Chapter 10
High-Order Symmetric Hermite–Birkhoff Time Integrators for Semilinear KG Equations

The computation of the Klein–Gordon equation featuring a nonlinear potential function is of great importance in a wide range of application areas in science and engineering. It represents major challenges because of the nonlinear potential. The main aim of this chapter is to present symmetric and arbitrarily high-order time-stepping integrators and analyse their stability, convergence and long-time behaviour for the semilinear Klein–Gordon equation. To achieve this, under the assumption of periodic boundary conditions, an abstract ordinary differential equation (ODE) and its operator-variation-of-constants formula are formulated on a suitable function space based on operator spectrum theory. By applying a two-point Hermite–Birkhoff interpolation to the nonlinear integrals that appear in the operator-variation-of-constants formula, as a result, a suitable spatial discretisation leads to the fully discrete scheme, which needs only a weak temporal smoothness assumption.

10.1 Introduction

It is well known that the nonlinear wave equation plays a prominent role in a wide range of applications in engineering and science, including nonlinear optics, solid state physics and quantum field theory [1]. Most importantly, the *Klein–Gordon (KG) equation,* a relativistic counterpart of the Schrödinger equation, is used to model diverse nonlinear phenomena, such as the propagation of dislocations in crystals and the behaviour of elementary particles and of Josephson junctions (see Chap. 2 in [2] for details). Numerical computations play an important role in the study of nonlinear waves. We here restrict ourselves to the one-dimensional case, although all ideas, algorithms and analysis described in this chapter can be easily extended to the solution of semilinear KG equations in a moderate number of space dimensions.

X. Wu, B. Wang, *Geometric Integrators for Differential Equations with Highly Oscillatory Solutions*, https://doi.org/10.1007/978-981-16-0147-7_10

We now consider the following semilinear KG equation in a single space variable:

$$\begin{cases} u_{tt} - a^2 \Delta u = f(u), & t_0 < t \leqslant T, \quad x \in \Omega, \\ u(x, t_0) = \varphi_1(x), \quad u_t(x, t_0) = \varphi_2(x), & x \in \bar{\Omega}, \end{cases} \tag{10.1}$$

where $u(x, t)$ represents the wave displacement at position x and time t, and the nonlinear function $f(u)$ is the negative derivative of a potential energy $V(u) \geqslant 0$. Here it is assumed that the initial value problem (10.1) is equipped with the periodic boundary conditions on the domain $\Omega = (-\pi, \pi)$,

$$u(x, t) = u(x + 2\pi, t), \qquad x \in (-\pi, \pi], \tag{10.2}$$

where 2π is the fundamental period with respect to x. The semilinear KG equation (10.1) is used to model many different nonlinear phenomena, including the propagation of dislocations in crystals and the behaviour of elementary particles and of Josephson junctions (see Chap. 8.2 in [2] for details). In general, it has also been the subject of detailed investigation in studies of solitons and in nonlinear science. In the literature, there are various choices of the potential $f(u)$. Among typical examples is the best known sine-Gordon equation

$$u_{tt} - a^2 \Delta u + \sin(u) = 0,$$

and it also appears with polynomial $f(u)$, and other nonlinear functions. Another point is that, if $u(\cdot, t) \in H^1(\Omega)$ and $u_t(\cdot, t) \in L^2(\Omega)$, the energy conservation is a key feature of the KG equation (10.1) with periodic boundary condition (10.2), that is

$$E(t) = \frac{1}{2} \int_{\Omega} \left(u_t^2 + a^2 |\nabla u|^2 + 2V(u) \right) \mathrm{d}x \equiv E(t_0). \tag{10.3}$$

This is an essential property in the theory of solitons. Therefore, it is also very important to test the effectiveness of a numerical method for (10.1) for the preservation of the corresponding discrete energy.

The KG equation has received much attention in both its numerical and analytical aspects. With regard to analytical issues, the initial value problem (10.1) was investigated by many authors (see, e.g. [3–7]). In particular, for the defocusing case, $V(u) \geqslant 0, u \in \mathbb{R}$, the global existence of solutions was established in [3], and for the focusing case, $V(u) \leqslant 0, u \in \mathbb{R}$, possible finite time blow-up was investigated. In numerical analysis, various solution procedures have been proposed and studied including classical finite difference methods such as explicit, semi-implicit, compact finite difference and symplectic conservative discretisations [8–12]. Other effective numerical methods, such as the finite element method and the spectral method were also studied in [13–16]. Although various numerical methods for the semilinear KG equation have been derived and investigated in the literature, their accuracy is

limited, and little attention has been paid to the special structure brought by spatial discretisations.

It is known that recent interest in exponential integrators for semilinear parabolic problems has led to the development of numerical schemes (see, e.g. [17–21]). Motivated by this and based on the operator spectrum theory (see, e.g. [22]), we first formulate the nonlinear KG equation (10.1)–(10.2) as an abstract second-order ordinary differential equation. Then, the operator-variation-of-constants formula (also is termed the *Duhamel Principle*) for the abstract equation is introduced, which is in fact an implicit expression of the solution of the semilinear KG equation (see [23]). In a similar way to the useful approach to dealing with the semiclassical Schrödinger equation in [24], we forego the standard steps, of first semidiscretising and then dealing with the semidiscretisation, in a totally different approach which greatly reduces the requirement of the smoothness with respect to time. Employing the operator-variation-of-constants formula, we interpolate the nonlinear integrators by two-point Hermite interpolation, and then a class of symmetric and arbitrarily high-order time integration formulae is derived and analysed. In fact, the space semidiscretisation is deferred to the very last moment, and this helps us take a subtle but powerful advantage of dealing with the undiscretised operator Δ and incorporate the special structure brought by spatial discretisations into the underlying numerical integrator.

10.2 The Symmetric and High-Order Hermite–Birkhoff Time Integration Formula

In this section, using operator theory (see, e.g. [22]), we firstly formulate the nonlinear problem (10.1)–(10.2) as an abstract ordinary differential equation on the Hilbert space $L^2(\Omega)$. Then, the operator-variation-of-constants formula for the abstract equation is presented, which is in fact an implicit expression of the solution for the system (see, e.g. [23, 25]). Keeping the eventual discretisation in mind and applying Hermite–Birkhoff interpolation to the operator-variation-of-constants formula, we will present a class of symmetric and arbitrarily high-order time integrators in a suitable infinite-dimensional function space.

10.2.1 The Operator-Variation-of-Constants Formula

In this subsection, we start with recalling the abstract second-order ordinary differential equation and its operator-variation-of-constants formula (see [23]) before considering the design of the numerical integrators,.

To formulate an abstract formulation for the problem (10.1)–(10.2), we first consider the differential operator $\mathscr{A}$ defined by

$$(\mathscr{A}v)(x) = -a^2 v_{xx}(x),$$

where $\mathscr{A}$ is a linear, unbounded positive semi-definite operator, whose domain is

$$D(\mathscr{A}) := \left\{ v \in H^1(\Omega) : v(x) = v(x + 2\pi) \right\}.$$

Clearly, the operator $\mathscr{A}$ has a complete system of orthogonal eigenfunctions $\left\{ \mathrm{e}^{\mathrm{i}kx} : k \in \mathbb{Z} \right\}$. The linear span of all these eigenfunctions

$$X := \mathrm{lin}\left\{ \mathrm{e}^{\mathrm{i}kx} : k \in \mathbb{Z} \right\} \tag{10.4}$$

is dense in the Hilbert space $L^2(\Omega)$. Thus, we obtain the orthonormal basis of eigenvectors of the operator $\mathscr{A}$ with the corresponding eigenvalues a^2k^2 for $k \in \mathbb{Z}$.

We next introduce the functions as follows:

$$\phi_j(x) := \sum_{k=0}^{\infty} \frac{(-1)^k x^k}{(2k+j)!}, \qquad j \in \mathbb{N} \quad \text{for} \quad \forall x \geqslant 0. \tag{10.5}$$

It is easy to see that the functions ϕ_j for $j = 0, 1, 2, \cdots$ are bounded for any $x \geqslant 0$. For instance,

$$\phi_0(x) = \cos(\sqrt{x}), \quad \phi_1(x) = \mathrm{sinc}(\sqrt{x}),$$

and it is obvious that $|\phi_j(x)| \leqslant 1$ for $j = 0, 1$ and $\forall x \geqslant 0$. These functions (10.5) can induce the bounded operators

$$\phi_j(t\mathscr{A}) : \ L^2(\Omega) \to L^2(\Omega)$$

for $j \in \mathbb{N}$ and $t_0 \leqslant t \leqslant T$:

$$\phi_j(t\mathscr{A})v(x) = \sum_{k=-\infty}^{\infty} \hat{v}_k \phi_j(ta^2k^2)\mathrm{e}^{\mathrm{i}kx} \quad \text{for} \quad v(x) = \sum_{k=-\infty}^{\infty} \hat{v}_k \mathrm{e}^{\mathrm{i}kx}. \tag{10.6}$$

The boundedness follows from the definition of the operator norm that

$$\|\phi_j(t\mathscr{A})\|_*^2 = \sup_{\|v\| \neq 0} \frac{\|\phi_j(t\mathscr{A})v\|^2}{\|v\|^2} \leqslant \sup_{t_0 \leqslant t \leqslant T} |\phi_j(ta^2k^2)|^2 \leqslant \gamma_j^2, \tag{10.7}$$

where $\|\cdot\|_*$ is the Sobolev norm $\|\cdot\|_{L^2(\Omega) \leftarrow L^2(\Omega)}$, and γ_j for $j \in \mathbb{N}$ are the bounds of the functions $|\phi_j(x)|$ for $j \in \mathbb{N}$ and $x \geqslant 0$.

In what follows, we define $u(t)$ as the function that maps x to $u(x,t)$:

$$u(t) := [x \mapsto u(x,t)].$$

The system (10.1)–(10.2) can be formulated as an abstract second-order ordinary differential equation

$$\begin{cases} u''(t) + \mathscr{A}u(t) = f\big(u(t)\big), \quad t_0 < t \leqslant T, \\ u(t_0) = \varphi_1(x), \quad u'(t_0) = \varphi_2(x), \end{cases} \tag{10.8}$$

on the closed subspace

$$\begin{aligned} \mathscr{X} :=& \Big\{ u(x,\cdot) \in X \;\Big|\; u(x,\cdot) \text{ satisfies the corresponding boundary conditions} \Big\} \\ & \subseteq L^2(\Omega). \end{aligned} \tag{10.9}$$

The next theorem characterizes the solution of the abstract second-order ordinary differential equation (10.8) (see [23]).

Theorem 10.1 *The solution of* (10.8) *and its derivative satisfy the following operator-variation-of-constants formula*

$$\begin{cases} u(t) = \phi_0\big((t-t_0)^2\mathscr{A}\big)u(t_0) + (t-t_0)\phi_1\big((t-t_0)^2\mathscr{A}\big)u'(t_0) \\ \qquad\quad + \displaystyle\int_{t_0}^{t} (t-\zeta)\phi_1\big((t-\zeta)^2\mathscr{A}\big) f\big(u(\zeta)\big)\mathrm{d}\zeta, \\ u'(t) = -(t-t_0)\mathscr{A}\phi_1\big((t-t_0)^2\mathscr{A}\big)u(t_0) + \phi_0\big((t-t_0)^2\mathscr{A}\big)u'(t_0) \\ \qquad\quad + \displaystyle\int_{t_0}^{t} \phi_0\big((t-\zeta)^2\mathscr{A}\big) f\big(u(\zeta)\big)\mathrm{d}\zeta, \end{cases} \tag{10.10}$$

for $t_0 \leqslant t \leqslant T$, *where both* $\phi_0\big((t-t_0)^2\mathscr{A}\big)$ *and* $\phi_1\big((t-t_0)^2\mathscr{A}\big)$ *are bounded operators.*

10.2.2 The Formulation of the Time Integrators

According to the operator-variation-of-constants formula (10.12) and the two-point Hermite interpolation, we develop a class of arbitrarily high-order and symmetric time integration formulae. We start with a few useful preliminaries.

Lemma 10.1 *The bounded functions* $\phi_j(\mathscr{A})$, $j \in \mathbb{N}$ *of the operator* $\mathscr{A}$ *reduced by* (10.5) *satisfy*

$$\begin{aligned}
&\int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{A}\big)z^j \mathrm{d}z = j!\phi_{j+2}(\mathscr{A}), \qquad j \in \mathbb{N},\\
&\int_0^1 \phi_0\big((1-z)^2\mathscr{A}\big)z^j \mathrm{d}z = j!\phi_{j+1}(\mathscr{A}), \qquad j \in \mathbb{N}.
\end{aligned} \tag{10.11}$$

Proof The first formula can be proved as follows

$$\begin{aligned}
\int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{A}\big)z^j \mathrm{d}z &= \sum_{k=0}^{\infty} \frac{(-1)^k \int_0^1 (1-z)^{2k+1} z^j \mathrm{d}z}{(2k+1)!}\mathscr{A}^k\\
&= \sum_{k=0}^{\infty} \frac{(-1)^k j!}{(2k+j+2)!}\mathscr{A}^k = j!\phi_{j+2}(\mathscr{A}).
\end{aligned}$$

Likewise, we can obtain the second formula. □

Corollary 10.1 *For every* $m, n \in \mathbb{N}$ *the operators* (10.6) *satisfy*

$$\begin{aligned}
&\int_0^1 (1-z)^{m+1}\phi_1\big((1-z)^2\mathscr{A}\big)z^n \mathrm{d}z = \sum_{i=0}^{m} C_m^i (-1)^{m-i}(m+n-i)!\phi_{m+n-i+2}(\mathscr{A}),\\
&\int_0^1 (1-z)^{m}\phi_0\big((1-z)^2\mathscr{A}\big)z^n \mathrm{d}z = \sum_{i=0}^{m} C_m^i (-1)^{m-i}(m+n-i)!\phi_{m+n-i+1}(\mathscr{A}),
\end{aligned}$$

where $C_m^i = \binom{m}{i}$ *is the binomial symbol.*

Proof We only prove the first formula

$$\begin{aligned}
&\int_0^1 (1-z)^{m+1}\phi_1\big((1-z)^2\mathscr{A}\big)z^n \mathrm{d}z\\
&= \sum_{i=0}^{m} C_m^i (-1)^{m-i} \int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{A}\big)z^{m+n-i} \mathrm{d}z\\
&= \sum_{i=0}^{m} C_m^i (-1)^{m-i}(m+n-i)!\phi_{m+n-i+2}(\mathscr{A}).
\end{aligned}$$

Likewise, the second formula can be obtained. □

It follows from Theorem 10.1 that the solution of (10.8) and its derivative at a time point $t_{n+1} = t_n + \Delta t$, $n \in \mathbb{N}$ are

$$\begin{cases} u(t_{n+1}) = \phi_0(\mathscr{V})u(t_n) + \Delta t\phi_1(\mathscr{V})u'(t_n) + \Delta t^2 \int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{V}\big)\tilde{f}(z)\mathrm{d}z, \\ u'(t_{n+1}) = -\Delta t\mathscr{A}\phi_1(\mathscr{V})u(t_n) + \phi_0(\mathscr{V})u'(t_n) + \Delta t \int_0^1 \phi_0\big((1-z)^2\mathscr{V}\big)\tilde{f}(z)\mathrm{d}z, \end{cases} \tag{10.12}$$

where $\mathscr{V} = \Delta t^2\mathscr{A}$ and $\tilde{f}(z) = f\big(u(t_n + z\Delta t)\big)$. Clearly, in order to obtain the time integration formula from (10.12), we need to consider efficient integrators for approximating the nonlinear integrals

$$\begin{aligned} I_1 &:= \int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{V}\big)\tilde{f}(z)\mathrm{d}z, \\ I_2 &:= \int_0^1 \phi_0\big((1-z)^2\mathscr{V}\big)\tilde{f}(z)\mathrm{d}z. \end{aligned} \tag{10.13}$$

Usually, the potential function $f(u)$ is nonlinear, and only the endpoints' information can be used directly when we deal with the two nonlinear integrals in (10.13) and design numerical methods. Accordingly, we are particularly concerned with fitting function values and derivatives at the two boundary points of the finite interval $[0, 1]$. This motivates us to interpolate $\tilde{f}(z)$ by a two-point Hermite interpolation $p_r(z)$ of degree $2r + 1$ (see, e.g. [26, 27]).

Lemma 10.2 *Assume that* $\tilde{f} \in C^{2r+2}\big([0, 1]\big)$. *Then there exists a Hermite interpolating polynomial* $p_r(z)$ *of degree* $2r + 1$

$$p_r(z) = \sum_{j=0}^{r} \Big[\beta_{r,j}(z)\tilde{f}^{(j)}(0) + (-1)^j\beta_{r,j}(1-z)\tilde{f}^{(j)}(1)\Big], \tag{10.14}$$

satisfying the interpolation conditions

$$p_r^{(j)}(0) = \tilde{f}^{(j)}(0), \quad p_r^{(j)}(1) = \tilde{f}^{(j)}(1), \qquad j = 0, 1, 2, \cdots, r,$$

where

$$\beta_{r,j}(z) = \frac{z^j}{j!}(1-z)^{r+1}\sum_{s=0}^{r-j} C_{r+s}^s z^s, \tag{10.15}$$

and the error on $[0,1]$ *is*

$$R_r = \tilde{f}(z) - p_r(z) = (-1)^{r+1} z^{r+1} (1-z)^{r+1} \frac{\tilde{f}^{(2r+2)}(\xi)}{(2r+2)!}, \quad \xi \in (0,1). \tag{10.16}$$

Replacing $\tilde{f}(z)$ in (10.13) by the Hermite interpolation $p_r(z)$ where $\tilde{f}(z) = f\big(u(t_n + z\Delta t)\big)$ and $\tilde{f}^{(j)}(z) = \Delta t^j f_t^{(j)}\big(u(t_n + z\Delta t)\big)$ yields

$$\begin{aligned}
\tilde{I}_1^r &= \sum_{j=0}^{r} \Delta t^j \Big[I_1[\beta_{r,j}(z)] f_t^{(j)}\big(u(t_n)\big) + (-1)^j I_1[\beta_{r,j}(1-z)] f_t^{(j)}\big(u(t_{n+1})\big) \Big], \\
\tilde{I}_2^r &= \sum_{j=0}^{r} \Delta t^j \Big[I_2[\beta_{r,j}(z)] f_t^{(j)}\big(u(t_n)\big) + (-1)^j I_2[\beta_{r,j}(1-z)] f_t^{(j)}\big(u(t_{n+1})\big) \Big],
\end{aligned} \tag{10.17}$$

where $f_t^{(j)}\big(u(t)\big)$ denotes the j-th derivative of $f\big(u(t)\big)$ with respect to t. In terms of the Hermite–Birkhoff quadrature formula (see, e.g. [28–30]), we will determine the coefficients $I_1[\beta_j(z)]$, $I_2[\beta_j(z)]$, $I_1[\beta_j(1-z)]$ and $I_2[\beta_j(1-z)]$. These coefficients are given by

$$\begin{aligned}
I_1[\beta_{r,j}(z)] &:= \int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{V}\big)\beta_{r,j}(z)\mathrm{d}z \\
&= \sum_{s=0}^{r-j}\sum_{i=0}^{r+1} (-1)^{r-i+1} C_{r+s}^s C_{r+1}^i \frac{(r+s+j-i+1)!}{j!} \phi_{r+s+j-i+3}(\mathscr{V}),
\end{aligned} \tag{10.18}$$

$$\begin{aligned}
I_2[\beta_{r,j}(z)] &:= \int_0^1 \phi_0\big((1-z)^2\mathscr{V}\big)\beta_{r,j}(z)\mathrm{d}z \\
&= \sum_{s=0}^{r-j}\sum_{i=0}^{r+1} (-1)^{r-i+1} C_{r+s}^s C_{r+1}^i \frac{(r+s+j-i+1)!}{j!} \phi_{r+s+j-i+2}(\mathscr{V}),
\end{aligned} \tag{10.19}$$

$$\begin{aligned}
I_1[\beta_{r,j}(1-z)] &:= \int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{V}\big)\beta_{r,j}(1-z)\mathrm{d}z \\
&= \sum_{s=0}^{r-j}\sum_{i=0}^{r+j} (-1)^{s+j-i} C_{r+s}^s C_{s+j}^i \frac{(r+s+j-i+1)!}{j!} \phi_{r+s+j-i+3}(\mathscr{V}),
\end{aligned} \tag{10.20}$$

$$\begin{aligned}
I_2[\beta_{r,j}(1-z)] &:= \int_0^1 \phi_0\big((1-z)^2\mathscr{V}\big)\beta_{r,j}(1-z)\mathrm{d}z \\
&= \sum_{s=0}^{r-j}\sum_{i=0}^{r+j} (-1)^{s+j-i} C_{r+s}^s C_{s+j}^i \frac{(r+s+j-i+1)!}{j!} \phi_{r+s+j-i+2}(\mathscr{V}).
\end{aligned} \tag{10.21}$$

From the definitions stated above, it is evident that the coefficients are bounded for any $j = 0, 1, \cdots, r$,

$$\begin{aligned}
&\|I_1[\beta_{r,j}(z)]\|_* \leqslant \max_{0\leqslant z\leqslant 1} |\beta_{r,j}(z)| \leqslant 1 \quad \text{and} \quad \|I_1[\beta_{r,j}(1-z)]\|_* \\
&\leqslant \max_{0\leqslant z\leqslant 1} |\beta_{r,j}(1-z)| \leqslant 1, \\
&\|I_2[\beta_{r,j}(z)]\|_* \leqslant \max_{0\leqslant z\leqslant 1} |\beta_{r,j}(z)| \leqslant 1 \quad \text{and} \quad \|I_2[\beta_{r,j}(1-z)]\|_* \\
&\leqslant \max_{0\leqslant z\leqslant 1} |\beta_{r,j}(1-z)| \leqslant 1.
\end{aligned}$$

Suppose that the following approximations have been given

$$u^n \approx u(t_n) \quad \text{and} \quad \mu^n \approx u'(t_n).$$

On the basis of the above analysis and the formula (10.12), we present the following time integration formula for the abstract ODE (10.8).

Definition 10.1 The Hermite–Birkhoff (HB) time integration formula for solving the abstract ODE (10.8) is defined by

$$\begin{cases}
u^{n+1} = \phi_0(\mathscr{V})u^n + \Delta t\phi_1(\mathscr{V})\mu^n \\
\qquad + \displaystyle\sum_{j=0}^{r} \Delta t^{j+2} \left\{ I_1[\beta_{r,j}(z)] f_t^{(j)}(u^n) + (-1)^j I_1[\beta_{r,j}(1-z)] f_t^{(j)}(u^{n+1}) \right\}, \\
\mu^{n+1} = -\Delta t\mathscr{A}\phi_1(\mathscr{V})u^n + \phi_0(\mathscr{V})\mu^n \\
\qquad + \displaystyle\sum_{j=0}^{r} \Delta t^{j+1} \left\{ I_2[\beta_{r,j}(z)] f_t^{(j)}(u^n) + (-1)^j I_2[\beta_{r,j}(1-z)] f_t^{(j)}(u^{n+1}) \right\},
\end{cases} \tag{10.22}$$

where $I_1[\beta_{r,j}(z)]$, $I_2[\beta_{r,j}(z)]$, $I_1[\beta_{r,j}(1-z)]$ and $I_2[\beta_{r,j}(1-z)]$ have been defined by (10.18)–(10.21), respectively.

Remark 10.1 The HB time integration formula (10.22) is derived by using a two-point Hermite interpolation to approximate the nonlinear function $\tilde{f}(z)$ appearing in the nonlinear integrals (10.13). Here, the high order derivatives $\dfrac{\mathrm{d}^m \tilde{f}(z)}{\mathrm{d}z^m}$ will be used. Fortunately, the high order derivative $u^{(m)}(t_n + z\Delta t)$ can be calculated from lower order derivative via the abstract equation (10.8), namely,

$$\frac{\mathrm{d}^m}{\mathrm{d}z^m} u(t_n + z\Delta t) = \frac{\mathrm{d}^{m-2}}{\mathrm{d}z^{m-2}} \Big(-\mathscr{A}u(t_n + z\Delta t) + f\big(u(t_n + z\Delta t)\big) \Big) \Delta t^2, \quad m \geqslant 2.$$

Hence, the high order derivatives $\dfrac{\mathrm{d}^m \tilde{f}(z)}{\mathrm{d}z^m}$ satisfy the following recursive relationship

$$\begin{aligned}
&\tilde{f}'(z) = f'\big(u(t_n + z\Delta t)\big)u'(t_n + z\Delta t)\Delta t,\\
&\frac{\mathrm{d}^m \tilde{f}(z)}{\mathrm{d}z^m} = \frac{\mathrm{d}^{m-2}}{\mathrm{d}z^{m-2}}\Big\{ f''\big(u(t_n + z\Delta t)\big)\Big(u'(t_n + z\Delta t)\Big)^2\\
&\qquad + f'\big(u(t_n + z\Delta t)\big)\Big(-\mathscr{A}u(t_n + z\Delta t) + f\big(u(t_n + z\Delta t)\big)\Big)\Big\}\Delta t^2, \quad m \geqslant 2.
\end{aligned}$$

This means that the high order derivatives $u^{(m)}(\cdot)$ for $m \geqslant 2$ will not be affected in the HB time integration formula (10.22).

Concerning the local error bounds of the formula (10.22), we have the following theorem.

Theorem 10.2 *Assume that* $f\big(u(\cdot,t)\big) \in C^{2r+2}([t_0,T])$ *and* $f_t^{(2r+2)}\big(u(x,\cdot)\big) \in L^2(\Omega)$. *Under the local assumptions of* $u^n = u(t_n)$, $\mu^n = u'(t_n)$, *the local error bounds of the HB time integration formula* (10.22) *are*

$$\|u(t_{n+1}) - u^{n+1}\| \leqslant C_1\Delta t^{2r+4} \quad \text{and} \quad \|u'(t_{n+1}) - \mu^{n+1}\| \leqslant C_2\Delta t^{2r+3}, \tag{10.23}$$

where the constants C_1 *and* C_2 *are given by*

$$C_1 = \frac{(r+2)!(r+1)!}{(2r+2)!(2r+4)!}\max_{t_0\leqslant t\leqslant T}\big\|f_t^{(2r+2)}\big(u(t)\big)\big\|$$

and

$$C_2 = \frac{\big[(r+1)!\big]^2}{(2r+2)!(2r+3)!}\max_{t_0\leqslant t\leqslant T}\big\|f_t^{(2r+2)}\big(u(t)\big)\big\|.$$

Proof Using (10.12) and (10.22), we obtain

$$u(t_{n+1}) - u^{n+1} = \Delta t^2\int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{V}\big)\Big[f\big(u(t_n + z\Delta t)\big) - p_r(z)\Big]\mathrm{d}z, \tag{10.24}$$

and

$$u'(t_{n+1}) - \mu^{n+1} = \Delta t\int_0^1 \phi_0\big((1-z)^2\mathscr{V}\big)\Big[f\big(u(t_n + z\Delta t)\big) - p_r(z)\Big]\mathrm{d}z. \tag{10.25}$$

As $\tilde{f}^{(j)}(z) = \Delta t^j f_t^{(j)}\big(u(t_n + z\Delta t)\big)$, it follows from Lemma 10.2 that

$$f\big(u(t_n + z\Delta t)\big) - p_r(z) = \Delta t^{2r+2}(-1)^{r+1} z^{r+1}(1-z)^{r+1} \frac{f_t^{(2r+2)}\big(u(t_n + \xi^n \Delta t)\big)}{(2r+2)!}. \tag{10.26}$$

Then inserting (10.26) into (10.24) and (10.25) yields

$$\begin{aligned} \|u(t_{n+1}) - u^{n+1}\| &\leqslant \Delta t^{2r+4} \frac{\big\| f_t^{(2r+2)}\big(u(t_n + \xi^n \Delta t)\big)\big\|}{(2r+2)!} \int_0^1 (1-z)^{r+2} z^{r+1} \mathrm{d}z \\ &\leqslant C_1 \Delta t^{2r+4}, \end{aligned}$$

and

$$\begin{aligned} \|u'(t_{n+1}) - \mu^{n+1}\| &\leqslant \Delta t^{2r+3} \frac{\big\| f_t^{(2r+2)}\big(u(t_n + \xi^n \Delta t)\big)\big\|}{(2r+2)!} \int_0^1 (1-z)^{r+1} z^{r+1} \mathrm{d}z \\ &\leqslant C_2 \Delta t^{2r+3}. \end{aligned}$$

The statement of this theorem is proved. □

Since the KG equation (10.1) is time symmetric and a most welcome feature of (10.22) is that it preserves time symmetry, in what follows, we show the symmetry of the formula (10.22). As a first step, we introduce some useful properties of the operator-valued functions $\phi_0(\mathscr{A})$, $\phi_1(\mathscr{A})$ and the coefficients defined by (10.18)–(10.21) in the following two lemmas.

Lemma 10.3 *The bounded operators $\phi_0(\mathscr{A})$ and $\phi_1(\mathscr{A})$ defined by (10.6) satisfy*

$$\phi_0^2(\mathscr{A}) + \mathscr{A}\phi_1^2(\mathscr{A}) = I, \tag{10.27}$$

where $\mathscr{A}$ is an arbitrary positive semi-definite operator or matrix.

Lemma 10.4 *The coefficients $I_1[\beta_{r,j}(z)]$, $I_2[\beta_{r,j}(z)]$, $I_1[\beta_{r,j}(1-z)]$ and $I_2[\beta_{r,j}(1-z)]$ from (10.22) satisfy*

$$\begin{aligned} &\phi_0(\mathscr{V}) I_1[\beta_{r,j}(z)] - \phi_1(\mathscr{V}) I_2[\beta_{r,j}(z)] = -I_1[\beta_{r,j}(1-z)], \\ &\mathscr{V}\phi_1(\mathscr{V}) I_1[\beta_{r,j}(z)] + \phi_0(\mathscr{V}) I_2[\beta_{r,j}(z)] = I_0[\beta_{r,j}(1-z)], \end{aligned} \tag{10.28}$$

where $\beta_{r,j}(z)$ for $j = 0, 1, \cdots, r$ are defined by (10.15) and $\mathscr{V} = \Delta t^2 \mathscr{A}$ with $\mathscr{A}$, an arbitrary positive semi-definite operator or matrix.

Proof According to the definitions of $I_1[\beta_{r,j}(z)]$ and $I_2[\beta_{r,j}(z)]$, we have

$$\begin{aligned}
&\phi_0(\mathscr{V})I_1[\beta_{r,j}(z)] - \phi_1(\mathscr{V})I_2[\beta_{r,j}(z)] \\
&= \int_0^1 \Big[(1-z)\phi_0(\mathscr{V})\phi_1\big((1-z)^2\mathscr{V}\big) - \phi_1(\mathscr{V})\phi_0\big((1-z)^2\mathscr{V}\big)\Big]\beta_{r,j}(z)\mathrm{d}z \\
&= \int_0^1 \Big[z\phi_0(\mathscr{V})\phi_1(z^2\mathscr{V}) - \phi_1(\mathscr{V})\phi_0(z^2\mathscr{V})\Big]\beta_{r,j}(1-z)\mathrm{d}z \\
&= -\int_0^1 (1-z)\phi_1\big((1-z)^2\mathscr{V}\big)\beta_j(1-z)\mathrm{d}z = -I_1[\beta_{r,j}(1-z)],
\end{aligned}$$

and

$$\begin{aligned}
&\mathscr{V}\phi_1(\mathscr{V})I_1[\beta_j(z)] + \phi_0(\mathscr{V})I_2[\beta_j(z)] \\
&= \int_0^1 \Big((1-z)\mathscr{V}\phi_1(\mathscr{V})\phi_1\big((1-z)^2\mathscr{V}\big) + \phi_0(\mathscr{V})\phi_0\big((1-z)^2\mathscr{V}\big)\Big)\beta_j(z)\mathrm{d}z \\
&= \int_0^1 \big(z\mathscr{V}\phi_1(\mathscr{V})\phi_1(z^2\mathscr{V}) + \phi_0(\mathscr{V})\phi_0(z^2\mathscr{V})\big)\beta_j(1-z)\mathrm{d}z \\
&= \int_0^1 \phi_0\big((1-z)^2\mathscr{V}\big)\beta_j(1-z)\mathrm{d}z = I_0[\beta_j(1-z)]. \qquad (10.29)
\end{aligned}$$

Hence, the theorem is proved. □

We note that Hairer et al. [31] have pointed out that symmetric methods have excellent long-time behaviour when solving reversible differential systems. Therefore, it is an important aspect of the design and analysis of symmetric integrators in numerical PDEs. We are now in a position to prove the time symmetry of (10.22).

Theorem 10.3 *The HB time integration formula* (10.22) *is symmetric with respect to the time variable.*

Proof Exchanging $u^{n+1} \leftrightarrow u^n$, $\mu^{n+1} \leftrightarrow \mu^n$ and replacing Δt by $-\Delta t$ in formula (10.22), we obtain

$$\begin{aligned}
u^n =& \phi_0\big(\mathscr{V}\big)u^{n+1} - \Delta t\phi_1\big(\mathscr{V}\big)\mu^{n+1} \\
&+ \sum_{j=0}^{r} \Delta t^{j+2}\left\{(-1)^j I_1[\beta_{r,j}(z)]f_t^{(j)}(u^{n+1}) + I_1[\beta_{r,j}(1-z)]f_t^{(j)}(u^n)\right\},
\end{aligned} \qquad (10.30)$$

$$\mu^n = \Delta t \mathscr{A} \phi_1(\mathscr{V}) u^{n+1} + \phi_0(\mathscr{V}) \mu^{n+1} - \sum_{j=0}^{r} \Delta t^{j+1} \left\{ (-1)^j I_2[\beta_{r,j}(z)] f_t^{(j)}(u^{n+1}) + I_2[\beta_{r,j}(1-z)] f_t^{(j)}(u^n) \right\}. \tag{10.31}$$

It follows from the calculation $\phi_0(\mathscr{V}) \times (10.30) + \Delta t \phi_1(\mathscr{V}) \times (10.31)$ that

$$\begin{aligned} u^{n+1} =& \phi_0(\mathscr{V}) u^n + \Delta t \phi_1(\mathscr{V}) \mu^n \\ & - \sum_{j=0}^{r} \Delta t^{j+2} \left\{ (-1)^j \left[\phi_0(\mathscr{V}) I_1[\beta_{r,j}(z)] - \phi_1(\mathscr{V}) I_2[\beta_{r,j}(z)] \right] f_t^{(j)}(u^{n+1}) \right. \\ & \left. + \left[\phi_0(\mathscr{V}) I_1[\beta_{r,j}(1-z)] - \phi_1(\mathscr{V}) I_2[\beta_{r,j}(1-z)] \right] f_t^{(j)}(u^n) \right\}. \end{aligned} \tag{10.32}$$

Likewise, the calculation $-\Delta t \mathscr{A} \phi_1(\mathscr{V}) \times (10.30) + \phi_0(\mathscr{V}) \times (10.31)$ results in

$$\begin{aligned} \mu^{n+1} =& -\Delta t \mathscr{A} \phi_1(\mathscr{V}) u^n + \phi_0(\mathscr{V}) \mu^n \\ & + \sum_{j=0}^{r} \Delta t^{j+1} \left\{ (-1)^j \left[\mathscr{V} \phi_1(\mathscr{V}) I_1[\beta_{r,j}(z)] + \phi_0(\mathscr{V}) I_2[\beta_{r,j}(z)] \right] f_t^{(j)}(u^{n+1}) \right. \\ & \left. + \left[\mathscr{V} \phi_1(\mathscr{V}) I_1[\beta_{r,j}(1-z)] + \phi_0(\mathscr{V}) I_2[\beta_{r,j}(1-z)] \right] f_t^{(j)}(u^n) \right\}. \end{aligned} \tag{10.33}$$

Then applying Lemma 10.4 to (10.32) and (10.33) yields the statement of the theorem. □

10.3 Stability of the Fully Discrete Scheme

This section will show the stability of the fully discrete scheme after the differential operator $\mathscr{A}$ is replaced by a suitable matrix A. Throughout this section $\|\cdot\|$ represents both the vector 2-norm and the matrix 2-norm (the spectral norm).

Under the assumption of the following *finite-energy condition* (see, e.g. [32–34])

$$\frac{1}{2}\|u'(t)\|^2 + \frac{\kappa^2}{2} u(t)^{\mathsf{T}} A u(t) \leqslant \frac{K^2}{2}, \tag{10.34}$$

where K is a constant, global error bounds of the Gaustchi-type method were proved to be independent on $\|A\|$. Consequently, the Gaustchi-type time integrator of order two coupled with suitable spatial discretisation is an excellent choice to solve nonlinear wave equations. Moreover, it is a most important result that the

long-time energy conservation for numerical methods can be achieved with the finite-energy condition (see, e.g. [33]). We here also suppose that the exact solution of the nonlinear system (10.8) after suitable spatial discretisation satisfies the finite-energy condition (10.34).

Assume that the perturbed problem of (10.8) is

$$\begin{cases} v''(t) + \mathscr{A} v(t) = f\big(v(t)\big), \quad t \in [t_0, T], \\ v(t_0) = \varphi_1(x) + \tilde{\varphi}_1(x), \quad v'(t_0) = \varphi_2(x) + \tilde{\varphi}_2(x), \end{cases} \tag{10.35}$$

where $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$ are perturbation functions. Let

$$\eta(t) = v(t) - u(t).$$

Subtracting (10.8) from (10.35) yields

$$\begin{cases} \eta''(t) + \mathscr{A}\eta(t) = f\big(v(t)\big) - f\big(u(t)\big), \quad t \in [t_0, T], \\ \eta(t_0) = \tilde{\varphi}_1(x), \quad \eta'(t_0) = \tilde{\varphi}_2(x). \end{cases} \tag{10.36}$$

We approximate the operator $\mathscr{A}$ by a symmetric positive semi-definite differentiation matrix A on an M-dimensional space since this assists in structure preservation. This implies that there exists an orthogonal matrix P and a diagonal matrix Λ with non-negative diagonal such that

$$A = P\Lambda P^{\mathsf{T}}.$$

Then $A = D^2$, where $D = P\Lambda^{\frac{1}{2}} P^{\mathsf{T}}$. Accordingly, the bounded operators $\phi_j(t^2\mathscr{A})$ are replaced by the matrix functions $\phi_j(t^2 A)$. Likewise, we also have

$$\|\phi_j(t^2 A)\| = \sqrt{\lambda_{\max}\Big(\phi_j^2(t^2 A)\Big)} \leqslant \gamma_j, \qquad j \in \mathbb{N}. \tag{10.37}$$

Moreover, it is clear that

$$\|D\alpha\|^2 = \alpha^{\mathsf{T}} A \alpha, \quad \forall \alpha \in \mathbb{R}^M,$$

because A is a symmetric positive semi-definite matrix.

We next analyse the stability for the HB time integrators (10.22). We assume that

$$\eta^n \approx \eta(t_n), \quad \zeta^n \approx \eta'(t_n) \qquad \text{and} \qquad v^n \approx v(t_n), \quad w^n \approx v'(t_n).$$

Applying HB time integration to (10.36) yields

$$\begin{cases} \eta^{n+1} = \phi_0(V)\eta^n + \Delta t\phi_1(V)\zeta^n + \sum_{j=0}^{r} \Delta t^{j+2} \Big\{ I_1[\beta_{r,j}(z)]\big[f_t^{(j)}(v^n) \\ \qquad - f_t^{(j)}(u^n)\big] + (-1)^j I_1[\beta_{r,j}(1-z)]\big[f_t^{(j)}(v^{n+1}) - f_t^{(j)}(u^{n+1})\big]\Big\}, \\ \zeta^{n+1} = -\Delta t A\phi_1(V)\eta^n + \phi_0(V)\zeta^n + \sum_{j=0}^{r} \Delta t^{j+1} \Big\{ I_2[\beta_{r,j}(z)]\big[f_t^{(j)}(v^n) \\ \qquad - f_t^{(j)}(u^n)\big] + (-1)^j I_2[\beta_{r,j}(1-z)]\big[f_t^{(j)}(v^{n+1}) - f_t^{(j)}(u^{n+1})\big]\Big\}, \end{cases} \tag{10.38}$$

where $V = \Delta t^2 A$, $I_1[\beta_{r,j}(z)]$, $I_2[\beta_{r,j}(z)]$, $I_1[\beta_{r,j}(1-z)]$ and $I_2[\beta_{r,j}(1-z)]$ are defined by (10.18)–(10.21), respectively. Similarly, we obtain

$$\|I_1[\beta_{r,j}(z)]\| \leqslant \max_{0\leqslant z\leqslant 1} |\beta_{r,j}(z)| \leqslant 1 \quad \text{and} \quad \|I_1[\beta_{r,j}(1-z)]\| \leqslant \max_{0\leqslant z\leqslant 1} |\beta_{r,j}(1-z)| \leqslant 1,$$

$$\|I_2[\beta_{r,j}(z)]\| \leqslant \max_{0\leqslant z\leqslant 1} |\beta_{r,j}(z)| \leqslant 1 \quad \text{and} \quad \|I_2[\beta_{r,j}(1-z)]\| \leqslant \max_{0\leqslant z\leqslant 1} |\beta_{r,j}(1-z)| \leqslant 1.$$

The schemes (10.38) can be rewritten in a compact form:

$$\begin{aligned} \begin{bmatrix} D\eta^{n+1} \\ \zeta^{n+1} \end{bmatrix} =& \Psi(V) \begin{bmatrix} D\eta^n \\ \zeta^n \end{bmatrix} + \sum_{j=0}^{r} \Delta t^{j+1} \int_0^1 \Psi_j\big(\beta(z), V\big)\mathrm{d}z \begin{bmatrix} 0 \\ f_t^{(j)}(v^n) - f_t^{(j)}(u^n) \end{bmatrix} \\ &+ \sum_{j=0}^{r} (-1)^j \Delta t^{j+1} \int_0^1 \Psi_j\big(\beta(1-z), V\big)\mathrm{d}z \begin{bmatrix} 0 \\ f_t^{(j)}(v^{n+1}) - f_t^{(j)}(u^{n+1}) \end{bmatrix}, \end{aligned} \tag{10.39}$$

where

$$\Psi(V) = \begin{bmatrix} \phi_0(V) & \Delta t D\phi_1(V) \\ -\Delta t D\phi_1(V) & \phi_0(V) \end{bmatrix} \tag{10.40}$$

and

$$\Psi_j\big(\beta(z), V\big) = \beta_{r,j}(z) \begin{bmatrix} \phi_0((1-z)^2 V) & \Delta t(1-z) D\phi_1((1-z)^2 V) \\ -\Delta t(1-z) D\phi_1((1-z)^2 V) & \phi_0((1-z)^2 V) \end{bmatrix}. \tag{10.41}$$

Before dealing with stability analysis, we should investigate the spectral norm of matrices $\Psi(V)$ and $\Psi_j\big(\beta(z), V\big)$ for $j = 0, 1, \cdots, r$.

Lemma 10.5 *Suppose that A is a symmetric positive semi-definite matrix and that $V = \Delta t^2 A$. Then the spectral norms of matrices $\Psi(V)$ and $\Psi_j\big(\beta(z), V\big)$ satisfy*

$$\|\Psi(V)\| = 1 \quad \textit{and} \quad \|\Psi_j\big(\beta(z), V\big)\| = |\beta_{r,j}(z)| \leqslant 1,$$
$$z \in [0, 1], \quad j = 0, 1, \cdots, r. \tag{10.42}$$

Proof It is trivial to verify the results based on Lemma 10.3, formulae (10.40) and (10.41) and the definition of the matrix 2-norm. The reader is referred to [23] for details. □

10.3.1 Linear Stability Analysis

We begin with the stability analysis of HB time integrators for the linear problem, i.e. $f(u) = u$. In this case, we have

$$f_t^{(2k)}\big(u(t)\big) = (I - \mathscr{A})^k u(t) \quad \text{and} \quad f_t^{(2k+1)}\big(u(t)\big) = (I - \mathscr{A})^k u'(t), \quad k \in \mathbb{N}. \tag{10.43}$$

Lemma 10.6 *Suppose that A is a symmetric matrix. Then*

$$\|(I - A)^k\| \leqslant \big[1 + \rho(A)\big]^k, \qquad k \in \mathbb{N},$$

where $\rho(A)$ is the spectral radius of A.

Proof It is immediately from the definition of the spectral norm that

$$\|(I - A)^k\| = \sqrt{\lambda_{\max}\big((I - A)^{2k}\big)} \leqslant \big(1 + \max_{1 \leqslant j \leqslant M} |\lambda_j|\big)^k = \big[1 + \rho(A)\big]^k,$$

where λ_j for $j = 1, 2, \cdots, M$ are the eigenvalues of A. □

Theorem 10.4 *Assume that the operator $\mathscr{A}$ is approximated by a symmetric positive semi-definite differentiation matrix A and let the finite energy condition (10.34) be satisfied. If the sufficiently small time stepsize Δt satisfies $\Delta t^2\big(1 + \rho(A)\big) \leqslant 1$ with $\Delta t \leqslant [4(r + 1)]^{-1}$, then we have the following stability results:*

$$\|\eta^n\| \leqslant \exp\big(2(2r + 3)T\big)\Big(\|\tilde{\varphi}_1\| + \sqrt{\tilde{\varphi}_1^{\mathsf{T}} A \tilde{\varphi}_1 + \|\tilde{\varphi}_2\|^2}\Big),$$
$$\|\zeta^n\| \leqslant \exp\big(2(2r + 3)T\big)\Big(\|\tilde{\varphi}_1\| + \sqrt{\tilde{\varphi}_1^{\mathsf{T}} A \tilde{\varphi}_1 + \|\tilde{\varphi}_2\|^2}\Big),$$

where $\tilde{\varphi}_l = \big(\tilde{\varphi}_l(x_0), \tilde{\varphi}_l(x_1), \cdots, \tilde{\varphi}_l(x_{M-1})\big)^{\intercal}$, *while* $\tilde{\varphi}_l(x_i)$ *for* $l = 1, 2$ *are the values of the perturbation functions* $\tilde{\varphi}_l$ *for* $l = 1, 2$, *at the grid points* $\{x_i\}_{i=0}^{M-1}$.

Proof Using the first formula in (10.38) and (10.39), we obtain

$$\|\eta^{n+1}\| \leqslant \|\eta^n\| + \Delta t\|\zeta^n\| + \sum_{j=0}^{r} \Delta t^{j+2}\big(\|f_t^{(j)}(v^n) - f_t^{(j)}(u^n)\| + \|f_t^{(j)}(v^{n+1}) - f_t^{(j)}(u^{n+1})\|\big),$$

and

$$\sqrt{(\eta^{n+1})^{\intercal} A\eta^{n+1} + \|\zeta^{n+1}\|^2} \leqslant \sqrt{(\eta^n)^{\intercal} A\eta^n + \|\zeta^n\|^2} + \sum_{j=0}^{r} \Delta t^{j+1}\Big(\|f_t^{(j)}(v^n) - f_t^{(j)}(u^n)\| + \|f_t^{(j)}(v^{n+1}) - f_t^{(j)}(u^{n+1})\|\Big).$$

Then summing up the above and using (10.43), we have

$$\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A\eta^{n+1} + \|\zeta^{n+1}\|^2} \leqslant \|\eta^n\| + \sqrt{(\eta^n)^{\intercal} A\eta^n + \|\zeta^n\|^2} + \Delta t\|\zeta^n\| + \Delta t(1+\Delta t)\sum_{j=0}^{r} \Delta t^j \big\|(I-A)^{[\frac{j}{2}]}\big\|\Big(\|\eta^n\| + \|\zeta^n\| + \|\eta^{n+1}\| + \|\zeta^{n+1}\|\Big). \tag{10.44}$$

Applying Lemma 10.6 to inequality (10.44) leads to

$$\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A\eta^{n+1} + \|\zeta^{n+1}\|^2} \leqslant \|\eta^n\| + \sqrt{(\eta^n)^{\intercal} A\eta^n + \|\zeta^n\|^2} + \Delta t\|\zeta^n\| + \Delta t(1+\Delta t)\sum_{j=0}^{r} \Delta t^j \big(1+\rho(A)\big)^{[\frac{j}{2}]}\Big(\|\eta^n\| + \|\zeta^n\| + \|\eta^{n+1}\| + \|\zeta^{n+1}\|\Big).$$

Under the assumption that the stepsize satisfies $\Delta t^2\big(1+\rho(A)\big) \leqslant 1$, we obtain

$$\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A\eta^{n+1} + \|\zeta^{n+1}\|^2} \leqslant \left[1 + \frac{\Delta t(2r+3)}{1 - 2\Delta t(r+1)}\right]\Big(\|\eta^n\| + \sqrt{(\eta^n)^{\intercal} A\eta^n + \|\zeta^n\|^2}\Big).$$

As $\Delta t \leqslant [4(r+1)]^{-1}$, we have

$$\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A \eta^{n+1} + \|\zeta^{n+1}\|^2} \leqslant \Big[1 + 2(2r+3)\Delta t\Big] \Big(\|\eta^n\| + \sqrt{(\eta^n)^{\intercal} A \eta^n + \|\zeta^n\|^2}\Big).$$

Then an inductive argument yields the following result

$$\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A \eta^{n+1} + \|\zeta^{n+1}\|^2} \leqslant \exp\big(2(2r+3)T\big) \Big(\|\tilde{\varphi}_1\| + \sqrt{\tilde{\varphi}_1^{\intercal} A \tilde{\varphi}_1 + \|\tilde{\varphi}_2\|^2}\Big).$$

Therefore, we have

$$\begin{aligned} \|\eta^n\| &\leqslant \exp\big(2(4r+5)T\big)\Big(\|\tilde{\varphi}_1\| + \sqrt{\|D\tilde{\varphi}_1\|^2 + \|\tilde{\varphi}_2\|^2}\Big), \\ \|\zeta^n\| &\leqslant \exp\big(2(4r+5)T\big)\Big(\|\tilde{\varphi}_1\| + \sqrt{\|D\tilde{\varphi}_1\|^2 + \|\tilde{\varphi}_2\|^2}\Big), \end{aligned} \tag{10.45}$$

and linear stability is proved. □

10.3.2 Nonlinear Stability Analysis

We will further analyse in this subsection the stability of HB time integrators for nonlinear problems. The analysis relies upon some assumptions.

Assumption 10.1 It is assumed that both (10.8) and (10.35) possess sufficiently smooth solutions and $f : D(\mathscr{A}) \to \mathbb{R}$ is sufficiently Fréchet differentiable in a strip along the exact solution.

As is known, it follows from Chap. 3 in [35] that

$$f_t^{(k)}\big(u(t)\big) = \sum_{\tilde{t} \in \mathrm{SENT}_{k+2}^f} \alpha(\tilde{t}) \mathscr{F}(\tilde{t})\big(u(t), u'(t)\big), \tag{10.46}$$

where $\mathrm{SENT}^f = \{\tau_2\} \cup \big\{\tilde{t} = [\tilde{t}_1, \cdots, \tilde{t}_m]_2 : \tilde{t}_i \in \mathrm{SENT}\big\}$ and SENT is the set of special extended Nyström trees defined in [35], $\alpha(\tilde{t})$ is the number of possible monotonic labellings of an extended Nyström tree $\tilde{t}$, and $\mathscr{F}(\tilde{t})\big(u, u'\big)$ is the corresponding elementary differential.

Assumption 10.2 We assume that $\mathrm{d}^k f(u)/\mathrm{d}u^k : D(\mathscr{A}) \to \mathbb{R}$ for $k = 0, 1, 2, \cdots, r$ are locally Lipschitz continuous in a strip along the exact solution u. Hence, there exist real numbers $L(R, \rho(A)^{\lfloor \frac{k}{2} \rfloor})$ such that

$$\begin{aligned}&\|\mathscr{F}(\tilde{t})\big(v(t), v'(t)\big) - \mathscr{F}(\tilde{t})\big(w(t), w'(t)\big)\| \\ &\leqslant L(R, \rho(A)^{\lfloor \frac{k}{2} \rfloor})\Big(\|v(t) - w(t)\| + \|v'(t) - w'(t)\|\Big), \quad \forall \tilde{t} \in \mathrm{SENT}^f_{k+2}\end{aligned}$$

for all $t \in [t_0, T]$ and $\max\Big(\|v - u(t)\|, \|w - u(t)\|, \|v' - u'(t)\|, \|w' - u'(t)\|\Big) \leqslant R$.

The next theorem shows the statement on nonlinear stability.

Theorem 10.5 *With Assumptions 10.1 and 10.2, suppose that the sufficiently small time stepsize satisfies*

$$\Delta t^2 L(R, \rho(A)) \leqslant 1 \qquad \textit{and} \qquad \Delta t \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}^f_{j+2}} \alpha(\tilde{t}) \leqslant \frac{1}{4}.$$

Then, if the operator $\mathscr{A}$ is approximated by a symmetric positive semi-definite matrix A, we have the following stability results,

$$\|\eta^n\| \leqslant \exp\left(2T\left(1 + 2\sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}^f_{j+2}} \alpha(\tilde{t})\right)\right)\left(\|\tilde{\varphi}_1\| + \sqrt{\tilde{\varphi}_1^{\mathsf{T}} A \tilde{\varphi}_1 + \|\tilde{\varphi}_2\|^2}\right),$$

$$\|\zeta^n\| \leqslant \exp\left(2T\left(1 + 2\sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}^f_{j+2}} \alpha(\tilde{t})\right)\right)\left(\|\tilde{\varphi}_1\| + \sqrt{\tilde{\varphi}_1^{\mathsf{T}} A \tilde{\varphi}_1 + \|\tilde{\varphi}_2\|^2}\right),$$

where $\tilde{\varphi}_l = \big(\tilde{\varphi}_l(x_0), \tilde{\varphi}_l(x_1), \cdots, \tilde{\varphi}_l(x_{M-1})\big)^{\mathsf{T}}$ and $\tilde{\varphi}_l(x_i)$ for $l = 1, 2$ are the values of the perturbation functions $\tilde{\varphi}_l$ for $l = 1, 2$, at the spatial grid points $\{x_i\}_{i=0}^{M-1}$.

Proof Using the first formula in (10.38) and (10.39), we obtain

$$\begin{aligned}&\|\eta^{n+1}\| \leqslant \|\eta^n\| + \Delta t\|\zeta^n\| + \sum_{j=0}^{r} \Delta t^{j+2}\Big[\|f_t^{(j)}(v^n) - f_t^{(j)}(u^n)\| \\ &+ \|f_t^{(j)}(v^{n+1}) - f_t^{(j)}(u^{n+1})\|\Big], \\ &\sqrt{(\eta^{n+1})^{\mathsf{T}} A \eta^{n+1} + \|\zeta^{n+1}\|^2} \leqslant \sqrt{(\eta^n)^{\mathsf{T}} A \eta^n + \|\zeta^n\|^2} \\ &+ \sum_{j=0}^{r} \Delta t^{j+1}\Big[\|f_t^{(j)}(v^n) - f_t^{(j)}(u^n)\| + \|f_t^{(j)}(v^{n+1}) - f_t^{(j)}(u^{n+1})\|\Big].\end{aligned} \tag{10.47}$$

Summing up (10.47) and inserting (10.46) into the right-hand side, we have

$$\begin{aligned}
&\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A \eta^{n+1} + \|\zeta^{n+1}\|^2} \leqslant \|\eta^{n}\| + \sqrt{(\eta^{n})^{\intercal} A \eta^{n} + \|\zeta^{n}\|^2} + \Delta t \|\zeta^{n}\| \\
&+ \Delta t(1+\Delta t) \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \Delta t^{j} \Big[\big\| \mathscr{F}(\tilde{t})(v^{n}, w^{n}) - \mathscr{F}(\tilde{t})(u^{n}, \mu^{n}) \big\| \\
&+ \big\| \mathscr{F}(\tilde{t})(v^{n+1}, w^{n+1}) - \mathscr{F}(\tilde{t})(u^{n+1}, \mu^{n+1}) \big\| \Big].
\end{aligned} \tag{10.48}$$

On the other hand, the use of Assumption 10.2 on the right-hand side of (10.48) gives

$$\begin{aligned}
&\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A \eta^{n+1} + \|\zeta^{n+1}\|^2} \\
&\leqslant \|\eta^{n}\| + \sqrt{(\eta^{n})^{\intercal} A \eta^{n} + \|\zeta^{n}\|^2} + \Delta t \|\zeta^{n}\| \\
&\quad + \Delta t(1+\Delta t) \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \Delta t^{j} L(R, \rho(A)^{\lfloor \frac{j}{2} \rfloor}) \\
&\quad \big(\|\eta^{n}\| + \|\zeta^{n}\| + \|\eta^{n+1}\| + \|\zeta^{n+1}\| \big).
\end{aligned} \tag{10.49}$$

As Δt satisfies $\Delta t^2 L(R, \rho(A)) \leqslant 1$, the inequality (10.49) results in

$$\begin{aligned}
&\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A \eta^{n+1} + \|\zeta^{n+1}\|^2} \\
&\leqslant \left\{ 1 + \frac{\Delta t \Big[1 + 2 \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \Big]}{1 - 2\Delta t \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})} \right\} \Big(\|\eta^{n}\| + \sqrt{(\eta^{n})^{\intercal} A \eta^{n} + \|\zeta^{n}\|^2} \Big).
\end{aligned}$$

Furthermore, as $\Delta t \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \leqslant \dfrac{1}{4}$, we obtain

$$\begin{aligned}
&\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\intercal} A \eta^{n+1} + \|\zeta^{n+1}\|^2} \\
&\leqslant \left[1 + 2\Delta t \Big(1 + 2 \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \Big) \right] \Big(\|\eta^{n}\| + \sqrt{(\eta^{n})^{\intercal} A \eta^{n} + \|\zeta^{n}\|^2} \Big).
\end{aligned}$$

Then an argument by induction gives the following result

$$\begin{aligned}&\|\eta^{n+1}\| + \sqrt{(\eta^{n+1})^{\mathsf{T}} A\eta^{n+1} + \|\zeta^{n+1}\|^2}\\ &\leqslant \exp\left(2T\left(1 + 2\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\right)\right)\left(\|\tilde{\varphi}_1\| + \sqrt{\tilde{\varphi}_1^{\mathsf{T}} A\tilde{\varphi}_1 + \|\tilde{\varphi}_2\|^2}\right).\end{aligned}$$

Consequently, we obtain

$$\|\eta^n\| \leqslant \exp\left(2T\left(1 + 2\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\right)\right)\left(\|\tilde{\varphi}_1\| + \sqrt{\tilde{\varphi}_1^{\mathsf{T}} A\tilde{\varphi}_1 + \|\tilde{\varphi}_2\|^2}\right),$$

$$\|\zeta^n\| \leqslant \exp\left(2T\left(1 + 2\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\right)\right)\left(\|\tilde{\varphi}_1\| + \sqrt{\tilde{\varphi}_1^{\mathsf{T}} A\tilde{\varphi}_1 + \|\tilde{\varphi}_2\|^2}\right),$$

This shows the stability of HB time integrators for nonlinear problems. □

10.4 Convergence of the Fully Discrete Scheme

10.4.1 Consistency

Under suitable assumptions on smoothness and the spatial discretisation, it is not difficult to obtain a spatial semidiscrete scheme which is consistent with the original system (10.1) or (10.8). We only require that the truncation error $\delta(\Delta x)$ in (10.58) satisfies $\delta(\Delta x) \to 0$ as $\Delta x \to 0$. In what follows, we analyse the consistency of the fully discrete scheme (10.62) or (10.63). To this end, we first analyse the truncation error of the fully discrete scheme (10.62) or (10.63).

Inserting the exact solution $U(t) = \big(u(x_0,t), u(x_1,t), \cdots, u(x_{M-1},t)\big)^{\mathsf{T}}$ into the fully discrete scheme (10.62), we obtain

$$\begin{cases}\mathscr{T}^n = U(t_{n+1}) - \phi_0(V)U(t_n) - \Delta t\phi_1(V)U'(t_n)\\ \quad - \sum_{j=0}^{r}\Delta t^{j+2}\Big\{I_1[\beta_{r,j}(z)]f_t^{(j)}\big(U(t_n)\big) + (-1)^j I_1[\beta_{r,j}(1-z)]f_t^{(j)}\big(U(t_{n+1})\big)\Big\},\\ \Gamma^n = U'(t_{n+1}) + \Delta t A\phi_1(V)U(t_n) - \phi_0(V)U'(t_n)\\ \quad - \sum_{j=0}^{r}\Delta t^{j+1}\Big\{I_2[\beta_{r,j}(z)]f_t^{(j)}\big(U(t_n)\big) + (-1)^j I_2[\beta_{r,j}(1-z)]f_t^{(j)}\big(U(t_{n+1})\big)\Big\},\end{cases} \tag{10.50}$$

where $\mathscr{T}^n$ and Γ^n are the truncation errors of the fully discrete scheme (10.62) at time t_n. Applying the variation-of-constants formula to (10.58) and comparing the result with (10.50) leads to

$$\begin{cases} \mathscr{T}^n = \Delta t^2 \displaystyle\int_0^1 (1-z)\phi_1\big((1-z)^2 V\big) f\big(U(t_n + z\Delta t)\big)\mathrm{d}z \| \\ \qquad - \displaystyle\sum_{j=0}^{r} \Delta t^{j+2} \Big\{ I_1[\beta_{r,j}(z)] f_t^{(j)}\big(U(t_n)\big) + (-1)^j I_1[\beta_{r,j}(1-z)] f_t^{(j)}\big(U(t_{n+1})\big) \Big\} \\ \qquad + \Delta t^2 \displaystyle\int_0^1 (1-z)\phi_1\big((1-z)^2 V\big)\delta(\Delta x)\mathrm{d}z, \\ \Gamma^n = \Delta t \displaystyle\int_0^1 \phi_0\big((1-z)^2 V\big) f\big(U(t_n + z\Delta t)\big)\mathrm{d}z - \displaystyle\sum_{j=0}^{r} \Delta t^{j+1} \Big\{ I_2[\beta_{r,j}(z)] f_t^{(j)}\big(U(t_n)\big) \\ \qquad + (-1)^j I_2[\beta_{r,j}(1-z)] f_t^{(j)}\big(U(t_{n+1})\big) \Big\} + \Delta t \displaystyle\int_0^1 \phi_0\big((1-z)^2 V\big)\delta(\Delta x)\mathrm{d}z. \end{cases} \tag{10.51}$$

Using the Hermite–Birkhoff interpolation polynomial (see Lemma 10.2) to approximate the nonlinear function $f\big(U(t_n + z\Delta t)\big)$ appearing in (10.51), we have

$$\begin{aligned} \mathscr{T}^n =& (-1)^{r+1}\Delta t^{2r+4} \int_0^1 (1-z)^{r+2}\phi_1\big((1-z)^2 V\big) z^{r+1}\mathrm{d}z \frac{f_t^{(2r+2)}\big(U(t_n + \xi^n \Delta t)\big)}{(2r+2)!} \\ &+ \Delta t^2 \int_0^1 (1-z)\phi_1\big((1-z)^2 V\big)\delta(\Delta x)\mathrm{d}z \end{aligned} \tag{10.52}$$

and

$$\begin{aligned} \Gamma^n =& (-1)^{r+1}\Delta t^{2r+3} \int_0^1 (1-z)^{r+1}\phi_0\big((1-z)^2 V\big) z^{r+1}\mathrm{d}z \frac{f_t^{(2r+2)}\big(U(t_n + \xi^n \Delta t)\big)}{(2r+2)!} \\ &+ \Delta t \int_0^1 \phi_0\big((1-z)^2 V\big)\delta(\Delta x)\mathrm{d}z. \end{aligned} \tag{10.53}$$

We next prove consistency, based on the truncation error analysis.

Theorem 10.6 *Suppose that the exact solution $u(x,t)$ of the original continuous equations (10.1) or (10.8) is sufficiently smooth such that $f\big(u(\cdot,t)\big) \in C^{2r+2}([t_0,T])$ and $f_t^{(2r+2)}\big(u(x,\cdot)\big) \in L^2(\Omega)$. Then, the fully discrete scheme (10.62) or (10.63) is consistent over a finite time interval $t \in [t_0,T]$, i.e.,*

$$\|\mathscr{T}^n\| \to 0 \quad \textit{and} \quad \|\Gamma^n\| \to 0 \quad \textit{as} \quad \Delta t, \Delta x \to 0. \tag{10.54}$$

Proof Taking the l_2-norm on both sides of the truncation errors (10.52) and (10.53) leads to

$$\|\mathscr{T}^n\| \leqslant \tilde{C}_1 \Delta t^{2r+4} + \frac{\Delta t^2}{2}\|\delta(\Delta x)\| \quad \text{and} \quad \|\Gamma^n\| \leqslant \tilde{C}_2 \Delta t^{2r+3} + \Delta t\|\delta(\Delta x)\|, \tag{10.55}$$

where the constants $\tilde{C}_1$ and $\tilde{C}_2$ are given by

$$\tilde{C}_1 = \frac{(r+2)!(r+1)!}{(2r+2)!(2r+4)!} \max_{t_0 \leqslant t \leqslant T} \left\| f_t^{(2r+2)}\big(u(\cdot,t)\big)\right\| \tag{10.56}$$

$$\tilde{C}_2 = \frac{\big[(r+1)!\big]^2}{(2r+4)!(2r+3)!} \max_{t_0 \leqslant t \leqslant T} \left\| f_t^{(2r+2)}\big(u(\cdot,t)\big)\right\|. \tag{10.57}$$

It can be confirmed from (10.55) and (10.56) that the constants $\tilde{C}_1$ and $\tilde{C}_2$ are only dependent on the bounds for derivatives of the exact solution $U(t)$ over a finite time interval $t \in [t_0, T]$. Then, the consistency of the fully discrete scheme (10.62) or (10.63) directly follows from the fact that

$$\|\mathscr{T}^n\| \to 0 \quad \text{and} \quad \|\Gamma^n\| \to 0 \quad \text{as} \quad \Delta t, \Delta x \to 0.$$

The proof of this theorem is complete. □

10.4.2 Convergence

Our next objective is to analyse convergence for the fully discrete schemes. It is well known that convergence of classical methods for linear partial differential equations is governed by the Lax equivalence theorem: convergence is equivalent to consistency plus stability [36]. The HB time integrators are consistent (see Theorem 10.6), and the stability of the fully discrete scheme for linear problems has been proved in Sect. 10.3.1. Consequently, the convergence of the HB time integrators for linear problems can be obtained by applying the Lax equivalence theorem. Unfortunately, however, the Lax equivalence theorem might be less useful for nonlinear problems.

In what follows, we analyse the convergence of the fully discrete scheme for nonlinear problems. Based on some suitable assumptions on smoothness and spatial discretisation strategies, the original continuous system (10.1) or (10.8) can be discretised as:

$$\begin{cases} U''(t) + AU(t) = f\big(U(t)\big) + \delta(\Delta x), \quad t \in [t_0, T], \\ U(t_0) = \varphi_1, \quad U'(t_0) = \varphi_2, \end{cases} \tag{10.58}$$

where $U(t) = \big(u(x_0,t), u(x_1,t), \cdots, u(x_{M-1},t)\big)^{\mathsf{T}}$, A is a positive semi-definite differentiation matrix and $\varphi_l = \big(\varphi_l(x_0), \varphi_l(x_1), \cdots, \varphi_l(x_{M-1})\big)^{\mathsf{T}}$ for $l = 1, 2$. Let $\delta(\Delta x) = \big(\delta_0(\Delta x), \delta_1(\Delta x), \cdots, \delta_{M-1}(\Delta x),\big)^{\mathsf{T}}$ be the truncation error brought by approximating the spatial differential operator $\mathscr{A}$ with a positive semi-definite matrix A, and the truncation error $\delta(\Delta x)$ satisfies $\delta_j(\Delta x) \to 0$ as $\Delta x \to 0$ for $j = 0, 1, \cdots, M-1$. For instance, if we replace the spatial derivative by the classical forth-order finite difference method (see, e.g. [37, 38]), the truncation error $\delta(\Delta x)$ is $\|\delta(\Delta x)\| = O(\Delta x^4)$.

Applying the time integration formula (10.22) to (10.58) leads to

$$\begin{cases} U(t_{n+1}) = \phi_0(V)U(t_n) + \Delta t\phi_1(V)U'(t_n) + \sum_{j=0}^{r} \Delta t^{j+2}\Big\{ I_1[\beta_{r,j}(z)] f_t^{(j)}\big(U(t_n)\big) \\ \qquad\qquad + (-1)^j I_1[\beta_{r,j}(1-z)] f_t^{(j)}\big(U(t_{n+1})\big)\Big\} + R^n, \\ U'(t_{n+1}) = -\Delta t A\phi_1(V)U(t_n) + \phi_0(V)U'(t_n) + \sum_{j=0}^{r} \Delta t^{j+1}\Big\{ I_2[\beta_{r,j}(z)] f_t^{(j)}\big(U(t_n)\big) \\ \qquad\qquad + (-1)^j I_2[\beta_{r,j}(1-z)] f_t^{(j)}\big(U(t_{n+1})\big)\Big\} + r^n, \end{cases} \tag{10.59}$$

where $R^n = \big(R_0^n, \cdots, R_{M-1}^n\big)^{\mathsf{T}}$ and $r^n = \big(r_0^n, \cdots, r_{M-1}^n\big)^{\mathsf{T}}$ are truncation errors, and

$$R_j^n = (-1)^{r+1}\Delta t^{2r+4}\frac{f_t^{(2r+2)}\big(u(x_j, t_n + \xi^n\Delta t)\big)}{(2r+2)!}\int_0^1 (1-z)^{r+2}\phi_1\big((1-z)^2V\big)z^{r+1}\mathrm{d}z$$
$$+ \Delta t^2\int_0^1 (1-z)\phi_1\big((1-z)^2V\big)\delta_j(\Delta x)\mathrm{d}z,$$

and

$$r_j^n = (-1)^{r+1}\Delta t^{2r+3}\frac{f_t^{(2r+2)}\big(u(x_j, t_n + \xi^n\Delta t)\big)}{(2r+2)!}\int_0^1 (1-z)^{r+1}\phi_0\big((1-z)^2V\big)z^{r+1}\mathrm{d}z$$
$$+ \Delta t\int_0^1 \phi_0\big((1-z)^2V\big)\delta_j(\Delta x)\mathrm{d}z$$

respectively. Under suitable assumptions of smoothness, the errors R_j^n and r_j^n satisfy

$$|R_j^n| \leqslant \frac{(r+2)!(r+1)!}{(2r+2)!(2r+4)!}\max_{t_0\leqslant t\leqslant T}\max_{x\in\bar{\Omega}}|f_t^{(2r+2)}\big(u(x,t)\big)|\Delta t^{2r+4} + \frac{\Delta t^2}{2}|\delta_j(\Delta x)|, \tag{10.60}$$

and

$$|r_j^n| \leqslant \frac{\big[(r+1)!\big]^2}{(2r+4)!(2r+3)!} \max_{t_0 \leqslant t \leqslant T} \max_{x \in \bar{\Omega}} |f_t^{(2r+2)}\big(u(x,t)\big)| \Delta t^{2r+3} + \Delta t |\delta_j(\Delta x)|. \tag{10.61}$$

Disregarding the small terms R^n and r^n in (10.59) and letting $u_j^n \approx u(x_j, t_n)$, $\mu_j^n \approx u_t(x_j, t_n)$, the following fully discrete scheme follows

$$\begin{cases} u^{n+1} = \phi_0(V)u^n + \Delta t \phi_1(V)\mu^n \\ \qquad + \sum_{j=0}^{r} \Delta t^{j+2} \Big\{ I_1[\beta_{r,j}(z)] f_t^{(j)}(u^n) + (-1)^j I_1[\beta_{r,j}(1-z)] f_t^{(j)}(u^{n+1}) \Big\}, \\ \mu^{n+1} = -\Delta t A \phi_1(V)u^n + \phi_0(V)\mu^n \\ \qquad + \sum_{j=0}^{r} \Delta t^{j+1} \Big\{ I_2[\beta_{r,j}(z)] f_t^{(j)}(u^n) + (-1)^j I_2[\beta_{r,j}(1-z)] f_t^{(j)}(u^{n+1}) \Big\}. \end{cases} \tag{10.62}$$

In terms of the notation in (10.46), we rewrite the fully discrete scheme (10.62) in the following form

$$\begin{cases} u^{n+1} = \phi_0\big(V\big)u^n + \Delta t \phi_1\big(V\big)\mu^n + \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^f} \alpha(\tilde{t}) \Delta t^{j+2} \\ \qquad \times \Big[I_1[\beta_{r,j}(z)] \mathscr{F}(\tilde{t})(u^n, \mu^n) + (-1)^j I_1[\beta_{r,j}(1-z)] \mathscr{F}(\tilde{t})(u^{n+1}, \mu^{n+1}) \Big], \\ \mu^{n+1} = -\Delta t A \phi_1\big(V\big)u^n + \phi_0\big(V\big)\mu^n + \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^f} \alpha(\tilde{t}) \Delta t^{j+1} \\ \qquad \times \Big[I_2[\beta_{r,j}(z)] \mathscr{F}(\tilde{t})(u^n, \mu^n) + (-1)^j I_2[\beta_{r,j}(1-z)] \mathscr{F}(\tilde{t})(u^{n+1}, \mu^{n+1}) \Big], \end{cases} \tag{10.63}$$

where $I_1[\beta_{r,j}(z)]$, $I_2[\beta_{r,j}(z)]$, $I_1[\beta_{r,j}(1-z)]$ and $I_2[\beta_{r,j}(1-z)]$ have been defined by (10.18)–(10.21).

We next consider from first principles the convergence of the fully discrete scheme (10.63) for nonlinear problems. We let $e_j^n = u(x_j, t_n) - u_j^n$ and $\omega_j^n = u_t(x_j, t_n) - \mu_j^n$ for $j = 0, 1, \cdots, M-1$, i.e., $e^n = U(t_n) - u^n$ and $\omega^n = U'(t_n) - \mu^n$. Subtracting (10.63) from (10.59), and inserting exact initial conditions, we get a

recurrence relation for the errors,

$$\begin{cases} e^{n+1} = \phi_0(V)e^n + \Delta t\phi_1(V)\omega^n \\ \quad + \sum_{j=0}^{r} \sum_{\tilde{t}\in \mathrm{SENT}^f_{j+2}} \alpha(\tilde{t})\Delta t^{j+2} \Big\{ I_1[\beta_{r,j}(z)]\Big[\mathscr{F}(\tilde{t})\big(U(t_n), U'(t_n)\big) - \mathscr{F}(\tilde{t})\big(u^n, \mu^n\big)\Big] \\ \quad + (-1)^j I_1[\beta_{r,j}(1-z)]\Big[\mathscr{F}(\tilde{t})\big(U(t_{n+1}), U'(t_{n+1})\big) - \mathscr{F}(\tilde{t})\big(u^{n+1}, \mu^{n+1}\big)\Big]\Big\} + R^n, \\ \omega^{n+1} = -\Delta t A\phi_1(V)e^n + \phi_0(V)\omega^n \\ \quad + \sum_{j=0}^{r} \sum_{\tilde{t}\in \mathrm{SENT}^f_{j+2}} \alpha(\tilde{t})\Delta t^{j+1} \Big\{ I_2[\beta_{r,j}(z)]\Big[\mathscr{F}(\tilde{t})\big(U(t_n), U'(t_n)\big) - \mathscr{F}(\tilde{t})\big(u^n, \mu^n\big)\Big] \\ \quad + (-1)^j I_2[\beta_{r,j}(1-z)]\Big[\mathscr{F}(\tilde{t})\big(U(t_{n+1}), U'(t_{n+1})\big) - \mathscr{F}(\tilde{t})\big(u^{n+1}, \mu^{n+1}\big)\Big]\Big\} + r^n, \end{cases} \tag{10.64}$$

with the initial conditions $e^0 = 0, \omega^0 = 0$.

For the convergence analysis, we quote the Gronwall's inequality (see, e.g. [39]), which plays an important role in the analysis.

Lemma 10.7 *Let λ be positive, a_k, b_k, $k \in \mathbb{N}$, be nonnegative and assume further that*

$$a_k \leqslant (1 + \lambda\Delta t)a_{k-1} + \Delta t b_k, \qquad k \in \mathbb{N}.$$

Then

$$a_k \leqslant \exp(\lambda k\Delta t)\left(a_0 + \Delta t \sum_{m=1}^{k} b_m\right), \qquad k \in \mathbb{N}.$$

Theorem 10.7 *With Assumptions 10.1 and 10.2, suppose that $u(x,t)$ satisfies suitable smoothness assumptions. If the time stepsize Δt satisfies*

$$\Delta t^2 L\big(R, \rho(A)\big) \leqslant 1 \qquad \text{and} \qquad \Delta t \sum_{j=0}^{r} \sum_{\tilde{t}\in \mathrm{SENT}^f_{j+2}} \alpha(\tilde{t}) \leqslant \frac{1}{4},$$

then there exists a constant C such that

$$\|e^n\| \leqslant CT\exp\left(2T\left(1+2\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}}\alpha(\tilde{t})\right)\right)\left(\Delta t^{2r+2}+\|\delta(\Delta x)\|\right),$$

$$\|\omega^n\| \leqslant CT\exp\left(2T\left(1+2\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}}\alpha(\tilde{t})\right)\right)\left(\Delta t^{2r+2}+\|\delta(\Delta x)\|\right).$$

Proof The error system (10.64) can be rewritten in a compact form

$$\mathscr{F}(\tilde{t})_n \equiv \mathscr{F}(\tilde{t})\big(U(t_n), U'(t_n)\big) - \mathscr{F}(\tilde{t})\big(u^n, \mu^n\big)$$

and

$$\mathscr{F}(\tilde{t})_{n+1} \equiv \mathscr{F}(\tilde{t})\big(U(t_{n+1}), U'(t_{n+1})\big) - \mathscr{F}(\tilde{t})\big(u^{n+1}, \mu^{n+1}\big),$$

$$\begin{aligned}\begin{bmatrix} De^{n+1} \\ \omega^{n+1}\end{bmatrix} &= \Omega(V)\begin{bmatrix} De^{n} \\ \omega^{n}\end{bmatrix} + \sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}}\Delta t^{j+1}\int_0^1 \Omega_j\big(\beta(z), V\big)\mathrm{d}z\begin{bmatrix} 0 \\ \mathscr{F}(\tilde{t})_n\end{bmatrix}\\ &+\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}}(-1)^j\Delta t^{j+1}\int_0^1 \Omega_j\big(\beta(1-z), V\big)\mathrm{d}z\begin{bmatrix} 0 \\ \mathscr{F}(\tilde{t})_{n+1}\end{bmatrix} + \begin{bmatrix} DR^n \\ r^n\end{bmatrix},\end{aligned} \tag{10.65}$$

where $\Omega(V)$ and $\Omega\big(\beta(z), V\big)$ were defined in (10.40) and (10.41), respectively. On the one hand, taking the l_2-norm on both sides of the first formula in (10.64) and (10.65) and summing up the outcomes, we have

$$\begin{aligned}&\|e^{n+1}\| + \sqrt{(e^{n+1})^{\intercal}Ae^{n+1} + \|\omega^{n+1}\|^2} \leqslant \|e^n\| + \sqrt{(e^n)^{\intercal}Ae^n + \|\omega^n\|^2} + \Delta t\|\omega^n\|\\ &+\Delta t(1+\Delta t)\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}}\Delta t^j\left[\left\|\mathscr{F}(\tilde{t})\big(U(t_n), U'(t_n)\big) - \mathscr{F}(\tilde{t})\big(u^n, \mu^n\big)\right\|\right.\\ &\left.+\left\|\mathscr{F}(\tilde{t})\big(U(t_{n+1}), U'(t_{n+1})\big) - \mathscr{F}(\tilde{t})\big(u^{n+1}, \mu^{n+1}\big)\right\|\right]\\ &+\|R^n\| + \sqrt{\|DR^n\|^2 + \|r^n\|^2}.\end{aligned} \tag{10.66}$$

On the other hand, applying Assumption 10.2 to the right-hand side of (10.66) results in

$$\begin{aligned}\|e^{n+1}\| + \sqrt{(e^{n+1})^{\intercal} A e^{n+1} + \|\omega^{n+1}\|^2} &\leqslant \|e^{n}\| + \sqrt{(e^{n})^{\intercal} A e^{n} + \|\omega^{n}\|^2} + \Delta t \|\omega^{n}\| \\ &+ \Delta t(1+\Delta t) \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \Delta t^{j} L\big(R, \rho(A)^{\lfloor \frac{j}{2} \rfloor}\big) \Big(\|e^{n}\| + \|\omega^{n}\| + \|e^{n+1}\| \\ &+ \|\omega^{n+1}\| \Big) + \|R^{n}\| + \sqrt{\|DR^{n}\|^2 + \|r^{n}\|^2}. \end{aligned} \tag{10.67}$$

As $\Delta t^2 L\big(R, \rho(A)\big) \leqslant 1$, the inequality (10.67) leads to

$$\begin{aligned}&\|e^{n+1}\| + \sqrt{(e^{n+1})^{\intercal} A e^{n+1} + \|\omega^{n+1}\|^2} \\ &\leqslant \left\{ 1 + \frac{\Delta t \left[1 + 2 \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \right]}{1 - 2\Delta t \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})} \right\} \Big(\|e^{n}\| + \sqrt{(e^{n})^{\intercal} A e^{n} + \|\omega^{n}\|^2} \Big) \\ &+ \frac{1}{1 - 2\Delta t \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})} \Big(\|R^{n}\| + \sqrt{\|DR^{n}\|^2 + \|r^{n}\|^2} \Big). \end{aligned} \tag{10.68}$$

If the time stepsize Δt also satisfies $\Delta t \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \leqslant \frac{1}{4}$, then the inequality (10.68) results in

$$\begin{aligned}&\|e^{n+1}\| + \sqrt{(e^{n+1})^{\intercal} A e^{n+1} + \|\omega^{n+1}\|^2} \\ &\leqslant \left\{ 1 + 2\Delta t \Big[1 + 2 \sum_{j=0}^{r} \sum_{\tilde{t} \in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) \Big] \right\} \Big(\|e^{n}\| + \sqrt{(e^{n})^{\intercal} A e^{n} + \|\omega^{n}\|^2} \Big) \\ &+ 2 \Big(\|R^{n}\| + \sqrt{\|DR^{n}\|^2 + \|r^{n}\|^2} \Big). \end{aligned} \tag{10.69}$$

Note that R_j^n and r_j^n satisfy (10.60) and (10.61), respectively. Hence, there exists a constant C such that

$$\|R^n\| + \sqrt{\|DR^n\|^2 + \|r^n\|^2} \leqslant C\Delta t\Big(\Delta t^{2r+2} + \|\delta(\Delta x)\|\Big).$$

Applying the Gronwall's inequality (Lemma 10.7) to (10.69) yields

$$\|e^n\| + \sqrt{(e^n)^{\intercal} A e^n + \|\omega^n\|^2} \leqslant \exp\left(2n\Delta t\left(1 + 2\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^f}\alpha(\tilde{t})\right)\right)$$
$$\times\left[\|e^0\| + \sqrt{(e^0)^{\intercal} A e^0 + \|\omega^0\|^2} + Cn\Delta t\Big(\Delta t^{2r+2} + \|\delta(\Delta x)\|\Big)\right].$$

Therefore, we obtain

$$\|e^n\| \leqslant CT\exp\left(2T\left(1 + 2\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^f}\alpha(\tilde{t})\right)\right)\Big(\Delta t^{2r+2} + \|\delta(\Delta x)\|\Big),$$
$$\|\omega^n\| \leqslant CT\exp\left(2T\left(1 + 2\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^f}\alpha(\tilde{t})\right)\right)\Big(\Delta t^{2r+2} + \|\delta(\Delta x)\|\Big). \tag{10.70}$$

Then the proof of this theorem is complete. □

Obviously, it follows from the analysis of Theorem 10.7 that the precision of the derived HB time integrators can be of order $(2r + 2)$ in time, provided the exact solution $u(x, t)$ of the semilinear KG equations (10.1) satisfies $u(\cdot, t) \in C^{2r+2}\big([t_0, T]\big)$. Unfortunately, however, existing numerical schemes, such as the finite difference method and the finite element method, have only limited accuracy for solving the semilinear KG equations (10.1). Here, in order to design high-order numerical methods, higher smoothness assumptions of the underlying problem are required. For instance, assume that $u(x, t)$ has appropriately continuous derivatives with respect to the temporal variable, and we use the following fourth-order finite difference approximation

$$\frac{\partial^2 u(x_j, t_n)}{\partial t^2} = \frac{-u(x_j, t_{n+2}) + 16u(x_j, t_{n+1}) - 30u(x_j, t_n) + 16u(x_j, t_{n-1}) - u(x_j, t_{n-2})}{\Delta t^4}$$
$$- \frac{\Delta t^4}{90}\frac{\partial^6 u(x_j, \hat{\xi}^n)}{\partial t^6}.$$

This implies that the approximation needs the solution to satisfy $u(\cdot,t) \in C^6\big([t_0,T]\big)$ at last. However, under the same smoothness assumption of $u(\cdot,t) \in C^6\big([t_0,T]\big)$, we can obtain a sixth-order HB time integrator by Theorem 10.7. In particular, as an important example, if the exact solution satisfies $u(\cdot,t) \in C^4\big([t_0,T]\big)$, the well-known *leap-frog scheme* or the *Störmer–Verlet formula* is of order two in time. Fortunately, the derived HB time integrator with $r = 1$ can achieve fourth-order convergence. This is definitely a major improvement.

Moreover, under the smoothness assumption of $u(\cdot,t) \in C^2\big([t_0,T]\big)$, and taking $r = 0$ in the time integration formula (10.22), we obtain an interesting scheme as follows:

$$\begin{cases} u^{n+1} = \phi_0\big(\mathscr{V}\big)u^n + \Delta t\phi_1\big(\mathscr{V}\big)\mu^n \\ \qquad\qquad + \Delta t^2\left\{I_1[\beta_{0,0}(z)]f(u^n) + (-1)^j I_1[\beta_{0,0}(1-z)]f(u^{n+1})\right\}, \\ \mu^{n+1} = -\Delta t\mathscr{A}\phi_1\big(\mathscr{V}\big)u^n + \phi_0\big(\mathscr{V}\big)\mu^n \\ \qquad\qquad + \Delta t\left\{I_2[\beta_{0,0}(z)]f(u^n) + (-1)^j I_2[\beta_{0,0}(1-z)]f(u^{n+1})\right\}. \end{cases} \tag{10.71}$$

The scheme (10.71) is of order two.

Remark 10.2 Compared with the well-known *Störmer–Verlet method*, the second-order HB time integrator needs a much weaker smoothness assumption, whereas the interesting second-order scheme (10.71) exhibits excellent numerical behaviour. This remarkable superiority will be shown in the numerical experiments.

10.5 Spatial Discretisation

As stated above, the symmetric and arbitrarily high-order time integration formula (10.22) has been presented in operatorial terms in an infinite-dimensional function space $\mathscr{X}$. In order to render them into proper numerical algorithms, we need to replace the differential operator $\mathscr{A}$ with an suitable differentiation matrix A. Keeping the stability and convergence analysis in mind, we approximate the differential operator $\mathscr{A}$ by a positive semi-definite matrix A. Fortunately, there exists a great body of research investigating the replacement of spatial derivatives of nonlinear system (10.1) with periodic boundary conditions (10.2), and it is not difficult to find positive semi-definite differentiation matrices in this setting. Here, we mainly consider two types of spatial discretisations: Symmetric finite difference and Fourier spectral collocation discretisations.

1. *Symmetric Finite Difference (SFD)* (see, e.g. [37])
   As is known, finite difference methods are achieved when approximating a function by local polynomial interpolation. Its derivatives are then approximated by differentiating this local polynomial, where 'local' refers to the use of nearby grid points to approximate the function or its derivative at a given point. In general, a finite difference approximation is of moderate order. For instance, we approximate the operator $\mathscr{A}$ by the following differentiation matrix

$$A_{\text{sfd}} = \frac{a^2}{12\Delta x^2}\begin{bmatrix} 30 & -16 & 1 & & & 1 & -16 \\ -16 & 30 & -16 & 1 & & & 1 \\ 1 & -16 & 30 & -16 & 1 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & 1 & -16 & 30 & -16 & 1 \\ 1 & & & 1 & -16 & 30 & -16 \\ -16 & 1 & & & 1 & -16 & 30 \end{bmatrix}_{M\times M}.$$

   The approximation is of order four and the differentiation matrix A_{sfd} is clearly positive semi-definite.
2. *Fourier Spectral Collocation (FSC)* (see, e.g. [40, 41])
   A distinctive feature of Spectral methods is their global nature, and the computation at any given point depends not only on the information at neighbouring points, but on the entire domain. The topic of spectral methods is very wide, and various methods and sub-methods have been proposed and are actively used. The Fourier spectral collocation method is our method of choice, which can be presented as a limit of local finite difference approximations of increasing order of accuracy (see [40]). We concentrate on differentiation being performed in the physical space. The key point here is to interpolate the solution at the nodal values using a trigonometric polynomial. The entries of the second-derivative Fourier differentiation matrix $A_{\text{fsc}} = (a_{kj})_{M\times M}$ are given by

$$a_{kj} = \begin{cases} \dfrac{(-1)^{k+j}}{2} a^2 \sin^{-2}\left(\dfrac{(k-j)\pi}{M}\right), & k \neq j, \\ a^2\left(\dfrac{M^2}{12} + \dfrac{1}{6}\right), & k = j. \end{cases} \tag{10.72}$$

   It is known that the main appeal of spectral methods is that they exhibit spectral convergence to $\mathscr{A}$: the error decays for C^∞ functions faster than $O(M^{-\alpha})\ \forall\, \alpha > 0$ for sufficiently large M. Another advantage is that the differentiation matrix A_{fsc} is positive semi-definite.

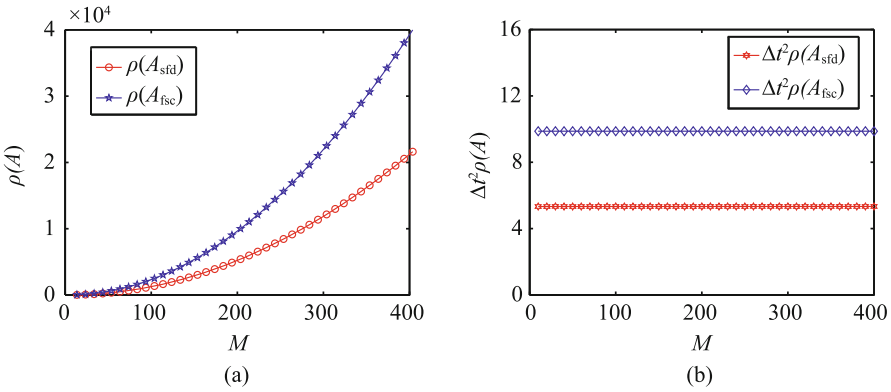


Fig. 10.1 Plots of $\rho(A)$ and $\Delta t^2\rho(A)$ for the differentiation matrices A_{sfd} and A_{fsc} for $M = 10i$ and $i = 1, 2, \cdots, 40$

Figure 10.1 illustrates the size of the spectral radius for the differentiation matrices A_{sfd} and A_{fsc}. In Fig. 10.1a, we show the spectral radius of A_{sfd} and A_{fsc} as a function of M. If we take the time stepsize $\Delta t = \dfrac{2\pi}{M}, M = 10i$ for $i = 1, 2, \cdots, 40$, Fig. 10.1b shows that $\Delta t^2\rho(A)$ is a constant. i.e. that $\rho(A) = O(\Delta t^{-2})$. Therefore, a small Δt $\left(\Delta t \leqslant \dfrac{2\pi}{M}\right)$ can be chosen to guarantee stability and convergence and obtain effective numerical methods.

We have already noted that energy conservation (10.3) is a crucial property of semilinear KG equations (10.1)–(10.2). As we approximate the operator $\mathscr{A}$ by a positive semi-definite differentiation matrix A, there is a corresponding discrete energy conservation law, which can be characterized by the following form:

$$\tilde{E}(t) = \frac{\Delta x}{2}\|u'(t)\|^2 + \frac{\Delta x}{2}\|Du(t)\|^2 + \Delta x \sum_{j=0}^{M-1} V\big(u_j(t)\big) \equiv \tilde{E}(t_0), \tag{10.73}$$

where the norm $\|\cdot\|$ is the standard vector 2-norm and $\Delta x = 2\pi/M$ is the spatial grid size. Actually, this energy can be thought of as an approximate energy (a semidiscrete energy) of the original continuous system. Consequently, discussing numerical experiments, we will also test the effectiveness of HB time integrators in preserving (10.73).

We are now concerned with how accurately the discrete energy conservation law (10.73) is preserved by the HB time integrators. We first rewrite the semidiscrete system of the nonlinear KG equation (10.1)–(10.2) in the form

$$\begin{bmatrix} u(t) \\ u'(t) \end{bmatrix}' = \begin{bmatrix} u'(t) \\ -Au(t) + f\big(u(t)\big) \end{bmatrix}. \tag{10.74}$$

On the one hand, we note that if we define

$$H\big(u(t), u'(t)\big) = \frac{1}{2}u'(t)^{\mathsf{T}}u'(t) + \frac{1}{2}u(t)^{\mathsf{T}}Au(t) + \tilde{V}\big(u(t)\big), \tag{10.75}$$

where $\tilde{V}\big(u(t)\big) = \sum_{j=0}^{M-1} V\big(u_j(t)\big)$, the discrete energy conservation law (10.73) is identical to

$$\Delta x H\big(u(t), u'(t)\big) \equiv \Delta x H\big(u(t_0), u'(t_0)\big). \tag{10.76}$$

By letting $y(t) = \Big[u(t)^{\mathsf{T}}, u'(t)^{\mathsf{T}}\Big]^{\mathsf{T}}$, where $u(t)$ and $u'(t)$ are the exact solution of (10.74) and its derivative, respectively, the system (10.74) can be further expressed as:

$$y'(t) = J^{-1}\nabla H\big(y(t)\big) \quad \text{with} \quad J = \begin{bmatrix} \mathbf{0} & -I_{M\times M} \\ I_{M\times M} & \mathbf{0} \end{bmatrix}. \tag{10.77}$$

On the other hand, if the numerical solutions u^{n+1} and μ^{n+1} are regarded as functions of Δt, and by denoting $z(t_n + \xi\Delta t) = \big[u^n(\xi\Delta t)^{\mathsf{T}}, \mu^n(\xi\Delta t)^{\mathsf{T}}\big]^{\mathsf{T}}$, it can be observed that the solutions of the HB time integration formula (10.22) satisfy

$$z'(t_n + \xi\Delta t) = \begin{bmatrix} \mu^n(\xi\Delta t) \\ \Upsilon^n(\xi\Delta t, u) \end{bmatrix}, \tag{10.78}$$

where

$$\Upsilon^n(\xi\Delta t, u) \equiv -Au^n(\xi\Delta t)$$
$$+\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\Delta t^j \Big[\beta_{r,j}(\xi)\mathscr{F}\big(u^n, \mu^n\big) + (-1)^j\beta_{r,j}(1-\xi)\mathscr{F}\big(u^{n+1}, \mu^{n+1}\big)\Big],$$

$\xi \in [0, 1]$, and $z(t_n + \xi\Delta t)$ satisfies:

$$z(t_n + \xi\Delta t)\big|_{\xi=0} = \begin{bmatrix} u^n \\ \mu^n \end{bmatrix} \quad \text{and} \quad z(t_n + \xi\Delta t)\big|_{\xi=1} = \begin{bmatrix} u^{n+1} \\ \mu^{n+1} \end{bmatrix}, \quad 0 \leqslant n \leqslant N.$$

Theorem 10.8 *Let u^n and μ^n be the solutions of the HB time integration formula (10.22). Then the discrete energy defined in (10.75) satisfies*

$$\max_{0\leqslant n\leqslant N} \big|H(u^n, u'^n) - H(u^0, u'^0)\big| = \mathscr{O}(\Delta t^{2r+2}), \tag{10.79}$$

and this implies the order of preservation of the discrete energy is $2r + 2$.

Proof Using (10.75) and (10.78), we obtain

$$
\begin{aligned}
&H\big(z(t_{n+1})\big) - H\big(z(t_n)\big) = \Delta t \int_0^1 \nabla H\big(z(t_n + \xi\Delta t)\big)^{\mathsf{T}} z'(t_n + \xi\Delta t)\mathrm{d}\xi \\
=&\Delta t \int_0^1 \left[\Big(Au^n(\xi\Delta t) - f\big(u^n(\xi\Delta t)\big)\Big)^{\mathsf{T}}, \mu^n(\xi\Delta t)^{\mathsf{T}}\right]\begin{bmatrix}\mu^n(\xi\Delta t)\\ \Upsilon^n(\xi\Delta t, u)\end{bmatrix}\mathrm{d}\xi \\
=&(-1)^{r+1}\Delta t^{2r+3}\left(\int_0^1 \mu^n(\xi\Delta t)\xi^{r+1}(1-\xi)^{r+1}\mathrm{d}\xi\right)^{\mathsf{T}} \frac{f_t^{(2r+2)}\big(u^n(\theta^n\Delta t)\big)}{(2r+2)!}, \\
&\theta^n \in [0,1].
\end{aligned}
$$

This leads to

$$
\big|H\big(z(t_{n+1})\big) - H\big(z(t_n)\big)\big| = \mathscr{O}(\Delta t^{2r+3}).
$$

It then follows from

$$
\big|H\big(z(t_n)\big) - H\big(z(t_0)\big)\big| \leqslant \sum_{j=0}^{n-1} \big|H\big(z(t_{j+1})\big) - H\big(z(t_j)\big)\big| = n\mathscr{O}(\Delta t^{2r+3}),
$$

that

$$
\max_{0\leqslant n\leqslant N} \big|H\big(z(t_n)\big) - H\big(z(t_0)\big)\big| = \mathscr{O}(\Delta t^{2r+2}).
$$

The proof of this is complete. □

10.6 Waveform Relaxation and Its Convergence

The previous sections derived and analysed the fully discrete scheme for (10.1)–(10.2) and presented its properties. However, the scheme (10.63) is implicit in general and iteration cannot be avoided in practical computation. In this section we introduce a *waveform relaxation method* as a suitable iterative procedure. The waveform relaxation method has been investigated by many authors (see, e.g. [42–46]).

For simplicity, in terms of the notation in (10.46), we first rewrite the fully discrete scheme (10.63),

$$\begin{cases} u^{n+1} = \phi_0(V)u^n + \Delta t\phi_1(V)\mu^n + \sum_{j=0}^{r} \sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\Delta t^{j+2} \\ \qquad \times \Big[I_1[\beta_j(z)]\mathscr{F}(\tilde{t})(u^n, \mu^n) + (-1)^j I_1[\beta_j(1-z)]\mathscr{F}(\tilde{t})(u^{n+1}, \mu^{n+1}) \Big], \\ \mu^{n+1} = -\Delta t A\phi_1(V)u^n + \phi_0(V)\mu^n + \sum_{j=0}^{r} \sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\Delta t^{j+1} \\ \qquad \times \Big[I_2[\beta_j(z)]\mathscr{F}(\tilde{t})(u^n, \mu^n) + (-1)^j I_2[\beta_j(1-z)]\mathscr{F}(\tilde{t})(u^{n+1}, \mu^{n+1}) \Big], \end{cases}$$

where $I_1[\beta_j(z)]$, $I_2[\beta_j(z)]$, $I_1[\beta_j(1-z)]$ and $I_2[\beta_j(1-z)]$ have been defined in (10.19)–(10.18). We then define the waveform relaxation method as follows:

$$\begin{cases} u_{[0]}^{n+1} = \phi_0(V)u^n + \Delta t\phi_1(V)\mu^n, \\ \mu_{[0]}^{n+1} = -\Delta t A\phi_1(V)u^n + \phi_0(V)\mu^n, \end{cases} \tag{10.80}$$

and subsequently iterate

$$\begin{cases} u_{[m+1]}^{n+1} = u_{[0]}^{n+1} + \sum_{j=0}^{r} \sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\Delta t^{j+2} \Big\{ I_1[\beta_{r,j}(z)]\mathscr{F}(\tilde{t})(u^n, \mu^n) \\ \qquad + (-1)^j I_1[\beta_{r,j}(1-z)]\mathscr{F}(\tilde{t})(u_{[m]}^{n+1}, \mu_{[m]}^{n+1}) \Big\}, \\ \mu_{[m+1]}^{n+1} = \mu_{[0]}^{n+1} + \sum_{j=0}^{r} \sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\Delta t^{j+1} \Big\{ I_2[\beta_{r,j}(z)]\mathscr{F}(\tilde{t})(u^n, \mu^n) \\ \qquad + (-1)^j I_2[\beta_{r,j}(1-z)]\mathscr{F}(\tilde{t})(u_{[m]}^{n+1}, \mu_{[m]}^{n+1}) \Big\} \end{cases} \tag{10.81}$$

for $m = 0, 1, \cdots$.

In what follows, we analyse the convergence of the algorithm (10.80)–(10.81).

Theorem 10.9 *Suppose that f satisfies Assumptions 10.1 and 10.2. Under the conditions*

$$\Delta t^2 L(R,\rho(A)) \leqslant 1 \qquad \text{and} \qquad \Delta t(1+\Delta t)\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) < 1,$$

the iterative procedure determined by (10.80)–(10.81) is convergent.

Proof According to Assumption 10.2 and (10.81), the following inequalities are true:

$$\begin{cases} \|u_{[m+1]}^{n+1} - u_{[m]}^{n+1}\| \\ \leqslant \Delta t^2 \sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\Delta t^j L(R,\rho(A)^{\lfloor \frac{j}{2}\rfloor})\Big(\|u_{[m]}^{n+1} - u_{[m-1]}^{n+1}\| + \|\mu_{[m]}^{n+1} - \mu_{[m-1]}^{n+1}\|\Big), \\ \|\mu_{[m+1]}^{n+1} - \mu_{[m]}^{n+1}\| \\ \leqslant \Delta t \sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\Delta t^j L(R,\rho(A)^{\lfloor \frac{j}{2}\rfloor})\Big(\|u_{[m]}^{n+1} - u_{[m-1]}^{n+1}\| + \|\mu_{[m]}^{n+1} - \mu_{[m-1]}^{n+1}\|\Big). \end{cases} \tag{10.82}$$

Summing up (10.82) and noting that $\Delta t^2 L(R,\rho(A)) \leqslant 1$, we obtain

$$\begin{aligned} &\|u_{[m+1]}^{n+1} - u_{[m]}^{n+1}\| + \|\mu_{[m+1]}^{n+1} - \mu_{[m]}^{n+1}\| \\ &\leqslant \Delta t(1+\Delta t)\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\Big(\|u_{[m]}^{n+1} - u_{[m-1]}^{n+1}\| + \|\mu_{[m]}^{n+1} - \mu_{[m-1]}^{n+1}\|\Big). \end{aligned}$$

An argument by induction then gives

$$\begin{aligned} &\|u_{[m+1]}^{n+1} - u_{[m]}^{n+1}\| + \|\mu_{[m+1]}^{n+1} - \mu_{[m]}^{n+1}\| \\ &\leqslant \left[\Delta t(1+\Delta t)\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t})\right]^m \Big(\|u_{[1]}^{n+1} - u_{[0]}^{n+1}\| + \|\mu_{[1]}^{n+1} - \mu_{[0]}^{n+1}\|\Big). \end{aligned}$$

The condition $\Delta t(1+\Delta t)\sum_{j=0}^{r}\sum_{\tilde{t}\in \mathrm{SENT}_{j+2}^{f}} \alpha(\tilde{t}) < 1$ results in

$$\lim_{m\to+\infty}\Big(\|u_{[m+1]}^{n+1} - u_{[m]}^{n+1}\| + \|\mu_{[m+1]}^{n+1} - \mu_{[m]}^{n+1}\|\Big) = 0. \tag{10.83}$$

Therefore, the iterative procedure (10.80)–(10.81) is convergent. □

10.7 Numerical Experiments

For the demonstration of the properties and performance of the HB time integrator, in this section, we derive three practical time integration formulae and use them to illustrate the solution of two semilinear wave equations.

The choice of $r = 0$ in (10.15) yields the first example of a symmetric time-stepping integrator for (10.1)–(10.2):

$$\beta_{0,0}(z) = (1 - z), \tag{10.84}$$

and the corresponding time integration formula, determined by (10.84) and (10.18)–(10.21), is defined by HB0.

As the second example, we take $r = 1$ in (10.15)

$$\beta_{1,0}(z) = (1 - z)^2(1 + 2z), \quad \beta_{1,1}(z) = z(1 - z)^2. \tag{10.85}$$

The time integration formula determined by (10.85) and (10.18)–(10.21) is denoted by HB1.

Letting $r = 2$ in (10.15) gives the third example:

$$\begin{aligned} &\beta_{2,0}(z) = (1 - z)^3(1 + 3z + 6z^2), \quad \beta_{2,1}(z) = z(1 - z)^3(1 + 3z), \\ &\beta_{2,2}(z) = \frac{1}{2}z^2(1 - z)^3. \end{aligned} \tag{10.86}$$

The corresponding time integration formula determined by (10.86) and (10.18)–(10.21) as HB2.

In order to compare different algorithms, we briefly describe a number of standard finite difference schemes and method-of-lines schemes for the semilinear KG equation (see, e.g. [9, 10, 39]).

1. *Standard Finite Difference Schemes*
   Let u_j^n be the approximation of $u(x_j, t_n)$ for $j = 0, 1, \cdots, M - 1$ and $n = 0, 1, \cdots, N$. We also introduce the standard central difference operators

$$\delta_t^2 u_j^n = \frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{\Delta t^2} \quad \text{and} \quad \delta_x^2 u_j^n = \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x^2}.$$

   We here consider three frequently used *finite difference schemes* to discretise the semilinear KG equation:

   - An explicit finite difference scheme Expt-FD

$$\delta_t^2 u_j^n - a^2 \delta_x^2 u_j^n = f(u_j^n);$$

- Semi-implicit finite difference scheme Simpt-FD

$$\delta_t^2 u_j^n - \frac{a^2}{2}\left(\delta_x^2 u_j^{n+1} + \delta_x^2 u_j^{n-1}\right) = f(u_j^n);$$

- Compact finite difference scheme Compt-FD

$$\left(I + \frac{\Delta x^2}{12}\delta_x^2\right)\delta_t^2 u_j^n - \frac{a^2}{2}\left(\delta_x^2 u_j^{n+1} + \delta_x^2 u_j^{n-1}\right) = \left(I + \frac{\Delta x^2}{12}\delta_x^2\right) f(u_j^n).$$

2. *Method-of-lines Schemes*
The method-of-lines approach to the approximation of (10.1)–(10.2) is composed of two stages: space and time discretisations. We first approximate the spatial differential operator $\mathscr{A}$ to obtain a semidiscrete scheme of the form

$$u''(t) + Au(t) = f\big(u(t)\big),$$

where A is a symmetric positive semi-definite matrix. We then use an ODE solver to deal with the semidiscrete scheme. Here, the time integrators we select for comparison are

- Gauss2s4: the two-stage Gauss method of order four from [31];
- Gauss3s6: the three-stage fourth-order Gauss method in [31];
- RKN3s4: the three-stage Runge–Kutta–Nyström (RKN) method of order four from [31];
- IRKN2s4: the two-stage implicit symplectic RKN method of order four derived in [47];
- IRKN3s6: the three-stage implicit symplectic RKN method of order six derived in [47];
- SV: classical Störmer–Verlet formula [31].

For the time integrators HB0, HB1 and HB2 derived in this chapter, we use the tolerance 10^{-15} and choose $m = 2$ in the waveform relaxation algorithm (10.80)–(10.81), which implies that just one iteration is needed at each step. Consequently, these two integrators can be implemented at lower cost. Here, it should be noted that when the error of a method under consideration is very large for some Δt, we do not plot the corresponding points in efficiency curves. Moreover, in order to compute the convergence order, we denote

$$\mathrm{EU}(\Delta x, \Delta t) = \max_{0\leqslant n\leqslant N} \sqrt{\Delta x \sum_{i=0}^{M-1} (U_i^n - u_i^n)^2}$$

and

$$\mathrm{EH}(\Delta x, \Delta t) = \max_{0 \leqslant n \leqslant N} \left| H(u^n, u'^n) - H(u^0, u'^0) \right|.$$

The computational order of the method is calculated with the following formulae:

$$\log_2 \left(\frac{\mathrm{EU}(\Delta x, \Delta t)}{\mathrm{EU}(\Delta x, \Delta t/2)} \right) \quad \text{and} \quad \log_2 \left(\frac{\mathrm{EH}(\Delta x, \Delta t)}{\mathrm{EH}(\Delta x, \Delta t/2)} \right).$$

Problem 10.1 We consider the semilinear KG equation

$$\frac{\partial^2 u(x,t)}{\partial t^2} - a^2 \frac{\partial^2 u(x,t)}{\partial x^2} + a u(x,t) - b u^3(x,t) = 0,$$

in the region $(x, t) \in [-20, 20] \times [0, T]$ with the initial conditions

$$u(x,0) = \sqrt{\frac{2a}{b}} \operatorname{sech}(\lambda x), \qquad u_t(x,0) = c\lambda \sqrt{\frac{2a}{b}} \operatorname{sech}(\lambda x) \tanh(\lambda x),$$

where $\lambda = \sqrt{a/(a^2 - c^2)}$ and $a, b, a^2 - c^2 > 0$. The exact solution of Problem 10.1 is

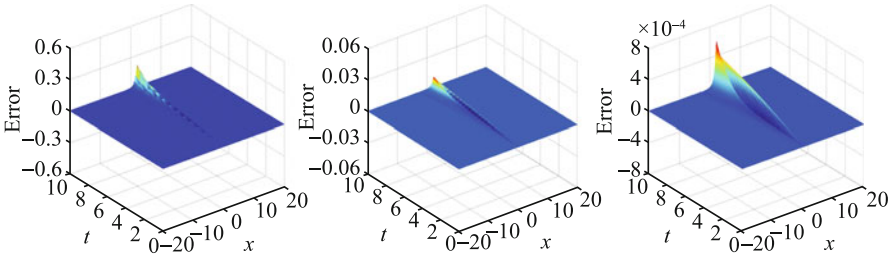
$$u(x,t) = \sqrt{\frac{2a}{b}} \operatorname{sech}(\lambda(x - ct)).$$

The real parameter $\sqrt{2a/b}$ represents the amplitude of a soliton which travels with velocity c. The potential function is $V(u) = au^2/2 - bu^4/4$. The problem can be found in [23]. We consider the parameters $a = 0.3$, $b = 1$ and $c = 0.25$ which are similar to those in [23].

In Figs. 10.2 and 10.3, we integrate the Problem 10.1 on the region $(x, t) \in [-20, 20] \times [0, 10]$ by using the time integrator HB2, coupled with the fourth-order symmetric finite difference (SFD) and Fourier spectral collocation (FSC). The graphs of errors are shown in Figs. 10.2 and 10.3 with the time stepsize $\Delta t = 0.01$ and different values of M. The numerical results demonstrate that the accuracy of the spatial discretisation is consistent with the theory presented in this chapter. It is evident that the Fourier spectral collocation method is the best choice to discretise the spatial variable.

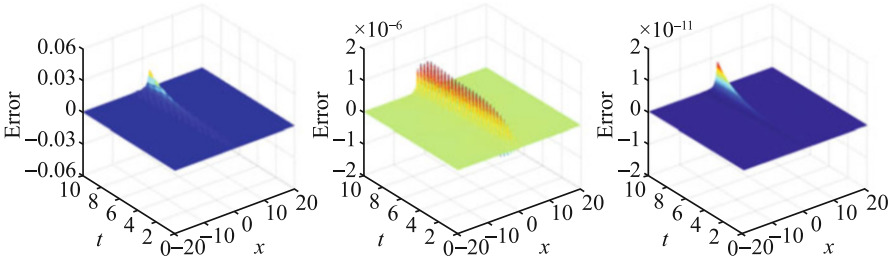
Table 10.1 provides the computational results with $M = 800$. The data demonstrate that the temporal convergence orders of HB0, HB1 and HB2 are second, fourth and sixth, respectively. The results show that the temporal accuracy is completely consistent with the theory presented in Theorem 10.7.

To compare the integrators presented in this chapter with classical finite difference and method-of-lines schemes, we integrate the problem in the region $(x, t) \in$



(a) M=200 (b) M=400 (c) M=800

Fig. 10.2 The errors for Problem 10.1 obtained by combining the time integrator HB2 with the fourth-order finite difference spatial discretisation for $\Delta t = 0.01$ with $M = 200, 400$ and 800



(a) M=200 (a) M=400 (c) M=800

Fig. 10.3 The errors for Problem 10.1 obtained by combining the time integrator HB2 with Fourier spectral collocation method for $\Delta t = 0.01$ with $M = 200, 400$ and 800

Table 10.1 Numerical convergence in time with different Δt, fixed $M = 800$ and up to $T = 10$

Δt	HB0		HB1		HB2	
	EU($\Delta x, \Delta t$)	Order	EU($\Delta x, \Delta t$)	Order	EU($\Delta x, \Delta t$)	Order
0.8	1.7941×10^{-1}	*	2.0990×10^{-3}	*	1.2742×10^{-4}	*
0.4	4.3627×10^{-2}	2.0400	2.6159×10^{-4}	3.0043	3.1967×10^{-6}	5.3168
0.2	1.0065×10^{-2}	2.1158	1.8650×10^{-5}	3.8101	5.3224×10^{-8}	5.9084
0.1	2.4614×10^{-3}	2.0318	1.2006×10^{-6}	3.9573	8.5113×10^{-10}	5.9666
0.05	6.1189×10^{-4}	2.0081	7.5579×10^{-8}	3.9896	1.2662×10^{-11}	6.0708

$[-20, 20] \times [0, 10]$ with different time stepsizes Δt, and the number of spatial nodal values is M. The numerical results are shown in Fig. 10.4. We compare the integrators presented in this chapter with the standard finite difference schemes with stepsizes $\Delta t = 0.01 \times 2^{3-j}$ for $j = 0, 1, 2, 3$ and $M = 1000$ for the finite difference schemes Expt-FD, Simpt-FD and Compt-FD and $M = 800$ for HB0-FSC, HB1-FSC and HB2-FSC. The logarithms of the global errors GE $= \|u(t_n) - u^n\|_\infty$ are plotted in Fig. 10.4a.

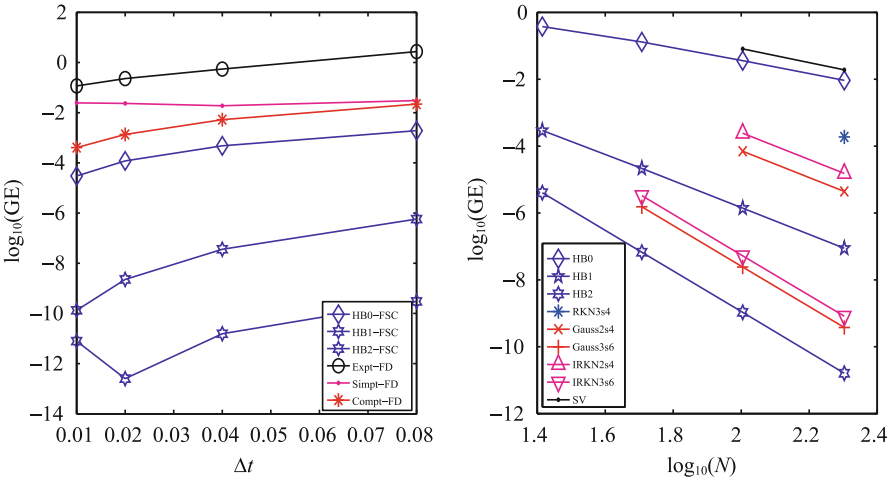


Fig. 10.4 The efficiency curves for Problem 10.1: (**a**) Comparison with standard finite difference schemes, (**b**) Comparison with method-of-lines schemes

Compared with the method-of-lines schemes, we discretise the spatial derivative using Fourier spectral collocation method with fixed $M = 800$ and integrate the KG equation with $\Delta t = 0.2/2^j$ for $j = 0, 1, 2, 3$. The efficiency curves (accuracy versus the computational cost measured by the number of function evaluations required by each method) are shown in Fig. 10.4b.

In conclusion, the numerical results in Fig. 10.4 demonstrate that the time integrators HB0, HB1 and HB2 derived in this chapter, combined with Fourier spectral collocation, have much better accuracy and are more efficient than those occurring in the literature.

The numerical results in Fig. 10.5 present the error of the semidiscrete energy conservation law as a function of the time-step calculated by $\tilde{E}(t)$, where EH $= |\tilde{E}(t) - \tilde{E}(t_0)|$. It can be observed form Fig. 10.5 that the error of HB0 is $\approx 10^{-4}$, for HB1 it is $\approx 10^{-11}$, while that of HB2 is $\approx 10^{-13}$. Moreover, the convergence orders of the preservation of the discrete energy by the HB time integrators are computed which are listed in Table 10.2. The numerical results show that the accuracy of discrete energy preservation by HB0 is of order two, by HB1 of order four and by HB2 of order six.

Problem 10.2 We consider the sine-Gordon equation

$$\frac{\partial^2 u}{\partial t^2}(x,t) - \frac{\partial^2 u}{\partial x^2}(x,t) + \sin(u(x,t)) = 0$$

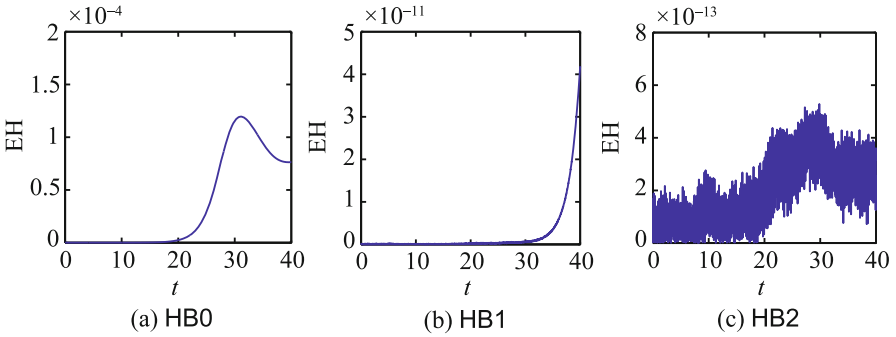


Fig. 10.5 Discrete energy conservation by HB0, HB1 and HB2 with the spatial discretisation by Fourier spectral collocation with $M = 800$ up to $T = 40$, using $\Delta t = 0.02$

Table 10.2 Numerical precision of the preservation of the semidiscrete energy up to $T = 40$ with various Δt and fixed $M = 200$

Δt	HB0		HB1		HB2	
	EH(Δx, Δt)	Order	EH(Δx, Δt)	Order	EH(Δx, Δt)	Order
0.16	1.8989×10^{-3}	*	5.2428×10^{-6}	*	1.1049×10^{-8}	*
0.08	4.7939×10^{-4}	1.9859	3.2875×10^{-7}	3.9952	1.7178×10^{-10}	6.0072
0.04	1.2005×10^{-4}	1.9976	2.0564×10^{-8}	3.9988	2.6613×10^{-12}	6.0123
0.02	3.0026×10^{-5}	1.9994	1.2860×10^{-9}	3.9992	5.0293×10^{-14}	5.7256
0.01	7.5071×10^{-6}	1.9999	8.0451×10^{-11}	3.9986	—	—

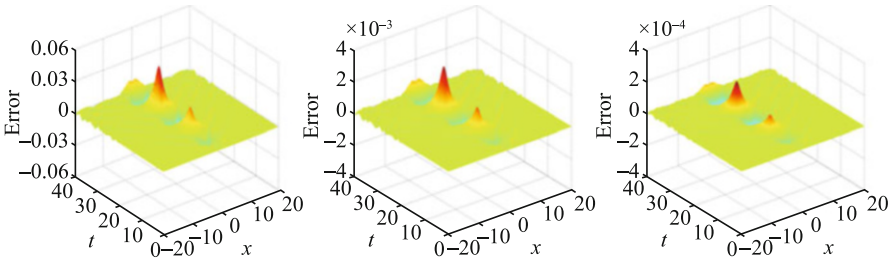
in the region $-20 \leqslant x \leqslant 20$, $0 \leqslant t \leqslant T$, subject to the initial conditions

$$u(x,0) = 0, \qquad u_t(x,0) = 4\,\mathrm{sech}(x/\sqrt{1+c^2})/\sqrt{1+c^2}.$$

The exact solution of Problem 10.2 is

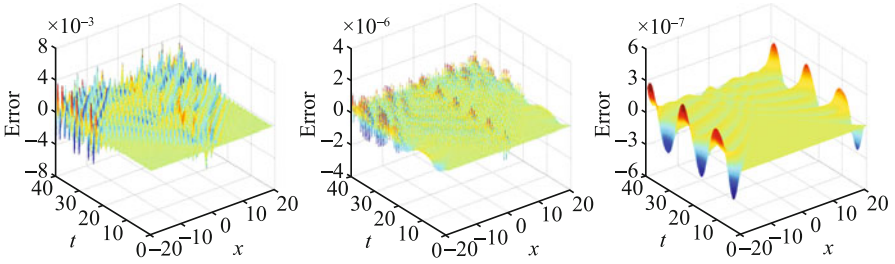
$$u(x,t) = 4\arctan\left(c^{-1}\sin(ct/\sqrt{1+c^2})\,\mathrm{sech}(x/\sqrt{1+c^2})\right).$$

This problem is known as *breather solution* of the sine-Gordon equation and represents a pulse-type structure of a soliton. The parameter c is the velocity and we choose $c = 0.5$. The potential function is $V(u) = 1 - \cos(u)$. Problem 10.2 is integrated by HB2, coupled either with the fourth-order symmetric finite difference SFD or Fourier spectral collocation FSC. The error graphs are shown in Figs. 10.6 and 10.7 with $\Delta t = 0.01$ and several values of M. They demonstrate how the accuracy of the spatial discretisation varies with M, and also indicate that the Fourier spectral collocation FSC is decisively superior to the fourth-order symmetric finite difference SFD.



(a) M=100 (b) M=200 (c) M=400

Fig. 10.6 The error for the sine-Gordon equation, blending the time integrator HB2 with fourth-order finite difference spatial discretisation for $\Delta t = 0.01$ and $M = 100, 200, 400$



(a) M=50 (b) M=100 (c) M=200

Fig. 10.7 The errors blending the time integrator HB2 with Fourier spectral method for $\Delta t = 0.01$ and $M = 50, 100, 200$

Table 10.3 Numerical convergence in time with different Δt, fixed $M = 200$ and up to $T = 40$

Δt	HB0		HB1		HB2	
	EU($\Delta x, \Delta t$)	Order	EU($\Delta x, \Delta t$)	Order	EU($\Delta x, \Delta t$)	Order
0.8	21.85583982	*	2.762406385	*	2.2240×10^{-1}	*
0.4	5.486416853	1.9941	1.2514×10^{-1}	4.4643	3.6983×10^{-3}	5.9101
0.2	1.176387217	2.2215	7.3631×10^{-3}	4.0871	4.8228×10^{-5}	6.2609
0.1	2.8235×10^{-1}	2.0588	4.5373×10^{-4}	4.0204	7.3285×10^{-7}	6.0402
0.05	6.9931×10^{-2}	2.0135	2.8264×10^{-5}	4.0048	–	–

The computational results in Table 10.3 demonstrate that the temporal convergence orders of HB0, HB1 and HB2 are of two, four and six, respectively. The results again verify the convergence accuracy in time is consistent with the theory in Theorem 10.7.

The efficiency curves are shown in Fig. 10.8. In order to compare the integrators with a standard finite difference scheme, in Fig. 10.8a we integrate the problem for $\Delta t = 0.04, 0.03, 0.02, 0.01$. We use $M = 1000$ for the finite difference scheme Expt-FD, Simpt-FD and Compt-FD, and $M = 200$ for the HB0-FSC, HB1-FSC and HB2-FSC.

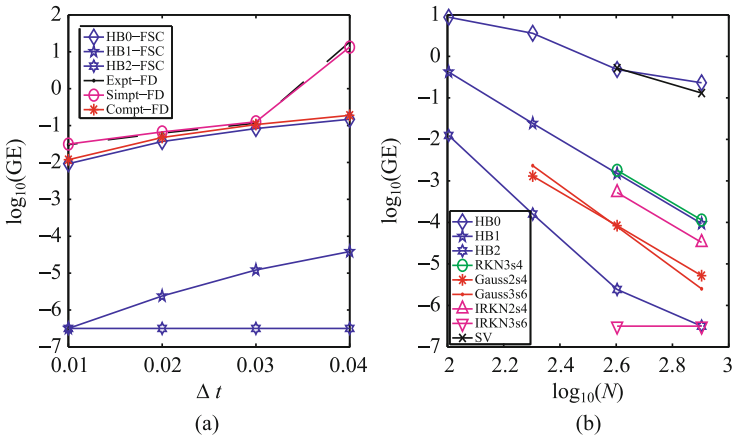


Fig. 10.8 Efficiency curves for Problem 10.2: (**a**) Comparison with standard finite difference schemes, (**b**) Comparison with method-of-lines schemes

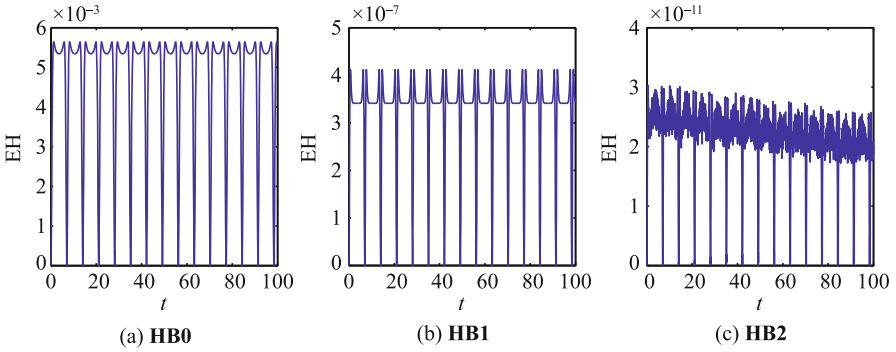


Fig. 10.9 Energy conservation by HB0, HB1 and HB2, both blended with FSC, using $M = 200$, $\Delta t = 0.02$ and $T = 100$

In Fig. 10.8b we compare the integrators presented in this chapter with method-of-lines schemes. The problem is integrated over the time interval $[0, 40]$ with fixed $M = 200$ and time stepsizes $\Delta t = 0.4/2^j$ for $j = 0, 1, 2, 3$. It can be observed that the time integrators HB0, HB1 and HB2, coupled with Fourier spectral collocation, are more efficient than other chosen methods.

The numerical results in Fig. 10.9 represent the error of the semidiscrete energy conservation law. It can be seen that the error does not grow with time. The errors obtained by HB0, HB1 and HB2 reach magnitudes of $\approx 10^{-3}$, $\approx 10^{-7}$ and $\approx 10^{-11}$, respectively. The precisions of the preservation of the discrete energy by the HB time integrators are listed in Table 10.4. It is shown that the accuracy of the discrete energy preservation by HB0 is of order two, by HB1 is of order four and by HB2 is of order six.

Table 10.4 Numerical precision of the preservation of the semidiscrete energy up to $T = 100$ with different Δt and fixed $M = 200$

Δt	HB0		HB1		HB2	
	EH($\Delta x, \Delta t$)	Order	EH($\Delta x, \Delta t$)	Order	EH($\Delta x, \Delta t$)	Order
0.16	3.6368×10^{-1}	*	1.7078×10^{-3}	*	8.0039×10^{-6}	*
0.08	9.0444×10^{-2}	2.0076	1.0578×10^{-4}	4.0130	1.2237×10^{-7}	6.0314
0.04	2.2580×10^{-2}	2.0020	6.5969×10^{-6}	4.0031	1.9184×10^{-9}	5.9952
0.02	5.6432×10^{-3}	2.0005	4.1208×10^{-7}	4.0008	3.0326×10^{-11}	5.9832
0.01	1.4107×10^{-3}	2.0001	2.5774×10^{-8}	3.9989	–	–

Below is an example of a high-dimensional problem.

Problem 10.3 We consider the 2D sine-Gordon equation (see, e.g. [23, 48–50]):

$$u_{tt} - (u_{xx} + u_{yy}) = -\sin(u), \quad t > 0, \tag{10.87}$$

in the spatial region $\Omega = [-14, 14] \times [-14, 14]$, with the initial conditions

$$u(x, y, 0) = 4 \arctan\left(\exp\left(3 - \sqrt{x^2 + y^2}\right)\right), \quad u_t(x, y, 0) = 0, \tag{10.88}$$

and the homogeneous Neumann boundary conditions

$$u_x(\pm 14, y, t) = u_y(x, \pm 14, t) = 0. \tag{10.89}$$

The exact solution of this problem is a phenomenon called a circular ring soliton (see, e.g. [48, 50]), and different initial conditions will result in different numerical phenomena. We here use the time integrators HB0, HB1 and HB2 coupled with the *discrete Fast Cosine Transformation* (see, e.g. [51, 52]) to simulate the particular circular ring solitons. In Figs. 10.10 and 10.11, we show the simulation results and the corresponding contour plots at the time points $t = 0, 2, 4, 6, 8$ and 10 with spatial stepsizes $\Delta x = \Delta y = 0.07$ and the time stepsize $\Delta t = 0.01$. The CPU time required to reach $t = 10$ is 1191.445350 s.

Likewise, to verify the theoretical results in Theorem 10.7, we fixed the spatial stepsizes as $\Delta x = \Delta y = 0.07$ and integrate the Problem 10.3 by the time integrators HB0, HB1 and HB2 with various time stepsizes. The data listed in Tables 10.5 and 10.6 demonstrate that the convergence order of the time integrators HB0, HB1 and HB2 are of order two, four and six, respectively. The results again verify the correctness of the theory presented in Theorem 10.7.

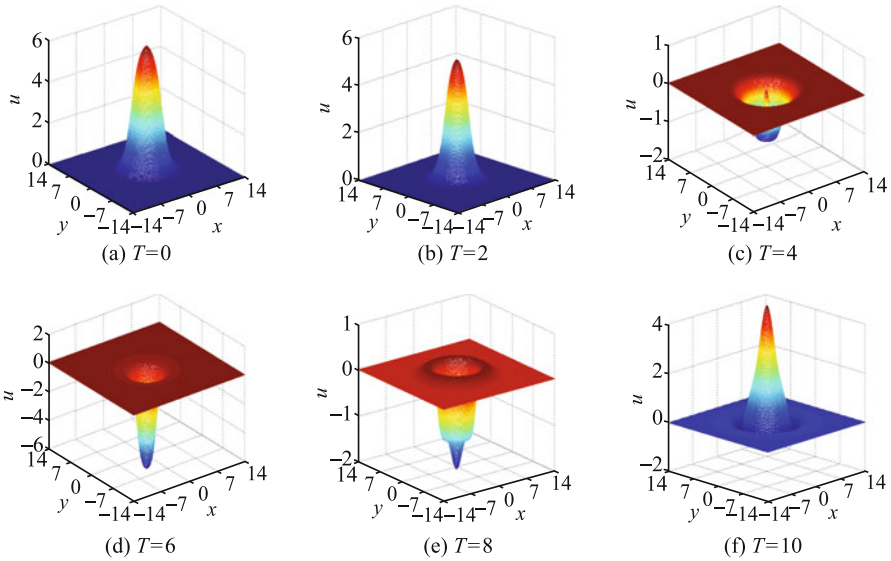


Fig. 10.10 Circular ring solitons: numerical solutions obtained by coupling the time integrator HB2 with the *discrete Fast Cosine Transformation* at the time points $t = 0, 2, 4, 6, 8$ and 10

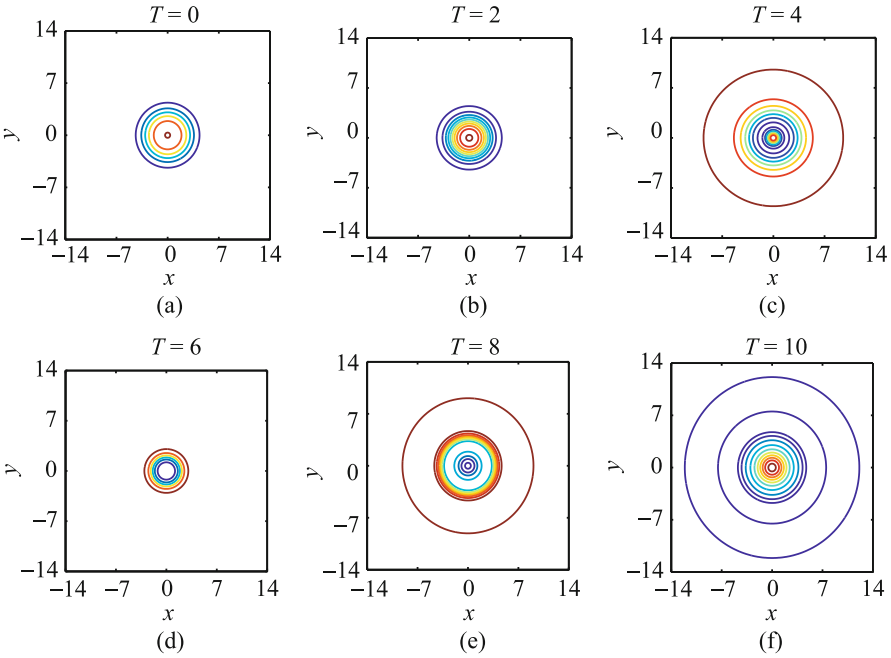


Fig. 10.11 Circular ring solitons: contours of the numerical solutions at the time points $t = 0, 2, 4, 6, 8$ and 10

Table 10.5 Numerical convergence "$u(x, y, t)$" in time with various Δt at time $T = 10$

$\Delta t_0 = 0.5$	HB0		HB1		HB2	
	EU($\Delta x, \Delta y, \Delta t$)	Order	EU($\Delta x, \Delta y, \Delta t$)	Order	EU($\Delta x, \Delta y, \Delta t$)	Order
$\frac{\Delta t_0}{2}$	7.9911×10^{-3}	$*$	7.9016×10^{-7}	$*$	4.4097×10^{-7}	$*$
$\frac{\Delta t_0}{4}$	1.9992×10^{-3}	1.9989	5.2878×10^{-7}	3.9014	9.4394×10^{-9}	5.5458
$\frac{\Delta t_0}{8}$	4.9990×10^{-4}	1.9997	3.3850×10^{-8}	3.9655	1.5975×10^{-10}	5.8848
$\frac{\Delta t_0}{16}$	1.2498×10^{-4}	1.9999	2.1293×10^{-9}	3.9907	2.5899×10^{-12}	5.9468

Table 10.6 Numerical convergence "$u_t(x, y, t)$" in time with different Δt at time $T = 10$

$\Delta t_0 = 0.5$	HB0		HB1		HB2	
	EU($\Delta x, \Delta y, \Delta t$)	Order	EU($\Delta x, \Delta y, \Delta t$)	Order	EU($\Delta x, \Delta y, \Delta t$)	Order
$\frac{\Delta t_0}{2}$	1.9101×10^{-2}	$*$	1.2757×10^{-4}	$*$	1.6356×10^{-5}	$*$
$\frac{\Delta t_0}{4}$	4.7841×10^{-3}	1.9973	1.1049×10^{-5}	3.5294	3.4931×10^{-7}	5.5491
$\frac{\Delta t_0}{8}$	1.1966×10^{-2}	1.9993	7.5552×10^{-7}	3.8703	5.9125×10^{-9}	5.8846
$\frac{\Delta t_0}{16}$	2.9919×10^{-4}	1.9998	4.8296×10^{-8}	3.9675	9.4309×10^{-11}	5.9702

10.8 Conclusions and Discussions

It is known that the KG equation and the Schrödinger equation are two important equations of Quantum Physics. We have derived and analysed a class of time integrators for the semilinear KG equation (10.1)–(10.2) in this chapter. As distinct from traditional approaches, these schemes are based on the operator-variation-of-constants formula (10.10) which is introduced on the Hilbert space $L^2(\Omega)$ using operator spectral theory, and it is in fact an implicit expression of the solution of the semilinear KG equation. Keeping the eventual discretisation in mind, a class of time integration formulae (10.22) has been designed by applying a two-point Hermite interpolation to the nonlinear integrals that appear in the operator-variation-of-constants formula. It has been shown that these formulae can have arbitrary order and are also symmetric. *A significant advantage of this approach is that the requirement of temporal smoothness is reduced compared with the traditional schemes for PDEs in the literature*. In order to approximate the unbounded positive semi-definite spatial differential operator $\mathscr{A}$, we have also discussed the importance of the choice of a positive semi-definite differentiation matrix. Moreover, stability and convergence for the fully discrete scheme have been proved in both linear and nonlinear settings. In particular, the long-time preservation of the discrete energy conservation law has been analysed. Since the fully discrete scheme is implicit, iteration is required, and we have applied the waveform relaxation algorithm (10.80)–(10.81) in practical computations and analysed the convergence of the iteration. Numerical experiments implemented in this chapter demonstrate that the

time integrators so constructed have excellent numerical behaviour in comparison with existing standard finite difference and method-of-lines schemes in the science literature.

Note that the methodology presented in this chapter can be extended to a range of other nonlinear wave equations. Some of the more immediate possible extensions are as follows:

1. *High Dimensional Problems.* Although Eq. (10.1) is one-dimensional, the method can be extended to KG equations in a moderate number, d, of space dimensions,

$$u_{tt} - a^2 \Delta u = f(u), \qquad t_0 \leqslant t \leqslant T, \quad x \in [-\pi, \pi]^d, \tag{10.90}$$

   where $u = u(x, t)$ and $\Delta' = \sum_{i=1}^{d} \dfrac{\partial^2}{\partial x_i^2}$, with periodic boundary conditions. A large dimension d requires combining the time integration formula (10.22) with other spatial approximate techniques, such as *sparse grids* [53] or *discrete FFT* [51, 54].
2. *Neumann and Dirichlet Boundary Problems.* In this chapter we only consider problems (10.1) subject to periodic boundary conditions (10.2). However, the approach presented in this chapter can be extended to problems with Neumann and Dirichlet boundary conditions with domain $\Omega = [0, \pi]^d$. The corresponding spatial discretisation could be the *discrete Fast Sine Transformation* for Dirichlet boundary conditions or *discrete Fast Cosine Transformation* for the Neumann boundary case. Fortunately, much related work on the *discrete Fast Cosine/Sine Transformation* has been widely published in the science literature (see, e.g. [55]). Therefore, we are hopeful of obtaining related new results.
3. Furthermore, the approach presented in this chapter also can be directly applied to the computation of the following problems:

   (a) *The damped semilinear KG equation*

$$\begin{cases} u_{tt} + \alpha(x) u_t - \beta \Delta u + u + f'(u) = 0, & (x, t) \in \Omega \times [t_0, +\infty), \\ u(x, t_0) = \varphi_1(x), \ u_t(x, t_0) = \varphi_2(x), & x \in \bar{\Omega}, \end{cases} \tag{10.91}$$

   where Ω is a C^1 domain in $\mathbb{R}^d$, β represents the amplitude of the diffusion and the damping coefficient $\alpha : \Omega \to [0, \infty)$ is effective uniform in the neighborhood of the spatial infinity,

$$\alpha(x) \geqslant 0, \qquad \alpha \in L^\infty(\Omega), \qquad \liminf_{|x| \to \infty} \alpha(x) > 0.$$

The damper $\alpha(x)$ satisfies appropriate conditions which guarantee that the total energy defined by

$$E(t) = \frac{1}{2}\int_{\Omega}\Big[|u_t|^2 + |\nabla u|^2 + |u|^2 + 2f(u)\Big]\mathrm{d}x$$

decays uniformly.

(b) *The hyperbolic telegraph equation*

$$\begin{cases} u_{tt} + 2\alpha u_t + \beta^2 u = \Delta u + f(x,t), & (x,t) \in \Omega \times [0,+\infty), \\ u(x,t_0) = \varphi_1(x),\ u_t(x,t_0) = \varphi_2(x), & x \in \{\bar{\Omega}, \end{cases} \tag{10.92}$$

where $\alpha > 0$ and $\beta > 0$ are known constants. This equation has been widely used in many different fields of science and mathematical engineering such as the vibration of structures, the transmission and propagation of electrical signals and random walk theory.

The material in this chapter is based on the work by Liu et al. [56].

References

1. Dodd, R.K., Eilbeck, I.C., Gibbon, J.D., et al.: Solitons and Nonlinear Wave Equations. Academic Press, London (1982)
2. Drazin, P.J., Johnson, R.S.: Solitons: An Introduction. Cambridge University Press, Cambridge (1989)
3. Brenner, P., von Wahl, W.: Global classical solutions of nonlinear wave equations. Math. Z. **176**, 87–121 (1981)
4. Ginibre, J., Velo, G.: The global Cauchy problem for the nonlinear Klein-Gordon equation. Math. Z. **189**, 487–505 (1985)
5. Ibrahim, S., Majdoub, M., Masmoudi, N.: Global solutions for a semilinear, two-dimensional Klein-Gordon equation with exponential-type nonlinearity. Commun. Pure Appl. Math. **59**, 1639–1658 (2006)
6. Kosecki, R.: The unit condition and global existence for a class of nonlinear Klein-Gordon equations. J. Differ. Equations **100**, 257–268 (1992)
7. Strauss, W.A.: Nonlinear Wave Equations. Regional Conference Series in Mathematics. Regional Conference Series in Mathematics, vol. 73 (American Mathematical Society, Providence, 1989)
8. Ablowitz, M.J., Kruskal, M.D., Ladik, J.F.: Solitary wave collisions. SIAM J. Appl. Math. **36**, 428–437 (1979)
9. Bao, W.Z., Dong, X.C.: Analysis and comparison of numerical methods for the Klein-Gordon equation in the nonrelativistic limit regime. Numer. Math. **120**, 189–229 (2012)
10. Duncan, D.B.: Symplectic finite difference approximations of the nonlinear Klein-Gordon equation. SIAM J. Numer. Anal. **34**, 1742–1760 (1997)
11. Li, S., Vu-Quoc, L.: Finite difference calculus invariant structure of a class of algorithms for the nonlinear Klein-Gordon equation. SIAM J. Numer. Anal. **32**, 1839–1875 (1995)
12. Pascual, P.J., Jiménezz, S., Vázquez, L.: Numerical Simulations of a Nonlinear Klein-Gordon Model. Applications. Lecture Notes and Physics **448**, 211–270 (1995)

13. Cao, W., Guo, B.: Fourier collocation method for solving nonlinear Klein-Gordon equation. J. Comput. Phys. **108**, 296–305 (1993)
14. Cohen, D., Hairer, E., Lubich, C.: Conservation of energy, momentum and actions in numerical discretizations of non-linear wave equations. Numer. Math. **110**, 113–143 (2008)
15. Guo, B.Y., Li, X., Vázquez, L.: A Legendre spectral method for solving the nonlinear Klein-Gordon equation. Comput. Appl. Math. **15**, 19–36 (1996)
16. Tourigny, Y.: Product approximation for nonlinear Klein-Gordon equations. IMA J. Numer. Anal. **9**, 449–462 (1990)
17. Cox, S., Matthews, P.: Exponential time differencing for stiff systems. J. Comput. Phys. **176**, 430–455 (2002)
18. Hochbruck, M., Ostermann, A.: Explicit exponential Runge-Kutta methods for semilinear parabolic problems. SIAM J. Numer. Anal. **43**, 1069–1090 (2005)
19. Hochbruck, M., Ostermann, A.: Exponential Runge-Kutta methods for parabolic problems. Appl. Numer. Math. **53**, 323–339 (2005)
20. Hochbruck, M., Ostermann, A.: Exponential integrators. Acta Numer. **19**, 209–286 (2010)
21. Kassam, A.K., Trefethen, L.N.: Fourth-order time stepping for stiff PDEs. SIAM J. Sci. Comput. **26**, 1214–1233 (2005)
22. Bátkai, A., Farkas, B., Csomós, P. et al.: Operator semigroups for numerical analysis. In: 15th Internet Seminar 2011/12 (2011)
23. Liu, C.Y., Wu, X.Y.: Arbitrarily high-order time-stepping schemes based on the operator spectrum theory for high-dimensional nonlinear Klein-Gordon equations. J. Comput. Phys. **340**, 243–275 (2017)
24. Bader, P., Iserles, A., Kropielnicka, K. et al.: Effective approximation for the semiclassical Schrödinger equation. Found. Comput. Math. **14**, 689–720 (2014)
25. Wu, X.Y., Wang, B.: Recent Developments in Structure-Preserving Algorithms for Oscillatory Differential Equations. Springer Nature Singapore Pte Ltd., Singapore (2018)
26. Grundy, R.E.: Hermite interpolation visits ordinary two-point boundary value problems. ANZIAM J. **48**, 533–552 (2007)
27. Phillips, G.M.: Explicit forms for certain Hermite approximations. BIT Numer. Math. **13**, 177–180 (1973)
28. Dyn, N.: On the existence of Hermite-Birkhoff quadrature formulas of Gaussian type. J. Approx. Theor. **31**, 22–32 (1981)
29. Jetter, K.: Uniqueness of Gauss-Birkhoff quadrature formulas. SIAM J. Numer. Anal. **24**, 147–154 (1987)
30. Nikolov, G.: Existence and uniqueness of Hermite-Birkhoff Gaussian quadrature formulas. Calcolo **26**, 41–59 (1989)
31. Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd edn. Springer, Berlin (2006)
32. Grimm, V.: On error bounds for the Gautschi-type exponential integrator applied to oscillatory second-order differential equations. Numer. Math. **100**, 71–89 (2005)
33. Hairer, E., Lubich, C.: Long-time energy conservation of numerical methods for oscillatory differential equations. SIAM J. Numer. Anal. **38**, 414–441 (2000)
34. Hochbruck, M., Lubich, C.: A Gautschi-type method for oscillatory second-order differential equations. Numer. Math. **83**, 403–426 (1999)
35. Wu, X.Y., You, X., Wang, B.: Structure-Preserving Algorithms for Oscillatory Differential Equations. Springer, Berlin (2013)
36. Iserles, A.: A First Course in the Numerical Analysis of Differential Equations, 2nd edn. Cambridge University Press, Cambridge (2008)
37. Bank, R., Graham, R.L., Stoer, J. et al.: High Order Difference Methods for Time Dependent PDEs. Springer, Berlin (2008)
38. Liu, C.Y., Shi, W., Wu, X.Y.: An efficient high-order explicit scheme for solving Hamiltonian nonlinear wave equations. Appl. Math. Comput. **246**, 696–710 (2014)
39. Sun, Z.Z.: Numerical Methods of Partial Differential Equations, 2nd edn. Science Press, Beijing (2012)

40. Hesthaven, J.S., Gottlieb, S., Gottlieb, D.: Spectral Methods for Time-Dependent Problems. Cambridge University Press, Cambridge (2007)
41. Shen, J., Tang, T., Wang, L.L.: Spectral Methods: Algorithms, Analysis, Applications. Springer, Berlin (2011)
42. Janssen, J., Vandewalle, S.: On SOR waveform relaxation methods. SIAM J. Numer. Anal. **34**, 2456–2481 (1997)
43. Khanamiryan, M.: Quadrature methods for highly oscillatory linear and nonlinear systems of ordinary differential equations, Part I. BIT Numer. Math. **48**, 743–762 (2008)
44. Lubich, C., Ostermann, A.: Multigrid dynamic iteration for parabolic equations. BIT Numer. Math. **27**, 216–234 (1987)
45. Vandewalle, S.: Parallel Multigrid Waveform Relaxation for Parabolic Problems. Teubner Scripts on Numerical Mathematics. Vieweg+Teubner Verlag, Wiesbaden (1993)
46. Wang, B., Liu, K., Wu, X.Y.: A Filon-type asymptotic approach to solving highly oscillatory second-order initial value problems. J. Comput. Phys. **243**, 210–223 (2013)
47. Tang, W.S., Ya, Y.J., Zhang, J.J. High order symplectic integrators based on continuous-stage Runge-Kutta-Nyström methods. Appl. Math. Comput. **361**, 670–679 (2019)
48. Bratsos, A.G.: A modified predictor-corrector scheme for the two-dimensional sine-Gordon equation. Numer. Algorithms **43**, 295–308 (2006)
49. Dehghan, M., Ghesmati, A.: Numerical simulation of two-dimensional sine-Gordon solitons via a local weak meshless technique based on the radial point interpolation method (RPIM). Comput. Phys. Commun. **181**, 772–786 (2010)
50. Sheng, Q., Khaliq, A.Q.M., Voss, D.A.: Numerical simulation of two-dimensional sine-Gordon solitons via a split cosine scheme. Math. Comput. Simul. **68**, 355–373 (2005)
51. Briggs, W.L., Henson, V.E.: The DFT: An Owner's Manual for the Discrete Fourier Transform. SIAM, Philadelphia (2000)
52. Britanak, V., Yip, P.C., Rao, K.R.: Discrete cosine and sine transforms: General properties, fast algorithms and integer approximations. IEEE Trans. Signal Process. **52**, 306–311 (2006)
53. Bungartz, H.J., Griebel, M.: Sparse grids. Acta Numer. **13**, 147–269 (2004)
54. Bueno-Orovio, A., Pérez-García, V.M., Fenton, F.H.: Spectral methods for partial differential equations in irregular domains: The spectral smoothed boundary method. SIAM J. Sci. Comput. **28**, 886–900 (2006)
55. Mulholland, L.S., Huang, W.Z., Sloan, D.M.: Pseudospectral solution of near-singular problems using numerical coordinate transformations based on adaptivity. SIAM J. Sci. Comput. **19**, 1261–1289 (1998)
56. Liu, C.Y., Iserles, A., Wu, X.Y.: Symmetric and arbitrarily high-order Birkhoff-Hermite time integrators and their long-time behaviour for solving nonlinear Klein-Gordon equations. J. Comput. Phys. **356**, 1–30 (2018)