A long-term numerical energy-preserving analysis of symmetric and/or symplectic extended RKN integrators for efficiently solving highly oscillatory Hamiltonian systems

Bin Wang · Xinyuan Wu

Received: 15 February 2020 / Accepted: 30 January 2021 / Published online: 25 June 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract
This paper presents a long-term analysis of one-stage extended Runge–Kutta–Nyström (ERKN) integrators for highly oscillatory Hamiltonian systems. We study the long-time numerical energy conservation not only for symmetric integrators but also for symplectic integrators. In the analysis, we neither assume symplecticity for symmetric methods, nor assume symmetry for symplectic methods. It turns out that these both types of integrators have a near conservation of the total and oscillatory energy over a long term. To prove the result for explicit integrators, a relationship between ERKN integrators and trigonometric integrators is established. For the long-term analysis of implicit integrators, the above approach does not work anymore and we use the technology of modulated Fourier expansion. By taking some adaptations of this technology for implicit methods, we derive the modulated Fourier expansion and show the near energy conservation.

Keywords Long-time energy conservation · Modulated Fourier expansions · Symmetric or symplectic methods · Extended RKN integrators · Highly oscillatory Hamiltonian systems

Communicated by Antonella Zanna Munthe-Kaas.

The research was supported in part by the National Natural Science Foundation of China under Grant 11671200.

Bin Wang
wangbinmaths@xjtu.edu.cn
Xinyuan Wu
xywu@nju.edu.cn

1 School of Mathematics and Statistics, Xi’an Jiaotong University, Xi’an 710049, Shaanxi, People’s Republic of China
2 Department of Mathematics, Nanjing University, Nanjing 210093, People’s Republic of China
3 School of Mathematical Sciences, Qufu Normal University, Qufu 273165, People’s Republic of China
1 Introduction

In this paper, we are concerned with the numerical energy-preserving analysis over a long term for extended Runge–Kutta–Nyström (ERKN) integrators when applied to the following highly oscillatory Hamiltonian system

\[
\begin{align*}
\dot{q} &= \nabla_p H(q, p), \quad q(0) = q^0, \\
\dot{p} &= -\nabla_q H(q, p), \quad p(0) = p^0,
\end{align*}
\]

with the Hamiltonian

\[
H(q, p) = \frac{1}{2}(\|p\|^2 + \|\Omega q\|^2) + U(q),
\]

where the vectors \( p = (p_1, p_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \) and \( q = (q_1, q_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \) are partitioned subject to the partition of the square matrix

\[
\Omega = \begin{pmatrix}
0_{d_1 \times d_1} & 0_{d_2 \times d_2} \\
0_{d_1 \times d_1} & \omega I_{d_2 \times d_2}
\end{pmatrix},
\]

and \( \omega \) is a large positive parameter. It is assumed in this paper that the initial values of (1) satisfy the condition

\[
\frac{1}{2} \|p(0)\|^2 + \frac{1}{2} \|\Omega q(0)\|^2 \leq E,
\]

where the constant \( E \) is independent of \( \omega \). As is known, the oscillatory energy of the system (1) is

\[
I(q, p) = \frac{1}{2} p_2^T p_2 + \frac{1}{2} \omega^2 q_2^T q_2.
\]

These highly oscillatory Hamiltonian systems frequently arise in a wide variety of applications including applied mathematics, quantum physics, classical mechanics, molecular biology, chemistry, astronomy, and electronics (see, e.g. [29,31,44,48]). Over the past two decades (and earlier), some novel approaches have been extensively studied for the highly oscillatory Hamiltonian system and we refer the reader to [11, 18,23,30,33,34,39–41,47] as well as the references contained therein.

The numerical long-time near-conservation of energy for such equations has already been researched for various numerical integrators, such as for symmetric trigonometric integrators in [5,6,21,29], for the Störmer–Verlet method in [20], for an implicit-explicit method in [35,38], and for heterogeneous multiscale methods in [37]. The essential technology used in the analysis is the modulated Fourier expansion, which was firstly developed in [21] to show long-time almost conservation properties of numerical methods for highly oscillatory Hamiltonian systems. A similar modulated
Fourier expansion was independently presented in [19] on the spectral formulation of the Nekhoroshev theorem for quasi-integrable Hamiltonian systems. Following those pioneering researches, modulated Fourier expansions have been developed as an important mathematical tool in studying the long-time behaviour of numerical methods and differential equations (see, e.g. [3,12,13,23–25,32]). With this technology, much work has been done about different numerical methods for various systems. Besides the work stated above for the highly oscillatory Hamiltonian system (2), the numerical energy conservation for other kinds of systems has also been researched in many publications, such as for multi-frequency Hamiltonian systems in [7,10], wave equations in [8,9,15,22], Schrödinger equations in [4,16,17], highly oscillatory Hamiltonian systems without any non-resonance condition in [14], Hamiltonian systems with a solution-dependent high frequency in [26], and charged-particle dynamics in [27,28].

On the other hand, in order to effectively solve (2) in the sense of structure-preservation, the authors in [46] formulated a standard form of extended Runge–Kutta–Nyström (ERKN) integrators and derived the corresponding order conditions by the B-series theory associated with the extended Nyström trees. In [42], the error bounds for explicit ERKN integrators were researched. Recently, symplecticity conditions of ERKN integrators were derived in [45] and symmetry conditions were derived in [48]. Based on these conditions, some practical symmetric or and symplectic ERKN integrators were constructed and analyzed in [43]. The results of numerical experiments appearing in the work mentioned above have shown that the symmetric or and symplectic ERKN integrators behave very well even in a long-time interval. However, the theoretical analysis of energy behaviour over a long term of symmetric or symplectic ERKN integrators has not been considered and researched yet in the literature, which motivates this paper.

It is noted that compared with explicit methods, implicit methods are usually more expensive. Fortunately, however, for implicit ERKN integrators, they can be solved by a rapidly convergent fixed-point iteration (sometimes one fixed-point iteration is enough) because the stiff part is the linear force $\Omega^2 q$. Moreover, for integrators with the same stage, implicit ones can have higher order and better stability than explicit ones (see [29]). In order to show this point more clearly, we choose three methods for comparison: IERKN (two-stage implicit ERKN method obtained by the Gauss collocation method of order four given in [36]), GAUSS (the well-known fourth order Gauss collocation) and EERKN (two stage explicit method SSMERKN2s2-1 of order two given in [48]). We apply these three methods to the problem given in Sect. 3.2. The system is solved on the interval $[0, 10]$ with $h = 1/2^k$ for $k = 5, \ldots, 11$. The errors of the solution against $h$ and CPU time are indicated in Fig. 1. The reference solution is obtained by using “ode45” of MATLAB. For the implicit methods, standard fixed point iteration is used and we set $10^{-16}$ as the error tolerance and 100 as the maximum number of each iteration. It follows from these results that the implicit ERKN integrator performs better than the other methods and its computational cost is not expensive. Therefore, implicit ERKN integrators are also important and then this paper is devoted to not only explicit ERKN integrators but also implicit ones.

The main contributions of this work are to show the long-time energy behaviour not only for one-stage explicit symmetric and/or symplectic ERKN integrators but also
for implicit symmetric and/or symplectic ERKN integrators. Similar results have been obtained only for symmetric explicit trigonometric integrators in [5,7,21,29]. However, in this paper we prove the long-time result for more diverse methods than that for those considered previously. In particular, we present the analysis for both explicit and implicit symmetric and/or symplectic ERKN integrators. We neither assume symplecticity for symmetric methods, nor assume symmetry for symplectic methods. It follows from the analysis that both symmetry and symplecticity can produce a good long-time energy conservation for ERKN integrators, which means that symmetry and symplecticity play a similar role in the numerical energy-preserving behaviour. This is a new discovery which is of great importance to geometric integration for highly oscillatory Hamiltonian systems. Moreover, in contrast to [5,7,21,29], in the long-term analysis of implicit integrators, the formulation of modulated Fourier expansions does not rely on the symmetry of the methods. This is of major importance in the context of long-term analysis of non-symmetric methods. It is also a main conceptual difference in comparison with [5,7,21,29].

For the analysis of explicit integrators, we have noted that some explicit ERKN integrators can be formulated as a Strang splitting method applied to an averaged equation (see, [1]). Very recently, the authors in [2] proved second-order error bounds of trigonometric integrators on the basis of the interpretation of trigonometric integrators as splitting methods for averaged equations. Following this approach, the long-term analysis of explicit ERKN integrators will be proved concisely by exploring the connection between ERKN integrators and symmetric trigonometric integrators researched in [5,7,21,29]. However, in the analysis of implicit ERKN integrators, we do not require the symmetry of methods and it is known that the symmetry plays an important role in the construction of modulated Fourier expansions and long-term analysis given in [5,7,21,29]. Moreover, the connection presented in [1] does not hold for implicit ERKN methods. Therefore, unfortunately, the approach used for explicit ERKN integrators does not apply to implicit ERKN integrators any more. In order to

![Fig. 1 The logarithm of the errors (err) against the logarithm of h (left). The logarithm of the errors (err) against the logarithm of CPU time (right).](image-url)
overcome this difficulty, we will use the technology of modulated Fourier expansion developed by Hairer and Lubich in [21] with some novel adaptations for implicit and non-symmetric methods. The modulated Fourier expansion of implicit ERKN integrators will be derived and two almost-invariants will be shown. Then the long-term result can be obtained. By using this technology, the result for explicit ERKN integrators can also be derived.

This paper is organized as follows. We first present some preparatories of ERKN integrators in Sect. 2. The main results as well as an illustrative numerical experiment are given in Sect. 3. Section 4 gives the proof of long-term result for implicit ERKN integrators, where the modulated Fourier expansion is constructed for implicit symmetric or/and symplectic integrators and two almost-invariants of the modulated Fourier expansions are studied. Then in Sect. 5, the result for explicit integrators is obtained immediately by the procedure of Sect. 4. Another proof is also given in Sect. 5 by exploring the connection between explicit ERKN integrators and trigonometric integrators and by using the previous results of symmetric trigonometric integrators shown in [5,7,21,29]. The concluding remarks are made in the last section.

2 Preliminaries to ERKN integrators

2.1 ERKN integrators

The highly oscillatory Hamiltonian system (2) can be rewritten as the following system of second-order differential equations

\[ q''(t) + \Omega^2 q(t) = g(q(t)), \quad q(0) = q^0, \quad q'(0) = p^0, \]

where \( g \) is the negative gradient of a real-valued function \( U(q) \). ERKN integrators were first formulated for integrating (6) in [46] and here we summarize the scheme of one-stage ERKN integrators as follows.

**Definition 1** (See [46]) A one-stage ERKN integrator for solving (6) is defined by

\[
\begin{align*}
Q^{n+c_1} &= \phi_0(c_1V)q^n + hc_1\phi_1(c_1V)p^n + h^2\bar{a}_{11}(V)g(Q^{n+c_1}), \\
q^{n+1} &= \phi_0(V)q^n + h\phi_1(V)p^n + h^2\bar{b}_1(V)g(Q^{n+c_1}), \\
p^{n+1} &= -h\Omega^2\phi_1(V)q^n + \phi_0(V)p^n + hb_1(V)g(Q^{n+c_1}),
\end{align*}
\]

(7)

where \( h \) is a stepsize, \( c_1 \in [0, 1] \) is a real constant, \( \bar{a}_{11}(V), b_1(V) \) and \( \bar{b}_1(V) \) are matrix-valued and bounded functions of \( V \equiv h^2\Omega^2 \), and

\[
\phi_j(V) := \sum_{k=0}^{\infty} \frac{(-1)^k V^k}{(2k+j)!}, \quad j = 0, 1.
\]

(8)

From (8), it is clear that

\[
\phi_0(V) = \cos(\sqrt{V}) = \cos(h\Omega), \quad \phi_1(V) = \sin(\sqrt{V})(\sqrt{V})^{-1} = \text{sinc}(h\Omega),
\]

\[\text{Springer}\]
where \( \text{sinc}(h\Omega) = \sin(h\Omega)/(h\Omega) \). Thence the scheme of one-stage ERKN integrators for (6) can be reformulated as follows.

**Definition 2** The one-stage ERKN integrator for integrating (6) is given by

\[
\begin{align*}
Q^{n+1} &= \cos(c_1 h\Omega)q^n + c_1 h \text{sinc}(c_1 h\Omega)p^n + h^2 \bar{a}_{11}(h\Omega)g(Q^{n+1}), \\
q^{n+1} &= \cos(h\Omega)q^n + h \text{sinc}(h\Omega)p^n + h^2 \bar{b}_1(h\Omega)g(Q^{n+1}), \\
p^{n+1} &= -h^2 \text{sinc}(h\Omega)q^n + \cos(h\Omega)p^n + h b_1(h\Omega)g(Q^{n+1}),
\end{align*}
\]

(9)

where the functions \( \bar{a}_{11}(h\Omega), \bar{b}_1(h\Omega) \) and \( b_1(h\Omega) \) are real-valued and bounded functions of \( h\Omega \).

### 2.2 Basic properties

As shown in [45,48], we obtain the following conditions for the integrator (9) to be symmetric and symplectic.

**Theorem 1** The ERKN integrator (9) is symmetric if and only if

\[
c_1 = 1/2, \quad (I + \cos(h\Omega))\bar{b}_1(h\Omega) = \text{sinc}(h\Omega)b_1(h\Omega).
\]

(10)

**Proof** It follows from the symmetry conditions given in [48] that this integrator is symmetric if and only if

\[
\begin{align*}
c_1 = 1/2, \quad \bar{b}_1(V) &= \phi_1(V)b_1(V) - \phi_0(V)\bar{b}_1(V), \\
\phi_0(c_1^2 V)\bar{b}_1(V) &= c_1 \phi_1(c_1^2 V)b_1(V).
\end{align*}
\]

(11)

By solving the second equation in (11), we obtain the second result of (10). It can also be verified that under the condition (10), the third equation of (11) is true. \( \square \)

**Theorem 2** For any real number \( d_1 \), if the coefficients are determined by

\[
b_1(h\Omega) = d_1 \cos((1 - c_1)h\Omega), \quad \bar{b}_1(h\Omega) = d_1(1 - c_1)\text{sinc}((1 - c_1)h\Omega),
\]

(12)

then the ERKN integrator (9) is symplectic.

**Proof** According to the symplectic conditions given in [45], we know that this method is symplectic if the following equations

\[
\begin{align*}
\phi_0(V)b_1(V) + V\phi_1(V)\bar{b}_1(V) &= d_1 \phi_0(c_1^2 V), & d_1 \in \mathbb{R}, \\
\phi_1(V)b_1(V) - \phi_0(V)\bar{b}_1(V) &= c_1 d_1 \phi_1(c_1^2 V),
\end{align*}
\]

are satisfied. The result is directly obtained by solving these two equations. \( \square \)

**Remark 1** These two theorems confirm the fact that an ERKN integrator can be symmetric and symplectic, or symmetric but not symplectic, or symplectic but not symmetric.
2.3 Practical examples

As some examples of ERKN integrators with certain structure characteristics, we present eight practical one-stage integrators and their coefficients and properties are listed in Table 1. The first four ones are explicit and the last four are implicit.

3 Main results and numerical examples

3.1 Main results of this paper

Before presenting the main results of this paper, we make the following assumptions.

Assumption 1  

– Assume that the initial values satisfy (4).
– The numerical solution is assumed to stay in a compact set.
– A lower bound on the stepsize is posed as:

$$h \omega \geq c_0 > 0. \quad (13)$$

– For a given $h$ and $\omega$, the numerical non-resonance condition is assumed to be held

$$\left| \sin \left( \frac{1}{2} kh \omega \right) \right| \geq c\sqrt{h} \quad \text{for} \quad k = 1, 2, \ldots, N \quad \text{with} \quad N \geq 2, \quad (14)$$

which imposes a restriction on $N$. How large of $N$ depends on the chosen stepsize $h$ and the value of $\omega$ appeared in the system. In the numerical experiment of Sect. 3.2, we will check the numerical non-resonance condition (14) for $k = 1, 2, \ldots$ to obtain the value of $N$. We also note here that usually $N$ is not needed to be very large because our main results given below will hold for $0 \leq nh \leq h^{-N+1}$. Even if $N$ is not large, $h^{-N+1}$ can be large enough to show the long time behaviour of the methods. In the following, $N$ is a fixed integer such that (14) holds.

– For the coefficients of the ERKN integrators, it is assumed that the function

$$\sigma(v) = \frac{\sin(v) \cos \left( \frac{1}{2} v \right)}{2b_1(v)} + v^2 \sin(v) \frac{\frac{1}{2}\sin \left( \frac{1}{2} v \right)}{2b_1(v)} \quad (15)$$

is bounded from below and above:

$$0 < c_1 \leq \sigma(v) \leq C_1 \quad \text{for} \quad v = 0, h \omega, \quad (16)$$

or the same estimate holds for $-\sigma$ instead of $\sigma$.

It is noted that the first four assumptions are considered by many publications in the energy analysis of symmetric trigonometric integrators for the Hamiltonian system (1) (see, e.g. [3,5,21,29]). The last assumption is obtained in the remainder analysis of this paper and it is similar to Assumption B proposed in [5].
### Table 1 Eight examples of one-stage explicit and implicit ERKN integrators

| Methods   | $c_1$ | $\tilde{b}_1(h\omega)$               | $b_1(h\omega)$               | $\tilde{a}_{11}(h\omega)$ | Symmetric | Symplectic |
|-----------|-------|--------------------------------------|--------------------------------|-----------------------------|-----------|------------|
| EERKN1    | $\frac{1}{2}$ | $\frac{1}{2} \sin^2(\frac{h\omega}{2})$ | $\cos(\frac{h\omega}{2})$ | 0                           | Non       | Non        |
| EERKN2    | $\frac{1}{2}$ | $\frac{1}{2} \sin(\frac{h\omega}{2})$ | $\cos(\frac{h\omega}{2})$ | 0                           | Symmetric | Symplectic |
| EERKN3    | $\frac{1}{2}$ | $\frac{1}{2} \sin(h\omega) \cos(\frac{h\omega}{2})$ | $\cos^3(\frac{h\omega}{2})$ | 0                           | Symmetric | Non        |
| EERKN4    | $\frac{2}{3}$ | $\frac{2}{3} \sin(\frac{2h\omega}{3})$ | $\cos(\frac{2h\omega}{3})$ | 0                           | Non       | Symplectic |
| IERKN1    | $\frac{1}{2}$ | $\frac{1}{2} \sin^2(\frac{h\omega}{2})$ | $\cos(\frac{h\omega}{2})$ | $\frac{1}{2} \sin^2(\frac{h\omega}{2})$ | Non       | Non        |
| IERKN2    | $\frac{1}{2}$ | $\frac{1}{2} \sin(\frac{h\omega}{2})$ | $\cos(\frac{h\omega}{2})$ | $\frac{1}{2} \sin(\frac{h\omega}{2})$ | Symmetric | Symplectic |
| IERKN3    | $\frac{1}{2}$ | $\frac{1}{2} \sin(h\omega) \cos(\frac{h\omega}{2})$ | $\cos^3(\frac{h\omega}{2})$ | $\frac{1}{2} \sin(h\omega) \cos(\frac{h\omega}{2})$ | Symmetric | Non        |
| IERKN4    | $\frac{1}{3}$ | $\frac{2}{3} \sin(\frac{2h\omega}{3})$ | $\cos(\frac{2h\omega}{3})$ | $\frac{2}{3} \sin(\frac{2h\omega}{3})$ | Non       | Symplectic |
With regard to the long-time total and oscillatory energy conservation along symmetric or symplectic ERKN integrators, we have the following two main results of this paper, which will be proved in detail in the next two sections, respectively.

**Theorem 3** (Main results for explicit methods) *Under the conditions given in Assumption 1, for one-stage explicit symmetric or/and symplectic ERKN integrators, it holds that*

\[
H(q^n, p^n) = H(q^0, p^0) + \mathcal{O}(h),
\]

\[
I(q^n, p^n) = I(q^0, p^0) + \mathcal{O}(h),
\]

*for* \(0 \leq nh \leq h^{-N+1}\). The constants symbolized by \(\mathcal{O}\) depend on \(N, T\) and the constants in the assumptions, but are independent of \(n, h, \omega\).

**Theorem 4** (Main results for implicit methods) *Under the conditions of Assumption 1, for one-stage implicit symmetric or/and symplectic ERKN integrators, we have*

\[
H(q^n, p^n) = H(q^0, p^0) + \mathcal{O}(h),
\]

\[
I(q^n, p^n) = I(q^0, p^0) + \mathcal{O}(h),
\]

*where* \(0 \leq nh \leq h^{-N+1}\). The constants symbolized by \(\mathcal{O}\) are independent of \(n, h, \omega\), but depend on \(N, T\) and the constants in the assumptions.

**Remark 2** At the first sight, the part \(\mathcal{O}(h)\) appearing in these two theorems has a conflict with the condition (13). The result \(\mathcal{O}(h)\) seems that \(h\) is arbitrarily small while (13) imposes a lower bound on \(h\). For this viability, we have the following comments.

For symplectic methods with stepsize \(h\) applied to the highly oscillatory Hamiltonian system (2), their long-time near-conservation of the energy can be derived with the help of backward error analysis. This technique interprets the numerical method as the (almost) exact flow of a modified Hamiltonian system, up to exponentially small terms \(\mathcal{O}(e^{-\frac{c}{\omega}h})\) (see [29]), which requires that \(h\omega \leq C\). Therefore, the stepsize \(h\) should be sufficiently small for \(\omega \gg 1\) and the backward error analysis is not applicable for methods with large stepsizes when applied to (2). In order to study the behaviour of a method with large stepsizes, the essential technology named as modulated Fourier expansion (MFE) was firstly developed in [21], and was further studied in [3,12,13,24,25]. With this technology, a method with the region that \(h\omega\) is bounded away from 0 (condition (13)) can be researched. Therefore, this paper is devoted to the study of methods with large stepsizes \(h (h < 1)\). The notation \(\mathcal{O}(h)\) used in this paper means that the bound of that corresponding part is no more than \(Ch\), where \(C\) is a constant independent of \(n, h, \omega\). The main results given in Theorem 3–4 will show that some considered methods preserve the total energy and oscillatory energy up to the accuracy \(\mathcal{O}(h)\). As the size of \(h\) increases, the error of the energy accordingly becomes large.

### 3.2 Numerical examples

As an important nonlinear model, we consider the Fermi–Pasta–Ulam problem which was researched in [29]. This model describes classical and quantum systems of
interacting particles in the physics of nonlinear phenomena. Denote by $x_i$ a scaled displacement of the $i$th stiff spring and by $x_{m+i}$ a scaled expansion or compression of the $i$th stiff spring. Their corresponding velocities are expressed in $y_i$ and $y_{m+i}$, respectively. Then the problem can be formulated by a Hamiltonian system with the Hamiltonian

$$H(y, x) = \frac{1}{2} \sum_{i=1}^{2m} y_i^2 + \frac{\omega^2}{2} \sum_{i=1}^{m} x_{m+i}^2 + \frac{1}{4} [(x_1 - x_{m+1})^4 + \sum_{i=1}^{m-1} (x_{i+1} - x_{m+i-1} - x_i - x_{m+i})^4 + (x_m + x_{2m})^4].$$

Following [29], we choose $m = 3$ and

$$x_1(0) = 1, \ y_1(0) = 1, \ x_4(0) = \frac{1}{\omega}, \ y_4(0) = 1$$

with zero for the remaining initial values. First, the system is integrated on the interval $[0, 10000]$ with $h = 0.02$ and $\omega = 50$. The errors of the total energy $H$ and oscillatory energy $I$ against $t$ for different ERKN integrators are shown in Fig. 2. Then we increase the size of $\omega$ to 200 and solve the problem with different stepizes $h = 0.02$ and $h = 0.01$. For each ERKN integrator, an identical behaviour as Fig. 2 can be obtained and we do not present these figures for brevity. It follows from Fig. 2 that all the explicit and implicit symmetric or/and symplectic methods approximately conserve $H$ and $I$ quite well over a long term for different stepizes and $\omega$. The non-symmetric and non-symplectic methods EERKN1 and IERKN1 do not approximately conserve $H$ and $I$. In the light of the results, it seems that the symmetry or symplecticity is essential for the conservation of $H$ and $I$. All these numerical behaviours can be explained by the theoretical conclusions. According to Theorems 3 and 4, it can be concluded that if ERKN integrators are symmetric or symplectic, they have a near conservation of $H$ and $I$ over a long term not only for explicit methods but also for implicit ones. This is the theoretical reason for the behaviour shown in this numerical experiment.

Meanwhile, we check the numerical non-resonance condition (14) by choosing $h = 0.02$ and $\omega = 50$. In Fig. 3, we display the results of NNRC := $\left| \frac{\sin \left( \frac{1}{2} kh\omega \right)}{\sqrt{h}} \right|$ as the function of $k$. It follows from this result that if $c$ is chosen to be one, the numerical non-resonance condition (14) holds for $k = 1, \ldots, 19$. This implies that $N = 18$ for this case and the main results are true for $0 \leq nh \leq 0.02^{-17} \approx 7.6294e + 28$.

Finally, we consider the energy exchange between stiff components that occurs for the oscillatory energies

$$I = I_1 + I_2 + I_3 \text{ with } I_j = \frac{1}{2} y_{3+j}^2 + \frac{1}{2} \omega^2 x_{3+j}^2,$$

where $x_{3+j}$ represents the elongation of the $j$th stiff spring. Following [29], we choose $\omega = 50$ for showing the energy exchange between stiff components. Figure 4 shows the oscillatory energies $I_1$, $I_2$, $I_3$ and their sum $I = I_1 + I_2 + I_3$ as functions of time on the interval $[0, 200]$ with $h = 0.01$ for the eight methods. It can be observed from
Fig. 2 The logarithm of the errors ($ERR$) of $H$ and $I$ against $t$ with $h = 0.02$ and $\omega = 50$
the results that all these eight methods show a good or not bad approximation of the energy exchanges of the highly oscillatory energies.

4 Proof for implicit integrators (Theorem 4)

In this section, we are devoted to the analysis of long-time conservation of the total and oscillatory energy along implicit ERKN integrators. In order to prove the long-time energy conservation result for implicit integrators, we will use the technology of modulated Fourier expansion (MFE) and make some adaptations for implicit methods. It is important to note that, although the analysis is based on the well-known MFE, there are some new technical difficulties of working with MFE for implicit ERKN integrators. The main difficulties and differences between the proof given in this section and the analysis of [5,21,29] are summarized as follows.

- It is known that symmetry plays an important role in the construction of MFE in [5,7,21,29]. By using symmetry, the considered method can be rewritten as a two-step method and a simple operator is defined to derive the MFE. In the proof of this section, we neither assume symplecticity for symmetric methods, nor assume symmetry for symplectic methods. Therefore, the methods cannot be rewritten as a two-step method and then further cannot be expressed by that simple operator. In order to overcome this difficulty, we will define three suitable operators with the needed properties. Based on them, MFE can be derived accordingly.

- All the methods researched in [5,7,21,29] are explicit. In the proof of this section, we will unify the technology of MFE and make it be applicable to a wide class of methods including implicit methods. Since the methods considered here contain not only updates but also internal stages, the proof will be more complicated than those given in [5,7,21,29]. Thus MFE should be used very carefully and some necessary adaptations should be made.
Fig. 4 The energy exchange between stiff components for the methods as a function of $t$ on the interval $[0, 200]$ (stepsize $h = 0.01$).
4.1 Modulated Fourier expansion

We now derive the MFE of implicit ERKN integrators and present the bounds of the modulated Fourier functions. We present the proof for IEIKN4 and it is easy to modify the proof for other implicit methods by changing the operator $L_2(hD)$.

**Theorem 5** Under the conditions of Theorem 4 and for $0 \leq t = nh \leq T$, the numerical solution of IERKN4 admits the following MFE

\[
q^n = \sum_{|k| < N} e^{ikt} \zeta_h^k(t) + R_{h,N}(t), \\
p^n = \sum_{|k| < N} e^{ikt} \eta_h^k(t) + S_{h,N}(t),
\]  

(17)

where the remainder terms are bounded by

\[
R_{h,N}(t) = O(thN), \quad S_{h,N}(t) = O(th^{N-1}).
\]  

(18)

The coefficient functions $\zeta_h^k$ as well as all their derivatives are bounded by

\[
\ddot{\zeta}_{h,1}^0 = O(1), \quad \dot{\zeta}_{h,1}^1 = O(h^2), \quad \zeta_{h,1}^k = O(h^{|k|+1}), \\
\ddot{\zeta}_{h,2}^0 = O(h^2), \quad \dot{\zeta}_{h,2}^1 = O(h), \quad \zeta_{h,2}^k = O(h^{|k|+1}),
\]  

(19)

for $|k| > 1$ and further bounded by

\[
\ddot{\zeta}_{h,1}^0 = O(1), \quad \ddot{\zeta}_{h,2}^0 = O(h). 
\]  

(20)

Moreover, we have the following results for coefficient functions $\eta_h^k$

\[
\eta_{h,1}^0 = \dot{\zeta}_{h,1}^0 + O(h), \quad \eta_{h,1}^1 = O(h), \quad \eta_{h,1}^k = O(h^{|k|}), \\
\eta_{h,2}^0 = O(h^2), \quad \eta_{h,2}^1 = i\omega \zeta_{h,2}^1 + O(h), \quad \eta_{h,2}^k = O(h^{|k|}),
\]  

(21)

where $|k| > 1$. Moreover, we have $\zeta^{-k} = \overline{\zeta^k}$ and $\eta^{-k} = \overline{\eta^k}$. The constants symbolized by the notation are independent of $h$ and $\omega$, but depend on the constants from Assumption 1 and the final time $T$.

**Proof** The proof follows that used in the modulated Fourier expansions of previous publications (see [5,7,21,29]) but with some novel adaptations for non-symmetric and implicit methods. The proof of this theorem does not rely on the symmetry of the methods, which is a main conceptual difference in comparison with that in [5,7,21,29].

We will prove that there exist two functions

\[
q_h(t) = \sum_{|k| < N} e^{ikt} \zeta_h^k(t), \quad p_h(t) = \sum_{|k| < N} e^{ikt} \eta_h^k(t)
\]  

(22)
with smooth (in the sense that all their derivatives are bounded independently of $h$ and $\omega$) coefficients $\zeta_k^h, \eta_k^h$, such that, for $t = nh$,

$$q^n = q_h(t) + O(h^N), \quad p^n = p_h(t) + O(h^{N-1}).$$

**Construction of the coefficients functions.** To do this, we begin with defining the operators

$$\mathcal{L}_1(hD) = \left( e^{hD} - \cos(h\Omega) - h^2\Omega^2 \text{sinc}(h\Omega) \hat{b}_1(h\Omega) \hat{b}_1^{-1}(h\Omega) \right),$$

$$\mathcal{L}_2(hD) = \cos(c_1h\Omega) + \hat{a}_{11}(h\Omega) \hat{b}_1(h\Omega)^{-1}(e^{hD} - \cos(h\Omega)),$$

$$\mathcal{L}(hD) = \left( e^{hD} - \cos(h\Omega) - \text{sinc}(h\Omega) \mathcal{L}_1(hD) \right) \left( \hat{b}_1(h\Omega) \mathcal{L}_2(hD) \right)^{-1},$$

where $D$ is the differential operator (see [26]).

According to the second and third formulae of (9), it is arrived at

$$q^{n+1} = \left( \cos(h\Omega) + h^2\Omega^2 \text{sinc}(h\Omega) \hat{b}_1(h\Omega) \hat{b}_1^{-1}(h\Omega) \right) q^n + h \hat{b}_1(h\Omega) \hat{b}_1^{-1}(h\Omega) p^{n+1} + h \left( \text{sinc}(h\Omega) - \hat{b}_1(h\Omega) \hat{b}_1^{-1}(h\Omega) \cos(h\Omega) \right) p^n.$$

Inserting (22) into this result, comparing the coefficients of $e^{ik\omega t}$ and considering the definition of $\mathcal{L}_1$ implies the relationship between $\zeta_k^h$ and $\eta_k^h$ as follows:

$$\mathcal{L}_1(hD)\zeta_0^h = h\eta_0^h,$$

$$\mathcal{L}_1(hD + ikh\omega)\zeta_k^h = h\eta_k^h \quad \text{for} \quad |k| > 0.$$

For the first formula of the ERKN integrator (9), we look for the function

$$\tilde{q}_h(t) := \sum_{|k| < N} e^{ik\omega t} \xi_k^h(t)$$

as the modulate Fourier expansion of $Q^{n+\frac{1}{2}}$ at $t = (n + c_1)h$. Inserting (22)–(24) into the first formula of (9) and comparing the coefficients of $e^{ik\omega t}$ yields

$$\mathcal{L}_2(hD)\xi_0^h = \xi_0^h,$$

$$\mathcal{L}_2(hD + ikh\omega)\xi_k^h = \xi_k^h \quad \text{for} \quad |k| > 0,$$

which gives the relationship between the coefficient functions $\xi_k^h$ and $\xi_k^h$. 

\[\text{Springer}\]
We insert the above expansions into the second equation of (9), expand the nonlinear function into its Taylor series and compare the coefficients of \(e^{ik\omega t}\). We then obtain

\[
\mathcal{L}(hD)\xi^0_h = h^2 \left( g(\xi^0_h) + \sum_{s(\alpha) = 0} \frac{1}{m!} g^{(m)}(\xi^0_h)(\xi^0_h)^\alpha \right),
\]

\[
\mathcal{L}(hD + ikh\omega)\xi^k_h = h^2 \sum_{s(\alpha) = k} \frac{1}{m!} g^{(m)}(\xi^0_h)(\xi^0_h)^\alpha,
\]

where \(|k| > 0\), the sum ranges over \(m \geq 0\), the multi-indices \(\alpha = (\alpha_1, \ldots, \alpha_m)\) with integer \(\alpha_i\) satisfying \(0 < |\alpha_i| < N\) have a given sum \(s(\alpha) = \sum_j \alpha_j\), and \((\xi^0_h)^\alpha\) is an abbreviation for the \(m\)-tuple \((\xi^0_{h1}, \ldots, \xi^0_{hm})\).

Comparing the dominate terms in the relations for the coefficients functions \(\xi^k_h\) motivates the following ansatz of the modulated Fourier functions:

\[
\dot{\xi}^0_{h1} = G_{10}(\cdot) + \cdots, \quad \dot{\xi}^0_{h2} = \frac{h^2(\cos(\frac{1}{2}h\omega) + \cos(\frac{1}{2}h\omega))}{2h\omega \sin(\frac{1}{2}h\omega)} (G_{20}(\cdot) + \cdots),
\]

\[
\dot{\xi}^1_{h1} = \frac{h^2(4\cos(h\omega) - 1)}{12 \sin^2(\frac{1}{2}h\omega)} (\mathcal{F}^1_{10}(\cdot) + \cdots), \quad \dot{\xi}^1_{h2} = \frac{h^2}{2h\omega} (\mathcal{F}^1_{20}(\cdot) + \cdots),
\]

\[
\dot{\xi}^k = \frac{h^2 \sin(\frac{h\omega}{2})(1 - 2\cos(\frac{4h\omega}{1}) + 4\cos(\frac{4h\omega}{2})\cos(kh\omega))}{-4h \sin\left(\frac{k\omega h + k\Omega}{2}\right) \sin\left(\frac{kh\omega - k\Omega}{2}\right)} (\mathcal{F}^k_0(\cdot) + \cdots) \quad \text{for } |k| > 1,
\]

where the dots stand for power series in \(\sqrt{h}\). Since the series in the ansatz usually diverge, in this paper we truncate them after the \(O(h^{N+1})\) terms (see [21]).

**Initial values.** In this part, we derive the initial values for the differential equations appearing in the ansatz (26). On the basis of the conditions \(p_h(0) = p^0\) and \(q_h(0) = q^0\), it can be deduced that

\[
p^0_1 = \eta^0_{h,1}(0) + \mathcal{O}(h) = \xi^0_{h,1}(0) + \mathcal{O}(h), \quad p^0_2 = 2\text{Re}(\eta^1_{h,2}(0)) + \mathcal{O}(h^{\frac{1}{2}}),
\]

\[
q^0_1 = \xi^0_{h,1}(0) + \mathcal{O}(h^2), \quad q^0_2 = 2\text{Re}(\xi^1_{h,2}(0)) + \mathcal{O}(h^{\frac{3}{2}}).
\]

On the other hand, we have

\[
p_{h,1}(h) = p^1_1, \quad p_{h,2}(h) = p^1_2, \quad q_{h,1}(h) = q^1_1, \quad q_{h,2}(h) = q^1_2.
\]

It follows from the scheme of the method (9) that

\[
q^1_2 - \cos(h\omega)q^0_2 = h\text{sinc}(h\omega)p^0_2 + h^2\mathcal{B}_1(h\omega)g_2(Q^{n+c_1}),
\]

where we have used the notation:

\[
g(Q^{n+c_1}) = (g_1(Q^{n+c_1})^T, g_2(Q^{n+c_1})^T)^T.
\]

\(\square\) Springer
By computing \( q_2^1 - \cos(h\omega)q_2^0 \), we get
\[
q_2^1 - \cos(h\omega)q_2^0 = q_{h,2}(h) - \cos(h\omega)q_{h,2}(0)
= \sum_{|k|<N} (e^{ik\omega h^k} - \cos(h\omega)^k(0))
= \zeta_{h,2}(h) + e^{i\omega h^1} \zeta_{h,2}(h) + e^{-i\omega h^{-1}}(h)
- \cos(h\omega)\left( \zeta_{h,2}^0(0) + \zeta_{h,2}^1(0) + \zeta_{h,2}^{-1}(0) \right) + O(h^2).
\]

Expanding the functions \( \zeta_{h,2}^0(h) \), \( \zeta_{h,2}^1(h) \), \( \zeta_{h,2}^{-1}(h) \) at \( h = 0 \) yields
\[
q_2^1 - \cos(h\omega)q_2^0 = (1 - \cos(h\omega))\zeta_{h,2}^0(0) + i \sin(h\omega)(\zeta_{h,2}^1(0) - \zeta_{h,2}^{-1}(0)) + O(h^2)
= 2 \sin^2(h\omega/2)\zeta_{h,2}^0(0) + i \sin(h\omega)(\zeta_{h,2}^1(0) - \zeta_{h,2}^{-1}(0)) + O(h^2)
= i \sin(h\omega)(\zeta_{h,2}^1(0) - \zeta_{h,2}^{-1}(0)) + O(h^2),
\]
where we have used the result of \( \zeta_{h,2}^0 \) presented in (26). Now the formula (28) becomes
\[
i \sin(h\omega)(\zeta_{h,2}^1(0) - \zeta_{h,2}^{-1}(0)) = h\sin(h\omega)\rho_2^0 + h^2\tilde{b}_1(h\omega)g_2(Q^{n+c1}) + O(h^2),
\]
which confirms that
\[
2\text{Im}(\zeta_{h,2}^1(0)) = \omega^{-1} \rho_2^0 + O(\omega^{-1}). \tag{29}
\]
Therefore, using the implicit function theorem, conditions (27) and (29) yield the desired initial values \( \zeta_{h,1}^0(0) \), \( \zeta_{h,1}^0(0) \), \( \zeta_{h,2}^1(0) \) for the differential equations appearing in the ansatz (26).

**Bounds of the coefficients functions.** Based on the ansatz and Assumption 1, it is easy to get the bounds (19). From the condition (4), it follows that \( q_2^0 = O(h) \). Then by this result, (29) and the fourth formula of (27), we get \( \zeta_{h,1}^1(t) = O(h) \). This implies that \( \zeta_{h,2}^1(t) = O(h) \) by considering \( \zeta_{h,2}^1(t) = O((\omega^{-1}) \). Similarly, one arrives at that \( \zeta_{h,1}^0(0) = O(1) \), \( \zeta_{h,1}^0(0) = O(1) \) and then \( \zeta_{h,1}^0(t) = O(1) \). Therefore (20) is obtained. The bounds (21) can be derived easily by considering (23) and the bounds of \( \zeta_{h}^k \).

**Defect.** As the last part of the proof, we analyze the defect. Firstly, define the components of the defect, for \( t = nh \),
\[
\begin{align*}
d_1(t+h) &= q_{h,1}(t+h) - q_{h,1}(t) - h\rho_{h,1}(t) - h^2\tilde{b}_1(0)g_1(\tilde{q}_h(t+c_1h)),
d_2(t+h) &= p_{h,1}(t+h) - p_{h,1}(t) - hb_1(0)g_1(\tilde{q}_h(t+c_1h)),
d_3(t+h) &= q_{h,2}(t+h) - \cos(h\omega)q_{h,2}(t) - h\sin(h\omega)\rho_{h,2}(t)
- h^2\tilde{b}_1(h\omega)g_2(\tilde{q}_h(t+c_1h)),
d_4(t+h) &= p_{h,2}(t+h) - \cos(h\omega)p_{h,2}(t) + \omega\sin(h\omega)q_{h,2}(t)
- hb_1(h\omega)g_2(\tilde{q}_h(t+c_1h)).
\end{align*}
\]
By the definition of the coefficient functions $\zeta_k^h$, $\eta_k^h$, $\xi_k^h$, the following results are true

$$d_1 = O(h^{N+1}), \quad d_2 = O(h^{N+1}), \quad d_3 = O(h^{N+1}), \quad d_4 = O(h^{N+1}).$$

We are now in a position to estimate the remainders (18). To do this, we begin with defining

$$R_k^n = p_k^n - p_{h,k}(t), \quad S_k^n = q_k^n - q_{h,k}(t)$$

for $k = 1, 2$, and the norm

$$\| (S_1^n, R_1^n, S_2^n, R_2^n) \|_* = \| (S_1^n, R_1^n, \omega S_2^n, R_2^n) \|.$$

We first estimate the remainder for the difference $Q^n + c_1 - \tilde{q}_h(t + c_1 h)$. By the first two equations in (9) and the first equation in (24), if the function $g$ satisfies a Lipschitz condition with the Lipschitz constant $L$, one obtains

$$\| Q^n + c_1 - \tilde{q}_h(t + c_1 h) \| \leq \| S^n \| + c_1 h \| R^n \| + h^2 L \| Q^n + c_1 - \tilde{q}_h(t + c_1 h) \|,$$

which yields that

$$\| Q^n + c_1 - \tilde{q}_h(t + c_1 h) \| \leq 2 \| S^n \| + 2 c_1 h \| R^n \| = \alpha^n$$

under the condition that $h \leq \frac{1}{\sqrt{2L}}$. For the remainders, on noticing the scheme of the ERKN integrators, we then have

$$\| (S_1^{n+1}, R_1^{n+1}, S_2^{n+1}, R_2^{n+1}) \|_* \leq \| (S_1^n, R_1^n, S_2^n, R_2^n) \|_* + h \kappa_1 \| (S_1^n, R_1^n, S_2^n, R_2^n) \|_* + h \kappa_2 \alpha^n + \kappa_3 h^{N+1},$$

(30)

where $\kappa_1$, $\kappa_2$, $\kappa_3$ are constants. From the definition of the initial values, it follows that $\| (S_1^0, R_1^0, S_2^0, R_2^0) \|_* = O(h^{N+1})$. By using the relation (30) repeatedly, we obtain the following estimate for the remainders $\| (S_1^n, R_1^n, S_2^n, R_2^n) \|_* \leq Cnh^{N+1}$, which yields (18).

We complete the proof of this theorem. \hfill \Box

4.2 The first almost-invariant

Let

$$\xi = (\xi_h^{-N+1}, \ldots, \xi_h^{-1}, \xi_h^0, \xi_h^1, \ldots, \xi_h^{N-1}),$$
$$\eta = (\eta_h^{-N+1}, \ldots, \eta_h^{-1}, \eta_h^0, \eta_h^1, \ldots, \eta_h^{N-1}).$$

We have the first almost-invariant of the modulated Fourier functions as follows.
Theorem 6 Under the conditions of Theorem 5, there exists a function \( \hat{\mathcal{H}}[\xi, \eta] \) such that the coefficient functions of the MFE of IERKN4 satisfy

\[
\hat{\mathcal{H}}[\xi, \eta](t) = \hat{\mathcal{H}}[\xi, \eta](0) + \mathcal{O}(th^N)
\]

for \( 0 \leq t \leq T \). Moreover, this can be expressed as

\[
\hat{\mathcal{H}}[\xi, \eta] = \frac{1}{2}(q_{h,1}^0)^T q_{h,1}^0 + 2\omega^2(q_{h,2}^{-1})^T q_{h,2}^1 + U(q_h^0(t)) + \mathcal{O}(h^2).
\]

Proof With Theorem 5, we obtain

\[
\mathcal{L}(hD)q_h(t) = h^2 g(q_h(t)) + \mathcal{O}(h^{N+2}),
\]

where the following denotations are used:

\[
q_h(t) = \sum_{|k|<N} q_h^k(t) \text{ with } q_h^k(t) = e^{ik\omega t} \xi_h^k(t).
\]

Considering the definition of \( q_h \) and comparing the coefficients of \( e^{ik\omega t} \) yields the resulting equations in terms of \( q_h^k \):

\[
\mathcal{L}(hD)q_h^k(t) = -h^2 \nabla_{q^{-k}} \mathcal{U}(q_h(t)) + \mathcal{O}(h^{N+2}),
\]

where \( \mathcal{U}(q_h(t)) \) is defined as

\[
\mathcal{U}(q_h(t)) = U(q_h^0(t)) + \sum_{s(\alpha)=0} \frac{1}{m!} U^{(m)}(q_h^0(t)) (q_h(t))^\alpha,
\]

and \( q_h(t) \) is given by

\[
q_h(t) = (q_h^{-N+1}(t), \ldots, q_h^1(t), q_h^0(t), q_h^1(t), \ldots, q_h^N(t)).
\]

Multiplying the equation (33) with \( (q_h^{-k})^T \) and summing up conforms

\[
\frac{1}{h^2} \sum_{|k|<N} (\dot{q}_h^{-k})^T \mathcal{L}(hD)q_h^k + \frac{d}{dt} \mathcal{U}(q_h) = \mathcal{O}(h^N).
\]

We switch to the quantities \( \xi_h^k(t) \) and get the equivalent relation

\[
\mathcal{O}(h^N) = \frac{1}{h^2} \sum_{|k|<N} (\xi_h^{-k} - i k \omega \xi_h^{-k})^T \mathcal{L}(hD + ihk\omega)\xi_h^k + \frac{d}{dt} \mathcal{U}(\xi) = \frac{1}{h^2} \sum_{|k|<N} (\dot{\xi}_h^{-k} - i k \omega \dot{\xi}_h^{-k})^T \mathcal{L}(hD + ihk\omega)\xi_h^k + \frac{d}{dt} \mathcal{U}(\xi).
\]
With the Taylor expansions of $L(hD)$ and $L(hD + i k h \omega)$ and the “magic formulas” on p. 508 of [29], we know that the following part appearing in (35) is a total derivative

$$\sum_{|k| < N} (\ddot{\zeta}_h - i k \omega \dot{\zeta}_h) ^T L(hD + i k h \omega) \dot{\zeta}_h.$$ 

Thus, the right-hand side of (35) is a total derivative. Therefore, by (35) and the above analysis, it can be confirmed that there exists a function $\hat{H}$ such that

$$d \frac{d}{dt} \hat{H}[\zeta, \eta](t) = O(h^N)$$

and an integration yields the statement (31) of the theorem.

4.3 The second almost-invariant

**Theorem 7** Under the conditions of Theorem 6, there exists a function $\hat{I}[\zeta, \eta]$ such that the coefficient functions of the modulated Fourier expansion of IERKN4 satisfy

$$\hat{I}[\zeta, \eta](t) = \hat{I}[\zeta, \eta](0) + O(th^N)$$

for $0 \leq t \leq T$. Moreover, this can be expressed as

$$\hat{I}[\zeta, \eta] = 2 \omega^2 (\zeta_{h,2}^{-1}) ^T \zeta_{h,2} + U(\zeta_h) + O(h^2).$$

**Proof** Define the vector function $q(\lambda, t)$ of $\lambda$ as

$$q(\lambda, t) = \left( e^{i(-N+1)\lambda \omega} q_h^{-N+1}(t), \ldots, q_h^0(t), \ldots, e^{i(N-1)\lambda \omega} q_h^{N-1}(t) \right).$$

Then it follows from the definition (34) that $U(q(\lambda, t))$ does not depend on $\lambda$. Thus, its derivative with respect to $\lambda$ yields

$$0 = \frac{d}{d\lambda} U(q(\lambda, t)) = \sum_{|k| < N} ik \omega e^{i k \lambda \omega} (q_h^k(t))^T \nabla_q U(q(\lambda, t)).$$
Letting $\lambda = 0$ yields
\[ \sum_{|k|<N} i k \omega (q_h^{-k}(t))^T \nabla q^k \mathcal{U}(q_h(t)) = 0. \]
Therefore, one obtains
\[ 0 = \sum_{|k|<N} i k \omega (q_h^{-k}(t))^T \nabla q^{-k} \mathcal{U}(q_h(t)) \]
\[ = \frac{i \omega}{h^2} \sum_{|k|<N} k (q_h^{-k}(t))^T \mathcal{L}(hD) q_h^k(t) + \mathcal{O}(h^N). \]

Rewritten in the $\zeta_h^k(t)$ variables, this becomes
\[ \frac{i \omega}{h^2} \sum_{|k|<N} k (\zeta_h^{-k}(t))^T \mathcal{L}(hD + i k \omega) \zeta_h^k(t) = \mathcal{O}(h^N). \]

As in the proof of Theorem 6, the left-hand expression of this equation can be written as the time derivative of a function. Therefore, we get $\frac{d}{dt} \hat{I}[\zeta, \eta](t) = \mathcal{O}(h^N)$ and an integration yields statement (36) of the theorem.

According to the above analysis and the bounds of Theorem 5, the construction of $\hat{I}[\zeta, \eta](t)$ is obtained, which concludes the proof.

### 4.4 Long-time near-conservation of total and oscillatory energy

Theorem 4 will be proved in this subsection. Before that, we give the following theorem.

**Theorem 8** Under the conditions of Theorem 6, it holds that
\[ \hat{H}[\zeta, \eta](nh) = H(q^n, p^n) + \mathcal{O}(h), \]
\[ \hat{I}[\zeta, \eta](nh) = I(q^n, p^n) + \mathcal{O}(h), \]
where the constants symbolized by $\mathcal{O}$ depend on $N$, $T$ and the constants in the assumptions.

**Proof** In terms of the analysis given in Section XIII of [29] and the bounds presented in Theorem 5, one arrives at
\[ H(q^n, p^n) = \frac{1}{2} (\eta_{h,1}^0)^T \eta_{h,1}^0 + 2 \omega^2 (\zeta_{h,2}^{-1})^T \zeta_{h,2}^1 + U(\zeta_h^0) + \mathcal{O}(h), \]
\[ I(q^n, p^n) = 2 \omega^2 (\zeta_{h,2}^{-1})^T \zeta_{h,2}^1 + \mathcal{O}(h). \]

A comparison between (32), (37) and (38) gives the stated relations of this theorem. \(\square\)

On the basis of previous analysis given in this section and following the approach used in Chapter XIII of [29], Theorem 4 of IERKN4 is easily proved by patching together the local near-conservation result. For other methods, we just need to modify the operator $L_2(hD)$ and use its Taylor expansion.
5 Proof for explicit integrators (Theorem 3)

5.1 Proof

It is noted that by letting $\bar{a}_{11} = 0$, the procedure given in the above section also holds for explicit ERKN integrators. This can be used as the proof of Theorem 3.

5.2 Proof in another way

In the rest part of this section, we prove Theorem 3 in another way. This will be achieved by firstly showing that an ERKN method with specific parameters is equivalent to a symmetric trigonometric integrator researched in [5,21,29] and then using thereby leveraged results of [5,21,29].

5.2.1 One important connection

We denote the numerical flow of a one-stage explicit symmetric ERKN integrator with the symmetry condition (10) by $\Phi_h$, i.e.,

\[
(q^{n+1}, p^{n+1}) = \Phi_h(q^n, p^n).
\]

We now consider a Strang splitting applied to an averaged equation. More precisely, let $\Phi_{h,L}$ be the time $h$ flow of the linear equation $\frac{d}{dt}(q, p) = (p, \Omega^2 q)$ and $\Phi_{h,NL}$ be the time $h$ flow of the nonlinear equation $\frac{d}{dt}(q, p) = (0, \Upsilon(h\Omega)g(q))$, where $\Upsilon$ is a function of $h\Omega$ which will be determined below.

We then consider the following Strang splitting method

1. $(q_+^n, p_+^n) = \Phi_{h/2, L}(q^n_+, p^n_+) :$
   
   \[
   \begin{pmatrix}
   q_+^n \\
   p_+^n
   \end{pmatrix} = \begin{pmatrix}
   \cos\left(\frac{h\Omega}{2}\right) & \Omega^{-1} \sin\left(\frac{h\Omega}{2}\right) \\
   \Omega \sin\left(\frac{h\Omega}{2}\right) & \cos\left(\frac{h\Omega}{2}\right)
   \end{pmatrix} \begin{pmatrix}
   q^n \\
   p^n
   \end{pmatrix}.
   \]

2. $(q_-^n, p_-^n) = \Phi_{h,NL}(q_+^n, p_+^n) :$
   
   \[
   \begin{pmatrix}
   q_-^n \\
   p_-^n
   \end{pmatrix} = \begin{pmatrix}
   q_+^n \\
   p_+^n + h\Upsilon(h\Omega)g(q_-^n)
   \end{pmatrix}.
   \]

3. $(q^{n+1}_-, p^{n+1}_-) = \Phi_{h/2, L}(q_-^n, p_-^n) :$
   
   \[
   \begin{pmatrix}
   q^{n+1}_- \\
   p^{n+1}_-
   \end{pmatrix} = \begin{pmatrix}
   \cos\left(\frac{h\Omega}{2}\right) & \Omega^{-1} \sin\left(\frac{h\Omega}{2}\right) \\
   \Omega \sin\left(\frac{h\Omega}{2}\right) & \cos\left(\frac{h\Omega}{2}\right)
   \end{pmatrix} \begin{pmatrix}
   q_-^n \\
   p_-^n
   \end{pmatrix}.
   \]

It can be straightly verified that this Strang splitting method is identical to the ERKN integrator $\Phi_h$, i.e.,

\[
\Phi_h = \Phi_{h/2, L} \circ \Phi_{h,NL} \circ \Phi_{h/2, L} \tag{39}
\]

if and only if the following conditions hold

\[
\frac{1}{2} \sin\left(\frac{1}{2} h\Omega\right) \Upsilon(h\Omega) = \bar{b}_1(h\Omega), \quad \cos\left(\frac{1}{2} h\Omega\right) \Upsilon(h\Omega) = b_1(h\Omega). \tag{40}
\]
For general ERKN integrators, these two conditions may not hold simultaneously. However, it is noted that under the symmetry condition \((10), (40)\) is true. In this particular case, the function \(\Upsilon\) has the form

\[
\Upsilon(h\Omega) = b_1(h\Omega)\cos^{-1}\left(\frac{1}{2}h\Omega\right) = 2\bar{b}_1(h\Omega)\sin^{-1}\left(\frac{1}{2}h\Omega\right).
\]  

(41)

On the other hand, if we consider the Strang splitting in another way:

\[
\hat{\Phi}_h = \Phi_{h/2,\text{NL}} \circ \Phi_{h,\text{L}} \circ \Phi_{h/2,\text{NL}},
\]

which yields a class of trigonometric integrators

\[
\begin{align*}
q^{n+1} &= \cos(h\Omega)q^n + h\text{sinc}(h\Omega)\, p^n + \frac{1}{2}h^2\text{sinc}(h\Omega)\, \Upsilon(h\Omega)\, g(q^n), \\
p^{n+1} &= -\Omega\, \sin(h\Omega)q^n + \cos(h\Omega)\, p^n \\
&\quad + \frac{1}{2}h\left(\cos(h\Omega)\, \Upsilon(h\Omega)\, g(q^n) + \Upsilon(h\Omega)\, g(q^{n+1})\right).
\end{align*}
\]  

(42)

It is noted that this is exactly a trigonometric integrator (a form of (XIII.2.7)–(XIII.2.8) given on p.481 of [29]) with the following coefficients

\[
\begin{align*}
\Phi(h\Omega) &= 1, \\
\Psi(h\Omega) &= \text{sinc}(h\Omega)\, \Upsilon(h\Omega), \\
\Psi_0(h\Omega) &= \cos(h\Omega)\, \Upsilon(h\Omega), \\
\Psi_1(h\Omega) &= \Upsilon(h\Omega).
\end{align*}
\]  

(43)

In terms of the symmetry condition proposed in [29], it is clear that this trigonometric integrator is symmetric.

On the basis of the above analysis, the following important connection between symmetric ERKN integrators and symmetric trigonometric integrators is obtained.

**Proposition 1** For the one-stage explicit symmetric ERKN integrator and the symmetric trigonometric integrator \((42)\), they are related by

\[
\Phi_h = \Phi_{-h/2,\text{L}} \circ \Phi_{-h/2,\text{NL}} \circ \hat{\Phi}_h \circ \Phi_{h/2,\text{NL}} \circ \Phi_{h/2,\text{L}}.
\]  

(44)

The function \(\Upsilon\) appearing in the averaged equation and in \((43)\) is determined by \((41)\). Moreover, it is true that

\[
\underbrace{\Phi_h \circ \cdots \circ \Phi_h}_{n \text{ times}} = \Phi_{-h/2,\text{L}} \circ \Phi_{-h/2,\text{NL}} \circ \left(\underbrace{\hat{\Phi}_h \circ \cdots \circ \hat{\Phi}_h}_{n \text{ times}}\right) \circ \Phi_{h/2,\text{NL}} \circ \Phi_{h/2,\text{L}}
\]

\[
= \Phi_{h/2,\text{L}} \circ \Phi_{h/2,\text{NL}} \circ \left(\underbrace{\hat{\Phi}_h \circ \cdots \circ \hat{\Phi}_h}_{n-1 \text{ times}}\right) \circ \Phi_{h/2,\text{NL}} \circ \Phi_{h/2,\text{L}}.
\]  

(45)

### 5.2.2 Proof for explicit symmetric integrators

Since the long-time behaviour of symmetric trigonometric integrators is well understood (see [5,21,29]), the long-time near conservation of total and oscillatory energy of one-stage explicit symmetric ERKN integrators can be derived by using the relation
(45). To be very precise, according to the analysis of [5,21,29], one has to verify the following four key points.

- I. If \((q^0, p^0)\) satisfies (4), the following initial value
  \[
  (\tilde{q}^0, \tilde{p}^0) := \Phi_{h/2, \text{NL}} \circ \Phi_{h/2, \text{L}}(q^0, p^0)
  \]
  for \((\hat{\Phi}_h \circ \cdots \circ \hat{\Phi}_h)\) satisfies the finite-energy condition
  \[
  \frac{1}{2} \| \tilde{p}^0 \|^2 + \frac{1}{2} \| \Omega \tilde{q}^0 \|^2 \leq E.
  \] (46)

- II. For any \(j \in \mathbb{N}^+\), \((\hat{\Phi}_h \circ \cdots \circ \hat{\Phi}_h) \circ \Phi_{h/2, \text{NL}} \circ \Phi_{h/2, \text{L}}(q^0, p^0)\) stays in a compact set if \((\Phi_h \circ \cdots \circ \Phi_h)(q^0, p^0)\) does.

- III. The two additional steps with \(\Phi_{h/2, \text{L}}\) and \(\Phi_{h/2, \text{NL}}\) only introduce an \(O(h)\) deviation in the total and oscillatory energy, provided that the corresponding initial values \((\tilde{q}, \tilde{p})\) are bounded and also \(\Omega \tilde{q}\) is bounded.

- IV. The condition of [5] on \(\Phi\) and \(\Psi\) has to be satisfied for the choice (43).

In what follows, we show that the above four conditions are completely true.

- For the point I, according to the splitting method, it is easy to obtain that

  \[
  \tilde{q}^0 = \cos \left( \frac{h \Omega}{2} \right) q^0 + \Omega^{-1} \sin \left( \frac{h \Omega}{2} \right) p^0, \\
  \tilde{p}^0 = -\Omega \sin \left( \frac{h \Omega}{2} \right) q^0 + \cos \left( \frac{h \Omega}{2} \right) p^0 + \frac{1}{2} h \Upsilon \left( \frac{h \Omega}{2} \right) g(\tilde{q}^0).
  \]

  Thus (46) is clear by considering (4).

- It follows from (44) that

  \[
  (\hat{\Phi}_h \circ \cdots \circ \hat{\Phi}_h) \circ \Phi_{h/2, \text{NL}} \circ \Phi_{h/2, \text{L}}(q^0, p^0)
  \]
  \[
  = \Phi_{h/2, \text{NL}} \circ \Phi_{h/2, \text{L}} \circ (\Phi_h \circ \cdots \circ \Phi_h)(q^0, p^0),
  \]

  which gives the statement of II.

- The third result is clear in the light of the definitions of \(\Phi_{h/2, \text{L}}\) and \(\Phi_{h/2, \text{NL}}\).

- We remark that the long-term analysis of [21,29] requires

  \[
  |\Phi(h \omega)| \geq C \left| \text{sinc} \left( \frac{1}{2} h \omega \right) \right|.
  \]
However, for the trigonometric integrator (42) with \( \Phi = 1 \), it does not satisfy the above requirement. Therefore, the long-term analysis of \([21,29]\) can not be used here. Recently, the authors in [5] improved the analysis and presented a long-term analysis of numerical integrators for oscillatory Hamiltonian systems under minimal non-resonance conditions. A more relaxed restriction on \( \Phi \) was given there. In terms of that restriction and the relations (40) and (44), the statement of IV holds provided (16) is true.

Based on the above analysis and the long-time result of symmetric trigonometric integrators given in [5,21,29], Theorem 3 is immediately obtained for explicit symmetric integrators.

5.2.3 Proof for explicit symplectic integrators

It can be checked that the explicit symplectic integrator \( \tilde{\Phi}_h \) can be rewritten as

\[
\tilde{\Phi}_h = \Phi_{c_1, h, L} \circ \Phi_{h, N, L} \circ \Phi_{(1-c_1)h, L}
\]

with \( \gamma = d_1 \). Then for this integrator and the one-stage explicit symmetric ERKN integrator \( \Phi_h \), they have the relationship

\[
\tilde{\Phi}_h \circ \cdots \circ \tilde{\Phi}_h = \Phi_{(c_1-1/2)h, L} \circ \left( \Phi_{h} \circ \cdots \circ \Phi_{h} \right) \circ \Phi_{(1-c_1-1/2)h, L}.
\]

From the result of \( \Phi_h \) and this connection, it follows that Theorem 3 is true for explicit symplectic integrators.

6 Conclusions

In this paper, the long-time total and oscillatory energy conservation behaviour of one-stage ERKN integrators was studied when applied to highly oscillatory Hamiltonian systems. It turned out that a good long-time energy conservation holds not only for symmetric integrators but also for symplectic integrators and not only for explicit schemes but also for implicit ones. A relationship between explicit ERKN integrators and symmetric trigonometric integrators was established and on the basis of which, the long-time conservation for explicit ERKN integrators was proved.

It is noted that the long-time behaviour of ERKN integrators for multi-frequency highly oscillatory Hamiltonian systems can be easily derived by using the technology proposed in this paper and in [7].

This is a preliminary research on the long-time behaviour of ERKN integrators for highly oscillatory Hamiltonian systems and the authors are clearly aware that there are still some issues which will be further considered.

- The energy conservation behaviour of symmetric or symplectic ERKN integrators in other ODEs such as Hamiltonian systems with a solution-dependent high frequency or without any non-resonance condition will also be considered.
– We only consider one-stage ERKN integrators in this paper. The extension of this paper’s analysis to higher-stage ERKN integrators is not obvious since there is the technical difficulty which needs to be overcome. This issue will be considered in future investigations.

Acknowledgements The authors are sincerely thankful to three anonymous reviewers for their valuable suggestions, which help improve the presentation of the manuscript significantly. We are grateful to Professor Christian Lubich for his helpful comments and discussions on the topic of modulated Fourier expansions. We also thank Ludwig Gauckler and one anonymous reviewer of FoCM for drawing our attention to the connection between explicit ERKN methods and symmetric trigonometric integrators.

References

1. Blanes, S.: Explicit symplectic RKN methods for perturbed non-autonomous oscillators: splitting, extended and exponentially fitting methods. Comput. Phys. Commun. 193, 10–18 (2015)
2. Buchholz, S., Gauckler, L., Grimm, V., Hochbruck, M., Jahnke, T.: Closing the gap between trigonometric integrators and splitting methods for highly oscillatory differential equations. IMA J. Numer. Anal. 38, 57–74 (2018)
3. Cohen, D.: Conservation properties of numerical integrators for highly oscillatory Hamiltonian systems. IMA J. Numer. Anal. 26, 34–59 (2006)
4. Cohen, D., Gauckler, L.: One-stage exponential integrators for nonlinear Schrödinger equations over long times. BIT 52, 877–903 (2012)
5. Cohen, D., Gauckler, L., Hairer, E., Lubich, C.: Long-term analysis of numerical integrators for oscillatory Hamiltonian systems under minimal non-resonance conditions. BIT 55, 705–732 (2015)
6. Cohen, D., Hairer, E., Lubich, C.: Modulated Fourier expansions of highly oscillatory differential equations. Found. Comput. Math. 3, 327–345 (2003)
7. Cohen, D., Hairer, E., Lubich, C.: Numerical energy conservation for multi-frequency oscillatory differential equations. BIT 45, 287–305 (2005)
8. Cohen, D., Hairer, E., Lubich, C.: Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions. Arch. Ration. Mech. Anal. 187, 341–368 (2008)
9. Cohen, D., Hairer, E., Lubich, C.: Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations. Numer. Math. 110, 113–143 (2008)
10. Cohen, D., Jahnke, T., Lorenz, K., Lubich, C.: Numerical integrators for highly oscillatory Hamiltonian systems: a review. In: Mielke, A. (ed.) Analysis, Modeling and Simulation of Multiscale Problems, pp. 553–576. Springer, Berlin (2006)
11. Franco, J.M.: New methods for oscillatory systems based on ARKN methods. Appl. Numer. Math. 56, 1040–1053 (2006)
12. García-Archilla, B., Sanz-Serna, J.M., Skeel, R.D.: Long-time-step methods for oscillatory differential equations. SIAM J. Sci. Comput. 20, 930–963 (1999)
13. Gauckler, L.: Geometric numerical integration of nonlinear Schrödinger and nonlinear wave equations), Habilitationsschrift (habilitation thesis), Technische Universität Berlin (2017). http://userpage.fu-berlin.de/gauckler/habil-web.pdf
14. Gauckler, L., Hairer, E., Lubich, C.: Energy separation in oscillatory Hamiltonian systems without any non-resonance condition. Comm. Math. Phys. 321, 803–815 (2013)
15. Gauckler, L., Hairer, E., Lubich, C.: Long-term analysis of semilinear wave equations with slowly varying wave speed. Comm. Part. Differ. Equ. 41, 1934–1959 (2016)
16. Gauckler, L., Lubich, C.: Splitting integrators for nonlinear Schrödinger equations over long times. Found. Comput. Math. 10, 275–302 (2010)
17. Gauckler, L., Lubich, C.: Nonlinear Schrödinger equations and their spectral semi-discretizations over long times. Found. Comput. Math. 10, 141–169 (2010)
18. Grimm, V., Hochbruck, M.: Error analysis of exponential integrators for oscillatory second-order differential equations. J. Phys. A Math. Gen. 39, 5495–5507 (2006)
19. Guzzo, M., Benettin, G.: A spectral formulation of the Nekhoroshev theorem and its relevance for numerical and experimental data analysis. Disc. Dyn. Syst. Ser. B 1, 1–28 (2001)
20. Hairer, E., Lubich, C.: Energy conservation by Störmer-type numerical integrators. In: Griffiths, D.F., Watson, G.A. (eds.) Numerical Analysis 1999, pp. 169–190. CRC Press, Boca Raton (2000)
21. Hairer, E., Lubich, C.: Long-time energy conservation of numerical methods for oscillatory differential equations. SIAM J. Numer. Anal. 38, 414–441 (2000)
22. Hairer, E., Lubich, C.: Spectral semi-discretisations of weakly nonlinear wave equations over long times. Found. Comput. Math. 8, 319–334 (2008)
23. Hairer, E., Lubich, C.: Oscillations over long times in numerical Hamiltonian systems, in Highly oscillatory problems. In: Engquist, B., Fokas, A., Hairer, E., Iserles, A. (eds.) London Mathematical Society Lecture Note Series, vol. 366. Cambridge University Press, Cambridge (2009)
24. Hairer, E., Lubich, C.: Modulated Fourier expansions for continuous and discrete oscillatory systems. In: Cucker, F., et al. (Eds.) Foundations of Computational Mathematics: Budapest 2011, pp. 113–128. Cambridge University Press, Cambridge (2012)
25. Hairer, E., Lubich, C.: Long-term control of oscillations in differential equations. Int. Math. Nachr. 223, 1–16 (2013)
26. Hairer, E., Lubich, C.: Long-term analysis of the Störmer–Verlet method for Hamiltonian systems with a solution-dependent high frequency. Numer. Math. 134, 119–138 (2016)
27. Hairer, E., Lubich, C.: Long-term analysis of a variational integrator for charged-particle dynamics in a strong magnetic field. Numer. Math. 144, 699–728 (2020)
28. Hairer, E., Lubich, C., Wang, B.: A filtered Boris algorithm for charged-particle dynamics in a strong magnetic field. Numer. Math. 144, 787–809 (2020)
29. Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd edn. Springer, Berlin (2006)
30. Hochbruck, M., Lubich, C.: A Gautschi-type method for oscillatory second-order differential equations. Numer. Math. 83, 403–426 (1999)
31. Hochbruck, M., Ostermann, A.: Exponential integrators. Acta Numer. 19, 209–286 (2010)
32. Iserles, A., Nørsett, S.P.: From high oscillation to rapid approximation I: modified Fourier expansions. IMA J. Numer. Anal. 28, 862–887 (2008)
33. Li, Y.W., Wu, X.: Exponential integrators preserving first integrals or Lyapunov functions for conservative or dissipative systems. SIAM J. Sci. Comput. 38, 1876–1895 (2016)
34. Li, Y.W., Wu, X.: Functionally fitted energy-preserving methods for solving oscillatory nonlinear Hamiltonian systems. SIAM J. Numer. Anal. 54, 2036–2059 (2016)
35. McLachlan, R.I., Stern, A.: Modified trigonometric integrators. SIAM J. Numer. Anal. 52, 1378–1397 (2014)
36. Mei, L., Wu, X.: Symplectic exponential Runge–Kutta methods for solving nonlinear Hamiltonian systems. J. Comput. Phys. 338, 567–584 (2017)
37. Sanz-Serna, J.M.: Modulated Fourier expansions and heterogeneous multiscale methods. IMA J. Numer. Anal. 29, 595–605 (2009)
38. Stern, A., Grinspun, E.: Implicit–explicit variational integration of highly oscillatory problems. Multiscale Model. Simul. 7, 1779–1794 (2009)
39. Wang, B., Iserles, A., Wu, X.: Arbitrary-order trigonometric Fourier collocation methods for multi-frequency oscillatory systems. Found. Comput. Math. 16, 151–181 (2016)
40. Wang, B., Wu, X.: Long-time momentum and actions behaviour of energy-preserving methods for semilinear wave equations via spatial spectral semi-discretizations. Adv. Comput. Math. 45, 2921–2952 (2019)
41. Wang, B., Wu, X.: The formulation and analysis of energy-preserving schemes for solving high-dimensional nonlinear Klein–Gordon equations. IMA J. Numer. Anal. 39, 2016–2044 (2019)
42. Wang, B., Wu, X., Xia, J.: Error bounds for explicit ERKN integrators for systems of multi-frequency oscillatory second-order differential equations. Appl. Numer. Math. 74, 17–34 (2013)
43. Wang, B., Yang, H., Meng, F.: Sixth order symplectic and symmetric explicit ERKN schemes for solving multi-frequency oscillatory nonlinear Hamiltonian equations. Calcolo 54, 117–140 (2017)
44. Wu, X., Wang, B.: Recent Developments in Structure-Preserving Algorithms for Oscillatory Differential Equations. Springer, Berlin (2018)
45. Wu, X., Wang, B., Xia, J.: Explicit symplectic multidimensional exponential fitting modified Runge–Kutta–Nyström methods. BIT 52, 773–795 (2012)
46. Wu, X., You, X., Shi, W., Wang, B.: ERKN integrators for systems of oscillatory second-order differential equations. Comput. Phys. Comm. 181, 1873–1887 (2010)
47. Wu, X., Wang, B., Mei, L.: Oscillation-preserving algorithms for efficiently solving highly oscillatory second-order ODEs. Numer. Algo. 86, 693–727 (2021)
48. Wu, X., You, X., Wang, B.: Structure-Preserving Algorithms for Oscillatory Differential Equations. Springer, Berlin (2013)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.