quad \underline{G}_l := h \sum_{i=-N}^{N} G(u_{i,l}).$$

The discrete energy expression (4.5ab) can then be recast into the following compact form:

$$\text{(4.14c)} \qquad \mathbf{E}_{N,l}^{(E1)}(\underline{u}) = \frac{1}{2}||\underline{v}_l||^2 + \frac{1}{2}\mathfrak{A}_x||\underline{w}_l||^2 + \underline{G}_l \quad \forall l \in \underline{M}^+.$$

LEMMA 4.1 (discrete Gronwall inequality). *Let $\Phi(t)$ be a nonnegative function defined on the discrete set $\underline{M}^+\tau := \{0, \tau, 2\tau, \ldots, l\tau, \ldots, M\tau\}$ with $\tau > 0$.*[14] *Further, let $C \geq 0$, $K \geq 0$ be some real constants. If*

$$\text{(4.15a)} \qquad \Phi(t) \leq C + K\tau \sum_{n=0}^{l-1} \Phi(n\tau) \quad \forall t = l\tau \in \underline{M}^+\tau,$$

then

$$\text{(4.15b)} \qquad \Phi(t) \leq C \exp(Kt) \quad \forall t \in \underline{M}^+\tau.$$

Proof. By assumption, $\forall t = l\tau \in \underline{M}^+\tau$, $\exists A(t) > 0$ such that $\Phi(t)$ can be expressed as

$$\text{(4.15c)} \qquad \Phi(t) = A(t) \exp(Kt).$$

Thereby, for $0 \leq i \leq l$, $\exists t_i \leq t$, such that

$$\text{(4.15d)} \qquad A(t_i) = A^l_{\max} := \max_{0 \leq j \leq l} \{A(j\tau)\}.$$

Then by virtue of (4.15a), we have

$$\text{(4.15e)} \qquad A(t_i) \exp(Kt_i) \leq C + KA(t_i)\tau \sum_{n=0}^{i-1} \exp(Kn\tau).$$

[14]Recall that the set $\underline{M}^+$ was defined in §3.1 to be $\underline{M}^+ := \{0, 1, 2, \ldots, M\}$.

The last term on the right-hand side of (4.15e) satisfies the following inequality in turn

$$\tau \sum_{n=0}^{i-1} \exp(Kn\tau) \leq \int_{s=0}^{s=t_i=i\tau} \exp(Ks)ds = \frac{1}{K}\left[\exp(Kt_i) - 1\right]. \tag{4.15f}$$

Thus,

$$A(t_i)\exp(Kt_i) \leq C + A(t_i)\left[\exp(Kt_i) - 1\right]. \tag{4.15g}$$

It then follows that

$$A(t) \leq A(t_i) \leq C, \tag{4.15h}$$

which leads to (4.15b) because t is an arbitrary element in the set $\underline{M}^{+}\tau$.

THEOREM 4.3 (discrete energy inequality). *Assume that the potential function $G(U)$ is nonnegative, and the initial discrete energy is finite. Let $k_0 < 2$ be a fixed bound on the time step size k. Then $\exists c_1$ constant such that $\forall k < k_0$ we have*[15]

$$\boxed{|||\underline{u}_l|||_1 \leq c_1 |||\underline{u}_0|||_1 \quad \forall l \in \underline{M}^{+},} \tag{4.16a}$$

where $|||\underline{u}_l|||_1$ is defined as follows:

(4.16b)

$$\boxed{\left(|||\underline{u}_l|||_1\right)^2 := h \sum_{i=-N}^{N} \left[(u_{i,l})^2 + (v_{i,l})^2 + \mathfrak{A}_x (w_{i,l})^2 + 2G(u_{i,l})\right] = ||\underline{u}_l||^2 + 2\mathbf{E}_{N,l}^{(E1)}.}$$

Proof. We begin with the identity

$$\begin{aligned} u_{i,l}^2 - u_{i,0}^2 &\equiv 2k \sum_{j=0}^{l-1} \frac{u_{i,j+1} - u_{i,j}}{k} \cdot \frac{u_{i,j+1} + u_{i,j}}{2} \\ &= 2k \sum_{j=0}^{l-1} \frac{1}{k} \mathfrak{C}_t u_{i,j+\frac{1}{2}} \cdot \mathfrak{A}_t u_{i,j+\frac{1}{2}}. \end{aligned} \tag{4.17a}$$

Applying the Schwartz inequalities

$$(a+b)^2 \geq 0 \Longrightarrow ab \leq \frac{1}{2}\left(a^2 + b^2\right) \Longrightarrow (a+b)^2 \leq 2\left(a^2 + b^2\right)$$

[15]The subscript 1 in $|||\underline{u}_l|||_1$ refers to Algorithm I (or **ALGO_E1**); the subscript l refers to the time step l.

to the right-hand side of (4.17a), one obtains

$$
\begin{aligned}
&2k\sum_{j=0}^{l-1}\frac{1}{k}\mathfrak{C}_t u_{i,j+\frac{1}{2}}\cdot\mathfrak{A}_t u_{i,j+\frac{1}{2}}\\
&\quad=\frac{k}{2}\sum_{j=0}^{l-1}\left(v_{i,j+1}+v_{i,j}\right)\left(u_{i,j+1}+u_{i,j}\right)\\
&\quad\le\frac{k}{4}\left[\sum_{j=0}^{l-1}\left(v_{i,j+1}+v_{i,j}\right)^2+\sum_{j=0}^{l-1}\left(u_{i,j+1}+u_{i,j}\right)^2\right]\\
&\quad\le\frac{k}{2}\left[\sum_{j=0}^{l-1}\left(\left(v_{i,j+1}\right)^2+\left(v_{i,j}\right)^2\right)+\sum_{j=0}^{l-1}\left(\left(u_{i,j+1}\right)^2+\left(u_{i,j}\right)^2\right)\right]\\
&\quad=\frac{k}{2}\left[\sum_{j=0}^{l-1}\left(\left\{\left(v_{i,j+1}\right)^2-\left(v_{i,j}\right)^2\right\}+2\left(v_{i,j}\right)^2\right)\right.\\
&\qquad\left.+\sum_{j=0}^{l-1}\left(\left\{\left(u_{i,j+1}\right)^2-\left(u_{i,j}\right)^2\right\}+2\left(u_{i,j}\right)^2\right)\right]\\
&\quad=\frac{k}{2}\left(\left(u_{i,l}\right)^2+\left(v_{i,l}\right)^2\right)+k\sum_{j=0}^{l-1}\left(v_{i,j}\right)^2+k\sum_{j=0}^{l-1}\left(u_{i,j}\right)^2,
\end{aligned}
\tag{4.17b}
$$

whence

$$
\left(1-\frac{k}{2}\right)\left(u_{i,l}\right)^2 A-\frac{k}{2}\left(v_{i,l}\right)^2\le\left(u_{i,0}\right)^2+k\sum_{j=0}^{l-1}\left(v_{i,j}\right)^2+k\sum_{j=0}^{l-1}\left(u_{i,j}\right)^2
\tag{4.17c}
$$

by virtue of (4.17a). A multiplication of (4.17c) by h and a summation of i from $-N$ to N yields

$$
\left(1-\frac{k}{2}\right)||\underline{u}_l||^2-\frac{k}{2}||\underline{v}_l||^2\le||\underline{u}_0||^2+k\sum_{j=0}^{l-1}\left(||\underline{u}_j||^2+||\underline{v}_j||^2\right).
\tag{4.17d}
$$

Recall that $\mathbf{E}_{N,l}^{(E1)}=\mathbf{E}_{N,0}^{(E1)}$ by the discrete energy conservation Theorem 4.2. Then one can balance each side of inequality (4.17c) as follows:

$$
\left(1-\frac{k}{2}\right)||\underline{u}_l||^2-\frac{k}{2}||\underline{v}_l||_l^2+2\mathbf{E}_{N,l}^{(E1)}\le2\mathbf{E}_{N,0}^{(E1)}+||\underline{u}_0||^2+k\sum_{j=0}^{l-1}\left(||\underline{u}_j||^2+||\underline{v}_j||^2\right),
\tag{4.17e}
$$

or, equivalently,

$$
\begin{aligned}
&\left(1-\frac{k}{2}\right)\left(||\underline{u}_l||^2+||\underline{v}_l||^2\right)+\mathfrak{A}_x||\underline{w}_l||^2+2\underline{G}_l\\
&\qquad\le\left(||\underline{u}_0||^2+||\underline{v}_0||^2+\mathfrak{A}_x||\underline{w}_0||^2+2\underline{G}_0\right)+k\sum_{j=0}^{l-1}\left[||\underline{u}_j||^2+||\underline{v}_j||^2\right].
\end{aligned}
\tag{4.17f}
$$

Since $k \geq 0$ and $G(\cdot) \geq 0$, (4.17f) leads to the following inequality by adding appropriate terms on each side of (4.17f) based on (4.16b):

$$\left(1 - \frac{k}{2}\right) |||\underline{u}_l|||_1^2 \leq |||\underline{u}_0|||_1^2 + k \sum_{j=0}^{l-1} |||\underline{u}_j|||_1^2. \tag{4.17g}$$

Furthermore, with $(1 - \frac{k}{2}) > 0 \; \forall k \leq k_0 < 2$, it follows that

$$|||\underline{u}_l|||_1^2 \leq \frac{2}{2-k} \left[|||\underline{u}_0|||_1^2 + k \sum_{j=0}^{l-1} |||\underline{u}_j|||_1^2 \right]. \tag{4.17h}$$

By the discrete Gronwall inequality of Lemma 4.1 and since $lk \leq T$, we obtain

$$\begin{aligned} |||\underline{u}_0|||_1^2 &\leq \frac{2}{2-k} |||\underline{u}_0|||_1^2 \exp\left(\frac{2}{2-k} lk\right) \\ &\leq \frac{2}{2-k} |||\underline{u}_0|||_1^2 \exp\left(\frac{2}{2-k} T\right), \end{aligned} \tag{4.17i}$$

which then leads to

$$|||\underline{u}_l|||_1 \leq c_1 |||\underline{u}_0|||_1, \tag{4.17j}$$

where

$$c_1 := \left[\frac{2}{2-k_0} \exp\left(\frac{2}{2-k_0} T \right) \right]^{\frac{1}{2}} \tag{4.17k}$$

is independent of j, h, and k, as $k \leq k_0 < 2$.

Before we discuss nonlinear stability further, we analyze below the Newton–Raphson procedure to search for the unknown solution $\underline{u}_{j+1}$ when using **ALGO_E1**. From a computational standpoint, **ALGO_E1** can be written conveniently as (see the details in Vu-Quoc and Li [1993])

$$v_{i,j+1} = \frac{2}{k} (u_{i,j+1} - u_{i,j}) - v_{i,j}, \tag{4.18a}$$

(4.18b)

$$\begin{aligned} -4hrv_{i,j} - \{ &[u_{i+1,j+1} - (2 + 4r^2) u_{i,j+1} + u_{i-1,j+1}] \\ &+ [u_{i+1,j} - (2 - 4r^2) u_{i,j} + u_{i-1,j}] \} + 2h^2 \frac{G(u_{i,j+1}) - G(u_{i,j})}{u_{i,j+1} - u_{i,j}} = 0, \end{aligned}$$

where $r := \frac{1}{\lambda} = \frac{h}{k}$. In (4.18a) and (4.18b), the quantity $\underline{x}_j := \{\underline{u}_j, \underline{v}_j\}$ is assumed known, whereas $\underline{x}_{j+1} := \{\underline{u}_{j+1}, \underline{v}_{j+1}\}$ is obtained by first solving (4.18b), which can be expressed in the following form:

$$\mathbf{P}_k (\underline{x}_j, \underline{u}_{j+1}) = 0 \qquad \forall j \in \underline{M}^+ \tag{4.18c}$$

for $\underline{u}_{j+1}$. The solution $\underline{u}_{j+1}$ does not necessarily exist. In order to furnish a basis for the nonlinear stability analysis that will follow shortly, we first establish below the conditions for the existence and uniqueness of the solution $\underline{u}_{j+1}$ of (4.18c) based on the implicit function theorem.

PROPOSITION 4.1. *Let a critical time step size k_c be defined as follows*

$$k_c := \begin{cases} \infty & \text{if } g \geq 0, \\ \sqrt{\frac{2}{|g|}} & \text{if } g < 0, \end{cases} \tag{4.19a}$$

where

$$\begin{aligned} g &:= \min_{i \in \underline{N}^{\pm}, j \in \underline{M}^{+}} \left\{ \frac{G'(u_{i,j+1})(u_{i,j+1} - u_{i,j}) - \{G(u_{i,j+1}) - G(u_{i,j})\}}{(u_{i,j+1} - u_{i,j})^2} \right\} \\ &= \min_{i \in \underline{N}^{\pm}, j \in \underline{M}^{+}} \{G''(u_{i,j+1}) + \mathcal{O}(k)\}. \end{aligned} \tag{4.19b}$$

Suppose that $G(\cdot)$, $G'(\cdot)$, and $G''(\cdot)$ are continuous, and that $G''(\cdot)$ is finite over $\overline{\Omega}$. Then $\forall k \leq k_c$, there exists a unique operator C_k such that

$$\underline{x}_{j+1} = C_k(\underline{x}_j). \tag{4.19c}$$

Proof. Since the potential function $G(\cdot)$ is well behaved in the sense of the above assumption, one only needs to show that the tangent mapping[16] $D_2\mathbf{P}_k\left(\underline{x}_j, \underline{u}_{j+1}\right)$ has a continuous inverse

$$\Gamma := \left[D_2\mathbf{P}_k\left(\underline{x}_j, \underline{u}_{j+1}\right)\right]^{-1} \in \mathbb{R}^{(2N+1)\times(2N+1)} \tag{4.20a}$$

to apply the implicit function theorem (e.g., Kantorovich and Akilov [1982, p. 518]). It can be easily shown (Vu-Quoc and Li [1993]) that

$$D_2\mathbf{P}_k\left(\underline{x}_j, \underline{u}_{j+1}\right) = \begin{bmatrix} d_{-N} & -1 & 0 & \dots & \dots & 0 & -1 \\ -1 & d_{-N+1} & -1 & \dots & \dots & \dots & 0 \\ 0 & -1 & d_{-N+2} & -1 & \dots & \dots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & -1 & d_{N-1} & -1 \\ -1 & 0 & \dots & \dots & \dots & -1 & d_N \end{bmatrix}, \tag{4.20b}$$

where

$$d_i := 4r^2 + 2 + 2h^2 \frac{G'(u_{i,j+1})(u_{i,j+1} - u_{i,j}) - (G(u_{i,j+1}) - G(u_{i,j}))}{(u_{i,j+1} - u_{i,j})^2} \quad \forall i \in \underline{N}^{\pm}. \tag{4.20c}$$

A sufficient condition for $D_2\mathbf{P}_k$ to be invertible is that all the eigenvalues of (4.20b) are positive, which can be readily proved by using the Gershgorin theorem (e.g., Golub and Van Loan [1989, p. 341]). By the particular structure of $D_2\mathbf{P}_k$ in (4.20b), all eigenvalues μ of $D_2\mathbf{P}_k$ lie within the union of circles of centers d_i, $i \in \underline{N}^{\pm}$, and of radius 2; i.e.,

$$|\mu - d_i| \leq 2. \tag{4.20d}$$

[16]Note that $D_2\mathbf{P}$ means the partial derivative of $\mathbf{P}$ with respect to the second argument of $\mathbf{P}$.

It follows that the criterion for all eigenvalues μ to be positive is that $d_i \geq 2$, $i \in \underline{N}^{\pm}$, or, equivalently, from (4.20c)

$$\frac{4}{k^2} + 2g_{i,j} \geq 0, \tag{4.20e}$$

where

$$g_{i,j} := \frac{G'(u_{i,j+1})(u_{i,j+1} - u_{i,j}) - (G(u_{i,j+1}) - G(u_{i,j}))}{(u_{i,j+1} - u_{i,j})^2}.$$

As long as $g_{i,j} \geq 0$, the inequality (4.20e) is automatically satisfied. If $g_{i,j} < 0$, the inequality (4.20e) leads to the second part of (4.19a) and (4.19b), which guarantees that there exists a unique operator F_k such that

$$\underline{u}_{j+1} = F_k(\underline{x}_j). \tag{4.20f}$$

Combining (4.20f) with (4.18a), one is then led to (4.19c).

Based on the above analysis, one can be assured that $\forall k \in (0, k_c)$ there exists a unique operator C_k for **ALGO_E1** such that

$$\underline{x}_{j+1} = C_k\left(\underline{x}_j\right) = \left(C_k \circ C_k\right)\left(\underline{x}_{j-1}\right) = \cdots = \left(C_k\right)^{j+1}\left(\underline{x}_0\right)\ , \tag{4.21}$$

which therefore provides a foundation for the computations in our previous work, Vu-Quoc and Li [1993]. Similar to (4.12b), let the norm of $\underline{x}$ be defined as

$$\mathfrak{N}(\underline{x}) := \max_{[l \in \underline{M}^+]} \left\{ \sqrt{\|\underline{u}_l\|^2 + \|\underline{v}_l\|^2} \right\}. \tag{4.22a}$$

THEOREM 4.4 (nonlinear stability). *Suppose the potential function $G(u) \geq 0$, and both discrete energy $\mathbf{E}_{N,1}^{(E1)}$ and $\|u_0\|$ are finite. For sufficiently small k (i.e., $k < k_0 = \min(k_c, 2)$), the implicit* **ALGO_E1** *in* (4.1a) *and* (4.1b) *is unconditionally stable*[17] *in the sense that*

$$\mathfrak{N}(\underline{x}) \leq c_1 |||\underline{u}_0|||_1, \tag{4.22b}$$

where c_1 is a constant independent of h and k.

Proof. By (4.16a) and (4.16b) of Theorem 4.3, we have that

$$\sqrt{\|\underline{u}_l\|^2 + \|\underline{v}_l\|^2} \leq |||\underline{u}_l|||_1 \leq c_1 |||\underline{u}_0|||_1, \tag{4.22c}$$

which, together with definition (4.22a), leads to (4.22b).[18]

Remark 4.4. (1) The definition of stability as employed in Theorem 4.4 is a generalization of the Lax–Richtmyer stability originally defined for linear operators (Richtmyer and Morton [1967], Smith [1985, p. 53]). A similar general definition for stability of nonlinear operators can be found in Lakshmikantham and Trigiante [1988].

(2) The assumption that $G(u) \geq 0$ is sufficiently general for most applications. In fact, for any function $G(u)$ bounded below, i.e., there exists a lower bound G_{lb} such that

$$G_{lb} = \inf_u \{G(u)\} > -\infty, \tag{4.23a}$$

---

[17] I.e., regardless of the values of $\lambda = \frac{k}{h}$.

[18] Recall from Theorem 4.3 that c_1 is independent of l, h, and k, whereas $|||\underline{u}_0|||_1$ depends on the initial conditions.

one can always adjust the constant to make $G(u)$ nonnegative by replacing $G(u)$ by $G(u)^*$ such that

$$G(u)^* := -\langle G_{lb} \rangle + G(u) \geq 0, \tag{4.23b}$$

where $\langle G_{lb} \rangle := \frac{1}{2}(G_{lb} - |G_{lb}|)$. Examples are $G(U) = 1 - \cos(U) \geq 0$ for the sine-Gordon equation, $G(U) = 2 - \cos(U) - \cos(\frac{1}{2}U) \geq 0$ for the double sine-Gordon equation, etc. Thus, the assumption of nonnegativity of $G(U)$ encompasses many practical NLKGEs. An interesting case regarding the "phi-four" equation will be discussed in Remark 5.2.

4.2. Energy-conserving Algorithm II.

4.2.1. Algorithm II (ALGO_E2) and its algebraic invariant. Instead of constructing the algorithm from the point $(i, j + \frac{1}{2})$ as in **ALGO_E1**, we now consider the evolution equation (2.2) at the point (i, j). First, let us define the following symbol:

$$v_{i,j+\frac{1}{2}} := \frac{1}{k}(u_{i,j+1} - u_{i,j}) \quad \forall j \in \underline{M}^+, \tag{4.24}$$

which then leads to (4.25a) below, the discrete counterpart of (2.2a). Next by applying the $\mathfrak{H}$ operator (3.5), the averaging rule (3.3b), and the central difference operators (3.4) to approximate the differential operators $\partial/\partial t$ and $\partial/\partial x$ in (2.2b), we obtain (4.25b) as the discrete counterpart of (2.2b). **ALGO_E2** is summarized below.

$$\boxed{\begin{aligned} &\frac{1}{2k}\mathfrak{H}_t u_{i,j} = \mathfrak{A}_t v_{i,j}, \\ &\frac{1}{k}\mathfrak{C}_t v_{i,j} - \frac{1}{h}\mathfrak{A}_t^2 \mathfrak{C}_x w_{i,j} + \frac{\mathfrak{H}_t G(u_{i,j})}{\mathfrak{H}_t u_{i,j}} = 0. \end{aligned}} \tag{4.25ab}$$

Since by the definitions of $v_{i,j}$ in (3.9) and of $w_{i,j}$ in (3.10b), we have

$$\mathfrak{C}_t v_{i,j} = \frac{1}{k}\mathfrak{C}_t^2 u_{i,j}, \quad \mathfrak{C}_x w_{i,j} = \frac{1}{h}\mathfrak{C}_x^2 u_{i,j},$$

and thus the algebraic evolution equation (4.25b) can be expressed in terms of the unknown u only as follows:

$$\boxed{\frac{1}{k^2}\mathfrak{C}_t^2 u_{i,j} - \frac{1}{4h^2}\left(\mathfrak{C}_x^2 u_{i,j+1} + 2\mathfrak{C}_x^2 u_{i,j} + \mathfrak{C}_x^2 u_{i,j-1}\right) + \frac{\mathfrak{H}_t G(u_{i,j})}{\mathfrak{H}_t u_{i,j}} = 0.} \tag{4.25c}$$

Remark 4.5. At the initial time step, i.e., $j = 0$, instead of (4.25a), we use

$$\frac{1}{2k}\mathfrak{H}_t u_{i,j} = \mathfrak{A}_t v_{i,0} \equiv v_{i,0}. \tag{4.26a}$$

In fact, the above identity amounts to the identification of the averaging operator $\mathfrak{A}$ and the identity operator $\mathbf{1}$. To obtain (4.25b), we also made use of a similar identification:

$$(\mathbf{1} - \mu_t^2)\,\mathfrak{C}_x^2 u_{i,j} = 0. \tag{4.26b}$$

As in Remark 4.1, we emphasize that the identifications made in (4.26a) and (4.26b) above are only made at the *design level* of **ALGO_E2**—i.e., to obtain expressions (4.17a) and (4.17b)—but *not* in the subsequent analysis of **ALGO_E2** in the context of exact finite difference calculus. The second-order accurate identifications (4.26a) and (4.26b) on the other hand are consistent with the second-order accuracy of **ALGO_E2** (Vu-Quoc and Li [1993]). Finally, we note that (4.25a) is not needed to obtain (4.25c); (4.25a) is used in the following theorem.

THEOREM 4.5 (algebraic invariant). **ALGO_E2** *is derivable from the following local algebraic invariant form:*

(4.27)

$$\boxed{\frac{1}{2k}\mathfrak{C}_t\left[(v_{i,j})^2+\mathfrak{A}_x\left(\mathfrak{A}_t w_{i,j}\right)^2+2\mathfrak{A}_t G(u_{i,j})\right]-\frac{1}{h}\mathfrak{C}_x\left[\left(\mathfrak{A}_t\mathfrak{A}_x v_{i,j}\right)\left(\mathfrak{A}_t^2 w_{i,j}\right)\right]=0.}$$

Proof. First, multiply (4.27) throughout by k. Then by applying the commutativity of $\mathfrak{C}$ and $\mathfrak{A}$ (Lemma 3.2), and the discrete Leibniz rule (3.15), one finds that

$$\begin{aligned}&\frac{1}{2}\mathfrak{C}_t v_{i,j}^2+\frac{\mathfrak{A}_x}{2}\mathfrak{C}_t\left(\mathfrak{A}_t w_{i,j}\right)^2+\mathfrak{A}_t\mathfrak{C}_t G(u_{i,j})-\frac{k}{h}\mathfrak{C}_x\left[\left(\mathfrak{A}_t\mathfrak{A}_x v_{i,j}\right)\left(\mathfrak{A}_t^2 w_{i,j}\right)\right]\\ &\quad=\mathfrak{C}_t v_{i,j}\mathfrak{A}_t v_{i,j}+\mathfrak{A}_x\left[\left(\mathfrak{C}_t\mathfrak{A}_t w_{i,j}\right)\left(\mathfrak{A}_t^2 w_{i,j}\right)\right]+\mathfrak{A}_t\mathfrak{C}_t G(u_{i,j})\\ &\quad\quad-\frac{k}{h}\left[\left(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x v_{i,j}\right)\left(\mathfrak{A}_x\mathfrak{A}_t^2 w_{i,j}\right)+\left(\mathfrak{A}_t\mathfrak{A}_x^2 v_{i,j}\right)\left(\mathfrak{A}_t^2\mathfrak{C}_x w_{i,j}\right)\right].\end{aligned} \tag{4.28a}$$

Next, apply the averaging rule for function products (Lemma 3.4) to the second term of the right-hand side of (4.28a) to have

(4.29)

$$\mathfrak{A}_x\left[\left(\mathfrak{C}_t\mathfrak{A}_t w_{i,j}\right)\left(\mathfrak{A}_t^2 w_{i,j}\right)\right]=\left(\mathfrak{A}_x\mathfrak{A}_t\mathfrak{C}_t w_{i,j}\right)\left(\mathfrak{A}_x\mathfrak{A}_t^2 w_{i,j}\right)+\frac{1}{4}\left(\mathfrak{A}_t\mathfrak{C}_x\mathfrak{C}_t w_{i,j}\right)\left(\mathfrak{A}_t^2\mathfrak{C}_x w_{i,j}\right).$$

Then by substituting (4.29) into (4.28a), together with the use of (3.10b), (3.9), Lemma 3.1 and the use of (3.22a), we arrive at[19]

(4.28b)

$$\begin{aligned}&\mathfrak{C}_t v_{i,j}\ \mathfrak{A}_t v_{i,j}+\left(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_t w_{i,j}\right)\left(\mathfrak{A}_x\mathfrak{A}_t^2 w_{i,j}\right)+\frac{k}{4h}\left(\mathfrak{A}_t\mathfrak{C}_x^2 v_{i,j}\right)\left(\mathfrak{A}_t^2\mathfrak{C}_x w_{i,j}\right)\\ &\quad+\mathfrak{A}_t\mathfrak{C}_t G(u_{i,j})-\frac{k}{h}\left[\left(\mathfrak{A}_x\mathfrak{A}_t\mathfrak{C}_x v_{i,j}\right)\left(\mathfrak{A}_x\mathfrak{A}_t^2 w_{i,j}\right)+\left(\mathfrak{A}_t\mathfrak{A}_x^2 v_{i,j}\right)\left(\mathfrak{A}_t^2\mathfrak{C}_x w_{i,j}\right)\right]\\ &\quad=\mathfrak{C}_t v_{i,j}\ \mathfrak{A}_t v_{i,j}+\frac{k}{4h}\left(\mathfrak{A}_t^2\mathfrak{C}_x w_{i,j}\right)\left[\left(\mathfrak{A}_t\mathfrak{C}_x^2 v_{i,j}\right)-4\left(\mathfrak{A}_t\mathfrak{A}_x^2 v_{i,j}\right)\right]+\mathfrak{A}_t\mathfrak{C}_t G(u_{i,j})\\ &\quad=\mathfrak{C}_t v_{i,j}\ \mathfrak{A}_t v_{i,j}-\frac{k}{h}\left(\mathfrak{A}_t^2\mathfrak{C}_x w_{i,j}\right)\ \mathfrak{A}_t v_{i,j}+\mathfrak{A}_t\mathfrak{C}_t G(u_{i,j})=0,\end{aligned}$$

which then leads to the final result (4.25b) by the use of identity (3.22c).

Energy conservation and nonlinear stability. By defining the discrete energy for **ALGO_E2** at half time step as follows,

$$\boxed{\mathbf{E}_{N,j+\frac{1}{2}}^{(E2)}:=h\sum_{i=-N}^{N}\mathcal{E}_{i,j+\frac{1}{2}}^{(E2)},} \tag{4.30a}$$

[19]The first equality in (4.28b) is obtained by using (3.10b), (3.9), and Lemma 3.1; the second equality in (4.28b) is obtained by using (3.22a).

where

$$\mathcal{E}^{(E2)}_{i,j+\frac{1}{2}} := \frac{1}{2}\left[\left(v_{i,j+\frac{1}{2}}\right)^2 + \mathfrak{A}_x\left(\mathfrak{A}_t w_{i,j+\frac{1}{2}}\right)^2 + 2\mathfrak{A}_t G(u_{i,j+\frac{1}{2}})\right], \tag{4.30b}$$

one obtains the following result.

THEOREM 4.6 (discrete energy conservation). **ALGO_E2** *defined in* (4.25ab) *conserves the discrete energy* $\mathbf{E}^{(E2)}_{N,\cdot}$ *defined at half time step in the sense that*

$$\boxed{\mathbf{E}^{(E2)}_{N,j+\frac{1}{2}} = \mathbf{E}^{(E2)}_{N,\frac{1}{2}} = \mathbf{E}^{(E2)}_{N,-\frac{1}{2}} \quad \forall j \in \underline{M}^+,} \tag{4.31a}$$

provided that the following boundary condition holds:

(4.31b)

$$\boxed{\left(\mathfrak{A}_t\mathfrak{A}_x v_{N+\frac{1}{2},j}\right)\left(\mathfrak{A}_t^2 w_{N+\frac{1}{2},j}\right) = \left(\mathfrak{A}_t\mathfrak{A}_x v_{-N-\frac{1}{2},j}\right)\left(\mathfrak{A}_t^2 w_{-N-\frac{1}{2},j}\right) \quad \forall j \in \underline{M}^+.}$$

Proof. By definition (4.30b) and the central difference operator (3.4), (4.27) can be rewritten as

$$\mathcal{E}^{(E2)}_{i,j+\frac{1}{2}} - \mathcal{E}^{(E2)}_{i,j-\frac{1}{2}} - \frac{k}{h}\mathfrak{C}_x\left[\left(\mathfrak{A}_t\mathfrak{A}_x v_{i,j}\right)\left(\mathfrak{A}_t^2 w_{i,j}\right)\right] = 0. \tag{4.32a}$$

Summing (4.32a) from $i = -N$ to $i = N$, we have

(4.32b)

$$\begin{aligned}&\mathbf{E}^{(E2)}_{N,j+\frac{1}{2}} - \mathbf{E}^{(E2)}_{N,j-\frac{1}{2}} \\ &\quad = \frac{k}{2}\left\{\left(\mathfrak{A}_t\mathfrak{A}_x v_{N+\frac{1}{2},j}\right)\left(\mathfrak{A}_t^2 w_{N+\frac{1}{2},j}\right) - \left(\mathfrak{A}_t\mathfrak{A}_x v_{-N-\frac{1}{2},j}\right)\left(\mathfrak{A}_t^2 w_{-N-\frac{1}{2},j}\right)\right\}.\end{aligned}$$

Application of the boundary condition (4.31b) yields

$$\mathbf{E}^{(E2)}_{N,j+\frac{1}{2}} = \mathbf{E}^{(E2)}_{N,j-\frac{1}{2}} \quad \forall j \in \underline{M}^+, \tag{4.32c}$$

and hence (4.30a).

A similar statement as in Corollary 4.1 can be made here. We present directly, however, a nonlinear stability result by first defining the following energy-related quantity at time $t = (l + \frac{1}{2})k$:

$$\boxed{\begin{aligned}\left(|||\underline{u}_{l+\frac{1}{2}}|||_2\right)^2 &:= ||\mathfrak{A}_t\underline{u}_{l+\frac{1}{2}}||^2 + ||\underline{v}_{l+\frac{1}{2}}||^2 + \mathfrak{A}_x||\mathfrak{A}_t\underline{w}_{l+\frac{1}{2}}||^2 + 2\mathfrak{A}_t\underline{G}_{l+\frac{1}{2}} \\ &= ||\mathfrak{A}_t\underline{u}_{l+\frac{1}{2}}||^2 + 2\mathbf{E}^{(E2)}_{N,l+\frac{1}{2}} \geq 0,\end{aligned}} \tag{4.33}$$

which[20] is the counterpart for **ALGO_E2** of $|||\underline{u}_l|||_1$ defined in (4.16b) for **ALGO_E1**. The reader is referred to (4.14a) and (4.14b) for the definition of various terms in (4.33).

[20]The right-hand side of (4.33) is nonnegative because of the nonnegativity assumption on $G(U)$.

THEOREM 4.7 (energy inequality). *For nonnegative potential function $G(u)$, finite initial discrete energy $\mathbf{E}_{N,0}^{(E2)}$, and time step size $k \in [0,2[$, the quasi-energy form* (4.33) *is bounded above.*

Proof. By using the definition (4.24) of $v_{i,j+1/2}$ in the identity (4.17a), and by using the Schwartz inequalities as in the proof of Theorem 4.3, we obtain

$$\begin{aligned} u_{i,l}^2 - u_{i,0}^2 &= 2k\sum_{j=0}^{l-1}\left(v_{i,j+\frac{1}{2}}\right)\cdot\left(\mathfrak{A}_t u_{i,j+\frac{1}{2}}\right) \\ &\le k\sum_{j=0}^{l-1}\left(v_{i,j+\frac{1}{2}}\right)^2 + k\sum_{j=0}^{l-1}\left(\mathfrak{A}_t u_{i,j+\frac{1}{2}}\right)^2 \end{aligned} \tag{4.34a}$$

and similarly
(4.34b)

$$\begin{aligned} u_{i,l+1}^2 - u_{i,0}^2 &\le k\sum_{j=0}^{l}\left(v_{i,j+\frac{1}{2}}\right)^2 + k\sum_{j=0}^{l}\left(\mathfrak{A}_t u_{i,j+\frac{1}{2}}\right)^2 \\ &= k\left[\left(v_{i,l+\frac{1}{2}}\right)^2 + \left(\mathfrak{A}_t u_{i,l+\frac{1}{2}}\right)^2 + \sum_{j=0}^{l-1}\left(v_{i,j+\frac{1}{2}}\right)^2 + \sum_{j=0}^{l-1}\left(\mathfrak{A}_t u_{i,j+\frac{1}{2}}\right)^2\right]. \end{aligned}$$

Now summing (4.34a) and (4.34b) leads to

$$\begin{aligned} & u_{i,l+1}^2 + u_{i,l}^2 - k\left(\mathfrak{A}_t u_{i,l+\frac{1}{2}}\right)^2 - k\left(v_{i,l+\frac{1}{2}}\right)^2 \\ &\quad \le 2\left(u_{i,0}\right)^2 + 2k\left[\sum_{j=0}^{l-1}\left(\mathfrak{A}_t u_{i,j+\frac{1}{2}}\right)^2 + \sum_{j=0}^{l-1}\left(\mathfrak{A}_t v_{i,j+\frac{1}{2}}\right)^2\right]. \end{aligned} \tag{4.34c}$$

The first two terms on the left-hand side of (4.34c) satisfy the inequality

$$(a+b)^2 \le 2\left(a^2+b^2\right) \Longrightarrow 2\left(\mathfrak{A}_t u_{i,l+\frac{1}{2}}\right)^2 \le \left(u_{i,l+1}\right)^2 + \left(u_{i,l}\right)^2, \tag{4.34d}$$

which when used in (4.34c) yields
(4.34e)

$$(2-k)\left(\mathfrak{A}_t u_{i,l+\frac{1}{2}}\right)^2 - k\left(v_{i,l+\frac{1}{2}}\right)^2 \le 2\left(u_{i,0}\right)^2 + 2k\sum_{j=0}^{l-1}\left[\left(\mathfrak{A}_t u_{i,j+\frac{1}{2}}\right)^2 + \left(v_{i,j+\frac{1}{2}}\right)^2\right].$$

Summing (4.34e) from $i=-N$ to $i=N$, we have

$$(2-k)\|\mathfrak{A}_t \underline{u}_{l+\frac{1}{2}}\|^2 - k\|\underline{v}_{l+\frac{1}{2}}\|^2 \le 2\|\underline{u}_0\|^2 + 2k\sum_{j=0}^{l-1}\left[\|\mathfrak{A}_t \underline{u}_{j+\frac{1}{2}}\|^2 + \|\underline{v}_{j+\frac{1}{2}}\|^2\right]. \tag{4.34f}$$

Next by the conservation of energy (4.31a), we can add to each side of the inequality (4.34f) the amount

$$4\mathbf{E}_{N,l+\frac{1}{2}}^{(E2)} = 4\mathbf{E}_{N,\frac{1}{2}}^{(E2)},$$

and then by using the definition (4.30) of the energy $\mathbf{E}^{(E2)}_{N,l+1/2}$, the nonnegativity assumption on $G(\cdot)$ with $k \geq 0$, and the definition (4.33) of $|||u|||^2_{2,l+1/2}$, we arrive at the following lower bound for the left-hand side of (4.34f):

$$\begin{aligned}(2-k)|||\underline{u}_{l+\frac{1}{2}}|||_2^2 \leq 2||\underline{u}_0||^2 + 2\left[||\underline{v}_{\frac{1}{2}}||^2 + \mathfrak{A}_x||\mathfrak{A}_t\underline{w}_{\frac{1}{2}}||^2 + 2\mathfrak{A}_t\underline{G}_{\frac{1}{2}}\right] \\ + 2k\sum_{j=0}^{l-1}\left[||\mathfrak{A}_t\underline{u}_{j+\frac{1}{2}}||^2 + ||\underline{v}_{j+\frac{1}{2}}||^2\right].\end{aligned} \tag{4.34g}$$

The right-hand side of (4.34g) can be reworked by first considering that

$$k^2\left(v_{i,\frac{1}{2}}\right)^2 + 4\left(\mathfrak{A}_t u_{i,\frac{1}{2}}\right)^2 = 2\left(u_{i,1}\right)^2 + 2\left(u_{i,0}\right)^2$$

or

$$2||\underline{u}_0||^2 + 2||\underline{u}_1||^2 = k^2||\underline{v}_{\frac{1}{2}}||^2 + 4||\mathfrak{A}_t\underline{u}_{\frac{1}{2}}||^2. \tag{4.34h}$$

Thus by adding $2||\underline{u}_1||^2$ to the right-hand side of (4.34g) and by using (4.34h) and the fact that $0 \leq k < 2$, and also by using the nonnegativity of $G(\cdot)$, we obtain an upper bound for the right-hand side of (4.34g) as follows:

$$(2-k)|||\underline{u}_{l+\frac{1}{2}}|||_2^2 \leq 6|||\underline{u}_{\frac{1}{2}}|||^2 + 2k\sum_{j=0}^{l-1}|||\underline{u}_{j+\frac{1}{2}}|||_2^2, \tag{4.34i}$$

which leads to

$$|||\underline{u}_{l+\frac{1}{2}}|||_2^2 \leq \frac{6}{2-k}|||\underline{u}_{\frac{1}{2}}|||_2^2 + \frac{2k}{2-k}\sum_{j=0}^{l-1}|||\underline{u}_{j+\frac{1}{2}}|||_2^2. \tag{4.34j}$$

Similar to (4.17i)–(4.17k), by the discrete Gronwall Lemma 4.1, and since $lk \leq T$ and $k \leq k_0 < 2$, we obtain the bound

$$|||\underline{u}_{l+\frac{1}{2}}|||_2 \leq c_2|||\underline{u}_{\frac{1}{2}}|||_2, \tag{4.34k}$$

where

$$c_2 := \left[\frac{6}{2-k_0}\exp\left(\frac{2T}{2-k_0}\right)\right]^{\frac{1}{2}} = \sqrt{3}c_1. \tag{4.34l}$$

Remark 4.6. From (4.34h), we have

$$2||\underline{u}_{l+1}||^2 + 2||\underline{u}_l||^2 = k^2||\underline{v}_{l+\frac{1}{2}}||^2 + 4||\mathfrak{A}_t\underline{u}_{l+\frac{1}{2}}||^2, \tag{4.34m}$$

which leads to the estimate

$$\begin{aligned}||\underline{u}_{l+1}||^2 &\leq \frac{1}{2}\left\{k^2||\underline{v}_{l+\frac{1}{2}}||^2 + 4||\mathfrak{A}_t\underline{u}_{l+\frac{1}{2}}||^2\right\} \\ &\leq 2\left\{||\underline{v}_{l+\frac{1}{2}}||^2 + ||\mathfrak{A}_t\underline{u}_{l+\frac{1}{2}}||^2\right\} \leq 2|||\underline{u}_{l+\frac{1}{2}}|||_2^2,\end{aligned} \tag{4.34n}$$

since $k < 2$ and by (4.33). It then follows from (4.34k) that

$$\mathfrak{N}(\underline{u}) := \max_{[l \in \underline{M}^+]} \|\underline{u}_l\| \leq \sqrt{2} \max_{[l \in \underline{M}^+]} |||\underline{u}_{l+\frac{1}{2}}|||_2 \leq \sqrt{2} c_2 |||\underline{u}_{\frac{1}{2}}|||_2. \tag{4.34o}$$

4.3. The Strauss–Vazquez algorithm. Strauss and Vazquez [1978] proposed an energy-conserving algorithm for the NLKGE and gave their discrete energy expression without specifying the corresponding boundary conditions. We re-examine here the Strauss–Vazquez algorithm from a more rigorous viewpoint; we give its local algebraic invariant, associated discrete energy, and corresponding boundary conditions. Although we do believe that the discrete energy expression given by Strauss and Vazquez is truly a global invariant with certain *special* boundary conditions, the algorithm (**ALGO_SV**) and its local algebraic invariant as proposed herein are well defined at each grid point in a unified manner and therefore are consistent with each other.[21]

Algorithm III (**ALGO_SV**). *At grid point* (i,j), *the Strauss–Vazquez algorithm reads as follows*:

$$\boxed{\begin{aligned} &\frac{1}{2k}\mathfrak{H}_t u_{i,j} =: \mathfrak{A}_t v_{i,j}, \\ &\frac{1}{k}\mathfrak{C}_t v_{i,j} - \frac{1}{h}\mathfrak{C}_x w_{i,j} + \frac{\mathfrak{H}_t G(u_{i,j})}{\mathfrak{H}_t u_{i,j}} = 0. \end{aligned}} \tag{4.35ab}$$

Note the difference between (4.35b) and (4.25b) in the second term. Since, by definition,

$$\frac{1}{k}\mathfrak{C}_t v_{i,j} := \frac{1}{k^2}\mathfrak{C}_t^2 u_{i,j}, \quad \frac{1}{h}\mathfrak{C}_x w_{i,j} := \frac{1}{h^2}\mathfrak{C}_x^2 u_{i,j},$$

the original **ALGO_SV** takes the form (obtained from (4.35b))

$$\boxed{\begin{aligned} \frac{1}{k^2}\left(u_{i,j+1} - 2u_{i,j} + u_{i,j-1}\right) - \frac{1}{h^2}\left(u_{i+1,j} - 2u_{i,j} + u_{i-1,j}\right) \\ - \frac{G(u_{i,j+1}) - G(u_{i,j-1})}{u_{i,j+1} - u_{i,j-1}} = 0. \end{aligned}} \tag{4.35c}$$

THEOREM 4.8 (algebraic invariant). **ALGO_SV** *admits the following local algebraic invariant:*

$$\boxed{\frac{1}{2k}\mathfrak{C}_t\left[\left(v_{i,j}\right)^2 + \mathfrak{A}_x\left(\mathfrak{G}_t w_{i,j}\right)^2 + 2\mathfrak{A}_t G(u_{i,j})\right] - \frac{1}{h}\mathfrak{C}_x\left[\left(\mathfrak{A}_t\mathfrak{A}_x v_{i,j}\right)\cdot w_{i,j}\right] = 0,} \tag{4.36}$$

from which it can be derived.

Proof. By (3.22b) of Lemma 3.5, one has

$$\left(\mathfrak{G}_t w_{i,j}\right)^2 = 2\left(\mathfrak{A}_t w_{i,j}\right)^2 - \mathfrak{A}_t\left(w_{i,j}\right)^2,$$

[21] See Remark 4.7.

which when used in (4.36) leads to

(4.37a)

$$\frac{1}{2}\mathfrak{C}_t\left[(v_{i,j})^2 + \mathfrak{A}_x(\mathfrak{G}_t w_{i,j})^2 + 2\mathfrak{A}_t G(u_{i,j})\right]$$
$$- \frac{k}{h}\mathfrak{C}_x\left[(\mathfrak{A}_t\mathfrak{A}_x v_{i,j})\cdot w_{i,j}\right]$$
$$= \frac{1}{2}\mathfrak{C}_t\left[v_{i,j}^2 + \mathfrak{A}_x\left(2(\mathfrak{A}_t w_{i,j})^2 - \mathfrak{A}_t w_{i,j}^2\right) + 2\mathfrak{A}_t G(u_{i,j})\right]$$
$$- \frac{k}{h}\mathfrak{C}_x\left[(\mathfrak{A}_t\mathfrak{A}_x v_{i,j})\cdot w_{i,j}\right]$$
$$= (\mathfrak{C}_t v_{i,j})(\mathfrak{A}_t v_{i,j})$$
$$+ \mathfrak{A}_x\left(2(\mathfrak{C}_t\mathfrak{A}_t w_{i,j})(\mathfrak{A}_t^2 w_{i,j}) - \mathfrak{A}_t(\mathfrak{C}_t w_{i,j})(\mathfrak{A}_t w_{i,j})\right) + \mathfrak{A}_t\mathfrak{C}_t G(u_{i,j})$$
$$- \frac{k}{h}\left[(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x v_{i,j})(\mathfrak{A}_x w_{i,j}) + (\mathfrak{A}_t\mathfrak{A}_x^2 v_{i,j})(\mathfrak{C}_x w_{i,j})\right] = 0,$$

with the help of (3.22b) and then Leibniz rule (3.15). Next, by (3.19) of Lemma 3.4,

$$\mathfrak{A}_t\left[(\mathfrak{C}_t w_{i,j})(\mathfrak{A}_t w_{i,j})\right] = (\mathfrak{A}_t\mathfrak{C}_t w_{i,j})(\mathfrak{A}_t^2 w_{i,j}) + \frac{1}{4}(\mathfrak{C}_t^2 w_{i,j})(\mathfrak{C}_t\mathfrak{A}_t w_{i,j}), \tag{4.38}$$

making (4.37a) into

(4.37b)

$$(\mathfrak{C}_t v_{i,j})(\mathfrak{A}_t v_{i,j}) + \mathfrak{A}_x\left[(\mathfrak{A}_t^2 w_{i,j})(\mathfrak{C}_t\mathfrak{A}_t w_{i,j}) - \frac{1}{4}(\mathfrak{C}_t^2 w_{i,j})(\mathfrak{C}_t\mathfrak{A}_t w_{i,j})\right] + \mathfrak{A}_t\mathfrak{C}_t G(u_{i,j})$$
$$- \frac{k}{h}\left[(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x v_{i,j})(\mathfrak{A}_x w_{i,j}) + (\mathfrak{A}_t\mathfrak{A}_x^2 v_{i,j})(\mathfrak{C}_x w_{i,j})\right] = 0,$$

by virtue of the commutativity between $\mathfrak{A}_t$ and $\mathfrak{C}_t$ established in Lemma 3.2. Since by (3.22a)

$$\mathfrak{A}_t^2 = \frac{1}{4}\mathfrak{C}_t^2 + \mathbf{1},$$

it follows that the second term in (4.37b) becomes

$$\left(\mathfrak{A}_t^2 w_{i,j} - \frac{1}{4}\mathfrak{C}_t^2 w_{i,j}\right)\cdot(\mathfrak{C}_t\mathfrak{A}_t w_{i,j}) = w_{i,j}\cdot(\mathfrak{C}_t\mathfrak{A}_t w_{i,j}). \tag{4.39}$$

A substitution of (4.36) into (4.37b) together with another application of Lemma 3.4 yields

(4.37c)

$$\mathfrak{C}_t v_{i,j}\cdot\mathfrak{A}_t v_{i,j} + \mathfrak{A}_x\left[w_{i,j}(\mathfrak{C}_t\mathfrak{A}_t w_{i,j})\right] + \mathfrak{A}_t\mathfrak{C}_t G(u_{i,j})$$
$$- \frac{k}{h}\left[(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x v_{i,j})(\mathfrak{A}_x w_{i,j}) + (\mathfrak{A}_t\mathfrak{A}_x^2 v_{i,j})(\mathfrak{C}_x w_{i,j})\right]$$
$$= \mathfrak{C}_t v_{i,j}\cdot\mathfrak{A}_t v_{i,j} + (\mathfrak{A}_x w_{i,j})(\mathfrak{A}_x\mathfrak{A}_t\mathfrak{C}_t w_{i,j}) + \frac{1}{4}(\mathfrak{C}_x w_{i,j})\cdot(\mathfrak{C}_x\mathfrak{C}_t\mathfrak{A}_t w_{i,j})$$
$$- (\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_t w_{i,j})(\mathfrak{A}_x w_{i,j}) - \frac{k}{h}(\mathfrak{A}_t\mathfrak{A}_x^2 v_{i,j})(\mathfrak{C}_x w_{i,j}) = 0,$$

in which the next to last term was obtained using (3.9), (3.10b), and the commutativity between $\mathfrak{C}_x$ and $\mathfrak{C}_t$ established in Lemma 3.1. Similarly, the last term in (4.37c) takes the form

$$(4.40)\quad \begin{aligned} -\frac{k}{h}\left(\mathfrak{A}_t\mathfrak{A}_x^2 v_{i,j}\right)\left(\mathfrak{C}_x w_{i,j}\right) &= -\frac{1}{h}\left(\mathfrak{A}_t\mathfrak{A}_x^2\mathfrak{C}_t u_{i,j}\right)\left(\mathfrak{C}_x w_{i,j}\right) \\ &= -\frac{1}{4}\left(\mathfrak{A}_t\mathfrak{C}_t\mathfrak{C}_x w_{i,j}\right)\left(\mathfrak{C}_x w_{i,j}\right) - \frac{k}{h}\left(\mathfrak{A}_t v_{i,j}\right)\left(\mathfrak{C}_x w_{i,j}\right), \end{aligned}$$

where the first equality is obtained by using (3.9), and the second equality by using (3.22a), the commutativity $\mathfrak{C}_x$ and $\mathfrak{C}_t$, (3.9) and (3.10b). Relation (4.40) enables us to simplify (4.37c) further to

$$(4.37d)\quad \mathfrak{C}_t v_{i,j}\cdot\mathfrak{A}_t v_{i,j} - \frac{k}{h}\left(\mathfrak{A}_t v_{i,j}\right)\left(\mathfrak{C}_x w_{i,j}\right) + \mathfrak{A}_t\mathfrak{C}_t G(u_{i,j}) = 0.$$

Finally, we obtain (4.35b) from (4.37d) by virtue of (4.35a), (3.22c), and by assuming that $\mathfrak{H}_t u_{i,j} \neq 0$.

By defining the discrete energy for **ALGO_SV** at half time step as

$$(4.41a)\quad \mathbf{E}^{(SV)}_{N,j+\frac{1}{2}} := h\sum_{j=-N}^{N}\mathcal{E}^{(SV)}_{i,j+\frac{1}{2}},$$

where the energy density $\mathcal{E}^{(SV)}_{i,j+1/2}$ is defined at grid point $(i, j+1/2)$ as

$$(4.41b)\quad \mathcal{E}^{(SV)}_{i,j+\frac{1}{2}} := \frac{1}{2}\left[\left(v_{i,j+\frac{1}{2}}\right)^2 + \mathfrak{A}_x\left(\mathfrak{G}_t w_{i,j+\frac{1}{2}}\right)^2 + 2\mathfrak{A}_t G(u_{i,j+\frac{1}{2}})\right],$$

which can be recognized as coming from the first term of (4.36), one obtains an immediate consequence of Theorem 4.8 in the following form.

COROLLARY 4.2. **ALGO_SV** *preserves the discrete energy* $\mathbf{E}^{(SV)}_{N,\cdot}$ *defined in* (4.41), *i.e.,*

$$(4.42a)\quad \mathbf{E}^{(SV)}_{N,j+\frac{1}{2}} = \mathbf{E}^{(SV)}_{N,\frac{1}{2}} \quad \forall j\in \underline{M}^+,$$

under the assumption at the boundaries that[22]

$$(4.42b)\quad \left(\mathfrak{A}_t\mathfrak{A}_x v_{N+\frac{1}{2},j}\right)\cdot w_{N+\frac{1}{2},j} = \left(\mathfrak{A}_t\mathfrak{A}_x v_{-N-\frac{1}{2},j}\right)\cdot w_{-N-\frac{1}{2},j}.$$

Remark 4.7. (1) With the conventional notation, (4.40a) reads as follows:

(4.43a)

$$\begin{aligned} \mathbf{E}^{(SV)}_{N,j+\frac{1}{2}} = \frac{h}{2}\sum_{i=-N}^{N}\Big\{ &\left(u_{i,j+1}-u_{i,j}\right)^2 + \frac{1}{2}[(u_{i+1,j+1}-u_{i,j+1})(u_{i+1,j}-u_{i,j}) \\ &+ (u_{i,j+1}-u_{i-1,j+1})(u_{i,j}-u_{i-1,j})] \\ &+ G(u_{i,j+1}) + G(u_{i,j})\Big\}, \end{aligned}$$

[22]Condition (4.42b) comes from the second term of (4.36).

whereas the energy expression given in Strauss and Vazquez [1978] is

$$\mathbf{E}_j = \frac{h}{2} \sum_{i=-N}^{N} \left\{ (u_{i,j+1} - u_{i,j})^2 + (u_{i+1,j+1} - u_{i,j+1})(u_{i+1,j} - u_{i,j}) + G(u_{i,j+1}) + G(u_{i,j}) \right\} . \tag{4.43b}$$

Comparing (4.43a) with (4.43b), one can see that the spatial derivative term in (4.43b) is not evaluated at point $(i, j + \frac{1}{2})$, but instead at point $(i + \frac{1}{2}, j + \frac{1}{2})$, which is not consistent with the other two terms of the discrete energy expression. Due to this inconsistency, difficulties will arise in an attempt to recover the local conservation form of the algorithm.

(2) Since $\mathfrak{A}_x(\mathfrak{G}_t w_{i,j+1/2})^2$ is not necessarily greater than zero,[23] the "energy" represented by either (4.43a) or (4.43b) is not necessarily positive definite; thus, the algorithm is not unconditionally stable, which confirms the stability analysis in the linear situation (Vu-Quoc and Li [1993]). This is the main drawback of **ALGO_SV**.

4.4. Linear-momentum-conserving algorithm. Instead of designing algorithms that preserve the local energy conservation law (2.2ab), one can also design algorithms that preserve another conservation law, such as the local linear-momentum conservation law (2.6ab). One such algorithm is presented below.

Algorithm IV (ALGO_LM). *We approximate the NLKGE at a grid point (i, j) by the algebraic equation*

$$\boxed{\begin{aligned} &\frac{1}{2h}\mathfrak{H}_x u_{i,j} =: \mathfrak{A}_x w_{i,j}, \\ &\frac{1}{k}\mathfrak{C}_t v_{i,j} - \frac{1}{h}\mathfrak{C}_x w_{i,j} + \frac{\mathfrak{H}_x G(u_{i,j})}{\mathfrak{H}_x u_{i,j}} = 0 \, . \end{aligned}} \tag{4.44ab}$$

Note the differences between (4.44ab) and (4.35a) and (4.35b). But **ALGO_LM** is very similar to **ALGO_SV**:

$$\boxed{\begin{aligned} &\frac{1}{k^2}(u_{i,j+1} - 2u_{i,j} + u_{i,j-1}) - \frac{1}{h^2}(u_{i+1,j} - 2u_{i,j} + u_{i-1,j}) \\ &\qquad + \frac{G(u_{i+1,j}) - G(u_{i-1,j})}{u_{i+1,j} - u_{i-1,j}} = 0. \end{aligned}} \tag{4.44c}$$

Comparing (4.44c) with (4.35c), the difference is only in the last term; we refer to Vu-Quoc and Li [1993] for a geometric interpretation of these algorithms. The counterpart of Theorem 4.7 for **ALGO_LM** is the following.

THEOREM 4.9. *The nonlinear algebraic operator of algorithm* **ALGO_LM** *described in* (4.44ab) *possesses the following discrete local linear-momentum conservation law at grid point* (i, j):

(4.45)

$$\boxed{\frac{1}{k}\mathfrak{C}_t \left[(\mathfrak{A}_t \mathfrak{A}_x w_{i,j}) \cdot v_{i,j}\right] - \frac{1}{2h}\mathfrak{C}_x \left[\mathfrak{A}_t (\mathfrak{G}_x v_{i,j})^2 + (w_{i,j})^2 - 2\mathfrak{A}_x G(u_{i,j})\right] = 0.}$$

[23] See Definition 3.4.

Proof. Applying Leibniz rule (3.15), (3.22b), and the commutativity (3.13a) between $\mathfrak{C}_x$ and $\mathfrak{A}_x$ to (4.45), we obtain

(4.46a)

$$
\begin{aligned}
&\frac{h}{k}\mathfrak{C}_t\left[(\mathfrak{A}_t\mathfrak{A}_x w_{i,j})\cdot v_{i,j}\right] - \frac{1}{2}\mathfrak{C}_x\left[\mathfrak{A}_t\left(\mathfrak{G}_x v_{i,j}\right)^2 + (w_{i,j})^2 - 2\mathfrak{A}_x G(u_{i,j})\right]\\
&\qquad = \frac{h}{k}\left[(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_t w_{i,j})(\mathfrak{A}_t v_{i,j}) + \left(\mathfrak{A}_t^2\mathfrak{A}_x w_{i,j}\right)(\mathfrak{C}_t v_{i,j})\right]\\
&\qquad\quad - \frac{1}{2}\mathfrak{C}_x\left\{\mathfrak{A}_t\left[2\left(\mathfrak{A}_x v_{i,j}\right)^2 - \mathfrak{A}_x\left(v_{i,j}\right)^2\right] + (w_{i,j})^2 - 2\mathfrak{A}_x G(u_{i,j})\right\}\\
&\qquad = \frac{h}{k}\left[(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_t w_{i,j})(\mathfrak{A}_t v_{i,j}) + \left(\mathfrak{A}_t^2\mathfrak{A}_x w_{i,j}\right)(\mathfrak{C}_t v_{i,j})\right]\\
&\qquad\quad - \mathfrak{A}_t\left[2\left(\mathfrak{A}_x\mathfrak{C}_x v_{i,j}\right)\left(\mathfrak{A}_x^2 v_{i,j}\right) - \mathfrak{A}_x\left(\mathfrak{C}_x v_{i,j}\right)\left(\mathfrak{A}_x v_{i,j}\right)\right]\\
&\qquad\quad - \left(\mathfrak{C}_x w_{i,j}\right)\left(\mathfrak{A}_x w_{i,j}\right) + \mathfrak{A}_x\mathfrak{C}_x G(u_{i,j}) = 0.
\end{aligned}
$$

Using the averaging rule (3.19) for function products, the commutativity (3.13a) between $\mathfrak{C}_x^n$ and $\mathfrak{A}_x$, (3.9) and (3.10b) so that $\frac{h}{k}\mathfrak{C}_t w_{i,j} \equiv \mathfrak{C}_x v_{i,j}$, the identity (3.22a), where $\mathfrak{A}^2 - \frac{1}{4}\mathfrak{C}^2 \equiv \mathbf{1}$, the averaging rule (3.19) of function products, one can then transform (4.46a) into

(4.46b)

$$
\begin{aligned}
&\frac{h}{k}\left[(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_t w_{i,j})\,(\mathfrak{A}_t v_{i,j}) + \left(\mathfrak{A}_t^2\mathfrak{A}_x w_{i,j}\right)(\mathfrak{C}_t v_{i,j})\right]\\
&\qquad - \mathfrak{A}_t\left[\left(\mathfrak{A}_x\mathfrak{C}_x v_{i,j}\right)\left(\left(\mathfrak{A}_x^2 v_{i,j}\right) - \frac{1}{4}\left(\mathfrak{C}_x^2 v_{i,j}\right)\right)\right]\\
&\qquad - \left(\mathfrak{C}_x w_{i,j}\right)\left(\mathfrak{A}_x w_{i,j}\right) + \mathfrak{A}_x\mathfrak{C}_x G(u_{i,j})\\
&\quad = \left(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x v_{i,j}\right)\left(\mathfrak{A}_t v_{i,j}\right) + \frac{h}{k}\left(\mathfrak{A}_t^2\mathfrak{A}_x w_{i,j}\right)\left(\mathfrak{C}_t v_{i,j}\right)\\
&\qquad - \mathfrak{A}_t\left[\left(\mathfrak{C}_x\mathfrak{A}_x v_{i,j}\right)\left(v_{i,j}\right)\right] - \left(\mathfrak{C}_x w_{i,j}\right)\left(\mathfrak{A}_x w_{i,j}\right) + \mathfrak{A}_t\mathfrak{C}_x G(u_{i,j})\\
&\quad = \left(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x v_{i,j}\right)\left(\mathfrak{A}_t v_{i,j}\right) + \frac{h}{4k}\left(\mathfrak{A}_x\mathfrak{C}_t^2 w_{i,j}\right)\left(\mathfrak{C}_t v_{i,j}\right)\\
&\qquad + \frac{h}{k}\left(\mathfrak{A}_x w_{i,j}\right)\left(\mathfrak{C}_t v_{i,j}\right) - \left(\mathfrak{A}_t\mathfrak{A}_x\mathfrak{C}_x v_{i,j}\right)\left(\mathfrak{A}_t v_{i,j}\right)\\
&\qquad - \frac{1}{4}\left(\mathfrak{C}_t\mathfrak{C}_x\mathfrak{A}_x v_{i,j}\right)\left(\mathfrak{C}_t v_{i,j}\right) - \left(\mathfrak{C}_x w_{i,j}\right)\left(\mathfrak{A}_x w_{i,j}\right) + \mathfrak{A}_x\mathfrak{C}_x G(u_{i,j})\\
&\quad = \frac{h}{k}\left(\mathfrak{A}_x w_{i,j}\right)\left(\mathfrak{C}_t v_{i,j}\right) - \left(\mathfrak{C}_x w_{i,j}\right)\left(\mathfrak{A}_x w_{i,j}\right) + \mathfrak{A}_x\mathfrak{C}_x G(u_{i,j}) = 0.
\end{aligned}
$$

Note that before reaching the last equation in (4.46b), the first and the fourth terms cancel each other, and similarly for the second and the fifth terms. By (4.44a) $\mathfrak{A}_x w_{i,j} := \frac{1}{2h}\mathfrak{H}_x u_{i,j}$. (4.46b) can next be recast into

$$
\left(\frac{1}{k}\mathfrak{C}_t v_{i,j}\right)\left(\frac{1}{2}\mathfrak{H}_x u_{i,j}\right) - \left(\frac{1}{h}\mathfrak{C}_x w_{i,j}\right)\left(\frac{1}{2}\mathfrak{H}_x u_{i,j}\right) + \mathfrak{A}_x\mathfrak{C}_x G(u_{i,j}) = 0. \tag{4.46c}
$$

Assuming that $\frac{1}{2}\mathfrak{H}_x u_{i,j} \neq 0$, we can divide (4.46c) throughout by $\frac{1}{2}\mathfrak{H}_x u_{i,j}$ and since $\mathfrak{H}_x = 2\mathfrak{A}_x\mathfrak{C}_x$ by (3.22c), we obtain (4.44b).

As pointed out in §2.2, with the appropriate boundary condition (2.8), the momentum integral exists. We now demonstrate that **ALGO_LM** can preserve the discrete version of this first integral using the discrete version of the boundary condition (2.8). First, let us define

$$
w_{i,j+\frac{1}{2}} := \frac{1}{h}\mathfrak{C}_x u_{i,j+\frac{1}{2}}. \tag{4.47}
$$

Employing (4.47) and the definition (4.24) of $v_{i,j+1/2}$, from the first term of (4.45), we can define the linear momentum density at point $(i, j+\frac{1}{2})$, which can be evaluated using finite difference solutions at the regular grid points as follows:

(4.48)
$$\begin{aligned} \mathcal{M}_{i,j+\frac{1}{2}} &:= v_{i,j+\frac{1}{2}} \left(\mathfrak{A}_t \mathfrak{A}_x w_{i,j+\frac{1}{2}}\right) \\ &= \frac{1}{k}(u_{i,j+1} - u_{i,j}) \cdot \frac{1}{2h}(\mathfrak{A}_t \mathfrak{H}_x u_{i,j+\frac{1}{2}}) \\ &= \frac{1}{4hk}(u_{i,j+1} - u_{i,j}) \left[(u_{i+1,j+1} - u_{i-1,j+1}) + (u_{i+1,j} - u_{i-1,j})\right], \end{aligned}$$

where we have made use of (3.22c) in the second factor of the second equation. Also, from the second term of (4.45), define $\mathcal{J}$ at point $(i+\frac{1}{2}, j)$ as

(4.49)
$$\begin{aligned} \mathcal{J}_{i+\frac{1}{2},j} &:= \frac{1}{2}\left[\mathfrak{A}_t \left(\mathfrak{G}_x v_{i+\frac{1}{2},j}\right)^2 + \left(w_{i+\frac{1}{2},j}\right)^2 - 2\mathfrak{A}_x G(u_{i+\frac{1}{2},j})\right] \\ &= \frac{1}{4k}\left[(u_{i+1,j+1} - u_{i+1,j})(u_{i,j+1} - u_{i,j}) + (u_{i+1,j} - u_{i+1,j-1})(u_{i,j} - u_{i,j-1})\right] \\ &\quad + \frac{1}{2h^2}(u_{i+1,j} - u_{i,j})^2 - \frac{1}{2}\left[G(u_{i+1,j}) + G(u_{i,j})\right], \end{aligned}$$

where the second equation is obtained using (3.3a), (3.3b), (4.47), (3.4a), and (4.24). Thus, (4.45) can be rewritten as

(4.50a)
$$\frac{\mathfrak{C}_t}{k}\mathcal{M}_{i,j} - \frac{\mathfrak{C}_x}{h}\mathcal{J}_{i,j} = 0$$

or

(4.50b)
$$\frac{1}{k}[\mathcal{M}_{i,j+\frac{1}{2}} - \mathcal{M}_{i,j-\frac{1}{2}}] - \frac{1}{h}[\mathcal{J}_{i+\frac{1}{2},j} - \mathcal{J}_{i-\frac{1}{2},j}] = 0.$$

Multiplying (4.50b) by h and summing i from $-N$ to N, we obtain

(4.51a)
$$h\sum_{i=-N}^{N}(\mathcal{M}_{i,j+\frac{1}{2}} - \mathcal{M}_{i,j-\frac{1}{2}}) = k\{-\mathcal{J}_{-N-\frac{1}{2},j} + \mathcal{J}_{N+\frac{1}{2},j}\}.$$

Denote

(4.52)
$$\mathbf{M}_{N,j\pm\frac{1}{2}} := h\sum_{i=-N}^{N}\mathcal{M}_{i,j\pm\frac{1}{2}};$$

then (4.51a) becomes

(4.51b)
$$\mathbf{M}_{N,j+\frac{1}{2}} - \mathbf{M}_{N,j-\frac{1}{2}} = k\left\{\mathcal{J}_{i+\frac{1}{2},j}\right\}\Big|_{i=-N-1}^{i=N}$$

which in turn leads to Theorem 4.10.

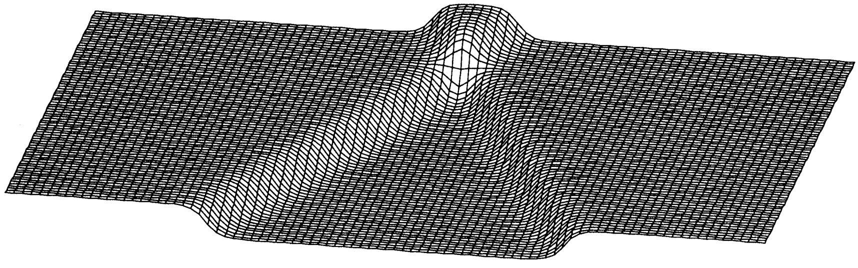

FIG. 5.1a. *Soliton collision for the sine-Gordon equation.*

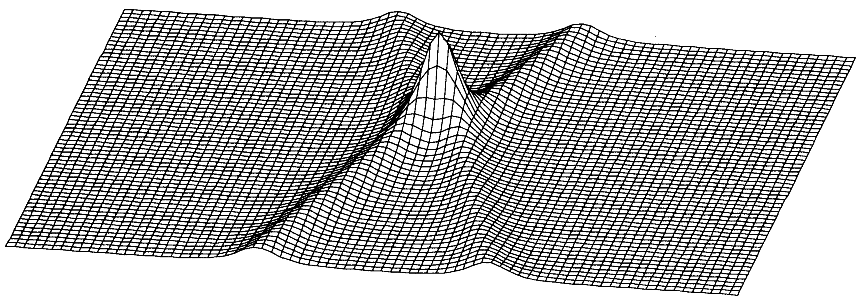

FIG. 5.1b. *Negative velocity profile—i.e., $-v(x,t)$—for the sine-Gordon kink-antikink pair.*

THEOREM 4.10 (linear momentum conservation). *The algorithm described in* (4.44ab) *conserves the discrete momentum* $\mathbf{M}_{N,j+1/2}$ *defined in* (4.52) *in the sense that*

$$\mathbf{M}_{N,j+\frac{1}{2}} = \mathbf{M}_{N,j-\frac{1}{2}} \quad \forall j, \tag{4.53a}$$

provided that the following boundary condition holds:

$$\mathcal{J}_{-N-\frac{1}{2},j} = \mathcal{J}_{N+\frac{1}{2},j} \qquad \forall j. \tag{4.53b}$$

5. Numerical experiments and discussions. We present numerical experiments that illustrate the conservation features of the proposed algorithms as presented herein. To this end, the elastic collision of two solitary waves of the sine-Gordon equation is analyzed using **ALGO_E1**. Detailed information on the computer implementation of the algorithms presented here and further numerical examples can be found in Vu-Quoc and Li [1993].[24]

Example 5.1. *Collision of sine-Gordon solitons.*

For the sine-Gordon equation, the potential function G takes the form

$$G(U) = 1 - \cos U, \tag{5.1}$$

[24]The proposed algorithms can be applied to other problems studied in Rodriguez-Plaza and Vazquez [1990], Sanchez, Vasquez, and Konotop [1991].

TABLE 5.1.
Collision of sine-Gordon solitons: Total discrete energy versus time.

Time	Energy
0.000000000000	16.769139559417
2.000000000000	16.769139559417
4.000000000000	16.769139559417
6.000000000000	16.769139559417
8.000000000000	16.769139559417
10.000000000000	16.769139559417
⋮	⋮
60.000000000000	16.769139559417
62.000000000000	16.769139559417
64.000000000000	16.769139559417
66.000000000000	16.769139559417
68.000000000000	16.769139559417
70.000000000000	16.769139559417
⋮	⋮
110.000000000000	16.769139559417
112.000000000000	16.769139559417
114.000000000000	16.769139559417
116.000000000000	16.769139559417
118.000000000000	16.769139559417
120.000000000000	16.769139559417

with the following soliton (kink) solution

$$U(x,t) = 4\tan^{-1}\left[\exp\left(\pm\frac{x-\beta t}{\sqrt{1-\beta^2}}\right)\right]. \tag{5.2}$$

The initial conditions are generated with the kink and antikink solutions moving toward each other with equal and opposite velocities. The result of the collision between these two solitary waves is shown in Fig. 5.1a with the time axis pointing toward the reader. In the computation, we employ the step sizes $h = 0.1$ and $k = 0.2$. It should be noted that $\lambda = \frac{k}{h} = 2 > 1$ here. The negative velocity profile—i.e., $-v(x,t)$ is plotted instead of $v(x,t)$—is shown in Fig. 5.1b.

Along with the computation for the displacement $u(x,t)$ and the velocity $v(x,t)$, we monitor the change in the total energy, linear momentum, and angular momentum. To verify further the numerical results and the robustness of the algorithms in the elastic collision of the sine-Gordon solitary waves, we recompute the numerical solution using the spatial step size $h = 0.1$ together with three different time step sizes $\lambda = \frac{k}{h} =$ 0.5, 1.0, and 2.0. The computed results remain virtually the same for the three cases regarding the displacement profile, velocity profile, and the system's total energy at each time step. We record below the numerical results obtained for the total discrete energy (Table 5.1), the linear momentum (Table 5.2), and the angular momentum (Table 5.3); these results are obtained using the ratio of time step size ($k = 0.2$) over spatial step size ($h = 0.1$) of $\lambda = \frac{k}{h} = 2$. From Table 5.1, it can be seen that the proposed **ALGO_E1** preserves the exact total energy.

Remark 5.1. We note here that the results in Table 5.1 are more accurate than the previously reported results in Vu-Quoc and Li [1993] due to the use of an equivalent algebraic expression to avoid numerical round-off error as follows:

$$\frac{G(u_{i,j+1}) - G(u_{i,j})}{u_{i,j+1} - u_{i,j}} = -\frac{\cos(u_{i,j+1}) - \cos(u_{i,j})}{u_{i,j+1} - u_{i,j}} \tag{5.3a}$$

TABLE 5.2.
Collision of sine-Gordon solitons: Linear momentum versus time.

Time	Linear momentum
0.000000000000	0.000000000000
2.000000000000	0.000000000000
4.000000000000	0.000000000000
6.000000000000	0.000000000000
8.000000000000	0.000000000000
10.000000000000	0.000000000000
⋮	⋮
60.000000000000	0.000000000000
62.000000000000	0.000000000000
64.000000000000	0.000000000000
66.000000000000	0.000000000000
68.000000000000	0.000000000000
70.000000000000	0.000000000000
⋮	⋮
110.000000000000	0.000000000001
112.000000000000	0.000000000001
114.000000000000	0.000000000001
116.000000000000	0.000000000001
118.000000000000	0.000000000001
120.000000000000	0.000000000001

(5.3b)
$$= -\frac{2\sin\frac{1}{2}(u_{i,j+1}+u_{i,j})\sin\frac{1}{2}(u_{i,j+1}-u_{i,j})}{u_{i,j+1}-u_{i,j}}.$$

In finite mathematics of computation, even with double precision, there is a significant difference, due to round-off error, between (5.3a) and (5.3b) as $(u_{i,j+1} - u_{i,j}) \to 0$; the results obtained from using (5.3b), which is approximated by $(-\sin\frac{1}{2}(u_{i,j+1}+u_{i,j}))$, are more accurate.

The fluctuation of the linear and angular momenta is also very small, as can be seen from Tables 5.2 and 5.3. In fact, these results indicate that **ALGO_E1** can also preserve the linear and angular momenta accurately (up to more than 10 digits in the present computation). Actually, we did not even expect this nice feature for **ALGO_E1**. The above numerical behavior of the proposed algorithms is verified in all numerical experiments that we performed. The results in Tables 5.1–5.3 with $\lambda = 2$ confirm the remarkable stability robustness of **ALGO_E1** as established in the paper.

Example 5.2. *Collision of the ϕ^4_- solitary waves.* The potential function for this example is

(5.4)
$$G(u) = -\frac{1}{2}m^2u^2 + \frac{\gamma}{4}u^4.$$

A soliton-like solution is (see Kudryavtsev [1975])

(5.5)
$$U(x,t) = \frac{m}{\sqrt{\gamma}}\tanh\left(\pm\frac{m(x-\beta t)}{\sqrt{2(1-\beta^2)}}\right).$$

The computed displacement $u(x,t)$ shown in Fig. 5.2 was produced with $\gamma = \frac{1}{\pi^2}$, $m = 1$, $x_0 = 3$, $\beta = 0.2$, together with the step sizes $h = 0.1$ and $k = 0.1$.

Remark 5.2. (1) It is noted that the potential function $G(u)$ here is not positive definite. In the actual numerical computation, we find that the energy $\mathbf{E}^{(E1)}_{N,j}$ is

TABLE 5.3.
Collision of sine-Gordon solitons: Angular momentum versus time.

Time	Angular momentum
0.000000000000	0.000000000000
2.000000000000	0.000000000000
4.000000000000	0.000000000000
6.000000000000	0.000000000001
8.000000000000	0.000000000001
10.000000000000	0.000000000001
⋮	⋮
60.000000000000	-0.000000000003
62.000000000000	-0.000000000002
64.000000000000	-0.000000000001
66.000000000000	-0.000000000005
68.000000000000	-0.000000000012
70.000000000000	-0.000000000012
⋮	⋮
110.000000000000	0.000000000033
112.000000000000	0.000000000029
114.000000000000	0.000000000040
116.000000000000	0.000000000041
118.000000000000	0.000000000037
120.000000000000	0.000000000030

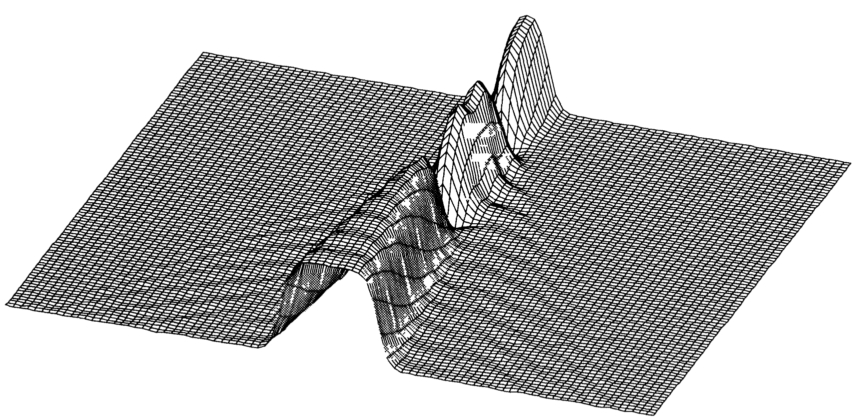

FIG. 5.2. *The soliton-like solution for the ϕ^4_- equation.*

negative for the above example. It is not difficult, however, to verify that under the condition $m^2 \leq 1$, the norm-like measure $|||\underline{u}_l|||_1$ defined in (4.16b) remains positive definite. In fact, when $m = 1$, it can be shown that

$$|||\underline{u}_l|||_1 \sim \frac{\sqrt{\gamma}}{2}||\underline{u}_l||^2, \quad \text{with } \gamma > 0.$$

Thus, the results for nonlinear stability hold for the present example.

(2) The function $G(u)$ in (5.2) is bounded below; i.e., $G_{lb} = -\frac{m^4}{\gamma}$ for $\gamma \neq 0$. By the argument provided in Remark 4.4, nonlinear stability is guaranteed. Indeed, numerical results do confirm a good stability property (see Fig. 5.2).

Finally, as mentioned, we attribute the invariant property of the proposed algorithms to the symmetry property of finite difference calculus. For the NLKGE, the conservation of energy is a direct consequence of the invariance of the system's Lagrangian under time variable translation, whereas the conservation of linear momentum is a consequence of the invariance of the Lagrangian under spatial variable translation (see Strauss [1978]). Thus, in general, to construct a discrete Lagrangian equation for NLKGE and to use the invariant condition under finite or periodic variation to derive invariant algorithms would be a fascinating task in the future.

6. Closure. We have presented an analysis of nonlinear stability of a class of invariant-conserving algorithms within the context of exact finite difference calculus. Discrete invariant expressions corresponding to each of the proposed algorithms are derived. Numerical results are presented to corroborate the theoretical findings on the stability robustness of these invariant-conserving algorithms.

Acknowledgments. The financial support of the National Science Foundation and the Space Research Assistantship Program of the Florida Space Grant Consortium is gratefully acknowleged.

REFERENCES

M. J. Ablowitz, M. D. Kruskal, and J. F. Ladik [1979], *Solitary wave collisions*, SIAM J. Appl. Math., 36, pp. 428–437.

A. Barone, F. Esposito, and J. Magee [1971], *Theory and applications of the sine-Gordon equation*, Riv. Del Nuovo Cimento, 1, pp. 227–267.

Y. S. Chin and C. Qin [1989], *Explicit energy-conserving schemes for the three-body problem*, J. Comput. Phys., 83, pp. 485–493.

R. Courant, K. O. Friedrichs, and H. Lewy [1928], *Über die partiellen differenzengleichungen der mathematischen physik*, Math. Ann., 100, pp. 32–74.

R. K. Dodd, J. C. Eilbeck, J. D. Gibbon, and H. C. Morris [1982], *Solitons and Nonlinear Wave Equations*, Academic Press, London.

Z. Fei and L. Vazquez [1991a], *Two energy-conserving numerical schemes for the sine-Gordon equation*, Appl. Math. Comput., 45, pp. 17–30.

——— [1991b], *Some conservative numerical schemes for an ordinary differential equation*, Appl. Comput. Math. (Matematica Aplicada e Computacional, Brazil), 10, pp. 59–69.

K. Feng [1984], *On difference schemes and symplectic geometry*, in Proc. Symposium on Differential Geometry and Differential Equations, F. Kang, ed., Science Press, Beijing, 1984, pp. 42–58.

K. Feng and M. Z. Qin [1986], *Difference schemes for Hamiltonian formalism and sympletic geometry*, J. Comput. Math., 4, pp. 279–288.

G. H. Golub and C. F. Van Loan [1989], *Matrix Computations*, Second ed., The Johns Hopkins University Press, Baltimore, London.

D. Greenspan [1984], *Conservative numerical methods for* $\ddot{x} = f(x)$, J. Comput. Phys., 56, pp. 28–41.

M. Huang [1991], *A Hamiltonian approximation to simulate solitary waves of the Korteweg-de Vries equation*, Math. Comp., 56, pp. 607–620.

T. J. R. Hughes, T. K. Caughey, and W. K. Liu [1978], *Finite element methods for nonlinear elastodynamics which conserve energy*, ASME J. Appl. Mech., 45, pp. 366–370.

S. Jimenez and L. Vazquez [1990], *Analysis of four numerical schemes for a nonlinear Klein-Gordon equation*, Appl. Math. Comput., 35, pp. 61–94.

C. Jordan [1950], *Calculus of Finite Differences*, 2nd ed., Chelsea, New York.

L. V. Kantorovich and G. P. Akilov [1982], *Functional Analysis*, 2nd ed., Pergamon Press, Elmsford, NY, Oxford.

A. E. Kudryavtsev [1975], *Soliton-like solutions for a Higgs scalar field,* JETP Lett., 22, pp. 82–83.

V. Lakshmikantham and D. Trigiante [1988], *Theory of Difference Equations: Numerical Methods and Applications*, Academic Press, New York.

G. L. Lamb, Jr [1980], *Elements of Soliton Theory*, John Wiley, New York.

L. M. Milne-Thomson [1933], *The Calculus of Finite Differences*, Macmillan and Co., Ltd., London.

J. P. Olver [1986], *Application of Lie Groups to Differential Equations*, Graduate Texts in Mathematics 107, Springer-Verlag, New York.

R. D. RICHTMYER AND K. W. MORTON [1967], *Difference Methods for Initial Value Problems*, 2nd ed., Wiley Interscience, New York.

M. J. RODRIGUEZ-PLAZA AND L. VAZQUEZ [1990], *Additive and multiplicative perturbations on Φ^4 kink,* Phys. Rev. B, 41, pp. 11437–11448.

A. SANCHEZ, L. VAZQUEZ, AND V. V. KONOTOP [1991], *Dynamics of a Φ^4 kink in the presence of strong potential fluctuations, dissipation, and boundaries*, Phys. Rev. A, 44, pp. 1086–1103.

J. C. SIMO AND K. K. WONG [1991], *Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum*, Internat. J. Numer. Methods Engrg., 31, pp. 19–52.

J. C. SIMO, M. S. RIFAI, AND D. D. FOX [1992], *On a stress resultant geometrically exact shell model. Part* VI: *Conserving algorithms for non-linear dynamics*, Internat. J. Numer. Methods Engrg., 34, pp. 117–164.

G. D. SMITH [1985], *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Oxford, UK.

H. STEUDEL [1975], *Noether's theorem and higher conservation laws in ultrashort pulse propagation*, Ann. Physik, pp. 217–226.

W. A. STRAUSS [1978], *Nonlinear invariant wave equations,* in Invariant Wave Equations, G. Velo and A. S. Wighman, eds., Springer-Verlag, New York, pp. 197–249.

W. STRAUSS AND L. VAZQUEZ [1978], *Numerical solution of a nonlinear Klein-Gordon equation,* J. Comput. Phys., 28, pp. 271–278.

L. VU-QUOC AND S. LI [1993], *Invariant-conserving finite difference algorithms for the nonlinear Klein-Gordon equation*, Comput. Methods Appl. Mech. Engrg., 107, pp. 341–391.

G. B. WHITHAM [1974], *Linear and Nonlinear Waves*, John Wiley, New York.