

## Approximate momentum conservation for spatial semidiscretizations of semilinear wave equations

Marcel Oliver<sup>1</sup>, Matthew West<sup>2</sup>, Claudia Wulff<sup>3</sup>

<sup>1</sup> School of Engineering and Science, International University Bremen, 28759 Bremen, Germany; e-mail: oliver@member.ams.org

<sup>2</sup> Department of Mathematics, University of California at Davis, Davis, CA 95616, USA; e-mail: west@math.ucdavis.edu

<sup>3</sup> Department of Mathematics & Statistics, University of Surrey, Guildford GU2 7XH, UK; e-mail: c.wulff@eim.surrey.ac.uk

Received September 17, 2002 / Revised version received May 15, 2003  
Published online February 4, 2004 – © Springer-Verlag 2004

**Summary.** We prove that a standard second order finite difference uniform space discretization of the semilinear wave equation with periodic boundary conditions, analytic nonlinearity, and analytic initial data conserves momentum up to an error which is exponentially small in the stepsize. Our estimates are valid for as long as the trajectories of the full semilinear wave equation remain real analytic.

The method of proof is that of backward error analysis, whereby we construct a modified equation which is itself Lagrangian and translation invariant, and therefore also conserves momentum. This modified equation interpolates the semidiscrete system for all time, and we prove that it remains exponentially close to the trigonometric interpolation of the semidiscrete system. These properties directly imply approximate momentum conservation for the semidiscrete system.

We also consider discretizations that are not variational as well as discretizations on non-uniform grids. Through numerical example as well as arguments from geometric mechanics and perturbation theory we show that such methods generically do not approximately preserve momentum.

*Mathematics Subject Classification (2000):* 65M20, 58J70, 70H33

## Contents

1	Introduction . . . . .	
2	Variational framework . . . . .	
3	Spatial semidiscretizations . . . . .	
	3.1 Variational semidiscretizations . . . . .	
	3.2 Non-variational semidiscretizations . . . . .	
	3.3 Interpolated momentum . . . . .	
4	Numerical results . . . . .	
5	Modified equations . . . . .	
6	Spatial discretizations and momentum maps . . . . .	
	6.1 Exact conservation of interpolated momenta . . . . .	
	6.2 Finite difference momenta . . . . .	
	6.3 Nonuniform space discretizations for linear waves . . . . .	
7	Momentum error bounds . . . . .	
	7.1 Spaces of analytic functions . . . . .	
	7.2 Functional setting for the wave equation . . . . .	
	7.3 Functional setting for the modified system . . . . .	
	7.4 Difference between wave equation and modified system . . . . .	
	7.5 Difference between discrete and modified system . . . . .	
	7.6 Approximate momentum conservation . . . . .	
8	Concluding remarks . . . . .	

## 1 Introduction

When differential equations possess special structure or symmetries, it is often desirable that numerical methods for such problems either preserve these structures or possess appropriate discrete analogs.

The theory of structure preserving discretizations is well-developed for ordinary differential equations. In particular, Hamiltonian ordinary differential equations can be solved by so-called symplectic or variational integrators [10, 18, 29]. In addition to exactly preserving the symplectic form, these integrators have been shown to conserve energy over exponentially long times [1, 8, 9, 27]. Such conservation properties are of crucial importance in celestial mechanics, molecular dynamics, and other application areas.

As a continuous system and its discretization will generally diverge arbitrarily far from each other as time advances, results on approximate energy conservation for symplectic methods usually rely on *backward error analysis*. To analyze the discrete system another continuous system is constructed, known as the modified equation, that can be proved to remain close to the discrete system for very long times. Then the properties of the modified system can be used to infer properties of the discrete system. The process of constructing the modified equation for time discretizations of ordinary differential equations can be seen as the averaging of a rapidly forced ordinary differential equation using a result of Neisthadt [22]. He was the first to

prove exponentially small error estimates for the embedding of a close to identity diffeomorphism (e.g., a one-step numerical method) into the flow of an autonomous differential equation.

Recently, similar ideas have been applied to construct and analyze structure preserving discretizations of partial differential equations. So-called multisymplectic integrators have been developed for Lagrangian systems by Marsden, Patrick, and Shkoller [17], and for Hamiltonian systems by Bridges and Reich [2, 28]. These have led to exciting new developments in methods for solid mechanics [14] and fluid systems [25]. In the case of vortex methods for fluids the numerical method itself has an interpretation as the “modified” pseudo partial differential equation [3] or even partial differential equation [23]. De Frutos and Sanz-Serna [4] studied energy and momentum conserving discretizations of KdV solitons, and found them to be much more accurate than non-conservative ones.

Most of the work in the context of partial differential equations focuses on exact preservation of conservation laws or their discrete analogs. Much less has been done for cases where one can obtain approximate conservation of some property under discretization, and there there are few rigorous results with, as far as we are aware, the following exceptions. Matthies [19] proves that a fully discrete reaction-diffusion system on a torus can be embedded into an autonomous ordinary differential equation with exponentially small error under a restrictive coupling assumption between spatial and temporal stepsize. In [31] we analyze approximate energy conservation of temporal semidiscretizations of semilinear wave equations – for a related result on averaging of rapidly forced Hamiltonian evolution equations see [20].

Moore and Reich [21] made a first step towards a formal backward error analysis for multisymplectic discretizations of semilinear wave equations and their corresponding energy and momentum conservation laws. They derive a modified higher order multisymplectic partial differential equation which is satisfied by the numerical solution with higher accuracy than the discretization error.

As shown in [17], variational multisymplectic discretizations can preserve vertical symmetries, i.e., symmetries which are not discretized. But horizontal symmetries, such as translations in space, are discretized by finite-difference type methods, so the question arises whether and how well a variational discretization will conserve the momentum corresponding to the discretized symmetry. In this paper we give an affirmative answer in a simple, yet prototypical situation: Conservation of momentum by the semilinear wave equation under a variational spatial semidiscretization on a uniform grid.

We study the semilinear 1 + 1 wave equation on the circle  $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ ,

$$(1.1) \quad \partial_{tt}u = \partial_{xx}u + f(u).$$

In addition to possessing a Lagrangian and Hamiltonian structure, the evolution preserves the momentum

$$(1.2) \quad J[u](t) = \int_{S^1} \partial_t u \partial_x u \, dx.$$

We find numerically that, among four different spatial semi-discretizations, only the variational spatial semi-discretization on a uniform grid preserves an interpolated momentum within computational accuracy, while all other schemes display significant momentum drift. This behavior is explained by the following theorem, an extended version of which is proved in Section 7, Theorem 7.9.

**Theorem 1.1** *Assume that  $f$  is entire, and that the exact solution  $u = u(t, x)$  to the semilinear wave equation is real analytic with uniform spatial radius of analyticity larger than some  $r > 0$  on  $[0, T] \times S^1$ . Further, let  $u_h(t)$  denote the approximate solution that corresponds to the standard symmetric space-semidiscretization with uniform grid size  $h$ , and let  $i_h(u_h)$  denote the trigonometric interpolant of  $u_h$ . Then there exists a constant  $\kappa = \kappa(u, T)$  independent of  $h$  such that*

$$(1.3) \quad |J[u](t) - J[i_h(u_h)](t)| \leq \kappa e^{-\pi r/h}$$

*uniformly on  $[0, T]$ .*

Note that if the nonlinearity is entire, the initial value problem for the semilinear wave equation is well-posed in certain spaces of space-analytic functions on some interval  $[0, T]$ , for details see [26, 6] and Section 7.2.

Our proof is based on a modified functional equation which is *not* a partial differential equation, but which interpolates the discrete system and so remains close to the discrete system for as long as the true solution of the wave equation remains sufficiently regular. To our knowledge, this is the first rigorous backward error analysis result for spatial semi-discretizations. We also show that the interpolated momentum for the linear wave equation is exactly conserved for the same variational uniform discretization (Theorem 6.2), but generically drifts on sufficiently long time scales when the grid is non-uniform (Theorem 6.4).

We believe that our result extends to higher order methods as well as to other equations and full space-time discretizations. Leaving these generalizations for future work, we restrict ourselves to a simple prototype model in this paper.

We begin the analysis of this model problem in Section 2 by recalling the variational structure of the wave equation. The spatial semidiscretizations are then defined in Section 3 and are studied numerically in Section 4. The behavior of these semidiscretizations is described by defining appropriate modified equations in Section 5.

The final two sections of the paper contain detailed proofs of the observations and claims in the earlier parts. In particular, Section 6 establishes why the discrete analogue of the momentum map is defined as it is, and why one cannot hope to prove a stronger result than Theorem 1.1. Finally, Section 7 contains a more general formulation and proof of this theorem.

## 2 Variational framework

The semilinear wave equation (1.1) has a Lagrangian formulation, where the configuration space  $Q$  is some space of sufficiently smooth functions, and the Lagrangian  $L: Q \times Q \rightarrow \mathbb{R}$  is given by

$$(2.1) \quad L(u, v) = \int_{S^1} \left[ \frac{1}{2} v^2 - \frac{1}{2} (\partial_x u)^2 + F(u) \right] dx,$$

where for most of the paper we assume that  $f = F'$  is entire, i.e., is an analytic function on all of  $\mathbb{C}$ . We also consider analytic functions  $f$  which are not defined on the whole of  $\mathbb{C}$  as well as  $C^k$  nonlinearities, but will focus on the case where  $f$  is entire.

Since  $v$  is the velocity of  $u$ , the corresponding Euler-Lagrange equations read

$$(2.2a) \quad \partial_t u = v,$$

$$(2.2b) \quad \partial_t v = \partial_{xx} u + f(u).$$

The semilinear wave equation is also Hamiltonian with

$$(2.3) \quad H(u, v) = \int_{S^1} \left[ \frac{1}{2} v^2 + \frac{1}{2} (\partial_x u)^2 - F(u) \right] dx.$$

Observe that the Lagrangian does not depend explicitly upon  $x$ . This means that it is invariant with respect to the group action

$$(2.4) \quad \tau_y: (u, v)(x) \mapsto (u, v)(x + y)$$

of  $\mathbb{R}$  on  $Q$ . The infinitesimal generator of this action on  $Q$  is  $\xi_Q: u \mapsto \partial_x u$ , and so the corresponding momentum map is

$$(2.5) \quad J(u, v) \cdot \xi = \frac{\partial L}{\partial v}(u, v) \cdot \xi_Q(u) = \xi \int_{S^1} v \partial_x u \, dx.$$

Thus,  $J: Q \times Q \rightarrow \mathbb{R}$  is given by

$$(2.6) \quad J(u, v) = \int_{S^1} v \partial_x u \, dx.$$

We know from Noether's theorem [16] that this quantity is an invariant of the motion, so that  $J(U(t)) = J(U(0))$  for all  $t$  and  $U(t) = (u(t), v(t)) \in Q \times Q$ .

Momentum conservation can also be checked by direct calculation,

$$\begin{aligned}
 \frac{d}{dt} \int_{S^1} v \partial_x u \, dx &= \int_{S^1} [\partial_t v \partial_x u + v \partial_{xt} u] \, dx \\
 &= \int_{S^1} [\partial_{xx} u \partial_x u + f(u) \partial_x u + \partial_t u \partial_{xt} u] \, dx \\
 &= \int_{S^1} \partial_x \left[ \frac{1}{2} (\partial_x u)^2 + F(u) + \frac{1}{2} (\partial_t u)^2 \right] \, dx \\
 (2.7) \qquad \qquad \qquad &= 0,
 \end{aligned}$$

where we have used integration by parts and the fact that  $S^1$  has no boundary. In Section 7.2 we review results on the existence of a flow of the semilinear wave equation on appropriate function spaces.

For our subsequent analysis we need to re-express the momentum in Fourier variables, where we adopt the following convention. The Fourier coefficients of a function  $u \in L^2(S^1)$  are denoted

$$(2.8) \qquad \hat{u}_k = \frac{1}{\sqrt{2\pi}} \int_{S^1} e^{-ikx} u(x) \, dx = \langle e_k, u \rangle_{L^2},$$

where

$$(2.9) \qquad e_k(x) = \frac{e^{ikx}}{\sqrt{2\pi}}$$

are the normalized eigenfunctions of the Laplacian on  $L^2(S^1)$ . The inverse transform then reads

$$(2.10) \qquad u(x) = \sum_{k \in \mathbb{Z}} \hat{u}_k e_k(x).$$

In Section 6.3 we discuss the effects of grid distortion and will need to define the trigonometric interpolant of a real function on a non-uniform mesh. For this purpose it is more natural to write the Fourier transform in terms of the normalized sine and cosine functions

$$(2.11) \qquad c_k(x) = \frac{\cos(kx)}{\sqrt{\pi}} \quad \text{and} \quad s_k(x) = \frac{\sin(kx)}{\sqrt{\pi}},$$

so that the inverse Fourier transform reads

$$(2.12) \qquad u(x) = \hat{u}_0 e_0 + \sum_{k=1}^{\infty} \hat{u}_k^c c_k(x) + \sum_{k=1}^{\infty} \hat{u}_k^s s_k(x)$$

Spatial semidiscretizations

where  $\hat{u}_k^c = \sqrt{2} \operatorname{Re} \hat{u}_k$  and  $\hat{u}_k^s = -\sqrt{2} \operatorname{Im} \hat{u}_k$ . Finally, by direct computation,

$$(2.13) \quad J(u, v) = \sum_{k \in \mathbb{Z}} ik \overline{\hat{v}_k} \hat{u}_k = \sum_{k=1}^{\infty} k (\hat{v}_k^c \hat{u}_k^s - \hat{v}_k^s \hat{u}_k^c).$$

### 3 Spatial semidiscretizations

Throughout this paper, we take  $N$  nodes in  $S^1$  at positions  $x_i$ , where  $i = 0, \dots, N-1$ , and denote approximations to  $u(x_i)$  and  $v(x_i)$  by  $u_{h,i}$  and  $v_{h,i}$ , respectively. Let  $h = 2\pi/N$  be the average grid spacing. When the grid is not uniform, the  $i$ th grid cell has width  $h_i = x_i - x_{i-1}$  for  $i = 1, \dots, N$ , and  $h_i \equiv h$  only in the special case of a uniform grid. It is understood that all indices are in  $\mathbb{Z}_N \equiv \mathbb{Z} \bmod N$ , so that  $i = -1$  is equivalent to  $i = N-1$  and, similarly,  $i = N$  is equivalent to  $i = 0$ . The discrete configuration manifold is  $Q_h = \mathbb{R}^N$ .

#### 3.1 Variational semidiscretizations

We replace the  $x$ -derivative in the Lagrangian (2.1) by a simple finite difference and approximate the integral in a way that will lead to a consistent finite difference scheme even when the grid is nonuniform. We obtain a semidiscrete Lagrangian  $L_h: Q_h \times Q_h \rightarrow \mathbb{R}$ , defined by

$$(3.1) \quad L_h(u_h, v_h) = \sum_{i=0}^{N-1} \frac{h_i + h_{i+1}}{2} \left( \frac{1}{2} v_{h,i}^2 + F(u_{h,i}) \right) - \frac{h_{i+1}}{2} \left( \frac{u_{h,i+1} - u_{h,i}}{h_{i+1}} \right)^2.$$

Noting that  $\partial_t u_h = v_h$ , we compute the semi-discrete Euler–Lagrange equations,

$$(3.2a) \quad \partial_t u_{h,i} = v_{h,i}$$

$$(3.2b) \quad \partial_t v_{h,i} = \frac{2}{h_{i+1} + h_i} \left( \frac{u_{h,i+1} - u_{h,i}}{h_{i+1}} - \frac{u_{h,i} - u_{h,i-1}}{h_i} \right) + f(u_{h,i}).$$

Writing  $\langle \cdot, \cdot \rangle$  to denote the standard Euclidean inner product on  $\mathbb{R}^N$  or  $\mathbb{C}^N$ , we now endow  $Q_h$  with the scalar product

$$(3.3) \quad \langle u_h, w_h \rangle_{Q_h} = \langle u_h, w_h \rangle_h = \sum_{i=0}^{N-1} \frac{h_i + h_{i+1}}{2} u_{h,i} w_{h,i},$$

and use this inner product on  $Q_h \times Q_h$ . This new scalar product corresponds to the metric which is generated by the kinetic energy of the discrete system

and converges to the  $L^2$ -scalar product on the continuous configuration space  $Q$  when  $h_1, \dots, h_N \rightarrow 0$ . Then the semidiscrete system is also Hamiltonian with energy

$$(3.4) \quad H_h(u_h, p_h) = \sum_{i=0}^{N-1} \frac{h_i + h_{i+1}}{2} \left( \frac{1}{2} p_{h,i}^2 - F(u_{h,i}) \right) + \frac{h_{i+1}}{2} \left( \frac{u_{h,i+1} - u_{h,i}}{h_{i+1}} \right)^2.$$

and with symplectic form  $\langle \cdot, \mathbb{J} \cdot \rangle_h \equiv \langle \cdot, \mathbb{J}_h \cdot \rangle$  on  $Q_h \times Q_h$ , where

$$(3.5) \quad \mathbb{J} = \begin{pmatrix} 0 & \text{id} \\ -\text{id} & 0 \end{pmatrix}.$$

We can then write

$$(3.6) \quad \dot{U}_h = \mathbb{J}_h^{-1} \nabla H_h(U_h) = A_h U_h + B_h(U_h),$$

where  $U_h = (u_h, v_h) \in Q_h \times Q_h$  and  $\nabla$  denotes the gradient, i.e.,  $\nabla H_h$  is a column vector. Moreover

$$(3.7) \quad A_h = \begin{pmatrix} 0 & \text{id} \\ \Delta_h & 0 \end{pmatrix}, \quad B_h(U_h) = \begin{pmatrix} 0 \\ f_h(u_h) \end{pmatrix},$$

$(f_h(u_h))_i \equiv f(u_{h,i})$ , where the discrete Laplacian  $\Delta_h$  is given by

$$(3.8) \quad (\Delta_h u_h)_i = \frac{2}{h_{i+1} + h_i} \left( \frac{u_{h,i+1} - u_{h,i}}{h_{i+1}} - \frac{u_{h,i} - u_{h,i-1}}{h_i} \right).$$

On a uniform grid, system (3.2) reduces to

$$(3.9a) \quad (u_{h,i})_t = v_{h,i}$$

$$(3.9b) \quad (v_{h,i})_t = \frac{u_{h,i+1} - 2u_{h,i} + u_{h,i-1}}{h^2} + f(u_{h,i}).$$

This system is  $\mathbb{Z}_N$ -equivariant, so that the vector field on the right of (3.9) commutes with the shift operator  $\tau_h: Q_h \times Q_h \rightarrow Q_h \times Q_h$  given by

$$(3.10) \quad (\tau_h u_h)_i = u_{h,i+1}.$$

We think of the grid symmetry on  $\mathbb{Z}_N$  as the remnant of the continuous translation symmetry on  $S^1$ .

### 3.2 Non-variational semidiscretizations

We are also interested in non-variational schemes, and introduce the family of semi-discrete wave equations

$$(3.11a) \quad \partial_t u_{h,i} = v_{h,i}$$

$$(3.11b) \quad \partial_t v_{h,i} = \frac{2}{h_{i+1} + h_i} \left( \frac{u_{h,i+1} - u_{h,i}}{h_{i+1}} - \frac{u_{h,i} - u_{h,i-1}}{h_i} \right) + \alpha f(u_{h,i+1}) + (1 - \alpha) f(u_{h,i})$$

with  $\alpha \in [0, 1]$ . This is a first order discretization in general, and it is second order when the grid is uniform and  $\alpha = 0$ . For  $\alpha = 0$  the system reduces to (3.2).

We call the discretization (3.11) *variational* if it is an Euler–Lagrange equation of a simple mechanical system. That is, we require that it possess a discrete Lagrangian  $L_h$  of the form  $L_h(u_h, v_h) = T(v_h) - V(u_h)$  with a kinetic energy  $T(v_h)$  and a potential energy  $V(u_h)$  where the kinetic energy reduces to  $T(v_h) = \frac{1}{2} \langle v_h, v_h \rangle_h$  in the case of a uniform grid. We then have the following.

**Lemma 3.1** *The discretization (3.11) is variational for arbitrary nonlinearities  $f$  and grid points  $x_0, \dots, x_{N-1}$  if and only if  $\alpha = 0$ .*

*Proof.* When  $\alpha = 0$ , (3.11) is variational with Lagrangian (3.1). When  $\alpha \neq 0$ , consider the linear uniform case where  $f(u) = u$  and  $x_i = hi$ . Since  $T(v_h) = \frac{1}{2} \langle v_h, v_h \rangle_h$ , the Euler–Lagrange equation can be written as a Hamiltonian vector field with structure matrix  $\mathbb{J}_h = h\mathbb{J}$  on  $Q_h \times Q_h$ . It is thus sufficient to show that the vector field in (3.11) is not variational with respect to the standard structure matrix  $\mathbb{J}$ , i.e., does not satisfy  $\mathbb{J}B = -B^T\mathbb{J}$ .

Indeed,  $B_h(U_h) = BU_h$  is linear with  $B$  given by

$$(3.12) \quad (Bu_h)_i = 0, \quad (Bv_h)_i = \alpha u_{h,i+1},$$

and therefore

$$(3.13) \quad \mathbb{J}B = \alpha \begin{pmatrix} \tau_h & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad -B^T\mathbb{J} = \alpha \begin{pmatrix} \tau_h^T & 0 \\ 0 & 0 \end{pmatrix},$$

where  $\tau_h$  is the grid translation defined in (3.10), and  $\tau_h^T = \tau_{-h}$ . Therefore,  $B$  is not infinitesimally symplectic with respect to the symplectic form  $\mathbb{J}_h$ . We conclude that (3.11) cannot be written as an Euler–Lagrange equation, and thus cannot be variational.  $\square$

*Remark 3.1* This lemma does not guarantee that the system is not symplectic, and hence variational, with respect to some non-canonical symplectic form. However, the numerically observed energy drift for  $\alpha \neq 0$  strongly suggests that the corresponding scheme is nonvariational with respect to any symplectic form – see Section 4.

### 3.3 Interpolated momentum

To study the conservation of the momentum  $J$  under discretization we must define a discrete analog of the momentum map. In principle, any consistent approximation to  $J$  will do – here we choose to first compute a trigonometric interpolant of the solution on the grid, and then apply the momentum  $J$  to the interpolated function. Later, when in Section 6 we establish the connection between the interpolated momentum and interpolations of discrete grid symmetries, this construction will turn out to be the natural choice.

We proceed as follows. Define a grid projection operator  $\pi_h: Q \rightarrow Q_h$  by

$$(3.14) \quad \pi_h(u) = (u(x_0), \dots, u(x_{N-1}))$$

and a trigonometric interpolation operator  $i_h: Q_h \rightarrow Q$  by

$$(3.15) \quad i_h(u_h)(x) = \frac{\hat{u}_{h,0}}{\sqrt{2\pi}} + \sum_{k=1}^{\lfloor N/2 \rfloor} \hat{u}_{h,k}^c c_k(x) + \sum_{k=1}^{\lfloor (N-1)/2 \rfloor} \hat{u}_{h,k}^s s_k(x),$$

where  $\lfloor a \rfloor$  denotes the greatest integer less than or equal to  $a$ , and the basis functions are the normalized sines and cosines

$$(3.16) \quad c_k(x) = \frac{\cos(kx)}{\sqrt{\pi}} \quad \text{and} \quad s_k(x) = \frac{\sin(kx)}{\sqrt{\pi}}.$$

The coefficients of the interpolation formula are real and can be chosen uniquely so that  $\pi_h \circ i_h = \text{id}$ . We will also write  $\pi_h: Q \times Q \rightarrow Q_h \times Q_h$  for the map  $(u_h, v_h) \rightarrow (\pi_h(u_h), \pi_h(v_h))$ , and similarly we write  $i_h: Q_h \times Q_h \rightarrow Q \times Q$ .

Defining the interpolated momentum map  $J_h \equiv J \circ i_h: Q_h \times Q_h \rightarrow \mathbb{R}$ , we see that it can be computed as a finite truncation of the Fourier expression for the continuous momentum map, equation (2.13),

$$(3.17) \quad J_h(u_h, v_h) = \sum_{k=1}^{\lfloor (N-1)/2 \rfloor} k (\hat{u}_{h,k}^s \hat{v}_{h,k}^c - \hat{u}_{h,k}^c \hat{v}_{h,k}^s).$$

When the grid is uniform, the interpolation coefficients  $\hat{u}_{h,0}$ ,  $\hat{u}_{h,k}^c$ , and  $\hat{u}_{h,k}^s$  are determined through the discrete Fourier transform [5, Section 7.2.2]. Namely,

$$(3.18) \quad \hat{u}_{h,k} = \frac{h}{\sqrt{2\pi}} \sum_{j=0}^{N-1} e^{-ikx_j} u_{h,j} \equiv \langle e_{h,k}, u_h \rangle_{Q_h}$$

for  $k = -\lfloor (N-1)/2 \rfloor, \dots, \lfloor N/2 \rfloor$ , where  $x_j = jh$  and

$$(3.19) \quad e_{h,k} = \pi_h \circ e_k = (e_k(x_0), \dots, e_k(x_{N-1}))$$

are the eigenfunctions of the discrete Laplacian  $\Delta_h$ , orthonormal with respect to the  $Q_h$  inner product. The inverse transform is given by

$$(3.20) \quad u_{h,j} = \frac{1}{\sqrt{2\pi}} \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} e^{ikx_j} \hat{u}_{h,k}.$$

By changing to normalized discrete sine and cosine functions

$$(3.21) \quad c_{h,k} = \pi_h \circ c_k \quad \text{and} \quad s_{h,k} = \pi_h \circ s_k,$$

the grid projection of the trigonometric interpolation formula coincides with the Fourier inversion formula

$$(3.22) \quad u_h = \hat{u}_{h,0} e_{h,0} + \sum_{k=1}^{\lfloor N/2\rfloor} \hat{u}_{h,k}^c c_{h,k} + \sum_{k=1}^{\lfloor(N-1)/2\rfloor} \hat{u}_{h,k}^s s_{h,k}$$

provided  $\hat{u}_{h,k}^c = \sqrt{2} \operatorname{Re} \hat{u}_{h,k}$  and  $\hat{u}_{h,k}^s = -\sqrt{2} \operatorname{Im} \hat{u}_{h,k}$ .

Finally, notice that the discrete Fourier transform is an orthogonal map (or isometry) from  $Q_h$  to  $\ell^2(\mathbb{Z}_N)$ , i.e.

$$(3.23) \quad \langle u_h, w_h \rangle_h = \langle \hat{u}_h, \hat{w}_h \rangle.$$

*Remark 3.2* Note that trigonometric interpolation is also used in von Neumann spectral analysis of uniform finite difference schemes. The reason behind this choice of interpolation is that the grid projection  $e_{h,k}$  of each Fourier mode  $e^{ikx}$  is an eigenfunction of the discrete Laplacian. Moreover, the one-dimensional space spanned by  $e^{ikx}$  is invariant under translations – it is a one-dimensional representation of the translation group. This distinguishes trigonometric interpolation from any other interpolation.

## 4 Numerical results

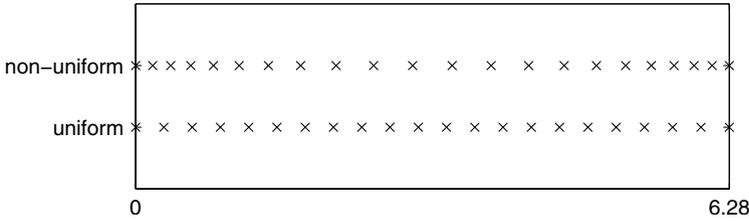
We now investigate numerically how well semidiscrete wave equations preserve the interpolated momentum. We compare four special cases of (3.11) with  $N = 21$  grid points on  $S^1$ .

(a) **Lagrangian, uniform:** The Lagrangian discretization (3.2) on a uniform grid specified by

$$(4.1) \quad x_i = hi \quad \text{for } i = 0, \dots, N-1.$$

(b) **Lagrangian, non-uniform:** The Lagrangian discretization (3.2) on a non-uniform grid (see Figure 1) specified by

$$(4.2) \quad x_i = hi - 0.4 \sin(hi) \quad \text{for } i = 0, \dots, N-1.$$



**Fig. 1.** Grids used for numerical examples. Both grids have  $N = 21$  points ( $x_0$  is identified with  $x_{21}$ ), and are specified by equations (4.1) and (4.2)

- (c) **Non-Lagrangian, uniform:** The non-Lagrangian discretization (3.11) with  $\alpha = 0.01$  on the uniform grid (4.1). Note that this is a second order discretization of  $\partial_x^2 u$ , but only a first order discretization of the nonlinear term.
- (d) **Non-Lagrangian, non-uniform:** The non-Lagrangian discretization (3.11) with  $\alpha = 0.01$  on the non-uniform grid (4.2).

In our example we take nonlinearity  $F(u) = 0.1 u^4$ , so that the semilinear wave equation reads

$$(4.3) \quad \partial_{tt} u = \partial_{xx} u + 0.4 u^3.$$

The initial conditions for all numerical simulations are taken to be the projection by  $\pi_h$  of the continuous functions

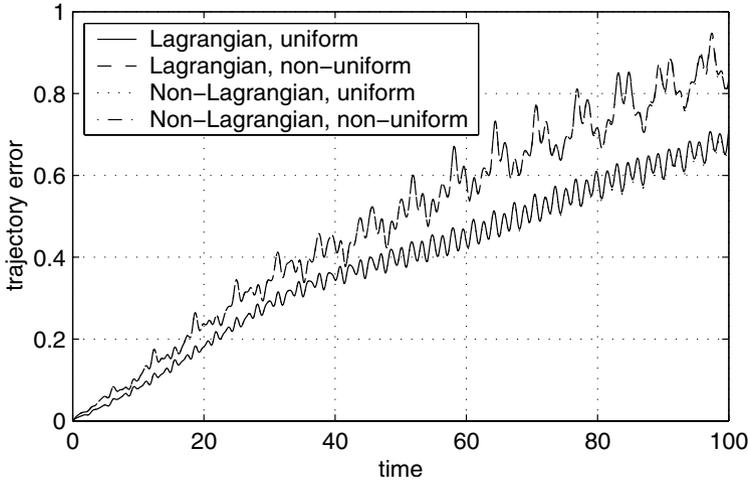
$$(4.4a) \quad u_0(x) = \frac{5}{2} \exp\left(\frac{\cos(x)}{5}\right) - \frac{5}{2}$$

$$(4.4b) \quad v_0(x) = \exp\left(\frac{\sin(x)}{5}\right) - 1$$

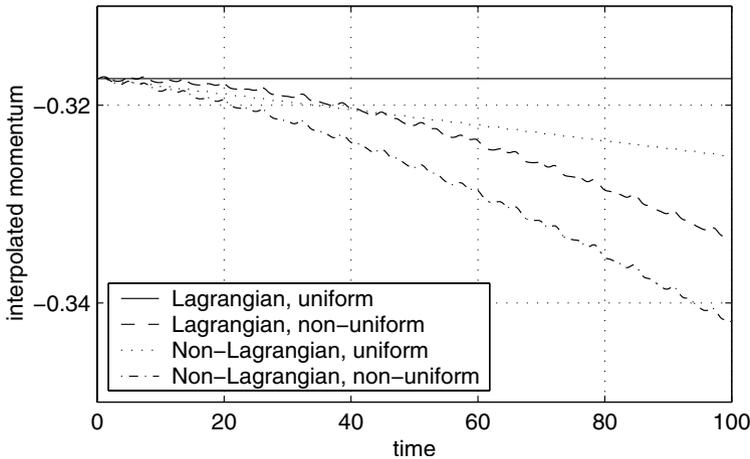
onto the appropriate grids. These initial conditions are clearly analytic, which is reflected in the exponential decay in their power spectra, as shown in Figure 5.

We simulate the four semidiscrete systems with a very accurate symplectic time integration scheme, so that there is very little error introduced by time discretization. The absolute error between the semidiscrete trajectories and the projection of the true solution to the semilinear wave equation is shown in Figure 2. From this figure it would appear that the four different semi-discretizations are behaving quite similarly.

If we now consider the evolution of the discrete momentum, however, as shown in Figure 3, then we see quite a different story. Here the Lagrangian semidiscretization on the uniform grid conserves the discrete momentum almost without error, while all of the other semidiscretizations experience substantial momentum drift.

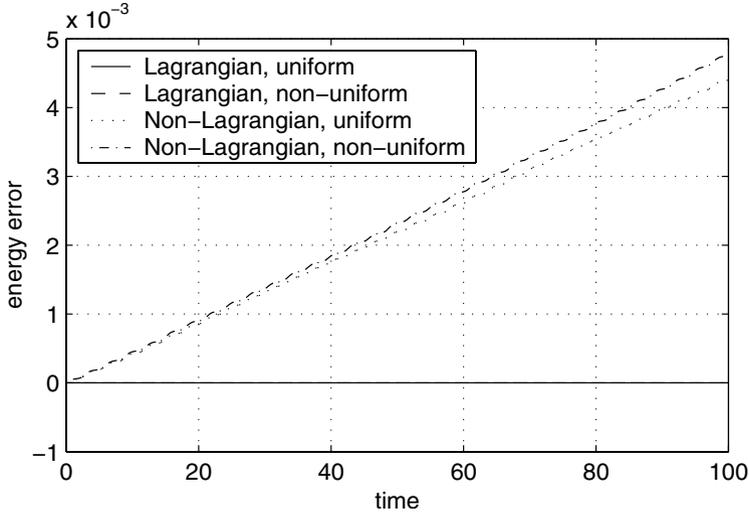


**Fig. 2.** Absolute trajectory error. All four semidiscretizations behave similarly. Due to the small value of  $\alpha$ , there is almost no difference between the Lagrangian and non-Lagrangian errors and consequently the corresponding error curves coincide. There is only a moderate difference between the errors of the solutions on uniform and non-uniform grids



**Fig. 3.** Interpolated momentum as defined in (3.17). Only the Lagrangian semidiscretization on a uniform grid has a nearly constant discrete momentum. All other semidiscretizations have significant momentum drift

The energy evolution of the various discretizations is shown in Figure 4. The situation here is different to that for momentum, as the Lagrangian semidiscretizations have exactly conserved energy functions, while the non-Lagrangian semidiscretizations do not have any conserved energies. This is



**Fig. 4.** Energy error. The Lagrangian semidiscretizations conserve energy, whereas the non-Lagrangian semidiscretizations do not

purely a product of the variational nature of the systems, and unrelated to uniformity of the grids.

Finally, the power spectra of the true solutions of the semilinear wave equation are shown in Figure 5. Note that the spectra decay exponentially, which will be important later when the regularity of solutions is considered.

## 5 Modified equations

To give a theoretical explanation for the numerical results of Section 4, we construct continuous equations which are closer to the semidiscrete systems than the original wave equation is – indeed, they interpolate the discrete solution. In Section 7 we analyze these modified equations to infer properties of the semi-discrete system.

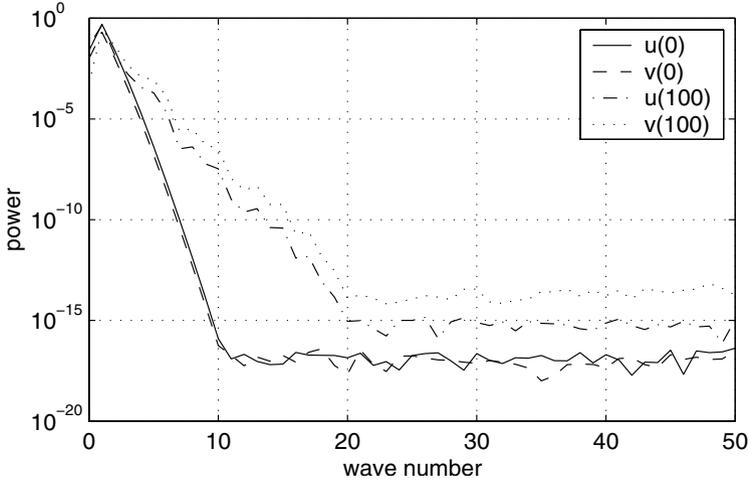
Let  $\tilde{Q}$  be a space of sufficiently smooth real valued functions over the circle  $S^1$  – for the precise functional setting see Section 7 – and let  $T\tilde{Q}$  be its tangent space.

Our grid is defined through a smooth, monotonic diffeomorphism  $g$  of  $S^1$  which leaves  $x = 0$  invariant. We set

$$(5.1) \quad x_i = g(hi)$$

for  $0 = 1, \dots, N - 1$  and  $h = 2\pi/N$ , and define the smooth, positive and  $2\pi$ -periodic grid increment function  $h^+$  via

$$(5.2) \quad h^+(g(z)) = g(z + h) - g(z)$$



**Fig. 5.** Power spectrum of exact solution (to numerical precision) of the semilinear wave equation with initial conditions (4.4) at  $t = 100$ . The exponential decay of the modes indicates that the exact solution is still in an appropriate Gevrey space. The power spectrum of  $u_h$  is given by  $(|u_{h,k}^c|^2 + |u_{h,k}^s|^2)^{1/2}$  for  $k \geq 1$  and  $|u_{h,0}|$ , with corresponding expressions for  $v_h$

for  $z \in S^1$ . It is convenient to also define the grid decrement  $h^-$  through

$$(5.3) \quad h^-(x + h^+(x)) = h^+(x)$$

for every  $x \in S^1$ . Finally, note that

$$(5.4) \quad \sum_{i=1}^N h^+(y_i) = 2\pi,$$

where  $y_i = x_i$  or, more generally,  $y_0 = y \in S^1$  and  $y_i = y_{i-1} + h^+(y_{i-1})$  for  $i = 1, \dots, N - 1$ . A natural grid interpolation corresponding to our numerical non-uniform grid function is the choice where  $h^+(x)$  and  $h^-(x)$  are computed from  $g(x) = x - \beta \sin(x)$ .

Now define the modified Lagrangian  $\tilde{L}_h: \tilde{Q} \times \tilde{Q} \rightarrow \mathbb{R}$  to be

$$(5.5) \quad \begin{aligned} \tilde{L}_h(\tilde{u}_h, \tilde{v}_h) &= \frac{1}{h} \int_{S^1} \frac{h^+(x) + h^-(x)}{2} \left( \frac{1}{2} \tilde{v}_h^2(x) + F(\tilde{u}_h(x)) \right) dg^{-1}(x) \\ &\quad - \frac{1}{h} \int_{S^1} \frac{h^+(x)}{2} \left( \frac{\tilde{u}_h(x + h^+(x)) - \tilde{u}_h(x)}{h^+(x)} \right)^2 dg^{-1}(x), \end{aligned}$$

which is consistent with the true Lagrangian (2.1). In the special case a uniform grid, this simplifies to

$$(5.6) \quad \tilde{L}_h(\tilde{u}_h, \tilde{v}_h) = \int_{S^1} \left[ \frac{1}{2} \tilde{v}_h^2(x) - \frac{1}{2} \left( \frac{\tilde{u}_h(x+h) - \tilde{u}_h(x)}{h} \right)^2 + F(\tilde{u}_h(x)) \right] dx.$$

When taking variations of the general modified Lagrangian, we need to make the change of variables  $y = x + h^+(x)$  in one of the terms. Note, therefore, that (5.2) implies that  $x + h^+(x) = g(g^{-1}(x) + h)$ , so that  $g^{-1}(y) = g^{-1}(x) + h$  and  $dg^{-1}(y) = dg^{-1}(x)$ . It is then straightforward to check that the Euler-Lagrange equations are

$$(5.7a) \quad \partial_t \tilde{u}_h(x) = \tilde{v}_h(x)$$

$$(5.7b) \quad \partial_t \tilde{v}_h(x) = \frac{2}{h^+(x) + h^-(x)} \left( \frac{\tilde{u}_h(x + h^+(x)) - \tilde{u}_h(x)}{h^+(x)} - \frac{\tilde{u}_h(x) - \tilde{u}_h(x - h^-(x))}{h^-(x)} \right) + f(\tilde{u}_h(x)).$$

For uniform grids, system (5.7) simplifies to

$$(5.8a) \quad \partial_t \tilde{u}_h(x) = \tilde{v}_h(x)$$

$$(5.8b) \quad \partial_t \tilde{v}_h(x) = \frac{\tilde{u}_h(x+h) - 2\tilde{u}_h(x) + \tilde{u}_h(x-h)}{h^2} + f(\tilde{u}_h(x)).$$

These equations are just an uncountable number of uncoupled ordinary differential equations of dimension  $N$ , which we regard as an ordinary differential equation on the function space  $\tilde{Q} \times \tilde{Q}$ .

The general modified equation corresponding to (3.11) is

$$(5.9a) \quad \partial_t \tilde{u}_h(x) = \tilde{v}_h(x)$$

$$(5.9b) \quad \partial_t \tilde{v}_h(x) = \frac{2}{h^+(x) + h^-(x)} \left( \frac{\tilde{u}_h(x + h^+(x)) - \tilde{u}_h(x)}{h^+(x)} - \frac{\tilde{u}_h(x) - \tilde{u}_h(x - h^-(x))}{h^-(x)} \right) + \alpha f(\tilde{u}_h(x + h^+(x))) + (1 - \alpha) f(\tilde{u}_h(x)).$$

Like the discrete system (3.11) in Lemma 3.1, the general modified equation (5.7) is variational if and only if  $\alpha = 0$ . Observe that the modified equations have the special property that they exactly coincide with the corresponding spatial semidiscretization on grid points.

We are interested in when the modified equation has the same spatial translation symmetry as the original wave equation. Recall that the group action has the infinitesimal generator  $\xi_{\tilde{Q}}: \tilde{u} \mapsto \partial_x \tilde{u}$ , which lifts to the generator  $\xi_{\tilde{Q} \times \tilde{Q}}: (\tilde{u}, \tilde{v}) \mapsto (\partial_x \tilde{u}, \partial_x \tilde{v})$ . The corresponding momentum map is

thus the same as that for the original wave equation. We will also write that  $J: \tilde{Q} \times \tilde{Q} \rightarrow \mathbb{R}$  is given by

$$(5.10) \quad J(\tilde{u}_h, \tilde{v}_h) = \int_{S^1} \tilde{v}_h \partial_x \tilde{u} \, dx.$$

In the special case of a uniform grid, the general modified equation (5.7) is translation invariant. For general grids, however, this is not the case as the following lemma shows.

**Lemma 5.1** *System (5.7) is equivariant under arbitrary space-shifts if and only  $h^\pm(x) \equiv h$ , i.e., the discretization is uniform.*

*Proof.* We have already argued that (5.7) is  $S^1$ -equivariant on a uniform grid. To prove the converse, assume (5.7) is  $S^1$ -equivariant under translations, i.e., the right hand side of (5.7b) commutes with the shift operator  $\tau_y$ , where  $(\tau_y u)(x) = u(x + y)$ . Fix  $z \in S^1$  and choose a smooth test function  $\tilde{u}(x)$  whose support is contained in a small neighborhood of  $z$  so that, in particular,  $\tilde{u}(z + h^+(z)) = \tilde{u}(z - h^-(z)) = 0$ . Then equivariance of (5.7b) implies that

$$(5.11) \quad q(x) \equiv \frac{2}{h^+(x) + h^-(x)} \left( \frac{-1}{h^+(x)} - \frac{1}{h^-(x)} \right) = \frac{-2}{h^+(x) h^-(x)}$$

is independent of  $x = z + y$  for  $y$  small. As the circle can be uniformly covered with such neighborhoods,  $q(x)$  is constant on all of  $S^1$ , and therefore

$$(5.12) \quad h^+(x) h^-(x) = C.$$

From (5.3) we see that  $[x, x + 2\pi]$  is divided into  $N$  intervals such that each of length  $h^-(x)$  is followed by one of length  $h^+(x)$ , followed again by one of length  $h^-(x)$ . So  $h^+(x) + h^-(x) = 2h$ . As a consequence,

$$(5.13) \quad \begin{aligned} (h^+(x) - h^-(x))^2 &= (h^+(x) + h^-(x))^2 - 4h^+(x)h^-(x) \\ &= 4h^2 - 4C \equiv C_2^2, \end{aligned}$$

so that  $h^+(x) - h^-(x) = \pm C_2$ . The sign, however, cannot depend on  $x$  as  $h^+(x)$  and  $h^-(x)$  are smooth functions. Noting that  $h^+(x) - h^-(x) = h^+(x) - h^+(x - h^-(x))$  while  $h^+$  is periodic, we see that the constant  $C_2$  must vanish.  $\square$

By Noether's theorem, the modified equation corresponding to the Lagrangian semidiscretization on a uniform grid exactly preserves the momentum map. That is, for any solution  $\tilde{U}_h(t) = (\tilde{u}_h(t), \tilde{v}_h(t))$  of (5.9) we have  $J(\tilde{U}_h(t)) = J(\tilde{U}_h(0))$  for all  $t$ . We can also check this by direct calculation:

$$\begin{aligned}
& \frac{d}{dt} \int_{S^1} \tilde{v}_h \partial_x \tilde{u}_h \, dx \\
&= \int_{S^1} (\partial_t \tilde{v}_h \partial_x \tilde{u}_h + \tilde{v}_h \partial_{tx} \tilde{u}_h) \, dx \\
&= \int_{S^1} \left[ \frac{\tilde{u}_h(x+h) - 2\tilde{u}_h(x) + \tilde{u}_h(x-h)}{h^2} \partial_x \tilde{u}_h(x) \right. \\
&\quad \left. + f(\tilde{u}_h(x)) \partial_x \tilde{u}_h(x) + \partial_t \tilde{u}_h(x) \partial_{xt} \tilde{u}_h(x) \right] \, dx \\
&= \int_{S^1} \left[ \frac{1}{h^2} \tilde{u}_h(x+h) \partial_x \tilde{u}_h(x) - \frac{1}{h^2} \partial_x \tilde{u}_h(x-h) \tilde{u}_h(x) \right. \\
&\quad \left. - \frac{1}{h^2} \partial_x \tilde{u}_h^2(x) + \partial_x F(\tilde{u}_h(x)) + \frac{1}{2} \partial_x (\partial_t \tilde{u}_h(x))^2 \right] \, dx \\
&= \int_{S^1} \left[ \frac{1}{h^2} \tilde{u}_h(x+h) \partial_x \tilde{u}_h(x) - \frac{1}{h^2} \partial_x \tilde{u}_h(x) \tilde{u}_h(x+h) \right] \, dx \\
(5.14) \quad &= 0,
\end{aligned}$$

where we have used integration by parts and the change of coordinates  $x \mapsto x+h$  in one term.

Observe that although the non-Lagrangian modified equation – (5.9) with  $\alpha \neq 0$  – on a uniform grid has the spatial translation symmetry, its non-Lagrangian nature means that it does not have the momentum map as a conserved quantity. This can be explicitly checked by computing the time derivative as above,

$$(5.15) \quad \frac{d}{dt} \int_{S^1} \tilde{v}_h \partial_x \tilde{u}_h \, dx = \alpha \int_{S^1} f(\tilde{u}_h(x+h)) \partial_x \tilde{u}_h \, dx,$$

which generally does not vanish.

In summary, the modified equations for the various spatial semidiscretizations are all exact at grid points, and their conservation properties are as follows.

- (a) **Lagrangian, uniform:** The modified equation is a Lagrangian system, which retains the spatial translation symmetry. The modified system thus preserves the continuous momentum function for all time. Note that this does not immediately provide any constraint on the evolution of the discrete momentum  $J_h$ , as the way in which the modified solution interpolates between the values at grid points will change over time, and the conserved quantity depends on this interpolation.
- (b) **Lagrangian, non-uniform:** The modified equation is a Lagrangian system, but it does not have the spatial translation symmetry. The momentum function is thus not conserved.

- (c) **Non-Lagrangian, uniform:** The modified equation does have the spatial translation symmetry, but it is not a Lagrangian system. The momentum is thus not conserved.
- (d) **Non-Lagrangian, non-uniform:** The modified equation is neither a Lagrangian system, nor does it have the spatial translation symmetry. The momentum function is thus not conserved.

We interpret the above properties as explaining the behavior which was numerically observed in Figure 3. Of course, this relies on the rigorous results of Section 7 to establish that properties of the modified equation do indeed imply properties of the semidiscrete system.

*Remark 5.1* A Taylor expansion of the modified system (5.8) in  $h$  gives higher order Hamiltonian PDEs with translation symmetry, as have previously been derived by Moore and Reich [21].

*Remark 5.2* Other choices for the modified system are conceivable, so long as they possess translation symmetry and accurately interpolate the discrete system. We have chosen a particularly simple modified equation for which the interpolation of the discrete system is exact.

## 6 Spatial discretizations and momentum maps

When space is discrete, there is no obvious sense in which the semidiscrete system (3.11) can have a continuous spatial translation symmetry with a corresponding conservation law. On the other hand, we have seen that there are numerical schemes – variational discretizations on an equispaced grid – for which the interpolated momentum  $J_h = J \circ i_h$  is conserved with very high accuracy. We show in Section 6.1 that on a uniform grid  $J_h$  is distinguished from other discrete momentum maps by generating a symmetry that interpolates the discrete grid shift, and where the interpolation coincides with the continuous  $S^1$  translation group for wavenumbers that are resolved on the grid. Moreover, we show that  $J_h$  belongs to a special class of interpolated momenta that are exactly conserved when the wave equation is linear. From Figure 3 it may appear that the momentum is also conserved in the semilinear case. While this turns out to be false, we shall see in Section 7 that even in the presence of nonlinearities momentum is approximately conserved.

Section 6.2 discusses finite difference momenta. We shall see that naive finite difference approximations to  $J$  are not conserved in any case, but will only oscillate with amplitude of  $O(h)$  about the interpolated momentum.

When the grid is non-uniform, the situation is radically different. In Section 6.3 we show that even for the linear wave equation any discrete momentum generically drifts by an  $O(1)$  amount within a sufficiently long interval of time.

### 6.1 Exact conservation of interpolated momenta

We now turn to the question of defining interpolated momenta that in the linear case are exactly conserved. Throughout this section we consider the Lagrangian semidiscretization on a uniform grid. For simplicity, we also assume that  $N$  is odd, and comment on necessary modifications for even  $N$  where appropriate.

We can interpolate the discrete symmetry group  $\mathbb{Z}_N$  of the grid which acts linearly on  $Q_h$ , i.e., the group of discrete shifts, to a linear  $S^1$ -action on  $Q_h$  such that for angles  $2\pi j/N$  we recover the grid group element which shifts the grid by  $j$  points. We will see that the continuous symmetry generated by the interpolated momentum  $J_h = J \circ i_h$  is such an interpolation that, moreover, corresponds to a translation of the trigonometric interpolant.

**Lemma 6.1** *Let  $N$  be odd and the grid be uniform. Then the infinitesimal generator  $\xi_h = \mathbb{J}_h \nabla J_h(U_h)$  on  $Q_h \times Q_h$  corresponding to the interpolated momentum  $J_h$  is given by*

$$(6.1) \quad \xi_h = \pi_h \circ \partial_x(i_h \circ U_h).$$

*Remark 6.1* Since  $\partial_x = \xi_Q$  is the infinitesimal generator of the translation on  $Q$ , the symmetry corresponding to the interpolated momentum acts identically to the continuous translation symmetry restricted to eigenmodes  $\pi_h \circ e_k$  with wave numbers  $|k| \leq [(N-1)/2]$ .

*Proof.* Using the antisymmetry of the discrete symplectic structure, the chain rule, the linearity of  $i_h$ , and the observations that  $i_h \circ \mathbb{J} = \mathbb{J}i_h$  and  $\mathbb{J}_h = h\mathbb{J}$ , we can write

$$(6.2) \quad \begin{aligned} \langle \xi_h(U_h), W_h \rangle_h &\equiv \mathbb{J}D J_h(U_h) W_h = -DJ_h(U_h) \mathbb{J}W_h \\ &= -DJ(i_h \circ U_h) i_h \circ \mathbb{J}W_h \\ &= -DJ(i_h \circ U_h) \mathbb{J}i_h \circ W_h. \end{aligned}$$

Since  $\partial_x$  is the infinitesimal generator corresponding to the momentum  $J$ , we know that  $\partial_x U = \mathbb{J} \nabla J(U)$ . Alternatively, we could compute the Fréchet derivative of  $J$  by direct computation and integrate by parts in one of the terms. Substituting either result into the last expression of (6.2), we obtain

$$(6.3) \quad \begin{aligned} \langle \xi_h(U_h), W_h \rangle_{Q_h \times Q_h} &= \langle \mathbb{J} \partial_x(i_h \circ U_h), \mathbb{J}i_h \circ W_h \rangle_{L^2} \\ &= \langle \partial_x(i_h \circ U_h), i_h \circ W_h \rangle_{L^2} \\ &= \langle \pi_h \circ \partial_x(i_h \circ U_h), \pi_h \circ i_h \circ W_h \rangle_h. \end{aligned}$$

The last equality is based on a key property of trigonometric interpolation: the  $L^2$  inner product can be expressed as a finite Fourier series involving

only wavenumbers with  $|k| \leq [N/2]$ . On this subspace, however, the discrete Fourier transform coincides with the continuous Fourier transform for  $S^1$  functions and, moreover, it is orthogonal by (3.23). We can thus use the discrete inverse Fourier transform to obtain the last expression in (6.3). Since  $\pi_h \circ i_h = \text{id}$ , the claim follows.  $\square$

*Remark 6.2* For  $N$  even, Lemma 6.1 remains true if we replace (6.1) by

$$(6.4) \quad \xi_h = \pi_h \circ P_{n-1} \partial_x (i_h \circ U_h),$$

where  $n = N/2$  and  $P_n$  denotes the projection onto the Fourier modes  $-n, \dots, n$ . Remark 6.1 also remains valid.

As an immediate consequence from Lemma 6.1 we see that  $\xi_h$  is diagonal with respect to the discrete Fourier basis  $e_{h,k} = \pi_h \circ e_k$  where  $k = -[N/2], \dots, [N/2]$ . The corresponding eigenvalues are  $ik$ , so that

$$(6.5) \quad (e^{t\xi_h} u_h)^\wedge_k = e^{ikt} \hat{u}_{h,k}.$$

In particular, for  $t = h$  we recover the eigenvalues and eigenvectors of the discrete shift operator  $\tau_h = \exp(\xi_h h)$ .

Vice versa, we can *define* a linear interpolation of the discrete shift,  $\exp(t\tilde{\xi}_h)$ , by requiring that  $\tau_h = \exp(\tilde{\xi}_h h)$ . Thus,  $\tilde{\xi}_h$  is also diagonal with respect to the discrete Fourier basis and has eigenvalues  $\nu_k = i(k \bmod N)$ . If we further require that the group action is real, then  $\nu_{-k} = \bar{\nu}_k$  and  $\nu_0 = 0$ . We can now ask whether the generalized interpolated momentum  $\tilde{J}_h$  corresponding to  $\tilde{\xi}_h$  is conserved. This is the content of the following theorem.

**Theorem 6.2** *Consider the variational spatial semidiscretization (3.9) on an equispaced grid with  $N$  odd. Further, let  $\tilde{\xi}_h \in \mathfrak{gl}(N)$  denote the infinitesimal generator of an arbitrary linear real interpolation of the discrete shift symmetry. Then the corresponding momentum  $\tilde{J}_h$  is a constant of motion if and only if the wave equation is affine, i.e., if  $f(u) = c_1 u + c_2$  for constants  $c_1$  and  $c_2$ .*

*Proof.* The proof is based on elementary facts from geometric mechanics. A function  $J: M \rightarrow \mathbb{R}$  on a symplectic space  $M$  with structure matrix  $\mathbb{J}$  is conserved by the flow of a Hamiltonian vector field  $X_H = \mathbb{J}DH(x)$  on  $M$  if and only if  $0 = \dot{J} = DJ\mathbb{J}\nabla H$ . Since  $DJ\mathbb{J}\nabla H = -DH(x)\xi(x)$  where  $\xi(x) \equiv \mathbb{J}\nabla J(x) = X_J(x)$  is the Hamiltonian vector field generated by  $J$ , conservation of  $J$  implies invariance of the Hamiltonian  $H(x)$  under the group  $\exp(t\xi)$  generated by  $\xi$ . It follows that the Hamiltonian vector field is equivariant under the group action; in other words, that  $X_H$  commutes with  $\exp(t\xi)$ .

If the wave equation is linear, then the Hamiltonian vector field of its discretization – the right side of (3.9) – is a linear affine combination of shifts.

We can therefore use the Fourier representation of the interpolation  $\tilde{\xi}_h$  of the shift symmetry  $\tau_h$  and check by direct calculation that  $DH_h(U_h)\tilde{\xi}_h U_h \equiv 0$ , where the semidiscrete Hamiltonian  $H_h$  is given by (3.4).

To prove the converse, assume that the Hamiltonian vector field commutes with  $\exp(t\tilde{\xi}_h)$ . This implies, in particular, that

$$(6.6) \quad f_h(e^{t\tilde{\xi}_h} u_h) = e^{t\tilde{\xi}_h} f_h(u_h),$$

for every  $u_h \in Q_h$ . Now choose  $u_{h,j} = \gamma \delta_{0j}$ , so that  $\hat{u}_{h,k} = \text{const} = \gamma \sqrt{2\pi}/N$  and

$$(6.7) \quad (e^{t\tilde{\xi}_h} u_h)_0 = \frac{1}{\sqrt{2\pi}} \sum_{k=-[N/2]}^{[N/2]} e^{vk t} \hat{u}_{h,k} \equiv \gamma \phi(t),$$

where

$$(6.8) \quad \phi(t) = \frac{1}{N} \sum_{k=-[N/2]}^{[N/2]} e^{vk t}.$$

On the one hand, we conclude that

$$(6.9) \quad (f_h(e^{t\tilde{\xi}_h} u_h))_0 = f(\gamma \phi(t)).$$

On the other hand, we write  $\exp(t\tilde{\xi}_h) f_h(u_h) = \exp(t\tilde{\xi}_h)(f_h(u_h) - f_h(0)) + \exp(t\tilde{\xi}_h) f_h(0)$ . Noting that the constant vector  $(1, \dots, 1)$  is the zero eigenvector of  $\tilde{\xi}_h$ , and thus eigenvector with eigenvalue 1 for  $\exp(t\tilde{\xi}_h)$ , we can immediately turn to the first term. Repeating the computation leading up to (6.7) with  $f_h(u_h) - f_h(0)$  in place of  $u_h$ , we find that

$$(6.10) \quad (e^{t\tilde{\xi}_h} f_h(u_h))_0 = (f(\gamma) - f(0)) \phi(t) + f(0).$$

Equivariance implies that the right sides of (6.9) and (6.10) coincide. In other words,

$$(6.11) \quad f(\gamma \phi(t)) - f(0) = \phi(t) (f(\gamma) - f(0)).$$

Since  $\phi$  is a continuous non-constant function and  $\gamma$  is arbitrary, we conclude that  $g(x) = f(x) - f(0)$  is linear, hence  $f$  is affine.  $\square$

*Remark 6.3* If  $N$  is even, the discrete shift  $\tau_h$  cannot be interpolated to an  $S^1$ -action on  $Q_h$  by real-valued matrices because  $-1$  is a simple eigenvalue of  $\tau_h$  and its eigenvector is the highest discrete Fourier mode  $(1, -1, \dots, 1, -1)$ . But it is still true that the momentum  $J_h$  is a conserved quantity of the discretization (3.9) if and only if the wave equation is affine. The proof is analogous to the proof of Theorem 6.2, but now  $\nu_{N/2} = 0$  and the summations in (6.7) and (6.8) run from  $-[(N-1)/2]$  to  $[N/2]$ .

## 6.2 Finite difference momenta

As a simpler way of computing an approximate momentum map, we could take finite difference approximations to the continuous momentum  $J$ . It will turn out, however, that low order finite difference momenta are neither conserved for equispaced discretizations of the linear wave equation, nor do they generate an  $S^1$  symmetry.

We define the first order *finite difference momentum*

$$\begin{aligned}
 J_h^{\text{fd},1}(u_h, v_h) &= \sum_{i=0}^{N-1} (x_{i+1} - x_i) \frac{u_{h,i+1} - u_{h,i}}{x_{i+1} - x_i} v_{h,i} \\
 (6.12) \qquad \qquad \qquad &= \sum_{i=0}^{N-1} (u_{h,i+1} - u_{h,i}) v_{h,i},
 \end{aligned}$$

and the *symmetric finite difference momentum* which is second order on a uniform grid by

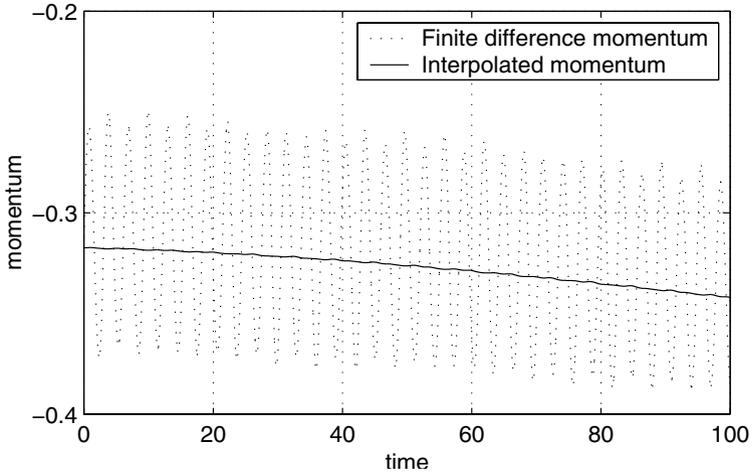
$$\begin{aligned}
 J_h^{\text{fd},2}(u_h, v_h) &= \sum_{i=0}^{N-1} \frac{x_{i+1} - x_{i-1}}{2} \frac{u_{h,i+1} - u_{h,i-1}}{x_{i+1} - x_{i-1}} v_{h,i} \\
 (6.13) \qquad \qquad \qquad &= \frac{1}{2} \sum_{i=0}^{N-1} (u_{h,i+1} - u_{h,i-1}) v_{h,i}.
 \end{aligned}$$

Figures 6 and 7 show that the first order finite difference momentum oscillates about, but does not drift from the interpolated momentum. The latter feature is a consequence of the consistency of the finite difference momentum as  $h \rightarrow 0$ . Moreover, when the grid is uniform and  $f = 0$ , it is easy to verify the following.

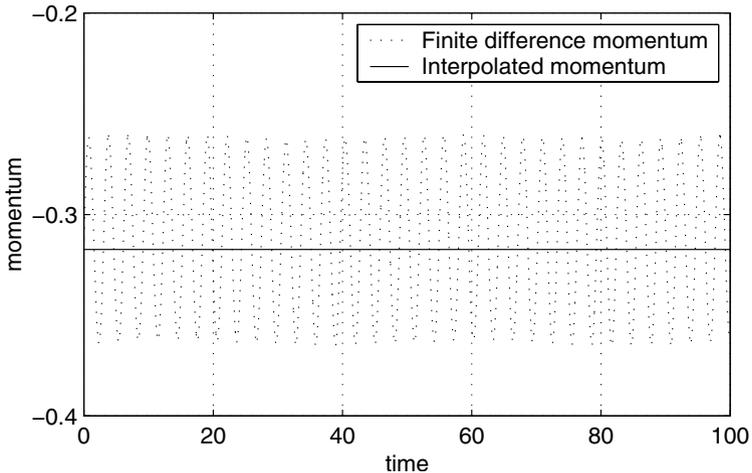
- (a) Neither  $J_h^{\text{fd},1}$  nor  $J_h^{\text{fd},2}$  are conserved by our variational semidiscretization on the uniform grid.
- (b) The group action  $\exp(t\xi_h^{\text{fd},1})$  on  $\mathcal{Q}_h \times \mathcal{Q}_h$  which is generated by  $\xi_h^{\text{fd},1} = \mathbb{J}\nabla J_h^{\text{fd},1}$  is not an orthogonal action for any scalar product, while the group action  $\exp(t\xi_h^{\text{fd},2})$  corresponding to  $J_h^{\text{fd},2}$  is orthogonal.
- (c) The group  $G_h^{\text{fd},1} \equiv \{\exp(t\xi_h^{\text{fd},1}) : t \in \mathbb{R}\}$  generated by  $\xi_h^{\text{fd},1}$  is isomorphic to  $\mathbb{R}$ , not to  $S^1$ . The group  $G_h^{\text{fd},2}$  generated by  $\xi_h^{\text{fd},2}$  is a torus group for  $N > 4$ ; i.e., it is isomorphic to  $S^1 \times \cdots \times S^1$ , not to  $S^1$ .

To verify claim (a), it suffices, as in the proof of Theorem 6.2, to show that the Hamiltonian vector field  $\xi_h^{\text{fd},1}$  generated by  $J_h^{\text{fd},1}$  is not an infinitesimal symmetry of  $H_h$ , i.e., that

$$(6.14) \qquad \qquad \qquad DH_h(x)\xi_h^{\text{fd},1}(x) \neq 0.$$



**Fig. 6.** First order finite difference momentum (6.12) for the Lagrangian discretization on a non-uniform grid, together with the interpolated momentum for comparison. The finite difference momentum oscillates about the value of the interpolated momentum, but tracks it as it drifts. The plots for the non-Lagrangian discretizations and for the symmetric finite difference momentum are similar

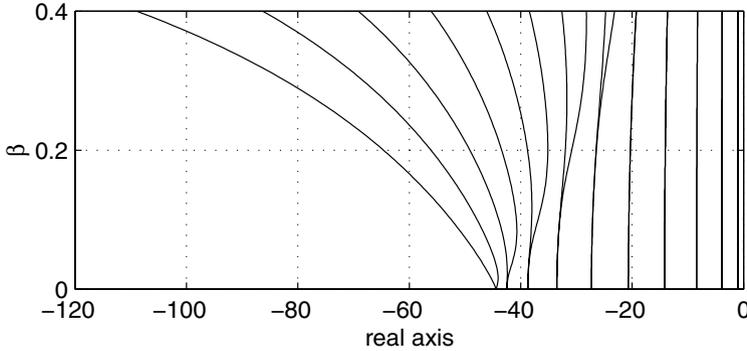


**Fig. 7.** First order finite difference momentum for the Lagrangian discretization on a uniform grid, with the interpolated momentum for reference. The plot for the symmetric finite difference momentum is similar. The finite different momentum does not drift, but merely oscillates about the true momentum value

We compute

$$(6.15) \quad \xi_h^{\text{fd},1} \begin{pmatrix} u_h \\ p_h \end{pmatrix} = \mathbb{J}_h \nabla J_h^{\text{fd},1}(u_h, p_h) = h^{-1} \begin{pmatrix} \tau_h u_h - u_h \\ p_h - \tau_{-h} p_h \end{pmatrix}.$$

Condition (6.14) can now be checked by direct computation.



**Fig. 8.** Splitting of the eigenvalues of the discrete Laplacian (3.8) induced by the non-uniform grid  $x_i = hi - \beta \sin(hi)$

To prove (b) and (c), note that the eigenvalues of  $\tau_h$  are  $e^{ikh}$ , where  $k = -[(N-1)/2], \dots, [N/2]$ . Hence, the eigenvalues of  $\xi_h^{\text{fd},1}$  are given by  $\lambda_{h,k,\pm}^{\text{fd},1} = \pm(e^{ikh} - 1)/h$ , and there is some  $k$  for which the real part of  $\lambda_{h,k,\pm}^{\text{fd},1}$  does not vanish. This implies that the group generated by  $\xi_h^{\text{fd},1}$  is noncompact and, since it is a one-parameter group, is isomorphic to  $\mathbb{R}$ . In particular,  $\xi_h^{\text{fd},1}$  is not a skew symmetric matrix and therefore the group action is not orthogonal.

For  $J_h^{\text{fd},2}$  we compute that  $\xi_h^{\text{fd},2} = \frac{1}{2h}(\tau_h - \tau_{-h})$ , and so, since  $\tau_h^T = \tau_{-h}$  we find that  $\xi_h^{\text{fd},2}$  is skew-symmetric. The eigenvalues of  $\xi_h^{\text{fd},2}$  are  $\lambda_{h,k}^{\text{fd},2} = \pm i \sin(kh)/h$  for  $k = 0, \dots, [N/2]$ . Since they are rationally independent for  $N > 4$ ,  $\xi_h^{\text{fd},2}$  generates a torus group. Still,  $\xi_h^{\text{fd},2}$  is not an infinitesimal symmetry of  $H_h$ , and therefore the corresponding momentum map is not a conserved quantity of (3.2).

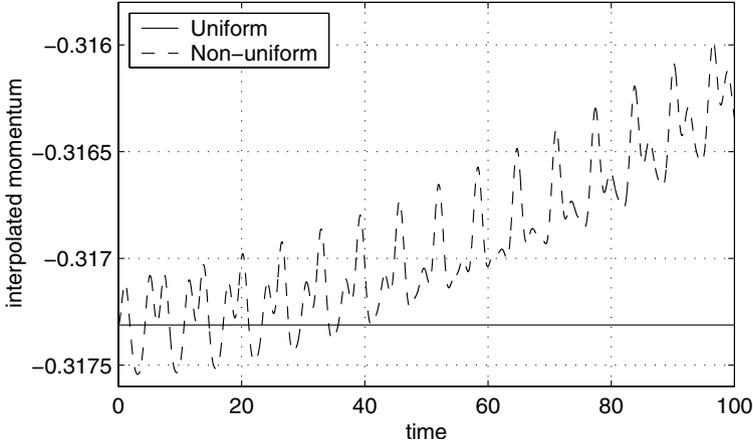
### 6.3 Nonuniform space discretizations for linear waves

In this section consider only the linear wave equation. We show that non-uniformity of the grid generically breaks conservation of any consistent approximate momentum.

When the grid is uniform and  $k \geq 1$ , the grid projections  $c_{h,k} = \pi_h \circ c_k$  and  $s_{h,k} = \pi_h \circ s_k$  span a two-dimensional eigenspace  $E_{h,k}$  of the discrete Laplacian  $\Delta_h$  with identical respective eigenvalues

$$(6.16) \quad \lambda_{h,k} = -\gamma_{h,k}^2 = -\sigma_{h,k}^2 = 2 \frac{\cos(hk) - 1}{h^2}.$$

In fact, when  $N$  is odd all eigenspaces of  $\Delta_h$  where the shift symmetry  $\mathbb{Z}_N$  acts nontrivially are necessarily two-dimensional, and these 1:1 resonances



**Fig. 9.** Error of the interpolated momentum of the discretized linear wave equation using a uniform grid and the non-uniform grid (4.2)

are essential for momentum conservation. If the discrete Laplacian  $\Delta_h$  is perturbed to a non-selfadjoint matrix which still commutes with  $\mathbb{Z}_N$ , then the 1:1 resonance between the real parts of the eigenvalues persists. When the shift symmetry is broken then generically the 1:1 resonances are destroyed, which implies momentum drift. This follows from the general theory of symmetric differential equations [7]. Figure 8 shows that the 1:1 resonances are indeed destroyed by a non-uniform grid distortion, and we will now derive an asymptotic formula for the resulting momentum drift, which can be seen in Figure 9.

First, note that for the uniform variational semidiscrete linear wave equation not only the actions

$$(6.17) \quad I_{h,k}^c \equiv |\hat{v}_{h,k}^c|^2 + \gamma_{h,k}^2 |\hat{u}_{h,k}^c|^2$$

$$(6.18) \quad I_{h,k}^s \equiv |\hat{v}_{h,k}^s|^2 + \sigma_{h,k}^2 |\hat{u}_{h,k}^s|^2,$$

but also the momenta

$$(6.19) \quad J_{h,k} = k (\hat{u}_{h,k}^s \hat{v}_{h,k}^c - \hat{u}_{h,k}^c \hat{v}_{h,k}^s)$$

are conserved. (In other words, the linear wave equation is super-integrable – there are more independent integrals than degrees of freedom.) It is therefore sufficient to study the momentum on only a single  $A_h$ -invariant subspace  $E_{h,k} \times E_{h,k}$ .

Let  $g^\varepsilon$  be a smooth family of diffeomorphisms on  $S^1$  with  $g^0 = \text{id}$ , and let  $x_i^\varepsilon = g^\varepsilon(hi)$  denote the nodes of a distorted grid. For the family of corresponding discrete Laplacians  $\Delta_h^\varepsilon$  there are smoothly varying  $\Delta_h^\varepsilon$ -invariant two-dimensional subspaces  $E_{h,k}^\varepsilon$ ,  $k = 1, \dots, \lfloor \frac{N-1}{2} \rfloor$ , with  $E_{h,k}^0 = E_{h,k}$ . Let

$c_{h,k}^\varepsilon$  and  $s_{h,k}^\varepsilon$  denote an orthonormal basis of  $E_{h,k}^\varepsilon$  which is smooth in  $\varepsilon$  such that  $c_{h,k}^0 = c_{h,k}$ ,  $s_{h,k}^0 = s_{h,k}$ . The  $2 \times 2$  matrix  $\Delta_{h,k}^\varepsilon = \Delta_h^\varepsilon|_{E_{h,k}^\varepsilon}$  is then also smooth in  $\varepsilon$ . The pseudo-momenta

$$(6.20) \quad J_{h,k}^\varepsilon = k (u_{h,k}^{s,\varepsilon} v_{h,k}^{c,\varepsilon} - u_{h,k}^{c,\varepsilon} v_{h,k}^{s,\varepsilon}),$$

defined in terms of the expansion coefficients with respect to the *distorted* basis functions will generally be different from, but  $O(\varepsilon)$ -close to, the interpolated momentum  $J_h|_{E_{h,k}^\varepsilon \times E_{h,k}^\varepsilon}$  for each distorted  $\Delta_{h,k}^\varepsilon$ -invariant subspace  $E_{h,k}^\varepsilon = \text{span}\{c_{h,k}^\varepsilon, s_{h,k}^\varepsilon\}$ .

The semi-discrete linear wave equation (3.2) is simply a collection of uncoupled harmonic oscillators for  $k \geq 1$  together with a subsystem for the constant eigenvector corresponding to  $k = 0$ , as is the original linear wave equation itself. Thus, restricting ourselves to  $E_{h,k}$  we have, for  $\varepsilon = 0$ ,  $\gamma_{h,k} = |\lambda_{h,k}|^{1/2}$ ,

$$(6.21) \quad u_{h,k}^c(t) = u_{h,k}^c(0) \cos(\gamma_{h,k}t) + \frac{v_{h,k}^c(0)}{\gamma_{h,k}} \sin(\gamma_{h,k}t)$$

$$(6.22) \quad v_{h,k}^c(t) = v_{h,k}^c(0) \cos(\gamma_{h,k}t) - u_{h,k}^c(0) \gamma_{h,k} \sin(\gamma_{h,k}t),$$

and corresponding solution formulas for  $u_{h,k}^s$  and  $v_{h,k}^s$ . Let  $\sigma_{h,k}^\varepsilon$  and  $\gamma_{h,k}^\varepsilon$  be the eigenvalues of  $\Delta_{h,k}^\varepsilon$  and let  $\lambda_{h,k}^\varepsilon = \frac{1}{2}(\sigma_{h,k}^\varepsilon + \gamma_{h,k}^\varepsilon)$ . We write

$$(6.23) \quad \Delta_{h,k}^\varepsilon = \begin{pmatrix} \lambda_{h,k}^\varepsilon + \delta_{h,k}^{\text{cc}} & \delta_{h,k}^{\text{cs}} \\ \delta_{h,k}^{\text{sc}} & \lambda_{h,k}^\varepsilon + \delta_{h,k}^{\text{ss}} \end{pmatrix},$$

so that the trace of  $\delta_{h,k} \equiv \Delta_{h,k}^\varepsilon - \lambda_{h,k}$  vanishes, i.e.,  $\delta_{h,k}^{\text{cc}} + \delta_{h,k}^{\text{ss}} = 0$ . We compute that

$$(6.24) \quad \begin{aligned} \frac{1}{k} \frac{d}{dt} J_{h,k}^\varepsilon(t) &= (\delta_{h,k}^{\text{cc}} - \delta_{h,k}^{\text{ss}}) u_{h,k}^{s,\varepsilon}(t) u_{h,k}^{c,\varepsilon}(t) \\ &\quad + \delta_{h,k}^{\text{cs}} (u_{h,k}^{s,\varepsilon}(t))^2 - \delta_{h,k}^{\text{sc}} (u_{h,k}^{c,\varepsilon}(t))^2. \end{aligned}$$

Choosing the parameterization of the grid projection such that  $|\text{Re}(\gamma_{h,k}^\varepsilon - \sigma_{h,k}^\varepsilon)| = \varepsilon$  and using that  $\delta_{h,k}^{\text{cc}} + \delta_{h,k}^{\text{ss}} = 0$ , we see that at least one of the coefficients  $\delta_{h,k}^{\text{cs}}$ ,  $\delta_{h,k}^{\text{sc}}$ , or  $(\delta_{h,k}^{\text{cc}} - \delta_{h,k}^{\text{ss}})$  is of order  $\varepsilon$ . Choosing  $u_{h,k}^s(0) = v_{h,k}^s(0) = v_{h,k}^c(0) = 0$  we have  $u_{h,k}^c(t) = u_{h,k}^c(0) \cos(\gamma_{h,k}t)$  at  $\varepsilon = 0$ , and inserting this initial condition into (6.24) we find

$$(6.25) \quad \frac{1}{k} \frac{d}{dt} J_{h,k}^\varepsilon(t) = -\delta^{\text{cs}} (u_{h,k}^c(0) \cos(\gamma_{h,k}t))^2 + O(\varepsilon^2).$$

Hence,  $J_{h,k}^\varepsilon(t) - J_{h,k}^\varepsilon(0) = O(1)$  after time  $O(1/\varepsilon)$ . We proceed similarly if  $\delta^{\text{cs}}$  grows quadratically in  $\varepsilon$ , but  $\delta^{\text{sc}}$  or  $(\delta^{\text{cc}} - \delta^{\text{ss}})$  grow linearly in  $\varepsilon$ . In summary, we have proved the following.

**Lemma 6.3** *Assume that the 1:1 resonance on  $E_{h,k}^\varepsilon$  is destroyed, let  $\gamma_{h,k}^\varepsilon$  and  $\sigma_{h,k}^\varepsilon$  be the eigenvalues of  $\Delta_{h,k}^\varepsilon$  and  $|\operatorname{Re}(\gamma_{h,k}^\varepsilon - \sigma_{h,k}^\varepsilon)| = \varepsilon$ . Then the momentum  $J_{h,k}^\varepsilon$  drifts by  $O(k)$  on a time scale of  $O(1/\varepsilon)$  for generic  $O(1)$  initial data in  $E_{h,k}^\varepsilon$ .*

*Remark 6.4* Assume that the grid consists of  $M$  equal blocks of  $N/M$  grid points for some  $M$  dividing  $N$ . Then the discrete Laplacian  $\Delta_h$  is  $\mathbb{Z}_M$ -equivariant where the  $\mathbb{Z}_M$ -symmetry is generated by the shift  $u_{h,j} \mapsto u_{h,j+N/M}$ . By representation theory [7], all eigenspaces of  $\Delta_h$  with a faithful  $\mathbb{Z}_\ell$ -action for some  $\ell \geq 3$  (where  $\ell$  divides  $M$ ) are two-dimensional. This  $\mathbb{Z}_\ell$ -action enforces an interpolating  $S^1$ -symmetry and corresponding conserved momentum maps on these eigenspaces. Since the shift  $u_{h,j} \mapsto u_{h,j+N/M}$  generates a faithful  $\mathbb{Z}_\ell$ -action with  $\ell \geq 3$  on the first  $M$  non-constant eigenmodes of  $\Delta_h$  (for  $N$  odd) the momentum maps  $J_{h,k}$  from (6.19),  $k = 1, \dots, M$ , are conserved. The form (3.17) of  $J_h$  then implies the conservation of the interpolated momentum  $J_h$  on the first  $M$  eigenmodes of  $\Delta_h$ . If the initial value of the continuous system is smooth so that its Fourier modes decay fast, then this implies approximate momentum conservation of the discretization.

Let  $J_h^\varepsilon: \mathcal{Q}_h \times \mathcal{Q}_h \rightarrow \mathbb{R}$  be defined by  $J_h^\varepsilon|_{E_{h,k}^\varepsilon \times E_{h,k}^\varepsilon} = J_{h,k}^\varepsilon$ . Applying Lemma 6.3 and noting that any consistent momentum is  $O(h + \varepsilon)$ -close to  $J_h^\varepsilon$ , we obtain the following (see also Figure 9).

**Theorem 6.4** *Consider a non-uniform discretization of the linear wave equation and let the stepsizes be selected such that the discrete Laplacian  $\Delta_h^\varepsilon$  has at least one pair of simple eigenvalues  $\gamma_{h,k}^\varepsilon \neq \sigma_{h,k}^\varepsilon$ ,  $\gamma_{h,k}^\varepsilon, \sigma_{h,k}^\varepsilon \approx \lambda_{h,k} \neq 0$ , and choose the grid distortion parameter  $\varepsilon$  proportional to the eigenvalue splitting, i.e., such that  $\varepsilon = |\gamma_{h,k}^\varepsilon - \sigma_{h,k}^\varepsilon|$ . Then for generic  $O(1)$  initial data, any consistent discrete momentum drifts by  $O(1)$  on a time scale of  $O(1/\varepsilon)$ .*

## 7 Momentum error bounds

In this section we provide rigorous upper bounds on the momentum drift for the semilinear wave equation under the variational semidiscretization (3.9) on a uniform grid and thereby prove Theorem 1.1. To this end, we first establish well-posedness of both the semilinear wave equation (2.2) and the modified system (5.8) in appropriate spaces of analytic functions. We then estimate the difference between the semidiscrete system and its modified equation. Since the modified system conserves momentum exactly, it follows that the discrete momentum is preserved up to exponentially small error. In this section we need only consider uniform grids, because we have shown in Theorem 6.4 that on non-uniform grids momentum is not even approximately conserved for linear systems.

### 7.1 Spaces of analytic functions

We work in spaces of functions that are analytic on the open strip of radius  $r$  in the complex plane about the real (mod  $2\pi$ ) axis. Such functions are characterized by the exponential decay of their Fourier coefficients (see, for example, the review in [13]). For real numbers  $m \geq 0$  and  $r \geq 0$ , let  $G^{r,m}(S^1, \mathbb{R}^n)$  denote the space of functions for which the norm

$$(7.1) \quad \|u\|_{G^{r,m}}^2 = |\hat{u}_0|^2 + \sum_{k \in \mathbb{Z} \setminus \{0\}} |k|^{2m} e^{2r|k|} |\hat{u}_k|^2$$

is finite. It is a Hilbert space with respect to the inner product

$$(7.2) \quad \langle u, v \rangle_{G^{r,m}} = \overline{\hat{u}_0} \hat{v}_0 + \sum_{k \in \mathbb{Z} \setminus \{0\}} |k|^{2m} e^{2r|k|} \overline{\hat{u}_k} \hat{v}_k.$$

In the main text we assume that the nonlinearity  $f$  is an entire function, and will remark on necessary modifications for non-analytic nonlinearities where appropriate. Let

$$(7.3) \quad f(z) = \sum_{n=0}^{\infty} a_n z^n$$

be the power series representation of  $f$  and let  $\phi: \mathbb{R} \rightarrow \mathbb{R}$  be its majorization

$$(7.4) \quad \phi(z) = \sum_{n=1}^{\infty} |a_n| z^n.$$

As  $f$  is entire, so must be  $\phi$ . Moreover,  $\phi$  is monotonically increasing on the positive real axis. We quote two results from Ferrari and Titi [6].

**Lemma 7.1** *The space  $G^{r,m}(S^1)$  is a topological algebra for every  $r \geq 0$  and  $m > 1/2$ . Specifically, there exists a constant  $c = c(m)$  such that for every  $u, v \in G^{r,m}(S^1)$  the product  $uv \in G^{r,m}(S^1)$  with*

$$(7.5) \quad \|uv\|_{G^{r,m}} \leq c \|u\|_{G^{r,m}} \|v\|_{G^{r,m}}.$$

Applying the algebra inequality to each term of the power series expansion of  $f$ , one immediately obtains the following.

**Lemma 7.2** *Let  $f$  be entire, and let  $\phi$  majorize  $f$  in the sense of (7.4). Assume further that  $u \in G^{r,m}(S^1)$  for  $r \geq 0$  and  $m > 1/2$ . Then  $f(u) \in G^{r,m}(S^1)$ , and*

$$(7.6) \quad \|f(u)\|_{G^{r,m}} \leq \phi(c \|u\|_{G^{r,m}}),$$

where  $c = c(m)$  is as in Lemma 7.1.

We prove the following extension of this lemma.

**Lemma 7.3** *Under the assumptions of Lemma 7.2, let  $u, v \in G^{r,m}(S^1)$ . Then*

$$(7.7) \quad \|f(u) - f(v)\|_{Gr,m} \leq \frac{1}{2} \left( \phi'(c \|u\|_{Gr,m}) + \phi'(c \|v\|_{Gr,m}) \right) \|u - v\|_{Gr,m},$$

where  $c = c(m)$  is as in Lemma 7.1.

*Remark 7.1* We use the letter  $c$  for universal constants, and  $K$  for constants which may depend on the initial data, but are independent of  $h$  and  $t$  at least for an interval of time on which a bounded analytic solution to the semilinear wave equations exists. Numbered constants are unique throughout the paper.

*Proof.* Observe that

$$(7.8) \quad u^n - v^n = (u - v) \sum_{k=0}^{n-1} u^k v^{n-1-k}.$$

Therefore,

$$(7.9) \quad f(u) - f(v) = (u - v) \sum_{n=1}^{\infty} a_n \sum_{k=0}^{n-1} u^k v^{n-1-k},$$

so that

$$(7.10) \quad \begin{aligned} \|f(u) - f(v)\|_{Gr,m} &\leq \|u - v\|_{Gr,m} \sum_{n=1}^{\infty} |a_n| c^{n-1} \sum_{k=0}^{n-1} \|u\|_{Gr,m}^k \|v\|_{Gr,m}^{n-1-k} \\ &\leq \|u - v\|_{Gr,m} \sum_{n=1}^{\infty} |a_n| c^{n-1} \\ &\quad \times \sum_{k=0}^{n-1} \left( \frac{k}{n-1} \|u\|_{Gr,m}^{n-1} + \frac{n-1-k}{n-1} \|v\|_{Gr,m}^{n-1} \right) \\ &= \|u - v\|_{Gr,m} \sum_{n=1}^{\infty} |a_n| c^{n-1} \frac{n}{2} (\|u\|_{Gr,m}^{n-1} + \|v\|_{Gr,m}^{n-1}), \end{aligned}$$

where, in the second step, we have used the Young inequality

$$(7.11) \quad x^\lambda y^{1-\lambda} \leq \lambda x + (1 - \lambda) y$$

for any  $\lambda \in [0, 1]$ . Comparing the last expression in (7.10) with the definition of  $\phi$ , we find (7.7).  $\square$

## 7.2 Functional setting for the wave equation

We rewrite the semilinear wave equation (2.2) in vector notation,

$$(7.12) \quad \partial_t U = AU + B(U),$$

where  $U = (u, v)$ ,

$$(7.13) \quad A = \begin{pmatrix} 0 & \text{id} \\ \Delta & 0 \end{pmatrix}, \quad B(U) = \begin{pmatrix} 0 \\ f(u) \end{pmatrix}.$$

We first collect some elementary facts about the linear operator  $A$ . Since  $\Delta$  is diagonal with respect to the Fourier basis  $e^{ikx}$  and has respective eigenvalues  $\lambda_k = -k^2$ , the restriction of  $A$  to the  $k$ th eigenspace, denoted  $A_k$ , has the Fourier representation

$$(7.14) \quad \hat{A}_k = \begin{pmatrix} 0 & 1 \\ \lambda_k & 0 \end{pmatrix}.$$

When  $\lambda_k \neq 0$ ,  $A_k$  has a pair of distinct complex conjugate eigenvalues  $\pm i\mu_k$  where  $\mu_k = \sqrt{|\lambda_k|} = |k|$ , and  $\exp(A_k t)$  has the Fourier representation

$$(7.15) \quad e^{\hat{A}_k t} = \begin{pmatrix} \cos \mu_k t & \mu_k^{-1} \sin \mu_k t \\ -\mu_k \sin \mu_k t & \cos \mu_k t \end{pmatrix}.$$

For  $k = 0$  we have  $\mu_k = 0$ , so that the Fourier representation of  $A_0$  is a Jordan block, and

$$(7.16) \quad e^{\hat{A}_0 t} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

We introduce the Hilbert space

$$(7.17) \quad Y^{r,m} = G^{r,m+1} \times G^{r,m},$$

where both components are endowed with the standard inner product (7.2). In particular,  $Y^{0,0} = H^1 \times L^2$ . As we will provide a mirror functional setting for the modified equation, it is useful to note that

$$(7.18) \quad G^{r,m+1} = \{u \in G^{r,m} : (-\Delta)^{1/2} u \in G^{r,m}\},$$

where the projection of  $(-\Delta)^{1/2}$  onto the  $k$ th eigenspace has Fourier multiplier  $\mu_k = |k|$ . We then write the inner product as

$$(7.19) \quad \langle u, v \rangle_{G^{r,m+1}} = \langle u, P^0 v \rangle_{G^{r,m}} - \langle u, \Delta v \rangle_{G^{r,m}}.$$

where  $P^0$  denotes the orthogonal projector onto the generalized eigenspace corresponding to the zero eigenvalue of  $\Delta$ . We set  $Q^0 = 1 - P^0$ .

The crucial observation is that the group generated by  $Q^0 A$  is unitary on any  $Y^{r,m}$ :

$$(7.20) \quad \|e^{Q^0 A t} U\|_{Y^{r,m}} = \|U\|_{Y^{r,m}}.$$

Note that the full group  $e^{A t}$  is not unitary because of the secular term from the Jordan block (7.16).

To make sense of the full semilinear wave equation (7.12), we introduce its *mild formulation*

$$(7.21) \quad U(t) = e^{tA} U_0 + \int_0^t e^{(t-s)A} B(U(s)) ds.$$

Since our nonlinearity is Lipschitz in  $G^{r,m}$  by Lemma 7.3, we obtain by direct application of the contraction mapping theorem [24, 11] that the semilinear wave equation is locally well posed.

**Theorem 7.4** *Let  $U_0 \in Y^{r,m}$  with  $r, m \geq 0$ . Then there exists a time  $T$  which depends only on the  $Y^{r,m}$  norm of the initial data and a solution  $U \in C([0, T]; Y^{r,m})$  to the mild formulation (7.21).*

*Remark 7.2* This theorem provides the functional setting in which the formal statements about momentum conservation in Section 2 can be justified. Specifically, if  $U_0 \in Y^{0,0} \equiv H^1 \times L^2$ , the momentum  $J$  is well-defined. We can thus approximate  $U_0$  by a sequence of smooth functions for which the manipulations in (2.7) can be carried out on a uniform interval of time, and finally pass to the limit.

*Remark 7.3* The statement remains true if  $f: S \rightarrow \mathbb{R}$  is analytic on some open convex subset  $S \subseteq \mathbb{R}$  provided the  $u$ -component of the initial data lies in  $S$ . Then  $f$  maps  $S^{r,m}$  to  $G^{r,m}$ , where

$$(7.22) \quad S^{r,m} = \{u \in G^{r,m}(S^1) : u(x) \in S\},$$

as can be seen by Taylor expanding  $f$ . Moreover  $f$  is still a Lipschitz map, since

$$(7.23) \quad \begin{aligned} \|f(u) - f(v)\|_{G^{r,m}} &\leq \int_0^1 \|f'(u + t(v-u))\|_{G^{r,m}} \|u - v\|_{G^{r,m}} dt \\ &\leq c(m) \sup_{t \in [0,1]} \|f'(u + t(v-u))\|_{G^{r,m}} \|u - v\|_{G^{r,m}}, \end{aligned}$$

where we used the algebra inequality (7.5). Note that pointwise evaluation of  $u$  makes sense because under the assumptions of Theorem 7.4  $u$  is at least  $H^1$  and therefore, by the Sobolev embedding theorem, continuous.

*Remark 7.4* If  $f$  is only of class  $C^{m+1}$ , all of the above holds with  $r = 0$ , and we find that the semilinear wave equation is locally well posed on  $Y^{0,m}$ . In particular, for  $r = m = 0$  we recover the well known existence of local weak solutions to the semilinear wave equation in  $Y^{0,0} = H^1 \times L^2$ , see [24].

### 7.3 Functional setting for the modified system

The above construction can be adapted to the modified equation by literally replacing all quantities by their discrete-interpolated counterparts. We write the general modified equation on an equispaced grid as

$$(7.24) \quad \partial_t \tilde{U} = \tilde{A}_h \tilde{U} + B_h(\tilde{U}),$$

where  $\tilde{U} = (\tilde{u}, \tilde{v})$ ,

$$(7.25) \quad \tilde{A}_h = \begin{pmatrix} 0 & \text{id} \\ \tilde{\Delta}_h & 0 \end{pmatrix}, \quad B_h(U) = \begin{pmatrix} 0 \\ (1 - \alpha) f(u) + \alpha f(\tau_h u) \end{pmatrix},$$

and  $\tilde{\Delta}_h$  denotes the discrete Laplacian

$$(7.26) \quad \tilde{\Delta}_h u(x) = \frac{u(x-h) - 2u(x) + u(x+h)}{h^2}.$$

It is clear that  $\tilde{\Delta}_h$  maps  $G^{r,m}(S^1)$  into itself because the translation operator  $\tau_h$  does. Moreover,  $-\tilde{\Delta}_h$  is a non-negative, self-adjoint operator with respect to the  $L^2$  inner product and has a complete orthonormal set of eigenfunctions – the Fourier modes  $e^{ikx}$ . Indeed, since  $(\tau_h u)_k = e^{ikh} \hat{u}_k$ , we have

$$(7.27) \quad \widehat{(\tilde{\Delta}_h u)}_k = \frac{e^{-ikh} - 2 + e^{ikh}}{h^2} \hat{u}_k = \frac{2}{h^2} (\cos(hk) - 1) \hat{u}_k.$$

Following the construction in Section 7.2, we write  $\tilde{P}_h^0$  to denote the spectral projector onto the – now infinite dimensional – generalized zero-eigenspace of  $\tilde{A}_h$ , set  $\tilde{Q}_h^0 = 1 - \tilde{P}_h^0$ , and introduce the Hilbert space

$$(7.28) \quad G_h^{r,m} = \{u \in G^{r,m} : (-\tilde{\Delta}_h)^{1/2} u \in G^{r,m}\}$$

endowed with inner product

$$(7.29) \quad \langle u, v \rangle_{G_h^{r,m}} = \langle u, \tilde{P}_h^0 v \rangle_{G^{r,m}} - \langle u, \tilde{\Delta}_h v \rangle_{G^{r,m}}.$$

When restricted to the grid,  $G_h^{r,m}$  is called the *discrete Gevrey space* [15].

*Remark 7.5* Since the spectrum of  $\tilde{\Delta}_h$  is bounded,  $G_h^{r,m} = G^{r,m}$  and the corresponding norms are equivalent, though not uniformly in  $h$ . In fact, it is easy to check that

$$(7.30) \quad \|u\|_{G^{r,m}} \leq \|u\|_{G_h^{r,m}} \leq \frac{\sqrt{2}}{h} \|u\|_{G^{r,m}}.$$

We now set  $Y_h^{r,m} = G_h^{r,m} \times G^{r,m}$  and note that, as in the fully continuous case, the group generated by  $\tilde{Q}_h^0 \tilde{A}_h$  is unitary on this space:

$$(7.31) \quad \|e^{\tilde{Q}_h^0 \tilde{A}_h t} U\|_{Y_h^{r,m}} = \|U\|_{Y_h^{r,m}}.$$

We immediately obtain the analog of Theorem 7.4.

**Theorem 7.5** *Let  $\tilde{U}_0 \in Y_h^{r,m}$  with  $r \geq 0$  and  $m > 1/2$ . Then there exists a time  $T$  which depends only on the  $Y_h^{r,m}$ -norm of the initial data and a solution  $\tilde{U} \in C([0, T]; Y_h^{r,m})$  to the mild formulation of the modified equation (7.24).*

*Remark 7.6* As a consequence of the above, the semilinear wave equation and the modified systems are well posed on different spaces, a necessary complication we need to overcome when estimating differences between exact and modified solutions.

*Remark 7.7* As for the semilinear wave equation (see Remarks 7.3 and 7.4), Theorem 7.5 remains true if  $f$  is only defined on an open subset  $S$  of  $\mathbb{R}$ , or if  $f$  is only  $C^{m+1}$  and  $r = 0$ .

#### 7.4 Difference between wave equation and modified system

We first estimate the  $G^{r,m}$ -difference between the true solution of the semilinear wave equation and the solution of the modified equation starting from the same initial data.

**Lemma 7.6** *Let  $m > 1/2$  and  $r \geq 0$ . Let  $U(t)$  be a solution of the semilinear wave equation (7.12) and  $\tilde{U}(t)$  be a solution of the modified system (7.24) with  $\tilde{U}(0) = U(0)$ . Assume that  $\|U(t)\|_{G^{r,m+4}} \leq K$  for  $t \in [0, T]$ . Then there exist constants  $K_1 = K_1(K)$ ,  $K_2 = K_2(K)$  and  $h_0(K, T) > 0$  such that*

$$(7.32) \quad \|U(t) - \tilde{U}(t)\|_{G^{r,m}} \leq K_1 (h^2 + \alpha h) (e^{K_2 t} - 1)$$

for  $0 \leq t \leq T$  and  $h \leq h_0(K, T)$ .

*Proof.* Let  $E(t) = U(t) - \tilde{U}(t)$  denote the ‘‘modification error’’, and subtract (7.24) from (7.12) to find

$$(7.33) \quad \begin{aligned} \partial_t E &= AU - \tilde{A}_h \tilde{U} + B(U) - B_h(\tilde{U}) \\ &= (A - \tilde{A}_h)U + \tilde{A}_h E + B(U) - B_h(\tilde{U}), \end{aligned}$$

and therefore

$$(7.34) \quad \begin{aligned} \partial_t (e^{-t\tilde{Q}_h^0 \tilde{A}_h} E) &= e^{-t\tilde{Q}_h^0 \tilde{A}_h} (A - \tilde{A}_h)U + \tilde{P}_h^0 \tilde{A}_h E \\ &\quad + e^{-t\tilde{Q}_h^0 \tilde{A}_h} (B(U) - B_h(\tilde{U})). \end{aligned}$$

Taking the  $Y_h^{r,m}$  inner product with  $\exp(-t\tilde{Q}_h^0 \tilde{A}_h)E$ , we find

$$(7.35) \quad \begin{aligned} \frac{1}{2} \frac{d}{dt} \|e^{-t\tilde{Q}_h^0 \tilde{A}_h} E\|_{Y_h^{r,m}}^2 &= \langle e^{-t\tilde{Q}_h^0 \tilde{A}_h} E, e^{-t\tilde{Q}_h^0 \tilde{A}_h} (A - \tilde{A}_h)U \rangle_{Y_h^{r,m}} \\ &\quad + \langle e^{-t\tilde{Q}_h^0 \tilde{A}_h} E, e^{-t\tilde{Q}_h^0 \tilde{A}_h} (B(U) - B_h(\tilde{U})) \rangle_{Y_h^{r,m}} \\ &\quad + \langle \tilde{P}_h^0 \tilde{A}_h E, e^{-t\tilde{Q}_h^0 \tilde{A}_h} E \rangle_{Y_h^{r,m}}. \end{aligned}$$

According to (7.31),  $\exp(-t \tilde{Q}_h^0 \tilde{A}_h)$  preserves the  $Y_h^{r,m}$  norm. We thus estimate

$$(7.36) \quad \frac{1}{2} \frac{d}{dt} \|E\|_{Y_h^{r,m}}^2 \leq \|E\|_{Y_h^{r,m}} \|(A - \tilde{A}_h)U\|_{Y_h^{r,m}} + \|E\|_{Y_h^{r,m}} \|(B(U) - B_h(\tilde{U}))\|_{Y_h^{r,m}} + \|E\|_{Y_h^{r,m}} \|\tilde{P}_h^0 \tilde{A}_h E\|_{Y_h^{r,m}}.$$

Since  $\|\tilde{P}_h^0 \tilde{A}_h E\|_{Y_h^{r,m}} \leq \|E\|_{Y_h^{r,m}}$ , we obtain

$$(7.37) \quad \frac{d}{dt} \|E\|_{Y_h^{r,m}} \leq \|(A - \tilde{A}_h)U\|_{Y_h^{r,m}} + \|B(U) - B_h(\tilde{U})\|_{Y_h^{r,m}} + \|E\|_{Y_h^{r,m}}.$$

The first term on the right is the local discretization error. It can be estimated straightforwardly by writing out a Taylor expansion with integral remainder for

$$(7.38) \quad \begin{aligned} \tilde{\Delta}_h u(x) - \Delta u(x) &= \frac{1}{h^2} \frac{1}{3!} \int_0^h u^{(4)}(x + \xi) (h - \xi)^3 d\xi \\ &+ \frac{1}{h^2} \frac{1}{3!} \int_0^{-h} u^{(4)}(x + \xi) (h + \xi)^3 d\xi, \end{aligned}$$

and taking its  $G^{r,m}$  norm, so that

$$(7.39) \quad \|(A - \tilde{A}_h)U\|_{Y_h^{r,m}} = \|\tilde{\Delta}_h u - \Delta u\|_{G^{r,m}} \leq \frac{2}{4!} h^2 \|u^{(4)}\|_{G^{r,m}} \leq h^2 K_3,$$

where  $K_3 = 2K/4!$ . The last term on the right of (7.37) is estimated using Lemma 7.3 and the monotonicity of  $\phi$ : We have

$$(7.40) \quad \begin{aligned} \|f(u) - f(\tilde{u})\|_{G^{r,m}} &\leq \frac{1}{2} (\phi(c \|u\|_{G^{r,m}}) + \phi(c \|\tilde{u}\|_{G^{r,m}})) \|u - \tilde{u}\|_{G^{r,m}} \\ &\leq \frac{1}{2} (\phi(c \|U\|_{G^{r,m}}) \\ &\quad + \phi(c \|U\|_{G^{r,m}} + c \|E\|_{G^{r,m}})) \|E\|_{G^{r,m}} \\ &\leq \phi(2cK) \|E\|_{G^{r,m}} \end{aligned}$$

for as long as  $\|E\|_{G^{r,m}} \leq K$ . Thus,

$$(7.41) \quad \begin{aligned} \|B(U) - B_h(\tilde{U})\|_{G^{r,m}} &\leq (1 - \alpha) \|f(u) - f(\tilde{u})\|_{G^{r,m}} \\ &\quad + \alpha \|f(u) - f(\tau_h \tilde{u})\|_{G^{r,m}} \\ &\leq \|f(u) - f(\tilde{u})\|_{G^{r,m}} + \alpha \|f(\tilde{u}) - f(\tau_h \tilde{u})\|_{G^{r,m}} \\ &\leq \phi(2cK) \|E\|_{G^{r,m}} + \alpha \phi(2cK) h \|u^{(1)}\|_{G^{r,m}} \\ &\leq \phi(2cK) \|E\|_{Y_h^{r,m}} + \alpha \phi(2cK) K h. \end{aligned}$$

Altogether, the differential inequality (7.37) reads

$$(7.42) \quad \frac{d}{dt} \|E\|_{Y_h^{r,m}} \leq K_3 h^2 + K_4 \alpha h + K_2 \|E\|_{Y_h^{r,m}},$$

where  $K_2 = 1 + \phi(2cK)$  and  $K_4 = K \phi(2cK)$ . Since  $E(0) = 0$ , integration in time yields

$$(7.43) \quad \|E\|_{Y_h^{r,m}} \leq \frac{K_3 h^2 + \alpha K_4 h}{K_2} (e^{K_2 t} - 1) \leq K_1 (h^2 + \alpha h) (e^{K_2 t} - 1)$$

with some appropriately chosen constant  $K_1$ . Provided we choose  $h$  small enough that

$$(7.44) \quad K_1 (h^2 + \alpha h) (e^{K_2 T} - 1) \leq K,$$

the estimate remains consistent with our *a priori* requirement  $\|E\|_{G^{r,m}} \leq K$ . This completes the proof.  $\square$

*Remark 7.8* An estimate similar to Lemma 7.6 holds for the difference between exact and numerical solution – they are  $O(h)$  close when  $\alpha \neq 0$  and  $O(h^2)$  close when  $\alpha = 0$ : Taking  $r = 0$  and  $m = 1$  in (7.32) and using the Sobolev embedding theorem  $\|u\|_{C^0} \leq c \|u\|_{H^1}$ , we estimate the discretization error

$$(7.45) \quad \|\pi_h U(t) - U_h(t)\|_{\mathbb{R}^{2N}} \leq c K_1 (h^2 + \alpha h) (e^{K_2 t} - 1)$$

for  $0 \leq t \leq T$  and  $h \leq h_0(K, T)$ . On the other hand, the difference between modified equation and numerical solution is much smaller – in fact exponentially small in the  $H^1$  norm for as long as the exact solution remains analytic. This will be proved in the following.

*Remark 7.9* Estimate (7.45) by no means implies that every invariant set of the discrete system has a continuous counterpart. For example, in [12] discrete traveling waves of the space-discretization are constructed which do not exist for the continuous system.

*Remark 7.10* Lemma 7.6 also holds if  $f$  is only defined on some open convex subset  $S$  of  $\mathbb{R}$ . The proof has to be modified as follows: (7.40) has to be replaced by

$$(7.46) \quad \|f(u) - f(\tilde{u})\|_{G^{r,m}} \leq K(r, m) \|E\|_{G^{r,m}}$$

where we use an estimate similar to (7.23) and

$$(7.47) \quad K(r, m) = c(m) \sup_{(u,v) \in \mathcal{K}^{r,m}} \|f'(u)\|_{G^{r,m}}.$$

Here  $\mathcal{K}^{r,m}$  is a bounded closed subset of  $\mathcal{S}^{r,m}$ , with  $\mathcal{S}^{r,m}$  the same as in (7.22), containing a tube of radius  $K_S > 0$  around  $\{U(t) : 0 \leq t \leq T\}$ . Such a set exists, because  $K_S$  satisfies  $\text{dist}(U(t), \partial \mathcal{S}^{r,m}) \geq K_S$ , and, since  $\{U(t) : 0 \leq t \leq T\}$  is closed and  $\mathcal{S}^{r,m}$  is an open subset of  $G^{r,m}$ , by the openness of  $S$  in  $\mathbb{R}$  we can always find such  $K_S > 0$  and consequently

also a set  $\mathcal{K}^{r,m}$  with the required properties. Moreover the constant  $K$  with  $\|E(t)\|_{G^{r,m}} \leq K$  which first appears in (7.40) has to be replaced by  $K_S$  and the constant  $\phi(2cK)$  appearing in (7.41) by  $K(r, m)$ . In this case we have  $K_2 = 1 + K(r, m)$  and  $K_4 = KK(r, m)$  in (7.42).

Likewise, Lemma 7.6 holds with  $r = 0$  if  $f$  is only  $C^{m+1}$ ,  $m \in \mathbb{N}$ .

### 7.5 Difference between discrete and modified system

We need the following version of the Shannon Sampling Theorem.

**Lemma 7.7** *Let  $u \in G^{r,m}$  with  $r \geq 0$  and  $m > 1/2$ . Then*

$$(7.48) \quad \|(1 - i_h \circ \pi_h)u\|_{H^1} \leq c_1(m)h^{m-3/2} \|u\|_{H^m}.$$

Moreover, if  $r > 0$ , then

$$(7.49) \quad \|(1 - i_h \circ \pi_h)u\|_{H^1} \leq c_2(r, m) h^{m-1} e^{-\pi r/h} \|u\|_{G^{r,m}},$$

for appropriate constants  $c_1(m)$  and  $c_2(r, m)$ .

*Proof.* By definition, the Fourier coefficients of the trigonometric interpolant of the grid projection  $i_h \circ \pi_h(u)$  are  $(i_h \circ \pi_h u)_k = \hat{u}_{h,k}$ . Recall the aliasing formula

$$(7.50) \quad \hat{u}_k - \hat{u}_{h,k} = \sum_{\substack{\ell \in \mathbb{Z} \\ \ell \neq 0}} \hat{u}_{k+\ell N}$$

which follows from

$$(7.51) \quad \begin{aligned} \hat{u}_{h,k} &= \frac{h}{\sqrt{2\pi}} \sum_{j=0}^{N-1} e^{-ikx_j} u(x_j) = \frac{1}{N} \sum_{j=0}^{N-1} e^{-ihjk} \sum_{\ell \in \mathbb{Z}} \hat{u}_\ell e^{ihj\ell} \\ &= \sum_{\ell \in \mathbb{Z}} \hat{u}_\ell \frac{1}{N} \sum_{j=0}^{N-1} e^{2\pi i(\ell-k)j/N} = \sum_{\ell \in \mathbb{Z}} \hat{u}_{k+\ell N}. \end{aligned}$$

Noting that  $\sum a_i^2 \leq (\sum a_i)^2$  when all the coefficients  $a_i$  are positive, we write

$$\begin{aligned} \|u - (i_h \circ \pi_h)u\|_{H^1}^2 &\leq \sum_{k=-[(N-1)/2]}^{[N/2]} (1 + |k|^2) |\hat{u}_{h,k} - \hat{u}_k|^2 \\ &\quad + \left( \sum_{k < -[(N-1)/2]} + \sum_{k \geq [N/2]} \right) (1 + |k|^2) |\hat{u}_k|^2 \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{k=-\lfloor(N-1)/2\rfloor}^{\lfloor N/2\rfloor} (1 + |k|^2) \left( \sum_{\substack{\ell \in \mathbb{Z} \\ \ell \neq 0}} |\hat{u}_{k+\ell N}| \right)^2 \\
&\quad + \sum_{|k| \geq \lfloor(N-1)/2\rfloor} (1 + |k|^2) |\hat{u}_k|^2 \\
(7.52) \quad &\leq \left(1 + \frac{N^2}{4}\right) \left( \sum_{|k| \geq n} |\hat{u}_k| \right)^2 + \sum_{|k| \geq n} (1 + |k|^2) |\hat{u}_k|^2,
\end{aligned}$$

where  $n = \lfloor(N-1)/2\rfloor$ . The second term on the right is the projection error, which has the straightforward upper bound

$$\sum_{\substack{|k| \geq n \\ (7.53)}} (1 + |k|^2) |\hat{u}_k|^2 \leq n^{-2(m-1)} e^{-2rn} \sum_{|k| \geq n} |k|^{2(m-1)} (1 + |k|^2) e^{2r|k|} |\hat{u}_k|^2.$$

Noting that  $h = 2\pi/N$ , we obtain a term of the form which can be estimated by the term on the right of (7.48) or (7.49) for  $r = 0$  and  $r > 0$ , respectively.

The first term on the right of (7.52) is the aliasing error. When  $r = 0$ , we estimate

$$\begin{aligned}
(7.54) \quad \left( \sum_{|k| \geq n} |\hat{u}_k| \right)^2 &= \left( \sum_{|k| \geq n} |k|^{-m} |k|^m |\hat{u}_k| \right)^2 \\
&\leq \sum_{|k| \geq n} |k|^{-2m} \sum_{|k| \geq n} |k|^{2m} |\hat{u}_k|^2 \\
&\leq \frac{2}{2m-1} (n-1)^{1-2m} \|u\|_{H^m}^2,
\end{aligned}$$

thereby obtaining the second term on the right of (7.48). The first inequality in (7.54) is a direct application of the Cauchy–Schwarz inequality, and the second inequality follows from an integral upper bound on the left sum. When  $r > 0$ , a similar argument yields

$$\begin{aligned}
(7.55) \quad \left( \sum_{|k| \geq n} |\hat{u}_k| \right)^2 &\leq \sum_{|k| \geq n} |k|^{-2m} e^{-2r|k|} \sum_{|k| \geq n} |k|^{2m} e^{2r|k|} |\hat{u}_k|^2 \\
&\leq 2n^{-2m} \int_{n-1}^{\infty} e^{-2r\kappa} d\kappa \|u\|_{G^{r,m}}^2 \\
&\leq n^{-2m} \frac{e^{2r(1-n)}}{r} \|u\|_{G^{r,m}}^2.
\end{aligned}$$

Substituting this expression back into (7.52), we obtain (7.49).  $\square$

**Lemma 7.8** *Let  $r > 0$  and let  $\tilde{U}$  denote the solution to the modified equation with initial data  $\tilde{U}(0)$ , and  $U_h$  the numerical solution with initial data  $U_h(0) = \pi_h \circ \tilde{U}(0)$ . Assume that  $\|\tilde{U}(t)\|_{G^{r,m}} \leq \tilde{K}$  for  $t \in [0, T]$ . Then there exists a constant  $K_5 = K_5(\tilde{K}, r)$  such that*

$$(7.56) \quad \|\tilde{U}(t) - i_h \circ U_h(t)\|_{H^1} \leq K_5 h^{m-1} e^{-\pi r/h},$$

for  $t \in [0, T]$ .

*Proof.* The modified solution is identical to the numerical solution on grid points. Hence,

$$(7.57) \quad \tilde{U} - i_h \circ U_h = (1 - i_h \circ \pi_h)\tilde{U} + i_h \circ (\pi_h \circ \tilde{U} - U_h) = (1 - i_h \circ \pi_h)\tilde{U}.$$

We use (7.49) to bound this expression, noting that  $\|\tilde{U}(t)\|_{G^{r,m}} \leq \tilde{K}$ .  $\square$

Note that there is no accumulation of error with time for as long as the modified solution remains in some bounded set of  $G^{r,m}$ . The only source of error in (7.56) is the exponentially small local interpolation error.

### 7.6 Approximate momentum conservation

We now use the bound on the distance between the semidiscrete and modified system to prove that the semidiscrete momentum will also remain close to the momentum of the modified system. Since the latter is constant in the variational case, i.e., when  $\alpha = 0$ , the corresponding semidiscrete system has exponentially small momentum drift.

**Theorem 7.9** *Let  $\tilde{U}$  denote the solution to the modified equation with initial data  $\tilde{U}(0)$ , and  $U_h$  denote the numerical solution with corresponding discrete initial data  $U_h(0) = \pi_h \circ \tilde{U}(0)$ . Assume that  $\|\tilde{U}(t)\|_{G^{r,m}} \leq \tilde{K}$  with  $m > 1/2$  and  $r > 0$  on some time interval  $[0, T]$ . Then*

$$(7.58) \quad |J \circ i_h \circ U_h(t) - J \circ \tilde{U}(t)| \leq K_6 h^{m-1} e^{-\pi r/h},$$

for  $t \in [0, T]$  and  $K_6 = K_6(\tilde{K})$ . When  $\alpha = 0$ , i.e., when the method is variational, this implies

$$(7.59) \quad |J \circ i_h \circ U_h(t) - J \circ U(0)| \leq K_6 h^{m-1} e^{-\pi r/h}.$$

*Proof.* Recall that  $J_h = J \circ i_h$  and use the mean value theorem to estimate

$$(7.60) \quad |J \circ \tilde{U} - J \circ (i_h \circ U_h)| \leq \sup_{\|V\|_{H^1} \leq 2\tilde{K}} \|DJ(V)\|_{\mathcal{L}(H^1, \mathbb{R})} \|\tilde{U} - i_h \circ U_h\|_{H^1} \leq K_6 h^{m-1} e^{-\pi r/h},$$

where the exponential error term comes from Lemma 7.8,  $DJ$  denotes the Fréchet derivative of  $J$ ,  $h$  is chosen small enough that  $\|\tilde{U} - i_h \circ U_h\|_{H^1} \leq \tilde{K}$ , and

$$(7.61) \quad |DJ(U)\tilde{U}| = \left| \int_{S^1} (v \partial_x \tilde{u} + \tilde{v} \partial_x u) dx \right| \leq \|U\|_{H^1} \|\tilde{U}\|_{H^1},$$

with  $U = (u, v)$  and  $\tilde{U} = (\tilde{u}, \tilde{v})$ . Since the momentum is conserved by the modified system for  $\alpha = 0$ ,  $J \circ \tilde{U}(t) = J \circ U(0)$ , and we obtain (7.59).  $\square$

*Proof of Theorem 1.1* The conclusions of Lemma 7.8 and Theorem 7.9 remain valid so long as  $U(t)$  has a  $G^{r,m+4}$  bound of the form  $\|U(t)\|_{G^{r,m+4}} \leq K$ , say. Then we know by Lemma 7.6 that  $U$  and  $\tilde{U}$  remain  $G^{r,m}$  close for some interval of time on which, in particular,  $\|\tilde{U}(t)\|_{G^{r,m}} \leq \tilde{K}$ . We have thus proved Theorem 1.1 with  $\kappa(u) = K_6$ .  $\square$

*Remark 7.11* We have seen in (5.15) that the modified system for the non-Lagrangian discretization on the uniform grid does not conserve momentum. The general momentum error estimate (7.58) thus implies that the non-Lagrangian semidiscrete system cannot approximately conserve momentum.

*Remark 7.12* Note that the closeness of the continuous solution and the solution of the modified system can only be guaranteed on finite time intervals with errors that grow exponentially in  $T$ , see Lemma 7.6. Consequently, if the norm of the continuous solution is used to establish the required bound on the solution of the modified system in Theorem 7.9 then the stepsizes  $h \leq h_0(K, T)$  must decrease exponentially with  $T$  according to (7.44).

*Remark 7.13* When  $r = 0$ , we need to estimate the sampling error using inequality (7.48). The statements of Lemma 7.8 and Theorem 7.9 remain true – with different constants – provided we replace  $h^{m-1}$  by  $h^{m-3/2}$ . Likewise Theorem 1.1 remains true with  $e^{-\pi r/h}$  replaced by  $h^{m-3/2}$ . Moreover Theorems 7.9 and 1.1 also apply if  $f$  is only defined on an open convex subset of  $\mathbb{R}$  or if  $f$  is only  $C^{m+1}$ ,  $m \in \mathbb{N}$ , and  $r = 0$ , see Remarks 7.3, 7.4, 7.7 and 7.10.

*Remark 7.14* The estimate in Theorem 7.9 can be re-written as

$$(7.62) \quad |J_h(U_h(t)) - J_h(U_h(0))| \leq K_6 h^{m-1} e^{-\pi r/h},$$

provided the solution  $U(t)$  of the semilinear wave equation with  $U(0) = i_h \circ U_h(0)$  satisfies  $\|U(t)\|_{G^{r,m+4}} \leq K$ . Now let  $\tilde{J}_h$  be another momentum discretization, consistent in the sense that

$$(7.63) \quad |J(U) - \tilde{J}_h(\pi_h U)| \leq c h \|U\|_{H^m}$$

for some  $m \in \mathbb{N}$ . When  $m = 3$  so that  $u \in H^3 \subseteq C^2$  we could, for example, take the first order or symmetric finite difference momenta from Section 6.2. Then  $\tilde{J}_h$  will remain within some bounded distance of the interpolated momentum  $J_h$  so long as  $\|U(t)\|_{G^{r,m+4}} \leq K$ , and will therefore be approximately conserved over that time interval. More specifically, by the triangle inequality, we have the estimate

$$(7.64) \quad |\tilde{J}_h(U_h(t)) - J_h(U_h(0))| \leq K_6 h^{m-1} e^{-\pi r/h} + 2ch\tilde{K},$$

with  $\|\tilde{U}(t)\|_{H^m} \leq \tilde{K}$ .

## 8 Concluding remarks

We have proved that for analytic initial values and Lagrangian uniform space discretizations of semilinear wave equations momentum is conserved up to an exponentially small error for as long as the modified solution remains reasonably bounded, a condition which can be verified for small stepsizes over finite time intervals if bounds on the continuous solution are given. Moreover, we have shown that for non-Lagrangian or non-uniform discretizations there is no approximate momentum conservation.

As mentioned in the introduction, our results are similar in spirit to the approximate conservation of energy of uniform stepsize Lagrangian integrators of ordinary differential equations. Here the approximate energy conservation is also destroyed by non-uniform stepsize selection [30] or non-symplectic integrators [29]. Both results, these and those in this paper, hold for uniform Lagrangian discretizations as long as the modified solution remains reasonably bounded. In both cases the norm of the modified solution can be estimated over finite time intervals via the norm of the exact solution. If this boundedness assumption is satisfied, however, then the approximate momentum conservation holds for all times, not only on exponentially long time scales as does the approximate energy conservation for Hamiltonian ODEs.

Indeed, let  $\Psi_h$  be a one-step method for an analytic autonomous Hamiltonian ODE and let  $\tilde{\Phi}_t^h$  be the flow of an  $h$ -dependent autonomous Hamiltonian vector field with Hamiltonian  $H_h$  approximating the dynamics of the one-step method. Then we have  $\|\tilde{\Phi}_t^h(U_0) - \Psi_h(U_0)\| \leq K_1 \exp(-K_2/h)$  for  $U_0$  in some bounded domain  $S$  of phase space [22, 1, 9]. For the Hamiltonian  $H_h$  of the Hamiltonian flow  $\tilde{\Phi}_t^h$  we obtain  $|H_h(\Psi_h(U_0)) - H_h(\tilde{\Phi}_t^h(U_0))| = K_3 \exp(-K_2/h)$  for  $U_0 \in S$ . Thus  $H_h$  is conserved over exponentially long times provided the iterates of  $\Psi_h(U_0)$  and the modified solution  $\tilde{\Phi}_t^h(U_0)$  stay in  $S$ .

We expect that the results in this paper can be extended to general semilinear Hamiltonian partial differential equations in arbitrary space dimensions,

general finite-difference discretizations, time-semidiscretizations (see [31]), and full space-time discretizations. Moreover, similar methods should be applicable to multisymplectic conservation laws for energy and momentum, as have been formally derived in [21]. However, note that our modified system (5.8) does not satisfy a local conservation law for the momentum density  $j(u, v) = u_x v$  of the conserved momentum  $J$ . Moreover, the interpolated momentum  $J_h = J \circ i_h$  is nonlocal in the grid variables. In future work we will construct discrete momentum maps with local densities for which approximate local momentum conservation laws hold on finite time intervals.

*Acknowledgements.* We thank Christian Lubich, Brian Moore, Sebastian Reich, Mark Roberts, and Steve Shkoller for interesting and helpful discussions, Mark Roberts in particular for inviting us to the Warwick Symposium on Geometric Mechanics and Symmetry during which this project was started. CW was employed by the University of Warwick and on leave from the Freie Universität Berlin during the time when this paper was written and acknowledges support and hospitality from both those institutions. The research of CW was supported by a European Union Marie Curie fellowship under contract number HPMF-CT-2000-00542, and MO acknowledges partial funding through SFB 386 at Tübingen University.

## References

1. Benettin, G., Giorgilli, A.: On the Hamiltonian interpolation of near-to-the-identity symplectic mappings with application to symplectic integration algorithms. *J. Statist. Phys.* **74**, 1117–1143 (1994)
2. Bridges, T., Reich, S.: Multi-symplectic intergrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity. *Phys. Lett. A* **284**, 184–193 (2001)
3. Cottet, G.H.: A new approach for the analysis of vortex methods in two and three dimensions. *Ann. Analysis Nonlineaire Henri Poinc.* **5**, 227–285 (1988)
4. de Frutos, J., Sanz-Serna, J.M.: Accuracy and conservation properties in numerical integration: the case of the Korteweg-de Vries equation. *Numer. Math.* **75**, 421–445 (1997)
5. Deuffhard, P., Hohmann, A.: *Numerical Analysis in Modern Scientific Computing: An Introduction*. Text in Applied Mathematics, Vol. **43**, 2nd edition, Springer 2003
6. Ferrari, A.B., Titi, E.S.: Gevrey regularity for nonlinear analytic parabolic equations. *Commun. Partial Diff. Eqs.* **23**, 1–16 (1998)
7. Golubitsky, M., Schaeffer, D., Stewart, I.: *Singularities and Groups in Bifurcation Theory: Vol. II*. Applied Mathematical Sciences **69**, Springer, New York, 1988
8. Hairer, E.: Backward analysis of numerical integrators and symplectic methods. *Ann. Numer. Math.* **1**, 107–132 (1994)
9. Hairer, E., Lubich, C.: The life-span of backward error analysis for numerical integrators. *Numer. Math.* **76**(4), 441–462 (1997)
10. Hairer, E., Lubich, C., Wanner, G.: *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations*. Springer, Berlin, 2002
11. Henry, D.: *Geometric Theory of Semilinear Parabolic Equations*. Lecture Notes in Mathematics **840**, Springer, New York, 1983
12. Iooss, G., Kirchgässner, K.: Travelling waves in a chain of coupled nonlinear oscillators. *Commun. Math. Phys.* **211**, 439–464 (2000)

13. Levermore, C.D., Oliver, M.: The Complex Ginzburg-Landau Equation as a Model Problem. In: Lectures in Applied Mathematics **31**, AMS, Providence, Rhode Island, 1996, pp. 141–190
14. Lew, A., Marsden, J.E., Ortiz, M., West, M.: Asynchronous variational integrators. *Arch. Rat. Mech. Anal.* (to appear), 2002
15. Lord, G.J., Stuart, A.M.: Discrete Gevrey regularity, attractors and upper-semicontinuity for a finite difference approximation to the Ginzburg-Landau equation. *Numer. Funct. Anal. Optimiz.* **16**, 1003–1047 (1995)
16. Marsden, J.E., Ratiu, T.S.: *Introduction to Mechanics and Symmetry*. Springer, New York, Berlin, Heidelberg, 1994
17. Marsden, J.E., Patrick, G., Shkoller, S.: Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Commun. Math. Phys.* **199**, 351–395 (1998)
18. Marsden, J.E., West, M.: Discrete mechanics and variational integrators. *Acta Num.* **10**, 357–514 (2001)
19. Matthies, K.: Backward error analysis of a full discretisation scheme for a class of parabolic partial differential equations. *Nonlinear Anal. Theory Methods Appl.* **52**, 805–826 (2003)
20. Matthies, K., Scheel, A.: Exponential averaging of Hamiltonian evolution equations. *Trans. AMS.* To appear
21. Moore, B., Reich, S.: Backward error analysis for Hamiltonian PDEs with applications to nonlinear wave equations. *Numer. Math.* To appear
22. Neishtadt, A.I.: On the separation of motions in systems with rapidly rotating phase. *J. Appl. Math. Mech.* **48**, 134–139 (1984)
23. Oliver, M., Shkoller, S.: The vortex blob method as a second-grade non-Newtonian fluid. *Commun. Partial Diff. Eqs.* **26**, 295–314 (2001)
24. Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Applied Mathematical Sciences **44**, Springer, New York, 1983
25. Pekarsky, S., West, M.: Discrete multisymplectic reduction and discrete Kelvin’s theorem for ideal fluids. In preparation
26. Reed, M.C.: Higher order estimates and smoothness of nonlinear wave equations. *Proc. Am. Math. Soc.* **51**, 79–85 (1975)
27. Reich, S.: Backward error analysis for numerical integrators. *SIAM J. Numer. Anal.* **36**(5), 1549–1570 (1999)
28. Reich, S.: Multi-symplectic Runge-Kutta collocation methods for Hamiltonian wave equations. *J. Com. Phys.* **156**, 1–27 (1999)
29. Sanz-Serna, J.: Symplectic integrators for Hamiltonian problems: An overview. *Acta Numerica* **1**, 243–286 (1992)
30. Skeel, R.D., Gear, C.W.: Does variable step size ruin a symplectic integrator? *Physica D* **60**, 311–313 (1992)
31. Wulff, C.: Approximate energy conservation for time semidiscretizations of semilinear wave equations. In preparation