Numerical conservation of energy, momentum and actions for extended RKN methods when applied to nonlinear wave equations via spatial spectral semi-discretizations

Bin Wang∗ Xinyuan Wu†

May 18, 2018

Abstract

This paper analyses the long-time behaviour of extended Runge–Kutta–Nyström (ERKN) methods when applied to nonlinear wave equations. It is shown that energy, momentum, and all harmonic actions are approximately preserved over a long time for a one-stage explicit ERKN method when applied to nonlinear wave equations via spectral semi-discretisations. The results are proved by deriving a multi-frequency modulated Fourier expansion of the ERKN method and showing three almost-invariants of the modulation system.

Keywords: nonlinear wave equations, extended RKN methods, multi-frequency modulated Fourier expansion, numerical conservation

MSC: 35L70, 65M70, 65M15

1 Introduction

In this paper, we study the long-time behaviour of extended Runge–Kutta–Nyström (ERKN) methods when applied to the one-dimensional non-linear wave equation

\[ u_{tt} - u_{xx} + \rho u + g(u) = 0, \quad t > 0, \quad -\pi \leq x \leq \pi, \]

where \( \rho > 0 \) and the non-linearity \( g \) is a smooth real function with \( g(0) = g'(0) = 0 \). We consider periodic boundary conditions and small initial data in appropriate Sobolev norms. It is required that the initial values \( u(\cdot, 0) \) and \( u_t(\cdot, 0) \) are bounded by a small parameter \( \epsilon \) (see [9]). This kind of equation frequently arises in various fields of scientific applications and many numerical methods have been developed and researched for solving the equation (see, e.g. [3], [5], [11], [13], [18], [20]).

∗School of Mathematical Sciences, Qufu Normal University, Qufu 273165, P.R. China; Mathematisches Institut, University of Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany. The research is supported in part by the Alexander von Humboldt Foundation and by the Natural Science Foundation of Shandong Province (Outstanding Youth Foundation) under Grant ZR2017JL003. E-mail: wang@ma.uni-tuebingen.de

†School of Mathematical Sciences, Qufu Normal University, Qufu 273165, P.R. China; Department of Mathematics, Nanjing University, Nanjing 210093, P.R. China. The research is supported in part by the National Natural Science Foundation of China under Grant 11671200. E-mail: xywu@nju.edu.cn
It is well known that the solution \((u(x, t), v(x, t))\) of (1) conserves several important quantities, where \(v = \partial_t u\). The solution of (1) exactly preserves the total energy

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

where the potential \(U(u)\) is defined as \(U(u) = \frac{1}{2} u'^2\). The momentum

\[
K(u, v) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \partial_x u(x)v(x) \, dx = -\sum_{k=-\infty}^{\infty} jkv_j
\]

is also conserved exactly by the solution of (1), where \(u_j = \mathcal{F}_ju\) and \(v_j = \mathcal{F}_jv\) are the Fourier coefficients in the series \(u(x) = \sum_{j=-\infty}^{\infty} u_j e^{ijx}\) and \(v(x) = \sum_{j=-\infty}^{\infty} v_j e^{ijx}\), respectively. It is noted that \(u_{-j} = \overline{u}_j\) and \(v_{-j} = \overline{v}_j\) since we consider only real solutions of (1). If we rewrite the equation (1) in terms of the Fourier coefficients, the equation \(\partial_t^2 u_j + \omega_j^2 u_j + \mathcal{F}_j g(u) = 0\) is obtained with the frequencies \(\omega_j = \sqrt{\rho + j^2}\) for \(j \in \mathbb{Z}\). The harmonic actions \(I_j(u, v) = \omega_j^2 |u_j|^2 + \frac{1}{2\omega_j}|v_j|^2\) are conserved for the linear wave equation. For the nonlinear equation (1), it has been shown in [1, 9] that the actions remain constant up to small deviations over a long time for almost all values of \(\rho > 0\) when the initial data is smooth and small.

For the methods applied to Hamiltonian systems of ordinary differential equations (ODEs), the long-time near conservation of the total energy and actions can be rigorously studied with the help of backward error analysis (see [19]). However, for highly oscillatory Hamiltonian systems with largest frequency \(\omega\), when the product \(h\omega\) is not small, no conclusion on the long-time behaviour can be drawn by the familiar backward error analysis. In order to overcome this restriction, Hairer and Lubich in [16] developed a novel technique called as modulated Fourier expansions for studying long-time conservation properties of numerical methods for highly oscillatory Hamiltonian systems of ODEs. It was extended to multi-frequency systems in [8] and has been used in the long-time analysis of different methods for various differential equations (see, e.g. [7, 18, 19, 23, 25, 28]). This technique has also been used in the analysis of wave equations, and we refer the reader to [9, 10, 12, 17].

On the other hand, the authors in [34] formulated a class of trigonometric integrators called as extended Runge–Kutta–Nyström (ERKN) methods for solving multi-frequency highly oscillatory second-order ODEs. Recently, further researches of trigonometric integrators on ODEs have been done and we refer to [2, 3, 26, 27, 30, 33]. The ERKN methods have also been well developed for wave equations on the basis of a novel operator-variation-of-constants formula (see, e.g. [21, 22, 24, 51, 52]). To our knowledge, however, the long-time conservation properties of ERKN methods when applied to wave equations have not been considered and researched yet in the literature. Under this background, in this paper, we devote to analysing the long-time behaviour of ERKN methods when solving the nonlinear wave equation (1). Our analysis here essentially follows the previous work about modulated Fourier expansions [10, 16, 28], with additional technical complications arising from the ERKN discretisation.

The rest of this paper is organised as follows. In Section 2, we consider a full discretisation of the nonlinear wave equation (1) by using spectral semi-discretisation in space and ERKN methods in time. Section 3 presents the long-time conservation properties of ERKN methods and reports some numerical results to support the theoretical results. The main result given in this section is proved.
in the remaining sections. A multi-frequency modulated Fourier expansion of ERKN methods is analysed in Section 4 and three almost-invariants are studied in Section 5. After showing the relationship between these almost-invariants with the actions, the momentum and the total energy of (1), we are then in a position to prove the main result by using the approach presented in [9, 10, 16, 19]. Finally, some concluding remarks are made in Section 6.

2 Full discretisation

In order to get a full discretisation of (1), we first discretise the wave equation in space by a spectral semi-discretisation and then in time by ERKN methods.

2.1 Spectral semi-discretisation in space

Following [10, 17], the pseudo-spectral semi-discretisation with equidistant collocation points

\[ x_k = k\pi/M \quad (\text{for } k = -M, -M+1, \ldots, M-1) \]

are chosen in space. Consider the following real-valued trigonometric polynomials as an approximation for the solution of (1)

\[ u^M(x, t) = \sum_{|j| \leq M} q_j(t) e^{ijx}, \quad v^M(x, t) = \sum_{|j| \leq M} p_j(t) e^{ijx}, \]

where \( p_j(t) = \frac{d}{dt} q_j(t) \) and the prime indicates that the first and last terms in the summation are taken with the factor 1/2. It is clear that the 2M-periodic coefficient vector \( q(t) = (q_j(t)) \) is a solution of the 2M-dimensional system of ODEs

\[ \frac{d^2 q}{dt^2} + \Omega^2 q = f(q), \] (5)

where \( \Omega \) is diagonal with entries \( \omega_j \) for \( |j| \leq M \), and \( f(q) = -F_{2M} g(F_{2M}^{-1} q) \). Here \( F_{2M} \) denotes the discrete Fourier transform

\[ (F_{2M} w)_j = \frac{1}{2M} \sum_{k=-M}^{M-1} w_k e^{-ijx_k} \text{ for } |j| \leq M. \]

We note that the non-linearity in (5) can be expressed as the form

\[ f_j(q) = -\frac{\partial}{\partial q_j} V(q) \] with \( V(q) = \frac{1}{2M} \sum_{k=-M}^{M-1} U((F_{2M} q)_k). \)

Therefore, the system (5) is a finite-dimensional complex Hamiltonian system with the energy

\[ H_M(q, p) = \frac{1}{2} \sum_{|j| \leq M} \left( |p_j|^2 + \omega_j^2 |q_j|^2 \right) + V(q). \]

The actions (for \( |j| \leq M \) and the momentum of (5) are respectively given by

\[ I_j(q, p) = \frac{\omega_j}{2} |q_j|^2 + \frac{1}{2\omega_j} |p_j|^2, \quad K(q, p) = -\sum_{|j| \leq M} ijq_j p_j, \]

where the double prime indicates that the first and last terms in the summation are taken with the factor 1/4. This paper is interested in real approximation (4) and thus we have \( q_{-j} = \bar{q}_j \), \( p_{-j} = \bar{p}_j \) and \( I_{-j} = I_j \).
In this paper, we use the same notations as those given in [10]. For \( k = (k_l)_{l=0}^\infty \) and \( \lambda = (\lambda_l)_{l=0}^\infty \), we denote

\[
|k| = (|k_l|)_{l=0}^\infty, \quad \|k\| = \sum_{l=0}^\infty |k_l|, \quad k \cdot \lambda = \sum_{l=0}^\infty k_l \lambda_l, \quad \lambda^{|k|} = \Pi_{l=0}^\infty \lambda_l^{|k_l|}
\]

for real \( \sigma \). The vector \((0, \ldots, 0, 1, 0, \ldots)\) with the only entry at the \( |j|\)-th position for \( j \in \mathbb{Z} \) is denoted by \( \langle j \rangle \). For \( s \in \mathbb{R}^+ \), we denote by \( H^s \) the Sobolev space of \( 2M \)-periodic sequences \( q = (q_j) \) endowed with the weighted norm \( ||q||_s = \left( \sum_{|j| \leq M} \omega_j^{2s} |q_j|^2 \right)^{1/2} \).

### 2.2 ERKN methods in time

As the time discretisation, we consider one-stage explicit extended Runge–Kutta–Nyström (ERKN) methods, which were first developed in [33].

**Definition 1** (See [33]) A one-stage explicit ERKN method for solving (14) is defined by

\[
\begin{align*}
Q^{n+1} &= \phi_0(c_1^2 V) q^n + h c_1 \phi_1(c_1^2 V) p^n, \\
q^{n+1} &= \phi_0(V) q^n + h \phi_1(V) p^n + h^2 b_\lambda(V) f(Q^{n+c_1}), \\
p^{n+1} &= -h \Omega^2 \phi_1(V) q^n + \phi_0(V) p^n + h b_1(V) f(Q^{n+c_1}),
\end{align*}
\]

where \( h \) is a stepsize, \( 0 \leq c_1 \leq 1 \), \( b_\lambda(V) \) and \( b_1(V) \) are matrix-valued and uniformly bounded functions of \( V \equiv h^2 \Omega^2 \), and \( \phi_j(V) := \sum_{k=0}^{\infty} \frac{(-1)^k V^k}{(2k+j)!} \) for \( j = 0, 1 \).

The following two theorems concerning the symmetry and symplecticity of ERKN methods will be used in this paper.

**Theorem 1** (Chap. 4 of [33]) The one-stage explicit ERKN method (14) is symmetric if and only if

\[
c_1 = 1/2, \quad \tilde{b}_1(V) = \phi_1(V) b_1(V) - \phi_0(V) \tilde{b}_1(V), \quad \phi_0(c_1^2 V) b_1(V) = c_1 \phi_1(c_1^2 V) b_1(V). \tag{8}
\]

**Theorem 2** (Chap. 4 of [33]) If there exists a real number \( d_1 \in \mathbb{R} \) such that

\[
\phi_0(V) b_1(V) + V \phi_1(V) \tilde{b}_1(V) = d_1 \phi_0(c_1^2 V), \quad \phi_1(V) b_1(V) - \phi_0(V) \tilde{b}_1(V) = c_1 d_1 \phi_1(c_1^2 V), \tag{9}
\]

then the one-stage explicit ERKN method (14) is symplectic.

It is noted that for \( V = h^2 \Omega^2 \), one has \( \phi_0(V) = \cos(h \Omega) \) and \( \phi_1(V) = \text{sinc}(h \Omega) \). Hence, in the remainder of this paper, we use the notations \( b_1(h \Omega) \) and \( b_1(h \Omega) \) to denote the coefficients appearing in the ERKN method (14).

### 3 Main result and numerical experiment

#### 3.1 Main result

In our analysis, we make the following assumptions, where the first five assumptions have been considered in [10].
Assumption 1 • The initial values $q(0)$ and $p(0)$ of (5) are assumed to satisfy

$$
\left( \|q(0)\|_{s+1}^2 + \|p(0)\|_{s+1}^2 \right)^{1/2} \leq \epsilon
$$

with a small parameter $\epsilon$.

• For a given stepsize $h$, we consider the non-resonance condition

$$
|\sin(h/2(\omega_j - k \cdot \omega)) \cdot \sin(h/2(\omega_j + k \cdot \omega))| \geq \epsilon^{1/2} h^2 |\omega_j + |k \cdot \omega||.
$$

(11)

If this condition is violated, we define a set of near-resonant indices

$$
\mathcal{R}_{\epsilon,h} = \{(j,k) : |j| \leq M, \|k\| \leq 2N, \ k \neq \pm (j), \text{ not satisfying (11)}\},
$$

(12)

where $N \geq 1$ is the truncation number of the expansion (19) which will be presented in the next section.

• There exist $\sigma > 0$ and a constant $C_0$ such that the following non-resonance condition is true

$$
\sup_{(j,k) \in \mathcal{R}_{\epsilon,h}} \frac{\omega_j^\sigma}{\omega_j^{|k|} \epsilon^{\|k\|/2}} \leq C_0 \epsilon^N.
$$

(13)

• The following numerical non-resonance condition is assumed further to be true

$$
|\sin(h\omega_j)| \geq h^{1/2} \text{ for } |j| \leq M.
$$

(14)

• For a positive constant $c > 0$, another non-resonance condition is considered

$$
|\sin(h/2(\omega_j - k \cdot \omega)) \cdot \sin(h/2(\omega_j + k \cdot \omega))| \geq c h^2 |\bar{b}_1(h\omega_j)|
$$

for $(j,k)$ of the form $j = j_1 + j_2$ and $k = \pm \langle j_1 \rangle \pm \langle j_2 \rangle$,

(15)

which leads to improved conservation estimates.

• The ERKN methods are required to satisfy the symmetry conditions (8) and that

$$
\text{sinc}(h\Omega)(2\bar{b}_1(h\Omega))^{-1}\cos(1/2h\Omega) + h^2\Omega^2\text{sinc}(h\Omega)(2\bar{b}_1(h\Omega))^{-1}1/2\text{sinc}(1/2h\Omega) = I,
$$

(16)

which gives that $\bar{b}_1(h\Omega) = \frac{1}{2}\text{sinc}(\frac{1}{2}h\Omega)$ and $b_1(h\Omega) = \cos(h\Omega/2)$.

The main result of this paper is stated as follows.

**Theorem 3** Under the conditions of Assumptions with $s \geq \sigma + 1$, the ERKN method (7) has the near-conservation estimates

$$
\sum_{l=0}^{M} \omega_l^{2s+1} \frac{|I_l(q^n, p^n) - I_l(q^0, p^0)|}{\epsilon^2} \leq C \epsilon,
$$

$$
\frac{|K(q^n, p^n) - K(q^0, p^0)|}{\epsilon^2} \leq C(\epsilon + M^{-s} + \epsilon M^{-s+1}),
$$

$$
\frac{|H_M(q^n, p^n) - H_M(q^0, p^0)|}{\epsilon^2} \leq C \epsilon
$$

for actions, momentum and energy, respectively, where $0 \leq t = nh \leq \epsilon^{-N+1}$. The constant $C$ depends on $s, N$, and $C_0$, but is independent of $\epsilon, M, h$ and the time $t = nh$. The bound $C \epsilon$ is weakened to $C \epsilon^{1/2}$ if condition (15) is not satisfied.
This theorem will be proved in the next two sections by using the technique of multi-frequency modulated Fourier expansions. We will present a multi-frequency modulated Fourier expansion of ERKN methods in Section 4 and then show that the expansion has three almost-invariants in Section 5. On the basis of these analysis, the result of this theorem will be obtained immediately by following the approach in [9, 10, 16, 19], which will be stated at the end of section 5.

3.2 Numerical experiment

As a numerical experiment, we consider the non-linear wave equation (1) with the following data:
\[ \rho = 0.5, \quad g(u) = -u^2, \quad \text{and the initial conditions } u(x, 0) = 0.1\left(\frac{x}{\pi} - 1\right)^3\left(\frac{x}{\pi} + 1\right)^2, \quad \partial_t u(x, 0) = 0.01\left(\frac{x}{\pi} - 1\right)^2\left(\frac{x}{\pi} + 1\right)^2 \quad \text{for } -\pi \leq x \leq \pi. \]
This problem has been considered in [10]. The spatial discretisation is (5) with the dimension \(2M = 2^7\). Consider two ERKN methods whose coefficients are displayed in Table 1. We apply them with the stepsize \(h = 0.5\) to this problem on \([0, 50000]\) and plot the errors of energy, momentum and actions against \(t\) in Figure 1. It follows from the results that the energy, momentum and actions are very well conserved by ERKN 2, which supports the results given in Theorem 3.

| Methods | \(c_1\) | \(b_1(h\Omega)\) | \(b_1(h\Omega)\) | Order | Symmetric | Symplectic |
|---------|---------|-----------------|-----------------|-------|-----------|------------|
| ERKN1   | \(\frac{1}{7}\) | \(\frac{1}{2}\) \text{sinc}^2(h\Omega/2) | \(\cos(h\Omega/2)\) | 2     | Non       | Non        |
| ERKN2   | \(\frac{1}{7}\) | \(\frac{1}{2}\) \text{sinc}(h\Omega/2) | \(\cos(h\Omega/2)\) | 2     | Yes       | Yes        |

Table 1: Two one-stage explicit ERKN methods.

4 Modulated Fourier expansion

Following the techniques and tools developed in [8, 9, 10, 17, 28], the long-time analysis of ERKN methods when applied to wave equation is based on a short-time multi-frequency modulated Fourier expansion, which will be constructed in the first part and proved in the rest parts of this section.
4.1 The result of the expansion

In this section we consider a weaker condition instead of the condition (16):

\[ |b_1(h\Omega)| \leq C|\text{sinc}(h\Omega)| \quad \text{for} \quad |j| \leq M, \quad (18) \]

and set \([|k|] = \begin{cases} (|k| + 1)/2, & k \neq 0, \\ 3/2, & k = 0. \end{cases} \]

**Theorem 4** Under the conditions of Theorem 3 (with the assumption (16) relaxed to (18)), the numerical solution \((q^n, p^n)\) given by the ERKN method (4) admits the following multi-frequency modulated Fourier expansion (with \(N\) from (12))

\[
\tilde{q}_j(t) = \sum_{|k| \leq 2N} e^{i(k-\omega)t}\zeta^k(\epsilon t), \quad \tilde{p}_j(t) = \sum_{|k| \leq 2N} e^{i(k-\omega)t}\eta^k(\epsilon t)
\]

(19)

such that

\[
\|q^n - \tilde{q}_j(t)\|_{s+1} + \|p^n - \tilde{p}_j(t)\|_s \leq C\epsilon^N \quad \text{for} \quad 0 \leq t = nh \leq \epsilon^{-1},
\]

(20)

where we use the notation introduced in (6) with \(k_1 = 0\) for \(l > M\). The expansion (19) is bounded by

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

(21)

It is further obtained that for \(|j| \leq M\),

\[
\tilde{q}_j(t) = \zeta_j^{(j)}(\epsilon t)e^{i\omega_j t} + \zeta_j^{(-j)}(\epsilon t)e^{-i\omega_j t} + \hat{r}_j \quad \text{with} \quad \|r\|_{s+1} \leq C\epsilon^2,
\]

\[
\tilde{p}_j(t) = \eta_j^{(j)}(\epsilon t)e^{i\omega_j t} + \eta_j^{(-j)}(\epsilon t)e^{-i\omega_j t} + \hat{\tilde{r}}_j \quad \text{with} \quad \|\tilde{r}\|_s \leq C\epsilon^{3/2}.
\]

(22)

(If the condition (15) fails to be satisfied, then the bound is \(\|r\|_{s+1} \leq C\epsilon^{3/2}\).) The modulation functions \(\zeta^k\) and \(\eta^k\) are bounded by

\[
\sum_{|k| \leq 2N} \left( \frac{\omega_{|k|}}{\epsilon|k|} \right)^2 \|\zeta^k(\epsilon t)\|_s \leq C, \quad \sum_{|k| \leq 2N} \left( \frac{\omega_{|k|}}{\epsilon|k|} \right)^2 \|\eta^k(\epsilon t)\|_{s-1} \leq C.
\]

(23)

The same bounds hold for any fixed number of derivatives of \(\zeta^k\) and \(\eta^k\) with respect to the slow time \(\tau = \epsilon t\). Moreover, we have \(\zeta_j^{(-k)} = \zeta_j^k\) and \(\eta_j^{(-k)} = \eta_j^k\). The constant \(C\) is independent of \(\epsilon, M, h\) and \(t \leq \epsilon^{-1}\).

This result will be proved in the rest parts of this section by following the approach in [10, 17]. We will just briefly highlight the main differences between this proof and those in [10, 17] and ignore the same derivations for brevity.

4.2 Modulation equations of the modulation functions

According to the first term of (7), we consider the function

\[
q^{n+\frac{1}{2}} := \tilde{q}_h(t + \frac{h}{2}) = \sum_{|k| \leq 2N} e^{i(k-\omega)t}\zeta^k(\epsilon t + \frac{h}{2}\epsilon)
\]

(24)
where

\[ \text{We next prove that there exist multi-frequency modulated Fourier expansions given by (19) such that a factor } \frac{1}{2} \text{ is included in the appearance of } Q^n + \frac{1}{2} \text{ as an ansatz of } \eta^n + \frac{1}{2}. \]

The right-hand side is obtained in a similar way to that in [10, 17]. The prime on the sum indicates that a factor 1/2 is included in the appearance of \( \zeta_{j_i}^k \) or \( \eta_{j_i}^k \) with \( j_i = \pm M \).
Similarly to the analysis in Sect. 6.2 of [10], for the case that \( k = \pm (j) \), the dominating term in (20) is \( \pm 2i \sin(h\omega_j)h\zeta^\pm_j(0) \) due to the fact that the first term vanishes and the condition (14).

When \( k \neq \pm (j) \), the first term in (20) is dominant under the condition (11). If the inequality (11) fails to be true, the non-resonance condition (13) ensures that the defect in simply setting \( \zeta^k_j = 0 \) is of size \( O(\epsilon^{N+1}) \) in an appropriate Sobolev-type norm. Same results hold for \( \eta^k_j \).

By the above analysis and (27), formal modulation equations are obtained for the modulated functions. On the other hand, we need to derive the initial values for the ODEs appearing in the formal modulation equations. From \( \tilde{q}(0) = q^0 \) and \( \tilde{p}(0) = p^0 \), it follows that

\[
\zeta^{(j)}_j(0) + \zeta^{-(j)}_j(0) = q^0_j - \sum_{k \neq \pm (j)} \zeta^k_j(0), \quad \eta^{(j)}_j(0) + \eta^{-(j)}_j(0) = p^0_j - \sum_{k \neq \pm (j)} \eta^k_j(0). \tag{28}
\]

Moreover, by the scheme of the ERKN method, we have

\[
\tilde{q}(t+h) - \tilde{q}(t-h) = 2h\sin(h\omega_j)\tilde{p}(t) + h^2\tilde{b}_1(h\omega_j)[f(\tilde{q}_h(t + \frac{h}{2})) - f(\tilde{q}_h(t - \frac{h}{2})], \tag{29}
\]

\[
\tilde{p}(t+h) - \tilde{p}(t-h) = -2\Omega \sin(h\omega_j)\tilde{q}(t) + h\tilde{b}_1(h\omega_j)[f(\tilde{q}_h(t + \frac{h}{2})) + f(\tilde{q}_h(t - \frac{h}{2})],
\]

which yields that for \( t = 0 \)

\[
\sum_{||k|| \leq 2N} \frac{(c^k_j(e^{i(k \cdot \omega)h} - c^{-k}_j)(-e^{-i(k \cdot \omega)h})}{2h\sin(h\omega_j)} = p^0_j + \frac{h^2\tilde{b}_1(h\omega_j)[f(\tilde{q}_h(\frac{h}{2})) - f(\tilde{q}_h(-\frac{h}{2})]}{2h\sin(h\omega_j)},
\]

\[
\sum_{||k|| \leq 2N} \frac{(\eta^k_j(e^{i(k \cdot \omega)h} - \eta^{-k}_j)(-e^{-i(k \cdot \omega)h})}{-2\omega_j \sin(h\omega_j)} = q^0_j + \frac{h\tilde{b}_1(h\omega_j)[f(\tilde{q}_h(\frac{h}{2})) + f(\tilde{q}_h(-\frac{h}{2})]}{-2\omega_j \sin(h\omega_j)}.
\]

The formulae (28) and (29) determine the initial values for \( \zeta^{\pm(j)}_j \) and \( \eta^{\pm(j)}_j \).

### 4.3 Reverse Picard iteration

In this subsection, we consider an iterative construction of the functions \( \zeta^k_j \) and \( \eta^k_j \) such that after \( 4N \) iteration steps, the defects in (27), (28) and (29) are of magnitude \( O(\epsilon^{N+1}) \) in the \( H^* \) norm.

In the iteration procedure, only the dominant terms are kept on the left-hand side, which is called as reverse Picard iteration by following [10, 17].

Denote the \( n \)th iterate by \([\cdot]^{(n)}\). For \( k = \pm (j) \), the iteration procedure is considered as

\[
\pm 2is_2k\hbar \left[ \zeta^{\pm(j)}_j \right]^{(n+1)} = \left[ -h^2\tilde{b}_1(h\omega_j) \sum_{m \geq 2} \frac{g^{(m)}(0)}{m!} \sum_{k^1 + \cdots + k^m = k_j \pmod{2M} \pmod{j \pmod{2}}} \sum_{j_{m+1} = 1}^{r} \left( (\xi^1_{j_1} \cdots \xi^m_{j_m})(te + \frac{\hbar}{2}e) + (\xi^1_{j_1} \cdots \xi^m_{j_m})(te - \frac{\hbar}{2}e) - (c_2 h^2 e^2 \zeta^{\pm(j)}_j + \cdots) \right)^{(n)}, \right.
\]

\[
\pm 2is_2k\hbar \left[ \eta^{\pm(j)}_j \right]^{(n+1)} = \left[ -h\tilde{b}_1(h\omega_j) \sum_{m \geq 2} \frac{g^{(m)}(0)}{m!} \sum_{k^1 + \cdots + k^m = k_j \pmod{2M} \pmod{j \pmod{2}}} \sum_{j_{m+1} = 1}^{r} \left( (\xi^1_{j_1} \cdots \xi^m_{j_m})(te + \frac{\hbar}{2}e) - (\xi^1_{j_1} \cdots \xi^m_{j_m})(te - \frac{\hbar}{2}e) - (c_2 h^2 e^2 \eta^{\pm(j)}_j + \cdots) \right)^{(n)} \right].
\]
For \( k \neq \pm \langle j \rangle \) and \( j \) satisfying the non-resonant \([11]\), we consider the following iteration procedure

\[
4s_{(j)} + k s_{(j)} - k [\zeta^k_{(j)}]^{(n+1)} = \left[ - h^2 b_1(h\omega_j) \sum_{m \geq 2} g^{(m)}(0) \sum_{\sum_{j=1}^m k_j = k} \sum_{m \equiv j \mod 2M} \right]^{(n)} \\
\left( (\xi^k_{j_1} \cdot \cdots \cdot \xi^k_{j_m}) (te + \frac{h}{2} \epsilon) + (\xi^k_{j_1} \cdot \cdots \cdot \xi^k_{j_m}) (te - \frac{h}{2} \epsilon) - (2is_2h e \xi^k_{j} + c_2h^2e^2 \zeta^k_{j} + \cdots) \right)^{(n)},
\]

(31)

\[
4s_{(j)} + k s_{(j)} - k [\eta^k_{j}]^{(n+1)} = \left[ - h b_1(h\omega_j) \sum_{m \geq 2} g^{(m)}(0) \sum_{\sum_{j=1}^m k_j = k} \sum_{m \equiv j \mod 2M} \right]^{(n)} \\
\left( (\xi^k_{j_1} \cdot \cdots \cdot \xi^k_{j_m}) (te + \frac{h}{2} \epsilon) - (\xi^k_{j_1} \cdot \cdots \cdot \xi^k_{j_m}) (te - \frac{h}{2} \epsilon) - (2is_2h e \eta^k_{j} + c_2h^2e^2 \eta^k_{j} + \cdots) \right)^{(n)}.
\]

For the initial values \([28]\), the iteration procedure becomes

\[
[\zeta^j_{(j)}(0) + \zeta^{-j}_{(j)}(0)]^{(n+1)} = \left[ q^0_j + \sum_{k \neq \pm \langle j \rangle} \zeta^k_{j}(0) \right]^{(n)},
\]

\[
[\eta^j_{j}(0) + \eta^{-j}_{j}(0)]^{(n+1)} = \left[ p^0_j + \sum_{k \neq \pm \langle j \rangle} \eta^k_{j}(0) \right]^{(n)}.
\]

(32)

With regard to the initial values \([29]\), we set

\[
\omega_j \left[ \zeta^j_{(j)}(0) - \zeta^{-j}_{(j)}(0) \right]^{(n+1)} = p^0_j - \frac{1}{2h\sin(h\omega_j)} \left[ \sum_{k \neq \pm \langle j \rangle} \zeta^k_{j}(0) (e^{i(k-\omega)h} - e^{-i(k-\omega)h}) \right]^{(n)}
\]

\[
+ \sum_{\|k\| \leq 2N} \left( (\zeta^k_{j}(eh) - \zeta^k_{j}(0)e^{i(k-\omega)h}) - (\zeta^k_{j}(-eh) - \zeta^k_{j}(0)e^{-i(k-\omega)h}) \right)^{(n)},
\]

\[
\omega_j \left[ \zeta^j_{(j)}(0) - \zeta^{-j}_{(j)}(0) \right]^{(n+1)} = p^0_j + \frac{1}{2\omega_j \sin(h\omega_j)} \left[ \sum_{k \neq \pm \langle j \rangle} \eta^k_{j}(0) (e^{i(k-\omega)h} - e^{-i(k-\omega)h}) \right]^{(n)}
\]

\[
+ \sum_{\|k\| \leq 2N} \left( (\eta^k_{j}(eh) - \eta^k_{j}(0)e^{i(k-\omega)h}) - (\eta^k_{j}(-eh) - \eta^k_{j}(0)e^{-i(k-\omega)h}) \right)^{(n)},
\]

(33)

Following \([10]\), we assume that \( \|k\| \leq K := 2N \) and \( \|k^i\| \leq K \) for \( i = 1, \ldots, m \) in these iterations.

Each iteration step involves a first-order problem of first-order ODEs for \( \zeta^\pm_{j}(\tau) \) and \( \eta^\pm_{j}(\tau) \) (for \( |j| \leq M \)) and algebraic equations for \( \zeta^k_{j} \) and \( \eta^k_{j} \) with \( k \neq \langle j \rangle \). We choose the starting iterates \((n = 0)\) as \( \zeta^k_{j}(\tau) = 0 \) and \( \eta^k_{j}(\tau) = 0 \) for \( k \neq \pm \langle j \rangle \), and \( \zeta^\pm_{j}(\tau) = \zeta^\pm_{j}(0) \) and \( \eta^\pm_{j}(\tau) = \eta^\pm_{j}(0) \), where \( \zeta^\pm_{j}(0) \) and \( \eta^\pm_{j}(0) \) are determined by the above formula. We remark that \( q_{0,j} = q^0_j \) and \( p_{0,j} = p^0_j \) for real initial data, and the above iteration implies \( [\zeta^{-k}_{-j}]^n = [\zeta^{-k}_{-j}]^0 \) and \( [\eta^{-k}_{-j}]^n = [\eta^{-k}_{-j}]^0 \) for all iterates \( n \) and all \( j, k \).
4.4 Rescaling and estimation of the nonlinear terms

Following Sect. 3.5 of [9] and Sect. 6.3 of [10], this subsection considers a more convenient rescaling

\[ c_\varepsilon^k = \frac{\omega^{[k]}}{\epsilon^{[k]}} \varepsilon^k, \quad c_\varepsilon k = (c_\varepsilon^k)_{|j| \leq M} = \frac{\omega^{[k]}}{\epsilon^{[k]}} \varepsilon^k \]

in the space \( H^* = (H^*)^K = \{ c_\varepsilon = (c_\varepsilon^k)_{k \in K} : c_\varepsilon^k \in H^* \} \). The norm of this space is chosen as \( ||| c_\varepsilon |||_2^2 = \sum_{k \in K} ||| c_\varepsilon^k |||_s^2 \) and the superscripts \( k \) are in the set \( K = \{ k = (k_l)_{l=0}^M \} \) with integers \( k_l : \|k\| \leq K \) with \( K = 2N \). Similarly, we have the notations \( c_0^k \in H^{*+1} \) and \( c_\varepsilon^k \in H^* \) with the same meaning.

In order to express the non-linearity in \( \hat{f}_j \) in these rescaled variables, we define the nonlinear function \( f = (f^k_j) \) by

\[ f^k_j(c_\varepsilon^k(\tau \pm \frac{h}{2} \epsilon)) = \frac{\omega^{[k]}}{\epsilon^{[k]}} \sum_{m \geq 2} g^{(m)}(0) \sum_{k_1+\cdots+k_m=k} \frac{\epsilon([k_1]+\cdots+[k_m])}{\omega^{([k_1]+\cdots+[k_m])}} \sum_{j_1+\cdots+j_m=j \mod 2M} (c_\varepsilon^{k_1}_{j_1} \cdots c_\varepsilon^{k_m}_{j_m})(\tau \pm \frac{h}{2} \epsilon). \]

By the results given in Sect. 3.5 of [9] and Sect. 6.4 of [10], it is easy to verify that

\[ \sum_{k \in K} \left| \left| f^k_*(c_\varepsilon^k) \right| \right|_{s,0}^2 \leq \epsilon P(|||c_\varepsilon^k|||_2^2), \quad \sum_{|j| \leq M} \left| \left| f^{\pm j}(c_\varepsilon^k) \right| \right|_{s,1}^2 \leq \epsilon^3 P_1(|||c_\varepsilon^k|||_2^2), \] \hspace{1cm} (34)

where \( P \) and \( P_1 \) are polynomials with coefficients bounded independently of \( \epsilon, h, \) and \( M \).

Likewise, define the following different rescaling

\[ \hat{c}_\varepsilon^k = \frac{\omega^{[k]}}{\hat{\epsilon}^{[k]}} \hat{\varepsilon}^k, \quad \hat{c}_\varepsilon^k = (\hat{c}_\varepsilon^k)_{|j| \leq M} = \frac{\omega^{[k]}}{\hat{\epsilon}^{[k]}} \hat{\varepsilon}^k \] \hspace{1cm} (35)

in \( H^1 = (H^1)^K \) with norm \( \|\| \hat{c}_\varepsilon^k \|\|_1^2 = \sum_{|k| \leq K} \left| \left| \hat{c}_\varepsilon^k \right| \right|_1^2 \), where \( \hat{f}_j^k \) is defined as \( f_j^k \) but with \( \omega^{[k]} \) replaced by \( \omega^{[k]} \). We have similar notations \( \hat{c}_0^k \in H^0 \) and \( \hat{c}_\varepsilon^k \in H^1 \), and similar bounds

\[ \sum_{k \in K} \left| \left| \hat{f}_j^k(\hat{c}_\varepsilon^k) \right| \right|_1^2 \leq \hat{\epsilon} \hat{P}(\|\|\hat{c}_\varepsilon^k\|\|_1^2), \quad \sum_{|j| \leq M} \left| \left| \hat{f}^{\pm j}(\hat{c}_\varepsilon^k) \right| \right|_1^2 \leq \hat{\epsilon}^3 \hat{P}_1(\|\|\hat{c}_\varepsilon^k\|\|_1^2) \]

with other polynomials \( \hat{P} \) and \( \hat{P}_1 \).

4.5 Reformulation of the reverse Picard iteration

Consider \( c_\varepsilon = (c_\varepsilon^k) \in H^* \) with \( c_\varepsilon^k = 0 \) for all \( k \neq \pm (j) \) with \( (j, k) \in \mathcal{R}_{+, h} \). In the light of the two cases: \( k = \pm (j) \) and \( k \neq \pm (j) \), the components of \( c_\varepsilon \) are split into \( a_\varepsilon = (a_\varepsilon^k) \in H^* \) and \( b_\varepsilon = (b_\varepsilon^k) \in H^* \), respectively:

\[
\begin{align*}
    a_\varepsilon^j & = c_\varepsilon^j, & \text{if } k = \pm (j), & \text{and 0 else,} \\
    b_\varepsilon^j & = c_\varepsilon^j, & \text{if } (11) & \text{is satisfied, and 0 else.}
\end{align*}
\]
It is noted that $a\zeta + b\zeta = c\zeta$ and $\|a\zeta\|^2 + \|b\zeta\|^2 = \|c\zeta\|^2$. The same property holds for the denotations $c_n$ and $c_\xi$.

Define differential operators $A, B$ respectively as

\[
(Aa\zeta)^{(j)}(\tau) = \frac{1}{\pm 2i \omega_j h e} (c_{2j} h^2 e_a a^{\pm(j)} + \cdots),
\]

\[
(Bb\zeta)^{(j)}(\tau) = \frac{1}{4 s_{(j)+k s_{(j)-k}} (2i \omega_j h e b^{\pm(j)} + c_{2j} h^2 e_b b^{\pm(j)} + \cdots) \text{ for } (j, k) \text{ satisfying } (11).
\]

We have the same notation as $Aa\zeta$ and $Bb\zeta$. According to the nonlinear function $f$ of the preceding subsection, we define the functions $F = (F_j), \tilde{F} = (\tilde{F}_j)$ and $G = (G_j), \tilde{G} = (\tilde{G}_j)$ with non-vanishing entries for $(j, k)$ satisfying $(11)$:

\[
F_j^{(j)}(a\zeta, b\zeta, a\eta, b\eta) = \frac{1}{\pm i \sigma_j \sinh(h\omega_j)} f_j^{(j)} (c\zeta(\tau + \frac{h}{2}e) + c\zeta(\tau - \frac{h}{2}e)),
\]

\[
\tilde{F}_j^{(j)}(a\zeta, b\zeta, a\eta, b\eta) = \frac{1}{\pm i \sigma_j \sinh(h\omega_j)} f_j^{(j)} (c\zeta(\tau + \frac{h}{2}e) - c\zeta(\tau - \frac{h}{2}e)),
\]

\[
G_j^{(j)}(a\zeta, b\zeta, a\eta, b\eta) = -\frac{h^2 (\omega_j + |k \cdot \omega|)}{4 s_{(j)+k s_{(j)-k}}} f_j^{(j)} (c\zeta(\tau + \frac{h}{2}e) + c\zeta(\tau - \frac{h}{2}e)),
\]

\[
\tilde{G}_j^{(j)}(a\zeta, b\zeta, a\eta, b\eta) = -\frac{h (\omega_j + |k \cdot \omega|)}{4 s_{(j)+k s_{(j)-k}}} f_j^{(j)} (c\zeta(\tau + \frac{h}{2}e) - c\zeta(\tau - \frac{h}{2}e)).
\]

Furthermore, let

\[
(\Omega x_j)^k = (\omega_j + |k \cdot \omega|) x_j^k, \quad (\tilde{b}_1(h\Omega) x_j)^k = \tilde{b}_1(h\omega_j) x_j^k, \quad (b_1(h\Omega) x_j)^k = b_1(h\omega_j) x_j^k.
\]

The iterations $(30)$ and $(31)$ then have the form

\[
a^{(n+1)} = \Omega^{-1} F(a^n, b^n, a^n, b^n) - Aa^n,
\]

\[
a^{(n+1)} = \Omega^{-1} \tilde{F}(a^n, b^n, a^n, b^n) - Aa^n,
\]

\[
b^{(n+1)} = \Omega^{-1} b_1(h\Omega) G(a^n, b^n, a^n, b^n) - Bb^n,
\]

\[
b^{(n+1)} = \Omega^{-1} b_1(h\Omega) \tilde{G}(a^n, b^n, a^n, b^n) - Bb^n.
\]

From the second formula of $(30)$ and condition $(13)$, it follows that $\|F\|_s \leq C \epsilon^{1/2}$. Moreover, according to $(11)$, we have $\|\tilde{b}_1(h\Omega)^{-1} \Omega^{-1} F\|_s \leq C$. In the light of the fact that $\|c\zeta(\tau + \frac{h}{2}e) - c\zeta(\tau - \frac{h}{2}e)\|_s \leq C e h$, and the analysis given in Sect.3.5, we obtain

\[
\sum_{k \in \mathcal{K}} \left\| f^k (c\zeta(\tau + \frac{h}{2}e) - c\zeta(\tau - \frac{h}{2}e)) \right\|_s^2 \leq C e^3 h^2,
\]

\[
\sum_{|j| \leq M} \left\| f^{(j)} (c\zeta(\tau + \frac{h}{2}e) - c\zeta(\tau - \frac{h}{2}e)) \right\|_s^2 \leq C e^3 h^2.
\]

From $(37)$ and $(11)$, we have $\|\tilde{F}\|_{s-1} \leq C \epsilon$ and $\|b_1(h\Omega)^{-1} \Omega^{-1} \tilde{F}\|_s \leq C \epsilon$. By $(11)$, one gets similar bounds $\|G\|_s \leq C$ and $\|G\|_s \leq C \epsilon$, which are valid uniformly in $\epsilon, h, M$ on bounded
subsets of $H^s$. For the derivatives of $F$, $\hat{F}$ and $G$, $\hat{G}$, the analogous bounds hold. The operators $A$ and $B$ are bounded as (Sect. 6.5 of \cite{10}):

$$
|||(Aa\zeta)(\tau)|||_s \leq C \sum_{l=2}^{N} \epsilon^{l-3/2} \left| \frac{d^l}{d\tau^l}(a\zeta)(\tau) \right|_s,
$$

$$
|||(Bb\zeta)(\tau)|||_s \leq C \epsilon^{1/2} |||((b\zeta)(\tau)|||_s + C \sum_{l=2}^{N} \epsilon^{l-1/2} \left| \frac{d^l}{d\tau^l}(b\zeta)(\tau) \right|_s. \quad (38)
$$

By the definitions of $a\zeta$ and $a\eta$, the initial value conditions (32)-(33) can be rewritten as

$$
\begin{align*}
&a^{(n+1)}\zeta(0) = v + Pb^{(n)}(0) + Q(a\zeta + b\zeta)(ch) + \Psi_1(c\xi)(ch), \\
&a^{(n+1)}\eta(0) = \hat{v} + Pb\eta^{(n)}(0) + Q(a\eta + b\eta)(ch) + \Psi_2(c\xi)(ch),
\end{align*}
$$

where $v$ and $\hat{v}$ are defined by, respectively:

$$
\begin{align*}
v^{\pm(j)}_j &= \frac{\omega_j}{\epsilon} \left( \frac{1}{2} \nu_j^0 + \frac{1}{2N} \tilde{\omega}_j^0 \right), \\
\hat{v}^{\pm(j)}_j &= \frac{\omega_j}{\epsilon} \left( \frac{1}{2} \nu_j^0 + \frac{\omega_j}{2N} \tilde{\omega}_j^0 \right).
\end{align*}
$$

(From (10), it follows that $v$ is bounded in $H^s$ and $\hat{v}$ is bounded in $H^{s-1}$.) The operators $P$ and $Q$ have been used in Sect. 6.5 of \cite{10}, which are defined by

$$
\begin{align*}
(Pb\zeta)^{\pm(j)}_j(0) &= -\frac{\omega_j}{2\epsilon \nu_j} \sum_{k \neq \pm j} \left( \sin(\omega_j h) \pm \sin((k \cdot \omega) h) \right) \frac{\epsilon^{[|k|]}}{\omega^{[|k|]}} b\zeta^k(0), \\
(Qc\zeta)^{\pm(j)}_j(\tau) &= \mp \frac{\omega_j}{4 \epsilon \nu_j} \sum_{\|k\| \leq K} \left( e^{i(k \cdot \omega)h} \frac{\epsilon^{[|k|]}}{\omega^{[|k|]}} (c\zeta^k_\tau - c\zeta^k_0) \\
&\quad - e^{-i(k \cdot \omega)h} \frac{\epsilon^{[|k|]}}{\omega^{[|k|]}} (c\zeta^k(-\tau) - c\zeta^k_0) \right).
\end{align*}
$$

The operators $\Psi_1$ and $\Psi_2$ are defined by

$$
\begin{align*}
\Psi_1(c\xi)(ch) &= \mp \frac{h^2 b_1(h \omega_j)}{4i \sin(h \omega_j)} f^{\pm(j)}_j (c\xi(\frac{h}{2}) - c\xi(-\frac{h}{2})), \\
\Psi_2(c\xi)(ch) &= \mp \frac{h b_1(h \omega_j)}{4i \sin(h \omega_j)} f^{\pm(j)}_j (c\xi(\frac{h}{2}) + c\xi(-\frac{h}{2})).
\end{align*}
$$

The bounds $|||(Pb\zeta)(0)|||_s \leq C$ and $|||(Qc\zeta)(0)|||_s \leq C\epsilon$ are given in Sect. 6.5 of \cite{10}, where $C$ is independent of $\epsilon, h$ and $M$, but depends on $K = 2N$. By the second formulæ of (37) and (34), the condition (14) gives the bounds $|||\Psi_1(c\xi)(ch)|||_s \leq C\epsilon h$ and $|||\Psi_2(c\xi)(ch)|||_s \leq C\epsilon$. The starting iterates of (39) are chosen as $a^{(0)}\zeta(\tau) = v$, $\eta^{(0)}(\tau) = \hat{v}$, and $b\zeta^{(0)}(\tau) = 0$, $b\eta^{(0)}(\tau) = 0$.

4.6 Bounds of the coefficient functions

By the non-resonance conditions (11) and (14), and the assumption (15), it is obtained by induction that the iterates $a\zeta^{(n)}$, $b\zeta^{(n)}$ and $a\eta^{(n)}$, $b\eta^{(n)}$ and their derivatives with respect to the slow time
\( \tau = ct \) are bounded in \( H^s \) and \( H^{s-1} \) for \( 0 \leq \tau \leq 1 \) and \( n \leq 4N \), respectively. The \((4N)\)-th iterates satisfy
\[
\|a\dot{\zeta}(0)\|_s \leq C, \quad \|\Omega a\dot{\zeta}(\tau)\|_s \leq Ce^{1/2}, \quad \|b_1(h\Omega)^{-1}a\dot{\zeta}(\tau)\|_s \leq C,
\]
and
\[
\|a\eta(0)\|_{s-1} \leq C, \quad \|\Omega a\eta(\tau)\|_{s-1} \leq Ce^{1/2}, \quad \|b_1(h\Omega)^{-1}a\eta(\tau)\|_{s-1} \leq Ce^{1/2},
\]
where \( C \) is independent of \( \epsilon, h, M, \) but depends on \( N \). The analogous bounds for higher derivatives of these functions with respect to \( \tau \) can also be obtained. These bounds together imply the bounds \((23)\). According to these bounds, it is also true that \( \|c\dot{\zeta}(\tau) - a\dot{\zeta}(0)\|_{s+1} \leq C \) and \( \|\eta(\tau) - a\eta(0)\|_{s} \leq C \). By Sect. 3.7 of \([9]\), Sect. 6.6 of \([10]\) and \((38)\), it can be confirmed that the bound \((21)\) is true.

From \((31)\) and \((36)\), it follows that
\[
\left( \sum_{\|k\|=1} \left\| (b_1(h\Omega)^{-1}\Omega b\dot{\zeta})^k \right\|_s \right)^{1/2} \leq C\epsilon \quad \text{for} \quad b\dot{\zeta} = (b\dot{\zeta})^{(4N)}.
\]
Moreover, in the light of \((15)\), one arrives that
\[
\sum_{|j| \leq M} \sum_{j_1+j_2=j} \sum_{k=\pm(j_1)\pm(j_2)} \omega_j^{2(s+1)} |b\dot{\zeta}_j^k|^2 \leq C\epsilon.
\]
These bounds as well as \((39)\) yield the first formula of \((22)\). Similarly, according to \((37)\) and \((36)\), we have
\[
\left( \sum_{\|k\|=1} \left\| (b_1(h\Omega)^{-1}\Omega b\eta)^k \right\|_{s-1} \right)^{1/2} \leq C\epsilon^2 \quad \text{and}
\]
\[
\sum_{|j| \leq M} \sum_{j_1+j_2=j} \sum_{k=\pm(j_1)\pm(j_2)} \omega_j^{2(s+1)} |b\eta_j^k|^2 \leq C\epsilon^2,
\]
which lead to the second statement of \((22)\).

The same bounds can be obtained for the alternative scaling \((55)\):
\[
\|a\dot{\zeta}(0)\|_1 \leq C, \quad \|\Omega a\dot{\zeta}(\tau)\|_1 \leq Ce^{1/2}, \quad \|b_1(h\Omega)^{-1}\Omega b\dot{\zeta}(\tau)\|_1 \leq C,
\]
\[
\|a\eta(0)\|_0 \leq C, \quad \|\Omega a\eta(\tau)\|_0 \leq Ce^{1/2}, \quad \|b_1(h\Omega)^{-1}\Omega b\eta(\tau)\|_0 \leq Ce^{1/2}.
\]
Moreover, the following bounds hold for this scaling:
\[
\left( \sum_{\|k\|=1} \left\| (b_1(h\Omega)^{-1}\Omega b\dot{\zeta})^k \right\|_1 \right)^{1/2} \leq C\epsilon, \quad \left( \sum_{\|k\|=1} \left\| (b_1(h\Omega)^{-1}\Omega b\eta)^k \right\|_0 \right)^{1/2} \leq C\epsilon^2.
\]

### 4.7 Defects

The defect in \((7)\) can be expressed as
\[
\delta_j(t) = \frac{\tilde{q}_j(t+h) - 2\cos(h\omega_j)\tilde{q}_j(t) + \tilde{q}_j(t-h)}{h^2 b_1(h\omega_j)} - \left[ f_j(\tilde{q}_h(t + \frac{h}{2})) + f_j(\tilde{q}_h(t - \frac{h}{2})) \right],
\]
\[
\delta_j(t) = \frac{\tilde{p}_j(t+h) - 2\cos(h\omega_j)\tilde{p}_j(t) + \tilde{p}_j(t-h)}{h b_1(h\omega_j)} - \left[ f_j(\tilde{q}_h(t + \frac{h}{2})) - f_j(\tilde{q}_h(t - \frac{h}{2})) \right],
\]
where \( f = (f_j) \) is defined by (1) and \( \tilde{q}(t) = (\tilde{q}_j(t)) \) and \( \tilde{p}(t) = (\tilde{p}_j(t)) \) are defined by (14) with \( \zeta_j^k = (\zeta_j^k)^{(4N)} \) and \( \eta_j^k = (\eta_j^k)^{(4N)} \) obtained after \( 4N \) iterations of the procedure in Subsection 4.3. These defects can be rewritten as

\[
\delta_j(t) = \sum_{\|k\| \leq NK} d_j^k(\epsilon t)e^{i(k\omega)t} + R_{N+1}(t),
\]

where

\[
\begin{align*}
\delta_j^*(t) &= \sum_{\|k\| \leq NK} d_j^k(\epsilon t)e^{i(k\omega)t} + R_{N+1}'(t),
\end{align*}
\]

and the remainder terms of the Taylor expansion of \( f \) after the \( \epsilon^N \) term. The functions \( R_{N+1} \) and \( R_{N+1}' \) contain the remainder terms of the Taylor expansion of \( f \) after \( N \) terms. We obtain \( \|R_{N+1}\|_s \leq C\epsilon^{N+1} \) and \( \|R_{N+1}'\|_s \leq C\epsilon^{N+1} \) by the bound (21) for the remainder in the Taylor expansion of \( f \) and the estimates (40) and (41) for the \( (N+1) \)-th derivative for \( \zeta_j^k \) and \( \eta_j^k \).

According to the analysis given in Sect. 3.8 of [9] and Sect. 6.7 of [10], it is obtained that

\[
\begin{align*}
\left\| \sum_{\|k\| \leq NK} d_j^k(\epsilon t)e^{i(k\omega)t} \right\|_s^2 \leq C \left\| \omega^{|k|} d_j^k(\epsilon t) \right\|_s^2, \\
\left\| \sum_{\|k\| \leq NK} d_j^k(\epsilon t)e^{i(k\omega)t} \right\|_s^2 \leq C \left\| \omega^{|k|} d_j^k(\epsilon t) \right\|_s^2.
\end{align*}
\]

In what follows, we prove the bounds

\[
\sum_{\|k\| \leq NK} \left\| \omega^{|k|} d_j^k(\epsilon t) \right\|_s^2 \leq C\epsilon^{2(N+1)}, \quad \sum_{\|k\| \leq NK} \left\| \omega^{|k|} d_j^k(\epsilon t) \right\|_s^2 \leq C\epsilon^{2(N+1)}
\]

for three cases: truncated, near-resonant and non-resonant modes, which will be shown by the following three parts.
4.7.1 Truncated mode

It is known that we set \( \xi_j^k = \eta_j^k = 0 \) for \( \|k\| > K := 2N \) (truncated modes) and for \( (j, k) \in \mathcal{R}_{\epsilon, h} \) (near-resonance modes). Thence for these both cases, the defect has the form

\[
d_j^k = \sum_{m \geq 2} \frac{g^{(m)}(0)}{m!} \sum_{k^1 + \cdots + k^m = k} \sum_{j_1, \ldots, j_m} \left[ (\xi_{j_1}^{k_1} \cdots \xi_{j_m}^{k_m})(\tau + \frac{h}{2}\epsilon) + (\xi_{j_1}^{k_1} \cdots \xi_{j_m}^{k_m})(\tau - \frac{h}{2}\epsilon) \right],
\]

\[
d_j^{k'} = \sum_{m \geq 2} \frac{g^{(m)}(0)}{m!} \sum_{k^1 + \cdots + k^m = k} \sum_{j_1, \ldots, j_m} \left[ (\xi_{j_1}^{k_1} \cdots \xi_{j_m}^{k_m})(\tau + \frac{h}{2}\epsilon) - (\xi_{j_1}^{k_1} \cdots \xi_{j_m}^{k_m})(\tau - \frac{h}{2}\epsilon) \right].
\]

For truncated modes we express the defect as

\[
d_j^k = e^{\|k\|} \omega^{-|k|} f_j^k (c\xi(\frac{h}{2}\epsilon) + c\xi(\frac{-h}{2}\epsilon)), \quad d_j^{k'} = e^{|k|} \omega^{-|k|} f_j^{k'} (c\xi(\frac{h}{2}\epsilon) - c\xi(\frac{-h}{2}\epsilon)).
\]

On the basis of \([34],[40]\) and \([41]\) with \( NK \) instead of \( K \), we have the bound \( \| ||f|| \|_{s}^2 \leq C\epsilon \), which yields

\[
\sum_{\|k\| > K} \sum_{|j| \leq M} \omega_j^{2s} |\omega^{|k|} d_j^k|^2 \leq \sum_{\|k\| > K} \sum_{|j| \leq M} \omega_j^{2s} |f_j^k|^2 \epsilon^{2|k|} \leq C\epsilon^{2(N+1)},
\]

\[
\sum_{\|k\| > K} \sum_{|j| \leq M} \omega_j^{2s} |\omega^{|k|} d_j^{k'}|^2 \leq \sum_{\|k\| > K} \sum_{|j| \leq M} \omega_j^{2s} |f_j^{k'}|^2 \epsilon^{2|k|} \leq C\epsilon^{2(N+1)},
\]

where we have used the fact that \( 2|k| = \|k\| + 1 \geq K + 2 = 2(N + 1) \).

4.7.2 Near-resonant mode

For the near-resonant modes we express the defect by the rescaling \([55]\) as follows

\[
d_j^k = e^{\|k\|} \omega^{-s|k|} f_j^k (c\xi(\frac{h}{2}\epsilon) + c\xi(\frac{-h}{2}\epsilon)), \quad d_j^{k'} = e^{\|k\|} \omega^{-s|k|} f_j^{k'} (c\xi(\frac{h}{2}\epsilon) - c\xi(\frac{-h}{2}\epsilon)).
\]

Therefore, it is obtained that

\[
\sum_{(j,k) \in \mathcal{R}_{\epsilon, h}} \omega_j^{2s} |\omega^{|k|} d_j^k|^2 = \sum_{(j,k) \in \mathcal{R}_{\epsilon, h}} \frac{\omega_j^{2(s-1)}}{\omega^{2(s-1)|k|}} \epsilon^{2|k|} \omega_j^{2|k|} |f_j^k|^2 \leq C \sup_{(j,k) \in \mathcal{R}_{\epsilon, h}} \frac{\omega_j^{2(s-1)}}{\omega^{2(s-1)|k|}} \epsilon^{2|k|} \leq C\epsilon^{2(N+1)}
\]

by considering \( \| ||f|| \|_{s}^2 \leq C\epsilon \) and the non-resonance condition \([13]\). In a same way, we arrive at

\[
\sum_{(j,k) \in \mathcal{R}_{\epsilon, h}} \omega_j^{2s} |\omega^{|k|} d_j^{k'}|^2 \leq C\epsilon^{2(N+1)},
\]

16
4.7.3 Non-resonant mode

Now consider the non-resonant mode, which means that $\|k\| > K$ and that $(j, k)$ satisfies the non-resonance condition (11). For this case, the defect is formulated in the scaled variables of Subsection 4.4 as

$$\omega^{[k]} d_j^k = e^{[k]} \left( \frac{1}{h^2 b_1 (h \omega_j)} \tilde{L}^k c_j^k + f_j^k (c \xi (\frac{h}{2} \epsilon) + c \xi (\frac{-h}{2} \epsilon)) \right),$$

$$\omega^{[k]} d_j^k = e^{[k]} \left( \frac{1}{h \omega_j} \tilde{L}^k c_n^k + f_j^k (c \xi (\frac{h}{2} \epsilon) - c \xi (\frac{-h}{2} \epsilon)) \right).$$

Split them into $k = \pm (j)$ and $k \neq \pm (j)$ and we obtain

$$\omega_j d_j^{\pm (j)} = e^{(\pm (j))} \left( \frac{1}{2} + \frac{2 \epsilon \omega_j \sin (h \omega_j)}{b_1 (h \omega_j)} (a_j^{\pm (j)} + (A \alpha_j)^{\pm (j)}) + f_j^{\pm (j)} (c \xi (\frac{h}{2} \epsilon) + c \xi (\frac{-h}{2} \epsilon)) \right),$$

$$\omega_j d_j^{\pm (j)} = e^{(\pm (j))} \left( \frac{1}{2} - \frac{2 \epsilon \omega_j \sin (h \omega_j)}{b_1 (h \omega_j)} (a_j^{\pm (j)} - (A \alpha_j)^{\pm (j)}) + f_j^{\pm (j)} (c \xi (\frac{h}{2} \epsilon) - c \xi (\frac{-h}{2} \epsilon)) \right),$$

$$\omega_j d_j^{\pm (j)} = e^{(\pm (j))} \left( \frac{1}{2} + \frac{2 \epsilon \omega_j \sin (h \omega_j)}{b_1 (h \omega_j)} (b_j^{\pm (j)} + (B \beta_j)^{\pm (j)}) + f_j^{\pm (j)} (c \xi (\frac{h}{2} \epsilon) + c \xi (\frac{-h}{2} \epsilon)) \right),$$

$$\omega_j d_j^{\pm (j)} = e^{(\pm (j))} \left( \frac{1}{2} - \frac{2 \epsilon \omega_j \sin (h \omega_j)}{b_1 (h \omega_j)} (b_j^{\pm (j)} - (B \beta_j)^{\pm (j)}) + f_j^{\pm (j)} (c \xi (\frac{h}{2} \epsilon) - c \xi (\frac{-h}{2} \epsilon)) \right),$$

We remark that the functions here are actually the $4N$-th iterates of the iteration in Subsection 4.6 and inserting them from (36) into this defect, we obtain

$$\omega_j d_j^{\pm (j)} = 2 \omega_j a_j^{\pm (j)} \left( \frac{1}{2} + \frac{2 \epsilon \omega_j \sin (h \omega_j)}{b_1 (h \omega_j)} (a_j^{\pm (j)} + (A \alpha_j)^{\pm (j)}) \right),$$

$$\omega_j d_j^{\pm (j)} = 2 \omega_j a_j^{\pm (j)} \left( \frac{1}{2} - \frac{2 \epsilon \omega_j \sin (h \omega_j)}{b_1 (h \omega_j)} (a_j^{\pm (j)} - (A \alpha_j)^{\pm (j)}) \right),$$

It is noted that these expressions are very similar to those in Sect. 6.9 of [10]. Thus, following the analysis given in [10], for $\tau \leq 1$, we have

$$\left( \sum_{\|k\| \leq K} \left\| \omega^{[k]} d^{k}(\tau) \right\|_s^2 \right)^{1/2} \leq Ce^{N+1}, \quad \left( \sum_{\|k\| \leq K} \left\| \omega^{[k]} d^{k}(\tau) \right\|_s^2 \right)^{1/2} \leq Ce^{N+1}. \quad (46)$$

Then the defect (44) becomes $\|\delta(t)\|_s \leq Ce^{N+1}$ and $\|\delta'(t)\|_s \leq Ce^{N+1}$ for $t \leq \epsilon^{-1}$. Considering the defect in the initial conditions (28) and (29), one has

$$\|q^0 - \tilde{q}(0)\|_{s+1} + \|p^0 - \tilde{p}(0)\|_s \leq Ce^{N+1}.$$
4.8 Remainders

Let \( \Delta q^n = \bar{q}(t_n) - q^n \), \( \Delta p^n = \bar{p}(t_n) - p^n \). We then have

\[
\begin{pmatrix}
\Delta q^{n+1} \\
\Omega^{-1} \Delta p^{n+1}
\end{pmatrix} =
\begin{pmatrix}
\cos(h\Omega) & \sin(h\Omega) \\
-\sin(h\Omega) & \cos(h\Omega)
\end{pmatrix}
\begin{pmatrix}
\Delta q^n \\
\Omega^{-1} \Delta p^n
\end{pmatrix}
+ \frac{h}{b_1(h\Omega)} \Omega^{-1}(\Delta f + \delta) + \frac{h}{b_1(h\Omega)} \Omega^{-1}(\Delta f + \delta'),
\]

where \( \Delta f = \left( f(\bar{q}(t_n + \frac{h}{2})) - f(q^{n+\frac{1}{2}}) \right) \). It follows from the Lipschitz bound (Sect. 4.2 in [17] and Sect. 6.10 in [10]) that

\[
\left\| \Omega^{-1} \Delta f \right\|_{s+1} = \left\| \Delta f \right\|_s \leq \epsilon \left\| \bar{q}(t_n + \frac{h}{2}) - q^{n+\frac{1}{2}} \right\|_s \leq \epsilon (\left\| \Delta q^n \right\|_s + \left\| \Delta p^n \right\|_{s-1}).
\]

Moreover, we have \( \left\| \Omega^{-1} \delta(t) \right\|_{s+1} = \left\| \delta(t) \right\|_s \leq C \epsilon^{N+1} \) and \( \left\| \Omega^{-1} \delta'(t) \right\|_{s+1} = \left\| \delta'(t) \right\|_s \leq C \epsilon^{N+1} \). Therefore, we obtain

\[
\left\| \begin{pmatrix}
\Delta q^{n+1} \\
\Omega^{-1} \Delta p^{n+1}
\end{pmatrix} \right\|_{s+1} \leq \left( \left\| \Delta q^n \right\|_{s+1} + \epsilon \left\| \Delta p^n \right\|_{s+1} + C \epsilon^{N+1} \right) + h \left( C \epsilon \left\| \Delta q^n \right\|_{s+1} + C \epsilon \left\| \Delta p^n \right\|_{s+1} + C \epsilon^{N+1} \right).
\]

Solving this inequality leads to \( \left\| \Delta q^n \right\|_{s+1} + \left\| \Omega^{-1} \Delta p^n \right\|_{s+1} \leq C(1 + t_n) \epsilon^{N+1} \) for \( t_n \leq \epsilon^{-1} \). This proves (20) and completes the proof of Theorem 4.

5 Long time near conservations

In this section, we will show that the modulated Fourier expansion [10] has three almost-invariants which are closed to the actions, the momentum and the total energy of (5).

According to the proof of Theorem 4, the defect formula (45) can be reformulated as

\[
\frac{1}{h^2 b_1(h\omega_j)} \tilde{L}_j \zeta^k_j + \nabla^{-k}_j \left[ \mathcal{U}(\xi(t + \frac{h}{2}\epsilon)) + \mathcal{U}(\xi(t - \frac{h}{2}\epsilon)) \right] = d^k_j,
\]

(48)

\[
\frac{1}{h b_1(h\omega_j)} \tilde{L}^k_j n^k_j + \nabla^{-k}_j \left[ \mathcal{U}(\xi(t + \frac{h}{2}\epsilon)) - \mathcal{U}(\xi(t - \frac{h}{2}\epsilon)) \right] = d^k_j,
\]

where \( \nabla^{-k}_j \mathcal{U}(y) \) is the partial derivative with respect to \( y^{-k}_j \) of the extended potential (10) (17)

\[
\mathcal{U}(\xi(t \pm \frac{h}{2}\epsilon)) = \sum_{l=-N}^{N} \mathcal{U}_l(\xi_l(t \pm \frac{h}{2}\epsilon)),
\]

\[
\mathcal{U}_l(\xi(t \pm \frac{h}{2}\epsilon)) = \frac{U^{(m+1)}(0)}{(m + 1)!} \sum_{\kappa^1 + \ldots + \kappa^{m+1} = 0} \sum_{j_1 + \ldots + j_{m+1} = 2M} \left( \zeta^{k^1}_{j_1} \ldots \zeta^{k^{m+1}}_{j_{m+1}} \right)(t \epsilon \pm \frac{h}{2}\epsilon).
\]

Here \( U \) is the potential appearing in (2) and \( \|k^l\| \leq 2N \) and \( |j_l| \leq M \).

Define \( S_\mu(\theta)y = (e^{i(\kappa_\mu)^T y})_\mu \|k\| \leq K \) and \( T(\theta)y = (e^{i\theta y})_\mu \|k\| \leq K \), where \( \mu = (\mu_l)_l \geq 0 \) is an arbitrary real sequence for \( \theta \in \mathbb{R} \). It is easy to see that the summation in the definition of \( \mathcal{U} \) is
over $k^1 + \ldots + k^{m+1} = 0$ and that in $U_0$ over $j_1 + \ldots + j_{m+1} = 0$, which yield $\mathcal{U}(S_\mu(\theta)y) = \mathcal{U}(y)$ and $\mathcal{U}_0(T(\theta)y) = \mathcal{U}_0(y)$ for $\theta \in \mathbb{R}$ (see [10]). Therefore, we have

$$0 = \frac{d}{d\theta} |_{\theta=0} \left[ \mathcal{U}(S_\mu(\theta)\xi(t + \frac{h}{2} \epsilon)) + \mathcal{U}(S_\mu(\theta)\xi(t - \frac{h}{2} \epsilon)) \right],$$

$$0 = \frac{d}{d\theta} |_{\theta=0} \left[ \mathcal{U}_0(T(\theta)\xi(t + \frac{h}{2} \epsilon)) + \mathcal{U}_0(T(\theta)\xi(t - \frac{h}{2} \epsilon)) \right].$$

(49)

### 5.1 Three almost-invariants

By the first formula of (49), we obtain

$$0 = \frac{1}{2} \frac{d}{d\theta} |_{\theta=0} \left[ \mathcal{U}(S_\mu(\theta)\xi(t + \frac{h}{2} \epsilon)) + \mathcal{U}(S_\mu(\theta)\xi(t - \frac{h}{2} \epsilon)) \right]$$

$$= \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i(k \cdot \mu) \left[ \xi_{\mu j}^{-k}(\tau + \frac{h}{2} \epsilon) \nabla_{-j}^{-k} \mathcal{U}(\xi(t + \frac{h}{2} \epsilon)) + \xi_{\mu j}^{-k}(\tau - \frac{h}{2} \epsilon) \nabla_{-j}^{-k} \mathcal{U}(\xi(t - \frac{h}{2} \epsilon)) \right]$$

$$= \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i(k \cdot \mu) \left[ \left( \cos\left(\frac{1}{2} h \omega_j\right) \xi_{\mu j}^{-k}(\tau) + \frac{1}{2} h \text{sinc}\left(\frac{1}{2} h \omega_j\right) \eta_{\mu j}^{-k}(\tau) \right) \nabla_{-j}^{-k} \mathcal{U}(\xi(t + \frac{h}{2} \epsilon)) 

+ \left( \cos\left(\frac{1}{2} h \omega_j\right) \xi_{\mu j}^{-k}(\tau) - \frac{1}{2} h \text{sinc}\left(\frac{1}{2} h \omega_j\right) \eta_{\mu j}^{-k}(\tau) \right) \nabla_{-j}^{-k} \mathcal{U}(\xi(t - \frac{h}{2} \epsilon)) \right].$$

According to (48) and the above formula, we have

$$\frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i(k \cdot \mu) \left[ \cos\left(\frac{1}{2} h \omega_j\right) \xi_{\mu j}^{-k}(\tau) L_j^{k \xi_j} + \frac{1}{2} h \text{sinc}\left(\frac{1}{2} h \omega_j\right) \eta_{\mu j}^{-k}(\tau) L_j^{k \eta_j} \right]$$

$$= \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i(k \cdot \mu) \left[ \cos\left(\frac{1}{2} h \omega_j\right) \xi_{\mu j}^{-k}(\tau) d_j^k + \frac{1}{2} h \text{sinc}\left(\frac{1}{2} h \omega_j\right) \eta_{\mu j}^{-k}(\tau) d_j^k \right].$$

(50)

By the expansion (26) of the operator $L_j^k$ and the analysis given in Sect. 7.3 of [10], we know that $\xi_{\mu j}^{-k}(\tau) L_j^{k \xi_j}$ and $\eta_{\mu j}^{-k}(\tau) L_j^{k \eta_j}$ are total derivatives. Therefore, the left-hand side of (51) is a total derivative of function $e^{f_0(\xi(t), \eta(t))}$ which depends on $\xi(t), \eta(t)$ and their up to $N - 1$th order derivatives. In this way, (51) becomes

$$- \epsilon \frac{d}{dt} \mathcal{J}_\mu(\xi, \eta)(\tau) = \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i(k \cdot \mu) \left[ \cos\left(\frac{1}{2} h \omega_j\right) \xi_{\mu j}^{-k}(\tau) d_j^k + \frac{1}{2} h \text{sinc}\left(\frac{1}{2} h \omega_j\right) \eta_{\mu j}^{-k}(\tau) d_j^k \right].$$

(51)

By the second formula of (10) and in a similar way, we obtain

$$\frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i \left[ \cos\left(\frac{1}{2} h \omega_j\right) \xi_{\mu j}^{-k}(\tau) L_j^{k \xi_j} + \frac{1}{2} h \text{sinc}\left(\frac{1}{2} h \omega_j\right) \eta_{\mu j}^{-k}(\tau) L_j^{k \eta_j} \right]$$

$$= \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i \left[ \cos\left(\frac{1}{2} h \omega_j\right) \xi_{\mu j}^{-k}(\tau) \left( d_j^k - \sum_{l \neq 0} \nabla_{-j}^{-k} \left( \mathcal{U}_l(\xi(t + \frac{h}{2} \epsilon)) + \mathcal{U}_l(\xi(t - \frac{h}{2} \epsilon)) \right) \right) 

+ \frac{1}{2} h \text{sinc}\left(\frac{1}{2} h \omega_j\right) \eta_{\mu j}^{-k}(\tau) \left( d_j^k - \sum_{l \neq 0} \nabla_{-j}^{-k} \left( \mathcal{U}_l(\xi(t + \frac{h}{2} \epsilon)) - \mathcal{U}_l(\xi(t - \frac{h}{2} \epsilon)) \right) \right) \right].$$

(52)
The left-hand side of this can be written as a total derivative of function $\epsilon K[\zeta, \eta](\tau)$ and becomes

$$-\epsilon \frac{d}{d\tau} K[\zeta, \eta](\tau) = \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i^j \left[ \cos \left( \frac{1}{2} \omega_j \right) \zeta^{-k}_j(\tau) \left( d^k_j - \sum_{l \neq 0} \nabla^{-l}_j (U_l(\xi(t + \frac{h}{2} \epsilon)) + U_l(\xi(t - \frac{h}{2} \epsilon))) \right) \right]$$

(53)

Another almost-invariant is obtained by considering

$$\frac{1}{2} \frac{d}{d\tau} \left[ U(\xi(t + \frac{h}{2} \epsilon)) + U(\xi(t - \frac{h}{2} \epsilon)) \right]$$

$$= \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} \left[ \xi^{-k}_j(\tau + \frac{h}{2} \epsilon) \nabla^{-l}_j U(\xi(t + \frac{h}{2} \epsilon)) + \xi^{-k}_j(\tau - \frac{h}{2} \epsilon) \nabla^{-l}_j U(\xi(t - \frac{h}{2} \epsilon)) \right]$$

$$= \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i(k \cdot \mu) \left[ \cos \left( \frac{1}{2} \omega_j \right) \zeta^{-k}_j(\tau) + \epsilon \zeta^{-k}_j(\tau) \right]$$

$$= \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i(k \cdot \mu) \left[ \cos \left( \frac{1}{2} \omega_j \right) (i(k \cdot \omega) \zeta^{-k}_j(\tau) + \epsilon \zeta^{-k}_j(\tau)) d^k_j \right]$$

(54)

Similarly to the analysis in Sect. 7.3 of [10], we know that there is a function $\epsilon H[\zeta, \eta](\tau)$ such that

$$-\epsilon \frac{d}{d\tau} H[\zeta, \eta](\tau) = \frac{1}{2} \sum_{\|k\| \leq K} \sum_{|j| \leq M} i(k \cdot \mu) \left[ \cos \left( \frac{1}{2} \omega_j \right) (i(k \cdot \omega) \zeta^{-k}_j(\tau) + \epsilon \zeta^{-k}_j(\tau)) d^k_j \right]$$

(54)

Consider the special case of $\mu = \langle l \rangle$ for $J_\mu$ and the following result is obtained on noticing the smallness of the right-hand sides in (51), (53) and (54).

**Theorem 5** Under the conditions of Theorem 4 for $\tau \leq 1$, it is true that

$$\sum_{l=1}^{M} \omega^{2s+1} \left| \frac{d}{d\tau} J[\zeta, \eta](\tau) \right| \leq C e^{N+1},$$

$$\left| \frac{d}{d\tau} K[\zeta, \eta](\tau) \right| \leq C (e^{N+1} + e^2 M^{-s+1}),$$

$$\left| \frac{d}{d\tau} H[\zeta, \eta](\tau) \right| \leq C e^{N+1}.$$
Proof. It follows from the bounds (42) and (47) that the first estimate is obtained by using a similar proof to Theorem 3 in [9]. Based on the bounds (23) and (40), the second estimate can be proved as in Theorem 5.2 of [17]. As in Theorem 6 of [11], the third estimate is got by the Cauchy-Schwarz inequality and the estimates (23) and (40).

5.2 Relationship with the quantities of (5)

Denote the harmonic action of the numerical solution by

\[ J_l = I_l + I_{-l} = 2I_l \quad \text{for} \quad 0 < l < M, \quad J_0 = I_0, \quad J_M = I_M. \]

It can be obtained that the almost-invariants \( J_l, H \) and \( K \) are close to the harmonic action \( J_l \), the Hamiltonian \( H_M \) and the momentum \( K \), respectively.

**Theorem 6** Under the conditions of Theorem 5 along the numerical solution \((q^n, p^n)\) of (9) and the associated modulation sequence \((\zeta(\epsilon t), \eta(\epsilon t))\), we have

\[
\begin{align*}
J_l[\zeta, \eta](\epsilon t_n) &= J_l(q^n, p^n) + \gamma_l(t_n)\epsilon^3, \\
K[\zeta, \eta](\epsilon t_n) &= K(q^n, p^n) + O(\epsilon^3) + O(\epsilon^2 M^{-s}), \\
H[\zeta, \eta](\epsilon t_n) &= H_M(q^n, p^n) + O(\epsilon^3),
\end{align*}
\]

where all the constants are independent of \( \epsilon, M, h, \) and \( n, \) and \( \sum_{l=0}^M \omega_l^{2s+1} \gamma_l(t_n) \leq C \) for \( t_n \leq \epsilon^{-1} \).

**Proof.** We only prove the second statement of this theorem. The others can be obtained in a similar way and we skip them for brevity.

By \( \tilde{q}_j(t) = \sum_{|k| \leq 2N} \hat{\epsilon}^{(k)}(\omega \cdot t) \zeta_k(\epsilon t) \) and

\[
2h\text{sinc}(h\Omega)\tilde{p}(t) = \tilde{q}(t + h) - \tilde{q}(t - h) - h^2\tilde{p}(h\Omega)[f(\tilde{q}_h(t + \frac{h}{2})) - f(\tilde{q}_h(t - \frac{h}{2}))],
\]

we obtain that \( \tilde{p}_j(t) = i\omega_j \left( \eta_j^{(j)}(\epsilon t)e^{i\omega_jt} - \eta_j^{(-j)}(\epsilon t)e^{-i\omega_jt} \right) + O(h^2) + O(\epsilon^2 h^2), \) where we have used the Taylor series of \( f \) at \( \tilde{q}_h(t) \). Thus we get \( \zeta_j^{(j)} = \frac{1}{2} \left( \tilde{q}_j + \frac{1}{i\omega_j} \tilde{p}_j \right) + O(\epsilon^2) \) and \( \zeta_j^{(-j)} = \frac{1}{2} \left( \tilde{q}_j - \frac{1}{i\omega_j} \tilde{p}_j \right) + O(\epsilon^2). \)

Moreover, it is obtained that \( |\eta_j^{(j)}| = |\omega_j| |\zeta_j^{(j)}| + O(\epsilon) \). According to (20), (22) and the “magic formulas” on p. 508 of [19], it is yielded that

\[
\begin{align*}
K[\zeta, \eta](\tau) &= \sum_{|k| \leq K} \sum_{|l| \leq M} \sum_{|l| \leq M} j \left[ \frac{\cos(\frac{1}{2}h\omega_j)}{b_1(h\omega_j)} ((k \cdot \omega)\text{sinc}(hk \cdot \omega) |\zeta_j^{(k)}|^2 + 2\epsilon c_{2k} \text{Im}(\zeta_j^{(k)})) \\
&\quad + \cdots \right] \left[ \frac{h^2\text{sinc}(\frac{1}{2}h\omega_j)}{b_1(h\omega_j)} ((k \cdot \omega)\text{sinc}(hk \cdot \omega) |\eta_j^{(k)}|^2 + 2\epsilon c_{2k} \text{Im}(\eta_j^{(k)})) + \cdots \right] \\
&= \sum_{|l| \leq M} j \left[ \frac{\cos(\frac{1}{2}h\omega_j)}{b_1(h\omega_j)} \omega_j \text{sinc}(h\omega_j) |\zeta_j^{(k)}|^2 + \frac{h^2\text{sinc}(\frac{1}{2}h\omega_j)}{2b_1(h\omega_j)} \omega_j \text{sinc}(h\omega_j) |\eta_j^{(k)}|^2 \right] + O(\epsilon^3).
\end{align*}
\]
From the above analysis, the condition (16) and the bounds (42)-(43), it follows that

\[ K[ζ, η](τ) = \sum_{|j| \leq M} jω_j \left( |ζ_j^{(j)}|^2 - |ζ_j^{(-j)}|^2 \right) + O(ε^3). \]

Therefore, the scheme of \( K \) becomes

\[ K[ζ, η](τ) = \sum_{|j| \leq M} \frac{jω_j}{4} \left( |\tilde{q}_j + \frac{1}{iω_j} \tilde{p}_j|^2 - |\tilde{q}_j - \frac{1}{iω_j} \tilde{p}_j|^2 \right) + O(ε^3) = \sum_{|j| \leq M} \frac{jω_j}{4} \frac{1}{ω_j} \tilde{q}_j \tilde{p}_j \]

\[ + O(ε^3) = K(\tilde{q}, \tilde{p}) + O(ε^3) + O(ε^2 M^{-s}) = K(q^n, p^n) + O(ε^3) + O(ε^2 M^{-s}), \]

where the results (22) and (24) are used.

### 5.3 Proof of Theorem 3

By the analysis given in this paper and following [9, 10], the statement of Theorem 3 can be proved by patching together many intervals of length \( ε^{-1} \).

### 6 Conclusions

In this work, we analysed the long-time conservation properties of ERKN methods when applied to nonlinear wave equations via spatial spectral semi-discretizations. We derived the main result concerning the long-time near conservations of actions, momentum and energy for ERKN methods in the full discretisation of nonlinear wave equations. The technique of multi-frequency modulated Fourier expansions was used to prove the main result by giving a modulated Fourier expansion of ERKN methods and showing three almost-invariants of the modulation system.

### Acknowledgements

The first author is grateful to Professor Christian Lubich for the useful comments on the topic of modulated Fourier expansions.

### References

[1] D. Bambusi, *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Phys., 234 (2003), pp. 253-285

[2] S. Blanes, *Explicit symplectic RKN methods for perturbed non-autonomous oscillators: splitting, extended and exponentially fitting methods*, Comput. Phys. Commun., 193 (2015), pp. 10-18.

[3] S. Buchholz, L. Gauckler, V. Grimm, M. Hochbruck, and T. Jahnke, *Closing the gap between trigonometric integrators and splitting methods for highly oscillatory differential equations*, IMA J. Numer. Anal., 38 (2018), pp. 57-74.
[4] B. Cano, *Conserved quantities of some Hamiltonian wave equations after full discretization*, Numer. Math., 103 (2006), pp. 197-223.

[5] B. Cano, *Conservation of invariants by symmetric multistep cosine methods for second-order partial differential equations*, BIT, 53 (2013), pp. 29-56.

[6] B. Cano and M. J. Moreta, *Multistep cosine methods for second-order partial differential systems*, IMA J. Numer. Anal., 30 (2010), pp. 431-461.

[7] D. Cohen and L. Gauckler, *One-stage exponential integrators for nonlinear Schrödinger equations over long times*, BIT, 52 (2012), pp. 877-903.

[8] D. Cohen, E. Hairer, and C. Lubich, *Numerical energy conservation for multi-frequency oscillatory differential equations*, BIT, 45 (2005), pp. 287-305.

[9] D. Cohen, E. Hairer, and C. Lubich, *Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions*, Arch. Ration. Mech. Anal., 187 (2008), pp. 341-368.

[10] D. Cohen, E. Hairer, and C. Lubich, *Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations*, Numer. Math., 110 (2008), pp. 113-143.

[11] L. Gauckler, *Error analysis of trigonometric integrators for semilinear wave equations*, SIAM J. Numer. Anal., 53 (2015), pp. 1082-1106.

[12] L. Gauckler, E. Hairer, and C. Lubich, *Long-term analysis of semilinear wave equations with slowly varying wave speed*, Comm. Part. Diff. Equa., 41 (2016), pp. 1934-1959.

[13] L. Gauckler, and D. Weiss, *Metastable energy strata in numerical discretizations of weakly nonlinear wave equations*, Disc. Contin. Dyn. Syst., 37 (2017), pp. 3721-3747

[14] L. Gauckler, J. Lu, J. Marzuola, F. Rousset, and K. Schratz, *Trigonometric integrators for quasilinear wave equations*, Math. Comp., (2018), doi:10.1090/mcom/3339

[15] V. Grimm, *On the use of the Gautschi-type exponential integrator for wave equations*, in Numerical Mathematics and Advanced Applications, Springer, Berlin, (2006), pp. 557-563.

[16] E. Hairer and C. Lubich, *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal., 38 (2000), pp. 414-441.

[17] E. Hairer and C. Lubich, *Spectral semi-discretisations of weakly nonlinear wave equations over long times*, Found. Comput. Math., 8 (2008), pp. 319-334.

[18] E. Hairer and C. Lubich, *Long-term analysis of the Störmer-Verlet method for Hamiltonian systems with a solution-dependent high frequency*, Numer. Math., 134 (2016), pp. 119-138.

[19] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd edn. Springer-Verlag, Berlin, Heidelberg, 2006.

[20] C. Liu, A. Iserles, and X. Wu, *Symmetric and arbitrarily high-order Birkhoff–Hermite time integrators and their long-time behaviour for solving nonlinear Klein–Gordon equations*, J. Comput. Phys., 356 (2018), pp. 1-30.
[21] C. Liu, and X. Wu, Arbitrarily high-order time-stepping schemes based on the operator spectrum theory for high-dimensional nonlinear Klein–Gordon equations, J. Comput. Phys., 340 (2017), pp. 243-275.

[22] C. Liu, and X. Wu, The boundness of the operator-valued functions for multidimensional nonlinear wave equations with applications, Appl. Math. Lett., 74 (2017), pp. 60-67.

[23] R. I. McLachlan and A. Stern, Modified trigonometric integrators, SIAM J. Numer. Anal., 52 (2014) 1378-1397.

[24] L. Mei, C. Liu, and X. Wu, An essential extension of the finite-energy condition for extended Runge-Kutta-Nyström integrators when applied to nonlinear wave equations, Commun. Comput. Phys., 22 (2017), pp. 742-764.

[25] J.M. Sanz-Serna, Modulated Fourier expansions and heterogeneous multiscale methods, IMA J. Numer. Anal., 29 (2009), pp. 595-605.

[26] B. Wang, A. Iserles, and X. Wu, Arbitrary-order trigonometric Fourier collocation methods for multi-frequency oscillatory systems, Found. Comput. Math., 16 (2016), pp. 151-181.

[27] B. Wang, F. Meng, and Y. Fang, Efficient implementation of RKN-type Fourier collocation methods for second-order differential equations, Appl. Numer. Math., 119 (2017), pp. 164-178.

[28] B. Wang and X. Wu, Long-time numerical energy conservation of extended RKN integrators for highly oscillatory Hamiltonian systems, (2017), arXiv:1712.04969.

[29] B. Wang, X. Wu, and F. Meng, Trigonometric collocation methods based on Lagrange basis polynomials for multi-frequency oscillatory second-order differential equations, J. Comput. Appl. Math., 313 (2017), pp. 185-201.

[30] B. Wang, H. Yang, and F. Meng, Sixth order symplectic and symmetric explicit ERKN schemes for solving multi-frequency oscillatory nonlinear Hamiltonian equations, Calcolo, 54 (2017), pp. 117-140.

[31] X. Wu and B. Wang, Recent Developments in Structure-Preserving Algorithms for Oscillatory Differential Equations, Springer Nature Singapore Pte Ltd, 2018.

[32] X. Wu, C. Liu, and L. Mei, A new framework for solving partial differential equations using semi-analytical explicit RK(N)-type integrators, J. Comput. Appl. Math., 301 (2016), pp. 74-90.

[33] X. Wu, X. You, and B. Wang, Structure-Preserving Algorithms for Oscillatory Differential Equations, Springer-Verlag, Berlin, Heidelberg, 2013.

[34] X. Wu, X. You, W. Shi, and B. Wang, ERKN integrators for systems of oscillatory second-order differential equations, Comput. Phys. Comm., 181 (2010), pp. 1873-1887.