

Spectral semi-discretisations of weakly nonlinear wave equations over long times

E. HAIRER¹ and CH. LUBICH²

April 18, 2007

¹ *Dept. de Mathématiques, Univ. de Genève, CH-1211 Genève 4, Switzerland.
email: Ernst.Hairer@math.unige.ch*

² *Mathematisches Institut, Univ. Tübingen, D-72076 Tübingen, Germany.
email: Lubich@na.uni-tuebingen.de*

Abstract

The long-time behaviour of spectral semi-discretisations of weakly nonlinear wave equations is analysed. It is shown that the harmonic actions are approximately conserved also for the semi-discretised system. This permits to prove that the energy of the wave equation along the interpolated semi-discrete solution remains well conserved over long times and close to the Hamiltonian of the semi-discrete equation. Although the momentum is no longer an exact invariant of the semi-discretisation, it is shown to be approximately conserved. All these results are obtained with the technique of modulated Fourier expansions.

1 Introduction

This paper is concerned with the long-time behaviour of spectral semi-discretisations of the one-dimensional nonlinear wave equation

$$u_{tt} - u_{xx} + \rho u + g(u) = 0 \tag{1}$$

for $t > 0$ and $-\pi \leq x \leq \pi$ subject to periodic boundary conditions. We assume $\rho > 0$ and a nonlinearity g that is a smooth real function with $g(0) = g'(0) = 0$. We consider small initial data: in appropriate Sobolev norms, the initial values $u(\cdot, 0)$ and $u_t(\cdot, 0)$ are bounded by a small parameter ε .

The near-conservation of actions and long-time regularity of exact solutions to the wave equation (1) have been studied by Bambusi [1] and Bourgain [2], and more recently in our paper [5]. There, we use the technique of modulated Fourier expansions to prove the almost-conservation properties. This approach is also chosen in the present paper on spatial semi-discretisations of (1) and in [4] for full discretisations.

In Section 2, we review the known results on the near-conservation of harmonic actions along exact solutions of (1). Section 3 describes spectral semi-discretisation in space and formulates the main result on the near-conservation of actions (and spatial regularity) along solutions of the semi-discrete equations over long times $t \leq \varepsilon^{-N}$ for any fixed $N \geq 1$. This holds under the same non-resonance condition as for the corresponding result for the wave equation.

As a consequence of this result, we further show that the continuous energy of the trigonometric polynomial determined by the semi-discretisation is well conserved and remains close to the discrete energy of the semi-discrete equations over long times. The exact solution conserves momentum, as a consequence of the shift invariance $x \rightarrow x + \xi$. There is no such invariance under a continuous group action in the semi-discretisation, and indeed momentum is not conserved. We will show, however, that momentum is approximately conserved. The proofs are given in Sections 4 to 6. Following [5] we study the modulated Fourier expansion in time of the semi-discretisation in Section 4 and its almost-invariants in Section 5. Conservation of energy and momentum are shown in Section 6.

Approximate momentum conservation for spatial semi-discretisations of semi-linear wave equations has previously been studied by Oliver, West and Wulff [7], for finite-difference discretisations on regular grids. They show almost-conservation with high accuracy of a modified momentum over short times. Their results do not extend to long times, however, because the regularity of solutions to modified equations is not under control. Another approach to almost-conservation properties of spatial (and full) discretisations of semi-linear wave equations within the framework of standard backward error analysis and modified equations has been given by Cano [3], where likewise the extension to long times rests on unverified regularity assumptions, which are formulated as conjectures.

2 The nonlinear wave equation with small data

Equation (1) has several conserved quantities. The *total energy* or Hamiltonian, defined for 2π -periodic functions u, v as

$$H(u, v) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{1}{2} (v^2 + (\partial_x u)^2 + \rho u^2)(x) + U(u(x)) \right) dx, \quad (2)$$

where the potential $U(u)$ is such that $U'(u) = g(u)$, and the *momentum*

$$K(u, v) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_x u(x) v(x) dx = - \sum_{j=-\infty}^{\infty} i j u_{-j} v_j \quad (3)$$

are exactly conserved along the solution $(u(\cdot, t), \partial_t u(\cdot, t))$ of (1). Here, $u_j = \mathcal{F}_j u$ are the Fourier coefficients in the series $u(x) = \sum_{j=-\infty}^{\infty} u_j e^{ijx}$. Since we consider only real solutions, we note that $u_{-j} = \overline{u_j}$. In terms of the Fourier coefficients, equation (1) reads

$$\partial_t^2 u_j + \omega_j^2 u_j + \mathcal{F}_j g(u) = 0, \quad j \in \mathbb{Z}, \quad (4)$$

with the frequencies

$$\omega_j = \sqrt{\rho + j^2}.$$

The *harmonic actions*

$$I_j(u, v) = \frac{\omega_j}{2} |u_j|^2 + \frac{1}{2\omega_j} |v_j|^2 \quad (5)$$

(note $I_{-j} = I_j$) are conserved for the linear wave equation ($g(u) \equiv 0$). In (1), they turn out to remain constant up to small deviations over long times for

almost all values of $\rho > 0$, when the initial functions are close to the equilibrium $u = 0$. Such a result is proved in Bambusi [1], Bourgain [2], and Cohen, Hairer, and Lubich [5]. We now give a precise statement of this result.

We consider the Sobolev space, for $s \geq 0$,

$$H^s = \{v \in L^2(\mathbb{T}) : \|v\|_s < \infty\}, \quad \|v\|_s = \left(\sum_{j=-\infty}^{\infty} \omega_j^{2s} |v_j|^2 \right)^{1/2},$$

where v_j denote the Fourier coefficients of a 2π -periodic function v . We assume that the initial position and velocity have small norms in H^{s+1} and H^s for suitably large s :

$$\left(\|u(\cdot, 0)\|_{s+1}^2 + \|\partial_t u(\cdot, 0)\|_s^2 \right)^{1/2} \leq \varepsilon. \quad (6)$$

This is equivalent to requiring $\sum_{j=-\infty}^{\infty} \omega_j^{2s+1} I_j(u(\cdot, 0), \partial_t u(\cdot, 0)) \leq \frac{1}{2} \varepsilon^2$.

To prepare for the formulation of a non-resonance condition, we consider sequences $\mathbf{k} = (k_\ell)_{\ell=0}^{\infty}$ with only finitely many integers $k_\ell \neq 0$. We denote $|\mathbf{k}| = (|k_\ell|)_{\ell=0}^{\infty}$, and we let

$$\|\mathbf{k}\| = \sum_{\ell=0}^{\infty} |k_\ell|, \quad \mathbf{k} \cdot \boldsymbol{\omega} = \sum_{\ell=0}^{\infty} k_\ell \omega_\ell, \quad \boldsymbol{\omega}^{\sigma|\mathbf{k}|} = \prod_{\ell=0}^{\infty} \omega_\ell^{\sigma|k_\ell|} \quad (7)$$

for real σ , where we use the notation $\boldsymbol{\omega} = (\omega_\ell)_{\ell=0}^{\infty}$. In particular, for $j \in \mathbb{Z}$, we write $\langle j \rangle = (0, \dots, 0, 1, 0, \dots)$ with the only entry at the $|j|$ -th position.

For a fixed integer N and for $\varepsilon > 0$, we consider the set of near-resonant indices

$$\mathcal{R}_\varepsilon = \{(j, \mathbf{k}) : j \in \mathbb{Z} \text{ and } \mathbf{k} \neq \pm \langle j \rangle, \|\mathbf{k}\| \leq 2N \text{ with } |\omega_j - |\mathbf{k} \cdot \boldsymbol{\omega}|| < \varepsilon^{1/2}\}. \quad (8)$$

We impose the following *non-resonance condition*: there are $\sigma > 0$ and a constant C_0 such that

$$\sup_{(j, \mathbf{k}) \in \mathcal{R}_\varepsilon} \frac{\omega_j^\sigma}{\boldsymbol{\omega}^{\sigma|\mathbf{k}|}} \varepsilon^{\|\mathbf{k}\|/2} \leq C_0 \varepsilon^N. \quad (9)$$

As is shown in [5], condition (9) is implied, for sufficiently large σ , by the non-resonance condition of Bambusi [1], which reads as follows: for every positive integer r , there exist $\alpha = \alpha(r) > 0$ and $c > 0$ such that for all combinations of signs,

$$|\omega_j \pm \omega_k \pm \omega_{\ell_1} \pm \dots \pm \omega_{\ell_r}| \geq c L^{-\alpha} \quad \text{for } j \geq k \geq L = \ell_1 \geq \dots \geq \ell_r \geq 0, \quad (10)$$

provided that the sum does not vanish unless the terms cancel pairwise. In [1] it is shown that for almost all (w.r.t. Lebesgue measure) ρ in a fixed interval of positive numbers there is a $c > 0$ such that condition (10) holds with $\alpha = 16r^5$.

Theorem 2.1 [5, Theorem 1] *Under the non-resonance condition (9) and assumption (6) on the initial data with $s \geq \sigma + 1$, the estimate*

$$\sum_{\ell=0}^{\infty} \omega_\ell^{2s+1} \frac{|I_\ell(t) - I_\ell(0)|}{\varepsilon^2} \leq C\varepsilon \quad \text{for } 0 \leq t \leq \varepsilon^{-N+1}$$

with $I_\ell(t) = I_\ell(u(\cdot, t), \partial_t u(\cdot, t))$ holds with a constant C which depends on s, N , and C_0 , but not on ε and t .

3 Spectral semi-discretisation in space

For the numerical solution of (1) we consider the “method of lines” approach. Pseudo-spectral semi-discretisation in space with equidistant collocation points $x_k = k\pi/M$ (for $k = -M, \dots, M-1$) yields an approximation by the real trigonometric polynomial

$$u^M(x, t) = \sum'_{|j| \leq M} q_j(t) e^{ijx}, \quad v^M(x, t) = \sum'_{|j| \leq M} p_j(t) e^{ijx}, \quad (11)$$

where the prime indicates that the first and last terms in the sum are taken with the factor $1/2$. Here, we have set $p_j(t) = \frac{d}{dt} q_j(t)$, and we note that $q_{-j} = \bar{q}_j$ and $p_{-j} = \bar{p}_j$. The $2M$ -periodic coefficient sequence $q(t) = (q_j(t))$ is a solution of the $2M$ -dimensional system of ordinary differential equations

$$\frac{d^2 q}{dt^2} + \Omega^2 q = f(q) \quad \text{with} \quad f(q) = -\mathcal{F}_{2M} g(\mathcal{F}_{2M}^{-1} q). \quad (12)$$

Here, Ω is the diagonal matrix with entries ω_j for $|j| \leq M$, and \mathcal{F}_{2M} denotes the discrete Fourier transform: $(\mathcal{F}_{2M} w)_j = \frac{1}{2M} \sum_{k=-M}^{M-1} w_k e^{-ijx_k}$. Since the nonlinearity in (12) has the components

$$f_j(q) = -\frac{\partial V(q)}{\partial q_{-j}} \quad \text{with} \quad V(q) = \frac{1}{2M} \sum_{k=-M}^{M-1} U((\mathcal{F}_{2M}^{-1} q)_k),$$

equation (12) is a finite-dimensional complex Hamiltonian system with the discrete energy

$$H_M(q, p) = \frac{1}{2} \sum'_{|j| \leq M} (|p_j|^2 + \omega_j^2 |q_j|^2) + V(q), \quad (13)$$

which is conserved along the solution $(q(t), p(t))$ with $p(t) = dq(t)/dt$, and differs from the continuous energy $H(u^M, v^M)$ evaluated at the trigonometric polynomials u^M, v^M of (11).

We consider the actions (for $|j| \leq M$) and the momentum

$$I_j(q, p) = \frac{\omega_j}{2} |q_j|^2 + \frac{1}{2\omega_j} |p_j|^2, \quad K(q, p) = -\sum''_{|j| \leq M} i j q_{-j} p_j, \quad (14)$$

where the double prime indicates that the first and last terms in the sum are taken with the factor $1/4$. These quantities are defined such that, with the trigonometric polynomials u^M, v^M of (11), we have

$$I_j(q, p) = I_j(u^M, v^M) \quad \text{and} \quad K(q, p) = K(u^M, v^M)$$

with the definitions of Section 2 used on the right-hand sides. The equality for I_j holds for $|j| < M$, whereas $I_{\pm M}(q, p) = 4I_{\pm M}(u^M, v^M)$. Since we are concerned with real approximations (11), the Fourier coefficients satisfy $q_{-j} = \bar{q}_j$ and $p_{-j} = \bar{p}_j$, so that $I_{-j} = I_j$.

For a $2M$ -periodic sequence $q = (q_j)$, we introduce the weighted norm

$$\|q\|_s = \left(\sum''_{|j| \leq M} \omega_j^{2s} |q_j|^2 \right)^{1/2}, \quad (15)$$

which is defined such that it equals the H^s norm of the trigonometric polynomial with coefficients q_j .

We assume that the initial data $q(0)$ and $p(0)$ satisfy a condition corresponding to (6):

$$\left(\|q(0)\|_{s+1}^2 + \|p(0)\|_s^2 \right)^{1/2} \leq \varepsilon. \quad (16)$$

Theorem 3.1 *Under the non-resonance condition (9) with exponent σ and the assumption (16) of small initial data with $s \geq \sigma + 1$, the estimate*

$$\sum_{\ell=0}^M \omega_\ell^{2s+1} \frac{|I_\ell(t) - I_\ell(0)|}{\varepsilon^2} \leq C\varepsilon \quad \text{for } 0 \leq t \leq \varepsilon^{-N+1}$$

with $I_\ell(t) = I_\ell(q(t), p(t))$ holds with a constant C which depends on s , N , and C_0 , but is independent of ε , M , and t .

We note that Theorem 3.1 implies long-time spatial regularity:

$$\left(\|u^M(\cdot, t)\|_{s+1}^2 + \|v^M(\cdot, t)\|_s^2 \right)^{1/2} \leq \varepsilon(1 + C\varepsilon) \quad \text{for } t \leq \varepsilon^{-N+1}. \quad (17)$$

The momentum is no longer an exactly conserved quantity in the semi-discretisation, but we have the following approximate-conservation result.

Theorem 3.2 *Under the assumptions of Theorem 3.1, the estimate*

$$\frac{|K(t) - K(0)|}{\varepsilon^2} \leq C t \varepsilon M^{-s-1} \quad \text{for } 0 \leq t \leq \varepsilon^{-N+1}$$

with $K(t) = K(q(t), p(t))$ holds with a constant C which depends on s , N , and C_0 , but is independent of ε , M , and t .

We do not know if the above estimate is optimal for large values of εt . In our numerical experiments we observed that on very long time intervals, the relative deviation of the momentum behaves like an almost-periodic function of $\varepsilon^2 t$, which depends on M and whose maximum decreases with a negative power of M .

The discrete energy (13) is not the same as the continuous energy (2) along the semi-discrete solution. However, since Theorem 3.1 controls the spatial regularity of the semi-discrete solution over long times, we have the following result.

Theorem 3.3 *Under the assumptions of Theorem 3.1, the estimate*

$$\frac{|H(t) - H(0)|}{\varepsilon^2} \leq C \varepsilon M^{-s-1} \quad \text{for } 0 \leq t \leq \varepsilon^{-N+1}$$

with $H(t) = H(u^M(\cdot, t), v^M(\cdot, t))$ holds with a constant C which depends on s , N , and C_0 , but is independent of ε , M , and t .

The proof of Theorem 3.3 also shows that, for $0 \leq t \leq \varepsilon^{-N+1}$,

$$\frac{|H(u^M(\cdot, t), v^M(\cdot, t)) - H_M(q(t), p(t))|}{\varepsilon^2} \leq C M^{-s-1}.$$

The rest of the paper is concerned with the proof of these results. The proof of Theorem 3.1 is a modification of the proof of the corresponding result for the continuous problem and is outlined in Sections 4 and 5. In parallel we give a proof of a variant of Theorem 3.2, which provides additional insight into the structure of the problem and has the advantage of being transferable to the fully discrete case (see [4]). A different proof, which yields the precise estimate of Theorem 3.2, is given in Section 6.1, where also Theorem 3.3 is proved.

4 Modulated Fourier expansion

The principal tool for the long-time analysis of the semi-discretised nonlinear wave equation is a modulated Fourier expansion as in [6, Chapter XIII]. The presentation follows closely the analysis of nonlinear wave equations in [5].

4.1 Estimates of modulation functions and remainder

In the following we use the abbreviations (7) concerning sequences $\mathbf{k} = (k_\ell)_{\ell \geq 0}$ with $k_\ell = 0$ for $\ell > M$ (because only the frequencies $\omega_0, \dots, \omega_M$ are present in the semi-discretisation), and we set

$$\llbracket \mathbf{k} \rrbracket = \begin{cases} \frac{1}{2}(\|\mathbf{k}\| + 1), & \mathbf{k} \neq 0 \\ \frac{3}{2}, & \mathbf{k} = 0. \end{cases}$$

Theorem 4.1 *Under the assumptions of Theorem 3.1 there exist truncated asymptotic expansions (with N from (9))*

$$\tilde{q}(t) = \sum_{\|\mathbf{k}\| \leq 2N} z^{\mathbf{k}}(\varepsilon t) e^{i(\mathbf{k} \cdot \boldsymbol{\omega})t}, \quad \tilde{p}(t) = \frac{d}{dt} \tilde{q}(t), \quad (18)$$

such that the solution $(q(t), p(t))$ of (12) satisfies

$$\|q(t) - \tilde{q}(t)\|_{s+1} + \|p(t) - \tilde{p}(t)\|_s \leq C\varepsilon^N \quad \text{for } 0 \leq t \leq \varepsilon^{-1}. \quad (19)$$

The truncated modulated Fourier expansion is bounded by

$$\|\tilde{q}(t)\|_{s+1} + \|\tilde{p}(t)\|_s \leq C\varepsilon \quad \text{for } 0 \leq t \leq \varepsilon^{-1}. \quad (20)$$

On this time interval, we further have, for $|j| \leq M$,

$$\tilde{q}_j(t) = z_j^{(j)}(\varepsilon t) e^{i\omega_j t} + z_j^{-(j)}(\varepsilon t) e^{-i\omega_j t} + r_j, \quad \text{with } \|r\|_{s+1} \leq C\varepsilon^2, \quad (21)$$

and the modulation functions $z^{\mathbf{k}}$ are bounded by

$$\sum_{\|\mathbf{k}\| \leq 2N} \left(\frac{\boldsymbol{\omega}^{|\mathbf{k}|}}{\varepsilon^{\llbracket \mathbf{k} \rrbracket}} \|z^{\mathbf{k}}(\varepsilon t)\|_s \right)^2 \leq C. \quad (22)$$

Bounds of the same type hold for any fixed number of derivatives of $z^{\mathbf{k}}$ with respect to the slow time $\tau = \varepsilon t$. Moreover, the modulation functions satisfy $z_{-j}^{-\mathbf{k}} = \overline{z_j^{\mathbf{k}}}$. The constants C are independent of ε , M , and of $t \leq \varepsilon^{-1}$.

The proof of this result follows closely that of Theorem 2 in [5]. We only outline the minor modifications that are necessary to treat the semi-discrete case.

4.2 Modifications in the proof for the analytic case

For an analysis it is convenient to rewrite equation (12) in the following notation: for a 2π -periodic function $w(x)$ we denote by $(\mathcal{Q}w)(x)$ the trigonometric interpolation polynomial to $w(x)$ in the points x_k . For a $2M$ -periodic coefficient sequence $q = (q_j)$ we denote by $(\mathcal{P}q)(x)$ the trigonometric polynomial with coefficients q_j , $(\mathcal{P}q)(x) = \sum'_{|j| \leq M} q_j e^{ijx}$. For the approximation given by (11) we then have $u^M = \mathcal{P}q$ with the solution $q(t)$ of (12), which is rewritten as

$$\partial_t^2 u^M - \partial_x^2 u^M + \rho u^M + \mathcal{Q}g(u^M) = 0. \quad (23)$$

Taylor expansion of the nonlinearity expresses it as

$$\mathcal{Q}g(u^M) = \sum_{m \geq 2} \frac{g^{(m)}(0)}{m!} \mathcal{Q}(\mathcal{P}q)^m. \quad (24)$$

For $w(x) = \sum_{j=-\infty}^{\infty} w_j e^{ijx}$, the interpolation polynomial is given by the aliasing formula

$$\mathcal{Q}w(x) = \sum'_{|j| \leq M} \left(\sum_{l=-\infty}^{\infty} w_{j+2Ml} \right) e^{ijx}. \quad (25)$$

We use this formula, insert the trigonometric polynomial $\mathcal{P}\tilde{q}$ with $\tilde{q}(t)$ from (18) into equation (23) with (24), and consider the j th Fourier coefficient. This yields the formal modulation equations as in Section 3.2 of [5], from which the modulation functions are obtained by a reverse Picard iteration:

$$\begin{aligned} & (\omega_j^2 - (\mathbf{k} \cdot \boldsymbol{\omega})^2) z_j^{\mathbf{k}} + 2i\varepsilon(\mathbf{k} \cdot \boldsymbol{\omega}) \dot{z}_j^{\mathbf{k}} + \varepsilon^2 \ddot{z}_j^{\mathbf{k}} \\ & + \sum_m \frac{g^{(m)}(0)}{m!} \sum_{\mathbf{k}^1 + \dots + \mathbf{k}^m = \mathbf{k}} \sum'_{j_1 + \dots + j_m \equiv j \pmod{2M}} z_{j_1}^{\mathbf{k}^1} \dots z_{j_m}^{\mathbf{k}^m} = 0. \end{aligned} \quad (26)$$

The only difference to the corresponding equation in [5] is the range $|j_i| \leq M$ and that the sum over the j_i is taken modulo $2M$. As in (11), the prime on the sum over j_1, \dots, j_m indicates that with every appearance of $z_{j_i}^{\mathbf{k}^i}$ with $j_i = \pm M$ a factor $\frac{1}{2}$ is included.

The nonlinearity in equation (26) now becomes the j th Fourier coefficient of the trigonometric polynomial

$$\sum_m \frac{g^{(m)}(0)}{m!} \sum_{\mathbf{k}^1 + \dots + \mathbf{k}^m = \mathbf{k}} \mathcal{Q}(\mathcal{P}z^{\mathbf{k}^1} \dots \mathcal{P}z^{\mathbf{k}^m}).$$

With the following simple (and known) lemma and noting that the norm (15) of q equals the H^s norm of $\mathcal{P}q$, $\|q\|_s = \|\mathcal{P}q\|_s$, we obtain the estimate

$$\|\mathcal{Q}(\mathcal{P}z^{\mathbf{k}^1} \dots \mathcal{P}z^{\mathbf{k}^m})\|_s \leq C \|z^{\mathbf{k}^1}\|_s \dots \|z^{\mathbf{k}^m}\|_s.$$

The proof of Theorem 4.1 is then identical to that of Theorem 2 in [5]. The bounds (20)-(22) follow from the estimates in Section 3.7 of [5].

Lemma 4.2 *There are constants C depending only on $s > \frac{1}{2}$, such that for all functions $v, w \in H^s$ the trigonometric interpolation operator satisfies*

$$\|\mathcal{Q}v\|_s \leq C \|v\|_s \quad (27)$$

$$\|\mathcal{Q}v - v\|_0 \leq CM^{-s} \|v\|_s. \quad (28)$$

Moreover, H^s is a normed algebra:

$$\|vw\|_s \leq C\|v\|_s\|w\|_s. \quad (29)$$

Proof. With the aliasing formula (25) and the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|Qv\|_s^2 &= \sum_{|j| \leq M}'' \omega_j^{2s} \left| \sum_{l=-\infty}^{\infty} v_{j+2Ml} \right|^2 \\ &\leq \sum_{|j| \leq M}'' \left(\sum_{l=-\infty}^{\infty} \frac{\omega_j^{2s}}{\omega_{j+2Ml}^{2s}} \right) \sum_{l=-\infty}^{\infty} \omega_{j+2Ml}^{2s} |v_{j+2Ml}|^2 \leq C_1 \|v\|_s^2. \end{aligned}$$

The bound (28) follows with the Cauchy-Schwarz inequality as

$$\begin{aligned} \|Qv - v\|_0^2 &\leq \sum_{|j| \geq M} |v_j|^2 + \sum_{|j| \leq M} \left| \sum_{l \neq 0} \omega_{j+2Ml}^{-s} \cdot \omega_{j+2Ml}^s v_{j+2Ml} \right|^2 \\ &\leq \sum_{|j| \geq M} \omega_j^{-2s} \cdot \omega_j^{2s} |v_j|^2 + \sum_{|j| \leq M} \left(\sum_{l \neq 0} \omega_{j+2Ml}^{-2s} \right) \left(\sum_{l \neq 0} \omega_{j+2Ml}^{2s} |v_{j+2Ml}|^2 \right) \\ &\leq C M^{-2s} \|v\|_s^2. \end{aligned}$$

Similarly, the inequality (29) follows with $\sum_{i+j=k} \omega_i^{-2s} \omega_j^{-2s} \leq C \omega_k^{-2s}$ and the Cauchy-Schwarz inequality. \square

4.3 Estimates of the defect

The modulation equations (26) are solved approximately by an iterative procedure [5, Section 3.3]. After $4N$ iterations this leaves a defect $\mathbf{d} = (d_j^{\mathbf{k}})$,

$$\begin{aligned} d_j^{\mathbf{k}} &= (\omega_j^2 - (\mathbf{k} \cdot \boldsymbol{\omega})^2) z_j^{\mathbf{k}} + 2i\varepsilon (\mathbf{k} \cdot \boldsymbol{\omega}) \dot{z}_j^{\mathbf{k}} + \varepsilon^2 \ddot{z}_j^{\mathbf{k}} \\ &\quad + \sum_{m=2}^N \frac{g^{(m)}(0)}{m!} \sum_{\mathbf{k}^1 + \dots + \mathbf{k}^m = \mathbf{k}} \sum_{j_1 + \dots + j_m \equiv j \pmod{2M}}' z_{j_1}^{\mathbf{k}^1} \dots z_{j_m}^{\mathbf{k}^m}. \end{aligned} \quad (30)$$

This is to be considered for $\|\mathbf{k}\| \leq NK$, where we set $z_j^{\mathbf{k}} = 0$ for $\|\mathbf{k}\| > K = 2N$. In Sections 3.8–3.11 of [5], the following bound is shown:

$$\left(\sum_{\|\mathbf{k}\| \leq NK} \|\boldsymbol{\omega}^{|\mathbf{k}|} d^{\mathbf{k}}(\tau)\|_s^2 \right)^{1/2} \leq C \varepsilon^{N+1} \quad \text{for } \tau \leq 1. \quad (31)$$

5 Conservation of actions and momentum

We now show that the system of equations determining the modulation functions has almost-invariants close to the actions and the momentum.

5.1 The extended potential

Corresponding to the modulation functions $z_j^{\mathbf{k}}(\varepsilon t)$ we introduce, for $\|\mathbf{k}\| \leq 2N$ and $2M$ -periodic in j ,

$$\mathbf{y} = (y_j^{\mathbf{k}}) \quad \text{with} \quad y_j^{\mathbf{k}}(t) = z_j^{\mathbf{k}}(\varepsilon t) e^{i(\mathbf{k} \cdot \boldsymbol{\omega})t}. \quad (32)$$

By construction, the functions $y_j^{\mathbf{k}}$ satisfy

$$\partial_t^2 y_j^{\mathbf{k}} + \omega_j^2 y_j^{\mathbf{k}} + \sum_{m=2}^N \frac{g^{(m)}(0)}{m!} \sum_{\mathbf{k}^1 + \dots + \mathbf{k}^m = \mathbf{k}} \sum'_{\substack{j_1 + \dots + j_m \\ \equiv j \pmod{2M}}} y_{j_1}^{\mathbf{k}^1} \dots y_{j_m}^{\mathbf{k}^m} = e_j^{\mathbf{k}}, \quad (33)$$

where $\|\mathbf{k}^i\| \leq 2N$ and $|j_i| \leq M$, and where the defects $e_j^{\mathbf{k}}(t) = d_j^{\mathbf{k}}(\varepsilon t) e^{i(\mathbf{k} \cdot \boldsymbol{\omega})t}$ are small. In (1), the nonlinearity $g(u)$ is the gradient of the potential $U(u) = \int_0^u g(v) dv$. The sum in (33) is recognised as the partial derivative with respect to $y_{-j}^{-\mathbf{k}}$ of the *extended potential* $\mathcal{U}(\mathbf{y})$ defined by

$$\mathcal{U}(\mathbf{y}) = \sum_{l=-N}^N \mathcal{U}_l(\mathbf{y}) \quad (34)$$

$$\mathcal{U}_l(\mathbf{y}) = \sum_{m=2}^N \frac{U^{(m+1)}(0)}{(m+1)!} \sum_{\mathbf{k}^1 + \dots + \mathbf{k}^{m+1} = \mathbf{0}} \sum'_{j_1 + \dots + j_{m+1} = 2Ml} y_{j_1}^{\mathbf{k}^1} \dots y_{j_{m+1}}^{\mathbf{k}^{m+1}},$$

where again $\|\mathbf{k}^i\| \leq 2N$ and $|j_i| \leq M$.

The modulation system (33) can now be rewritten as

$$\partial_t^2 y_j^{\mathbf{k}} + \omega_j^2 y_j^{\mathbf{k}} + \nabla_{-j}^{-\mathbf{k}} \mathcal{U}(\mathbf{y}) = e_j^{\mathbf{k}}, \quad (35)$$

where $\nabla_{-j}^{-\mathbf{k}} \mathcal{U}$ is the partial derivative of \mathcal{U} with respect to $y_{-j}^{-\mathbf{k}}$.

5.2 Invariance under group actions

For an arbitrary real sequence $\boldsymbol{\mu} = (\mu_\ell)_{\ell \geq 0}$ and for $\theta \in \mathbb{R}$, let

$$(S_{\boldsymbol{\mu}}(\theta)\mathbf{y})_j^{\mathbf{k}} = e^{i(\mathbf{k} \cdot \boldsymbol{\mu})\theta} y_j^{\mathbf{k}}, \quad (T(\theta)\mathbf{y})_j^{\mathbf{k}} = e^{ij\theta} y_j^{\mathbf{k}}. \quad (36)$$

Since the sum in the definition of \mathcal{U} is over $\mathbf{k}^1 + \dots + \mathbf{k}^{m+1} = \mathbf{0}$ and that of \mathcal{U}_0 over $j_1 + \dots + j_{m+1} = 0$, we have

$$\mathcal{U}(S_{\boldsymbol{\mu}}(\theta)\mathbf{y}) = \mathcal{U}(\mathbf{y}), \quad \mathcal{U}_0(T(\theta)\mathbf{y}) = \mathcal{U}_0(\mathbf{y}) \quad \text{for } \theta \in \mathbb{R}.$$

Differentiating these relations with respect to θ yields

$$0 = \frac{d}{d\theta} \Big|_{\theta=0} \mathcal{U}(S_{\boldsymbol{\mu}}(\theta)\mathbf{y}) = \sum_{\|\mathbf{k}\| \leq K} i(\mathbf{k} \cdot \boldsymbol{\mu}) \sum'_{|j| \leq M} y_j^{\mathbf{k}} \nabla_j^{\mathbf{k}} \mathcal{U}(\mathbf{y}) \quad (37)$$

$$0 = \frac{d}{d\theta} \Big|_{\theta=0} \mathcal{U}_0(T(\theta)\mathbf{y}) = \sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} ij y_j^{\mathbf{k}} \nabla_j^{\mathbf{k}} \mathcal{U}_0(\mathbf{y}). \quad (38)$$

5.3 Almost-invariants of the modulation system

We multiply (35) with $i(\mathbf{k} \cdot \boldsymbol{\mu})y_{-j}^{-\mathbf{k}}$ for $\mu = \langle \ell \rangle = (0, \dots, 0, 1, 0, \dots)$ with the only entry at the ℓ th position and sum over \mathbf{k} and j . Expressing the $y_j^{\mathbf{k}}$ of (32) in terms of $z_j^{\mathbf{k}}$, the invariance property (37) then implies that

$$\mathcal{J}_\ell(\mathbf{z}, \dot{\mathbf{z}}) := - \sum_{\|\mathbf{k}\| \leq K} ik_\ell \sum'_{|j| \leq M} z_{-j}^{-\mathbf{k}} \left(i(\mathbf{k} \cdot \boldsymbol{\omega})z_j^{\mathbf{k}} + \varepsilon \dot{z}_j^{\mathbf{k}} \right) \quad (39)$$

satisfies

$$\varepsilon \frac{d}{d\tau} \mathcal{J}_\ell(\mathbf{z}, \dot{\mathbf{z}}) = - \sum_{\|\mathbf{k}\| \leq K} ik_\ell \sum'_{|j| \leq M} z_{-j}^{-\mathbf{k}} d_j^{\mathbf{k}}. \quad (40)$$

As in Theorem 3 of [5] we obtain the following result.

Theorem 5.1 *Under the conditions of Theorem 4.1,*

$$\sum_{\ell=0}^M \omega_\ell^{2s+1} \left| \frac{d}{d\tau} \mathcal{J}_\ell(\mathbf{z}(\tau), \dot{\mathbf{z}}(\tau)) \right| \leq C \varepsilon^{N+1} \quad \text{for } \tau \leq 1.$$

We now proceed similarly, multiplying (35) with $ijy_{-j}^{-\mathbf{k}}$, summing over \mathbf{k} and j , and using (38):

$$\sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} ij y_{-j}^{-\mathbf{k}} \partial_t^2 y_j^{\mathbf{k}} = \sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} ij y_{-j}^{-\mathbf{k}} \left(e_j^{\mathbf{k}} - \sum_{l \neq 0} \nabla_{-j}^{-\mathbf{k}} \mathcal{U}_l(\mathbf{y}) \right).$$

The negative left-hand side is recognised as the time derivative of

$$- \sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} ij y_{-j}^{-\mathbf{k}} \partial_t y_j^{\mathbf{k}}$$

which, in terms of the variables \mathbf{z} of (32), equals

$$\mathcal{K}(\mathbf{z}, \dot{\mathbf{z}}) = - \sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} ij z_{-j}^{-\mathbf{k}} \left(i(\mathbf{k} \cdot \boldsymbol{\omega})z_j^{\mathbf{k}} + \varepsilon \dot{z}_j^{\mathbf{k}} \right). \quad (41)$$

We thus obtain

$$\varepsilon \frac{d}{d\tau} \mathcal{K}(\mathbf{z}(\tau), \dot{\mathbf{z}}(\tau)) = - \sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} ij z_{-j}^{-\mathbf{k}} \left(d_j^{\mathbf{k}} - \sum_{l \neq 0} \nabla_{-j}^{-\mathbf{k}} \mathcal{U}_l(\mathbf{z}) \right). \quad (42)$$

Theorem 5.2 *Under the conditions of Theorem 4.1,*

$$\left| \frac{d}{d\tau} \mathcal{K}(\mathbf{z}(\tau), \dot{\mathbf{z}}(\tau)) \right| \leq C(\varepsilon^{N+1} + \varepsilon^2 M^{-s+1}) \quad \text{for } \tau \leq 1.$$

Proof. With the Cauchy–Schwarz inequality and the bound $|j| \leq \omega_j$, we obtain

$$\left| \sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} ij z_{-j}^{-\mathbf{k}} d_j^{\mathbf{k}} \right| \leq \left(\sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} \omega_j^2 |z_j^{\mathbf{k}}|^2 \right)^{1/2} \left(\sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} |d_j^{\mathbf{k}}|^2 \right)^{1/2}.$$

The first factor on the right-hand side is bounded by $\mathcal{O}(\varepsilon)$ in view of (22), and the second factor is $\mathcal{O}(\varepsilon^{N+1})$ by (31).

The remaining expression of (42) contains terms of the form

$$\begin{aligned} & \sum_{\|\mathbf{k}\| \leq K} \sum'_{|j| \leq M} i_j z_{-j}^{-\mathbf{k}} \nabla_{-j}^{-\mathbf{k}} \mathcal{U}_l(\mathbf{z}) \\ &= \sum_{m=2}^N \frac{U^{(m+1)}(0)}{m!} \sum_{\mathbf{k}^1 + \dots + \mathbf{k}^{m+1} = \mathbf{0}} \sum'_{j_1 + \dots + j_{m+1} = 2Ml} z_{j_1}^{\mathbf{k}^1} \dots z_{j_m}^{\mathbf{k}^m} \cdot i_{j_{m+1}} z_{j_{m+1}}^{\mathbf{k}^{m+1}}, \end{aligned}$$

which is the $2Ml$ -th Fourier coefficient of the function

$$w(x) := \sum_{m=2}^N \frac{U^{(m+1)}(0)}{m!} \sum_{\mathbf{k}^1 + \dots + \mathbf{k}^{m+1} = \mathbf{0}} \mathcal{P}z^{\mathbf{k}^1}(x) \dots \mathcal{P}z^{\mathbf{k}^m}(x) \cdot \frac{d}{dx} \mathcal{P}z^{\mathbf{k}^{m+1}}(x).$$

Since H^{s-1} is a normed algebra for $s > 3/2$, the H^{s-1} norm of w is bounded by

$$\sum_{m=2}^N \frac{|U^{(m+1)}(0)|}{m!} \left(\sum_{\|\mathbf{k}\| \leq K} \|\mathbf{z}^{\mathbf{k}}\|_{s-1} \right)^m \left(\sum_{\|\mathbf{k}\| \leq K} \|\mathbf{z}^{\mathbf{k}}\|_s \right).$$

The terms in this sum are estimated using the Cauchy-Schwarz inequality,

$$\sum_{\|\mathbf{k}\| \leq K} \|\mathbf{z}^{\mathbf{k}}\|_s \leq \left(\sum_{\|\mathbf{k}\| \leq K} \omega^{-2|\mathbf{k}|} \right)^{1/2} \left(\sum_{\|\mathbf{k}\| \leq K} \|\omega^{|\mathbf{k}|} z^{\mathbf{k}}\|_s^2 \right)^{1/2}.$$

The first factor on the right-hand side is a finite constant by Lemma 2 of [5], and the second factor is $\mathcal{O}(\varepsilon)$ by (22). Hence we have

$$\|w\|_{s-1} \leq C\varepsilon^3,$$

and therefore the $2Ml$ -th Fourier coefficient of w is bounded by $C\varepsilon^3 \omega_{2Ml}^{-s+1} \leq C\varepsilon^3 (2Ml)^{-s+1}$. In this way the result follows from (42). \square

5.4 Relationship with actions and momentum

The almost-invariants \mathcal{J}_ℓ of the modulated Fourier expansion turn out to be close to the corresponding harmonic actions (5) of the solution of the nonlinear wave equation,

$$J_\ell = I_\ell + I_{-\ell} = 2I_\ell \quad \text{for } 0 < \ell < M, \quad J_0 = I_0, \quad J_M = I_M,$$

and \mathcal{K} is shown to be close to the momentum K .

With the same argument as in [5, Theorem 4] we obtain the following result.

Theorem 5.3 *Under the conditions of Theorem 4.1, along the semi-discrete solution $(q(t), p(t))$ of (12) and the associated modulation sequence $\mathbf{z}(\varepsilon t)$, it holds that*

$$\mathcal{J}_\ell(\mathbf{z}(\varepsilon t), \dot{\mathbf{z}}(\varepsilon t)) = J_\ell(q(t), p(t)) + \gamma_\ell(t) \varepsilon^3$$

with $\sum_{\ell=0}^M \omega_\ell^{2s+1} \gamma_\ell(t) \leq C$ for $t \leq \varepsilon^{-1}$. All appearing constants are independent of ε , M , and t .

For the momentum we have the following.

Theorem 5.4 *Under the conditions of Theorem 4.1, along the semi-discrete solution $(q(t), p(t))$ of (12) and the associated modulation sequence $\mathbf{z}(\varepsilon t)$, it holds that*

$$\mathcal{K}(\mathbf{z}(\varepsilon t), \dot{\mathbf{z}}(\varepsilon t)) = K(q(t), p(t)) + \mathcal{O}(\varepsilon^3) + \mathcal{O}(\varepsilon^2 M^{-s}).$$

Proof. Separating in (41) the terms with $\mathbf{k} = \pm\langle j \rangle$ and applying the bound (21) to the remaining terms, we find

$$\mathcal{K}(\mathbf{z}, \dot{\mathbf{z}}) = \sum'_{|j| \leq M} j \omega_j \left(|z_j^{(j)}|^2 - |z_j^{-(j)}|^2 \right) + \mathcal{O}(\varepsilon^3).$$

In terms of the Fourier coefficients of the modulated Fourier expansion $\tilde{q}_j(t) = \sum_{\|\mathbf{k}\| \leq K} z_j^{\mathbf{k}}(\varepsilon t) e^{i(\mathbf{k} \cdot \boldsymbol{\omega})t}$ and $\tilde{p}_j(t) = \frac{d}{dt} \tilde{q}_j(t)$, we have

$$\begin{aligned} \mathcal{K}[\mathbf{z}] &= \sum'_{|j| \leq M} j \frac{\omega_j}{4} \left(|\tilde{q}_j + (i\omega_j)^{-1} \tilde{p}_j|^2 - |\tilde{q}_j - (i\omega_j)^{-1} \tilde{p}_j|^2 \right) + \mathcal{O}(\varepsilon^3) \\ &= K(\tilde{q}, \tilde{p}) + \mathcal{O}(\varepsilon^3) + \mathcal{O}(\varepsilon^2 M^{-s}) \\ &= K(q, p) + \mathcal{O}(\varepsilon^3) + \mathcal{O}(\varepsilon^2 M^{-s}), \end{aligned}$$

where we have used the bound (21). The $\mathcal{O}(\varepsilon^2 M^{-s})$ comes from the boundary terms in the sum. The last equality is a consequence of the remainder bound of Theorem 4.1. \square

With an identical argument to that of [5, Section 4.5], Theorems 5.1–5.4 yield the statement of Theorem 3.1 by patching together many intervals of length ε^{-1} . For the momentum, the same argument gives the bound

$$\frac{|K(t) - K(0)|}{\varepsilon^2} \leq C(\varepsilon + M^{-s} + \varepsilon t M^{-s+1}) \quad \text{for } 0 \leq t \leq \varepsilon^{-N+1}$$

instead of that of Theorem 3.2.

6 Consequences of long-time spatial regularity

In this section we provide proofs of Theorems 3.2 and 3.3, which are based on the regularity estimate (17).

6.1 Conservation of momentum

Inserting the exact solution $\tilde{u}(x, t)$ of (1) with starting values $\tilde{u}(x, 0) = u^M(x, 0)$ and $\partial_t \tilde{u}(x, 0) = v^M(x, 0)$ into equation (23) yields

$$\partial_t^2 \tilde{u} - \partial_x^2 \tilde{u} + \rho \tilde{u} + \mathcal{Q}g(\tilde{u}) = d$$

with a defect $d = \mathcal{Q}g(\tilde{u}) - g(\tilde{u})$. Under condition (16) it is known from [5] that $\|\tilde{u}(\cdot, t)\|_{s+1} \leq C\varepsilon$ on intervals of length ε^{-1} . With the variation of constants formula, it then follows as in [5, Section 3.13] that, with $\tilde{v} = \partial_t \tilde{u}$,

$$\|u^M(\cdot, t) - \tilde{u}(\cdot, t)\|_1 + \|v^M(\cdot, t) - \tilde{v}(\cdot, t)\|_0 \leq C t \max_{0 \leq \sigma \leq t} \|d(\cdot, \sigma)\|_0$$

for $t \leq \varepsilon^{-1}$ and, together with Lemma 4.2,

$$\|d(\cdot, t)\|_0 \leq CM^{-s-1} \|g(\tilde{u}(\cdot, t))\|_{s+1}.$$

For g analytic in a neighbourhood of 0, the bound (17) implies, via $g(0) = g'(0) = 0$ and (29), that $\|g(\tilde{u}(\cdot, t))\|_{s+1} \leq C \|\tilde{u}(\cdot, t)\|_{s+1}^2 \leq C\varepsilon^2$. Hence,

$$\|d(\cdot, t)\|_0 \leq C\varepsilon^2 M^{-s-1} \quad \text{for } t \leq \varepsilon^{-1}.$$

This implies that on the short interval $0 \leq t \leq \varepsilon^{-1}$,

$$|K(u^M(\cdot, t), v^M(\cdot, t)) - K(\tilde{u}(\cdot, t), \tilde{v}(\cdot, t))| \leq Ct\varepsilon^3 M^{-s-1}.$$

To get the momentum conservation over a longer time interval we introduce the grid $t_n = n\varepsilon^{-1}$, and we consider the local solution $(\tilde{u}_n, \tilde{v}_n)$ of (1) corresponding to initial values $(u^M(\cdot, t_n), v^M(\cdot, t_n))$. Since K is exactly conserved along $(\tilde{u}_n, \tilde{v}_n)$, we have for $t_{n+1} \leq \varepsilon^{-N+1}$,

$$\begin{aligned} & |K(u^M(\cdot, t_{n+1}), v^M(\cdot, t_{n+1})) - K(u^M(\cdot, t_n), v^M(\cdot, t_n))| \\ &= |K(u^M(\cdot, t_{n+1}), v^M(\cdot, t_{n+1})) - K(\tilde{u}_n(\cdot, t_{n+1}), \tilde{v}_n(\cdot, t_{n+1}))| \\ &\leq C(t_{n+1} - t_n)\varepsilon^3 M^{-s-1}. \end{aligned}$$

The last estimate holds uniformly in n because of the regularity estimate (17). Summing up the telescoping sum yields the estimate of Theorem 3.2.

6.2 Conservation of energy

We finally prove Theorem 3.3. We note that by (2), (5), (11), and (13),

$$\begin{aligned} H(u^M(\cdot, t), v^M(\cdot, t)) &= H_M(q(t), p(t)) - \omega_M I_M(u^M(\cdot, t), v^M(\cdot, t)) \\ &\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(U(u^M(x, t)) - \mathcal{Q}U(u^M(x, t)) \right) dx. \end{aligned}$$

By the Cauchy–Schwarz inequality and Lemma 4.2, the last term is bounded by $CM^{-s-1} \|U(u^M(\cdot, t))\|_{s+1}$. For U analytic in a neighbourhood of 0, the bound (17) implies, via $U(0) = U'(0) = U''(0) = 0$ and (29),

$$\|U(u^M(\cdot, t))\|_{s+1} \leq C \|u^M(\cdot, t)\|_{s+1}^3 \leq C\varepsilon^3 \quad \text{for } t \leq \varepsilon^{-N+1}.$$

By Theorem 3.1,

$$|\omega_M I_M(u^M(\cdot, t), v^M(\cdot, t)) - \omega_M I_M(u^M(\cdot, 0), v^M(\cdot, 0))| \leq C\varepsilon^3 \omega_M^{-2s}.$$

Since H_M is conserved exactly along the solution of (12), these estimates yield the statement of Theorem 3.3.

Acknowledgement

This work was partially supported by the Fonds National Suisse, project No. 200020-113249/1, and by DFG, Project LU 532/4-1. A large part of this work was carried out when the authors visited the Isaac Newton Institute in Cambridge.

References

- [1] D. Bambusi, *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Phys. **234** (2003), 253–285.
- [2] J. Bourgain, *Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations*, Geom. Funct. Anal. **6** (1996), 201–230.
- [3] B. Cano, *Conserved quantities of some Hamiltonian wave equations after full discretization*, Numer. Math. **103** (2006), 197–223.
- [4] D. Cohen, E. Hairer, and C. Lubich, *Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations*, Tech. report, Isaac Newton Institute, 2007.
- [5] ———, *Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions*, Arch. Ration. Mech. Anal. (2007).
- [6] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations*, second ed., Springer Series in Computational Mathematics 31, Springer-Verlag, Berlin, 2006.
- [7] M. Oliver, M. West, and C. Wulff, *Approximate momentum conservation for spatial semidiscretizations of semilinear wave equations*, Numer. Math. **97** (2004), 493–535.