Version of 23 January 2012

On the energy distribution in Fermi–Pasta–Ulam lattices

Ernst Hairer[1], **Christian Lubich**[2]

[1] Section de mathématiques, 2-4 rue du Lièvre, Université de Genève, CH-1211 Genève 4, Switzerland. E-mail: Ernst.Hairer@unige.ch
[2] Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle, D-72076 Tübingen, Germany. E-mail: Lubich@na.uni-tuebingen.de

Abstract: For FPU chains with large particle numbers, the formation of a packet of modes with geometrically decaying harmonic energies from an initially excited single low-frequency mode and the metastability of this packet over longer time scales are rigorously studied in this paper. The analysis uses modulated Fourier expansions in time of solutions to the FPU system and exploits the existence of almost-invariant energies in the modulation system. The results and techniques apply to the FPU α- and β-models as well as to higher-order nonlinearities. They are valid in the regime of scaling between particle number and total energy in which the FPU system can be viewed as a perturbation to a linear system, considered over time scales that go far beyond standard perturbation theory. Weak non-resonance estimates for the almost-resonant frequencies determine the time scales that can be covered by this analysis.

1. Introduction

> This report is intended to be the first one in a series dealing with the behavior of certain nonlinear physical systems where the non-linearity is introduced as a perturbation to a primarily linear problem. The behavior of the systems is to be studied for times which are long compared to the characteristic periods of the corresponding linear problem.
>
> E. Fermi, J. Pasta, S. Ulam (1955)

The numerical experiment by Fermi, Pasta and Ulam [11], which showed unexpected recurrent behaviour instead of relaxation to equipartition of energy in a chain of weakly nonlinearly coupled particles, has incited a wealth of research in both mathematics and physics and continues to do so; see the recent volume edited by Gallavotti [14] and the review by Berman and Izrailev [7] as well as the accounts on the earlier FPU history by Ford [13] and Weissert [20].

The present paper contributes to the vast literature on FPU with an analysis, for large particle numbers, of the questions of the formation of a packet of modes with geometrically decaying energies, starting from a single excited mode, and of the metastability shown in the perseverance of the packet over longer time scales (see Benettin, Carati, Galgani and Giorgilli [3] for a review of the FPU problem from the metastability perspective). Only recently, such questions have been addressed analytically in an impressive paper by Bambusi and Ponno [1]. Here we present an approach to these questions that differs substantially: in the scaling between particle number N and total energy E considered, in the time scales of metastability obtained, and in the kind of analysis employed. The scaling considered here is such that the nonlinearity can indeed be viewed as a perturbation to the linear problem (cf. the citation above), which is the case for $E \ll N^{-3}$ in the FPU α-model (cubic potential), whereas in [1] the scaling $E \sim N^{-3}$ is studied. In the FPU β-model (quartic potential) we require $E \ll N^{-1}$. We derive our results using a non-resonant *modulated Fourier expansion* in time of the solution to the FPU system, as opposed to the use of an integrable resonant normal form in [1]. The results and techniques developed here apply to FPU α- and β-models as well as to FPU systems with higher-order nonlinearities, whereas those of [1] are so far restricted to the α-model. Integrability is an essential concept in [1], but plays no role here.

The modulated Fourier expansion employed here is a multiscale expansion whose coefficient functions are constructed from a modulation system that retains a Hamiltonian structure. The modulation system is shown to have almost-invariants that majorize the normal mode energies of the FPU system. This permits us to explain the preservation of the low-frequency packet over time scales that are much longer than the time scale for the formation of the packet. We mention that modulated Fourier expansions have been used similarly before in the analysis of numerical methods for oscillatory Hamiltonian differential equations [17,8] and in the long-time analysis of non-linearly perturbed wave equations [9] and Schrödinger equations [15,16].

For previous numerical experiments that are in relation to the present work, we refer to De Luca, Lichtenberg and Ruffo [10], Berchialla, Galgani and Giorgilli [6], and Flach, Ivanchenko and Kanakov [12].

The paper is organized as follows: In Section 2 we introduce the necessary notation and formulate the problem. Section 3 presents numerical experiments that illustrate the main result, Theorem 1, which is stated in Section 4. This theorem provides rigorous bounds on the geometric decay of the energies of all modes and on the long-time preservation of the packet, in the case of the FPU α-model with sufficiently small initial excitation in the first mode. The proof of Theorem 1 is given in Sections 5 to 8. Section 5 provides the necessary weak non-resonance estimates for the frequencies of the FPU system. These are needed for the construction of the non-resonant modulated Fourier expansion given in Section 6. Some further bounds for the modulation functions are derived in Section 7. In Section 8 we construct the almost-invariant energies of the modulation system and show that they bound the energies in the modes of the FPU system. We are then able to bound the mode energies over longer time scales than the validity of the modulated Fourier expansion and complete the proof of Theorem 1. We obtain even longer time scales by including certain high-order resonances among frequencies in the construction of the modulated Fourier expansion, which is done for the first appearing resonance in Section 9. Finally, the extension of Theorem 1 to the FPU β-model and higher-order models is given in Section 10.

2. Formulation of the problem

The periodic Fermi–Pasta–Ulam lattice with $2N$ particles has the Hamiltonian

$$H = \sum_{n=1}^{2N} \frac{1}{2} p_n^2 + \sum_{n=1}^{2N} \Big(\frac{1}{2} (q_{n+1} - q_n)^2 + V(q_{n+1} - q_n) \Big),$$

where the real sequences (p_n) and (q_n) are $2N$-periodic, and has the equations of motion

$$\ddot{q}_n - (q_{n+1} - 2q_n + q_{n-1}) = V'(q_{n+1} - q_n) - V'(q_n - q_{n-1})\,. \qquad (1)$$

The non-quadratic potential is typically taken as $V(x) = \alpha x^3/3 + \beta x^4/4$. With the discrete Fourier coefficients[1]

$$\mathbf{u} = \big(u_j\big)_{j=-N}^{N-1}, \qquad q_n = \sum_{j=-N}^{N-1} u_j\, \mathrm{e}^{\mathrm{i}jn\pi/N}, \qquad (2)$$

we obtain, for the special case $\alpha = 1$ and $\beta = 0$, the system

$$\ddot{u}_j + \omega_j^2\, u_j \;=\; -\,\mathrm{i}\,\omega_j \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)}\, \omega_{j_1}\, \omega_{j_2}\; u_{j_1}\, u_{j_2} \qquad (3)$$

with frequencies

$$\omega_j = 2\sin\Big(\frac{j\pi}{2N}\Big), \qquad j = -N, \dots, N-1. \qquad (4)$$

Equation (3) is a complex Hamiltonian system of the form

$$\ddot{u}_j + \omega_j^2 u_j = -\nabla_{-j} U(\mathbf{u}) \qquad (5)$$

(∇_{-j} denotes derivative with respect to u_{-j}) with potential

$$U(\mathbf{u}) \;=\; -\,\frac{\mathrm{i}}{3} \sum_{j_1+j_2+j_3=0 \bmod 2N} (-1)^{(j_1+j_2+j_3)/(2N)}\, \omega_{j_1}\, \omega_{j_2}\, \omega_{j_3}\; u_{j_1}\, u_{j_2}\, u_{j_3}. \qquad (6)$$

The *total energy* of the FPU system is

$$E = \sum_{j=-N}^{N-1} E_j + 2N\, U, \qquad (7)$$

where E_j is the jth *normal mode energy*,

$$E_j = 2N\, \epsilon_j \quad \text{with} \quad \epsilon_j = \frac{1}{2}\big|\dot{u}_j\big|^2 + \frac{1}{2}\omega_j^2\big|u_j\big|^2. \qquad (8)$$

Our interest in this paper is in the time evolution of these mode energies, and we write $E_j(t) = E_j(\mathbf{u}(t), \dot{\mathbf{u}}(t))$. Since the q_n are real, we have $u_{-j} = \overline{u}_j$, and hence $E_{-j} = E_j$ for all j. Occasionally it will be convenient to work with the specific energy

$$\epsilon = \frac{E}{2N}\,. \qquad (9)$$

We are mainly interested in small initial data that are different from zero only in a single pair of modes $\pm j_0 \neq 0$:

$$u_j(0) = \dot{u}_j(0) = 0 \quad \text{for} \quad j \neq \pm j_0. \qquad (10)$$

[1] In contrast to much of the FPU literature we omit the normalization factor $1/\sqrt{2N}$ in (2). With this scaling there is no factor $\sqrt{2N}$ in the system (3) and no factor $2N$ in the potential (6), but the factor $2N$ appears in the energies (7) and (8).

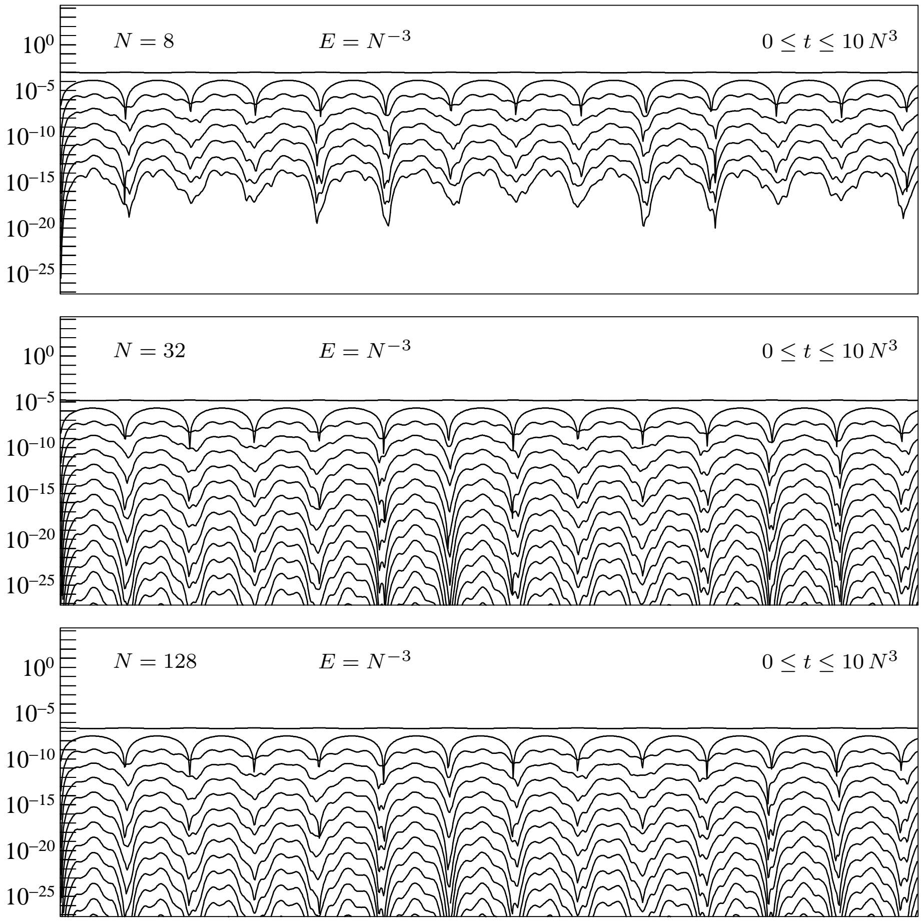

Fig. 1. Normal mode energies $E_j(t)$ as functions of time for the FPU α-problem with initial data of Section 3; increasing j corresponds to decreasing values of $E_j(t)$.

Equivalently, only the j_0th mode energy is non-zero initially. Because of $\ddot{u}_0 = 0$ (which follows from $\omega_0 = 0$) this also implies that $u_0(t) = 0$ for all t. This property is valid for general potentials V, because summing up the equation (1) over all n shows that the second derivative of $\sum_n q_n$ vanishes identically.

Remark 1. The original article [11] considers the differential equation (1) with fixed boundaries $q_N = q_0 = 0$. Extending such data by $q_{-j} = -q_j$ to a $2N$-periodic sequence, we notice that the extension is still a solution of (1), so that all statements of the present article remain valid also in this situation.

3. Numerical experiment

For our numerical experiment we consider the potential $V(x) = x^3/3$, i.e., $\alpha = 1$ and $\beta = 0$. For N an integral power of 2, we consider initial values

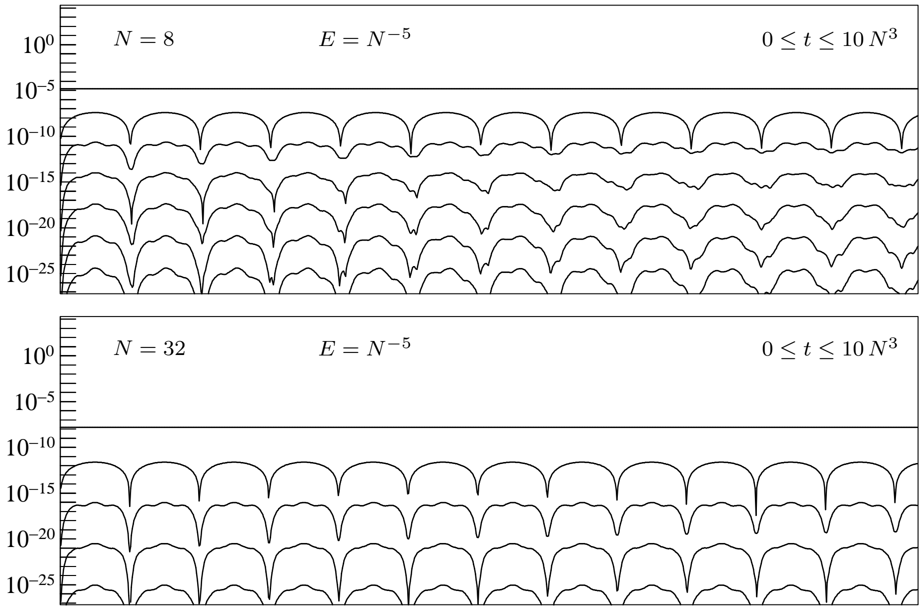

Fig. 2. Normal mode energies $E_j(t)$ as functions of time for the FPU α-problem with initial data as in Fig. 1, but with smaller total energy $E = N^{-5}$.

$$q_n(0) = \frac{\sqrt{2\,\epsilon}}{\omega_1}\,\sin x_n, \qquad \dot{q}_n(0) = \sqrt{2\,\epsilon}\,\cos x_n \tag{11}$$

with $x_n = n\pi/N$, so that (10) is satisfied with $j_0 = 1$ and the total energy is $E = 2N\epsilon$. The solution is computed numerically by a trigonometric method (treating without error the linear part of the differential equation) with step size $h = 0.01$ for $N = 8$, with $h = 0.1$ for $N = 32$, and with $h = 1$ for $N = 128$. Figure 1 shows the mode energies $E_j(t)$ as functions of time for $E = N^{-3}$. This is the situation treated in [1]. Figure 2 shows the mode energies for $E = N^{-5}$. In the figures we have chosen time intervals that are proportional to N^3, which looks like a natural time scale for the slow changes in the mode energies. Further numerical experiments with many different values of E and N indicate that, with small

$$\delta = \sqrt{E\,N^3} \ll 1\,, \tag{12}$$

the mode energies behave like

$$E_j(t) \approx E\,\delta^{2(j-1)} f_j(N^{-3}t) \qquad \text{for} \quad j \ge 2 \tag{13}$$

with functions $f_j(\tau)$ that do not depend significantly on E and N and which are of size $\approx c^{j-1}$ with $|c| < 1$. We have $E_1(0) = E_{-1}(0) = E/2$ and we observe that $E_1(t) - E_1(0) \approx -E\,\delta^2 f_2(N^{-3}t)$.

4. Main result

The proof of our main result on the long-time behaviour of normal mode energies is confronted with small denominators, and relies on lower bounds of the form

$$\Big|\omega_j^2 - \Big(\omega_{j-r} + \sum_{\ell=1}^{N} k_\ell \omega_\ell\Big)^2\Big| \geq \gamma\,\frac{\pi^3\,\omega_j}{N^3}, \tag{14}$$

where r and $\sum_{\ell=1}^N \ell |k_\ell|$ are small integers, but j can be arbitrarily large. Because of the special form (4) of the frequencies ω_j, such an estimate can be obtained for nearly all relevant situations, as will be elaborated in Section 5 below. The only difficulty arises when there exists an integer s such that the expression

$$\Big|\cos\Big(\frac{(2j-r)\pi}{4N}\Big) - \frac{s}{r}\Big| \tag{15}$$

is very small. We therefore restrict the admissible dimensions N to those belonging to the non-resonance set $\mathcal{N}(M,\gamma)$ of the following definition, where the integer M will determine the time scale $t \leq cN^2\delta^{-M-1}$ (with $\delta \ll 1$ of (12)) over which we obtain a geometric decay of the mode energies, while $\gamma > 0$ is from the non-resonance condition (14).

Definition 1. *Let an integer M and a constant $\gamma > 0$ be given. The dimension N belongs to the non-resonance set $\mathcal{N}(M,\gamma)$ if for all pairs of integers (s,r) with $0 < s < r \leq M$ and even $r - s$, for the integer j minimizing (15), and for all $(k_1,\dots,k_N)$ satisfying $\sum_{\ell=1}^N \ell\,|k_\ell| \leq M$ and $\sum_{\ell=1}^N \ell\,k_\ell = s$, condition (14) holds true.*

The following statements on the non-resonance set $\mathcal{N}(M,\gamma)$ are obtained by numerical investigation:

$M = 2$: All dimensions N belong to $\mathcal{N}(2,\gamma)$ for every γ, because for the only candidate $(s,r) = (1,2)$ the difference $r - s$ is not an even number.

$M = 3$: We have to consider $(s,r) = (1,3)$. For $\gamma = 0.1$ the set $\mathcal{N}(3,\gamma)$ contains all N between 100 and 1 000 000 except for $N = 728$. This value of N belongs to $\mathcal{N}(3,\gamma)$ for $\gamma = 0.04$.

$M = 4$: In addition to the pair $(1,3)$ we have to consider $(s,r) = (2,4)$. Since $\cos(\frac{\pi}{3}) = \frac{1}{2}$, the integer j minimizing (15) is the integer that is closest to $\frac{2N}{3} + s$. Whenever N is an integral multiple of 3 we have exact resonance, and condition (14) cannot be fulfilled with a positive γ. All other N (with the exception of $N = 728$) belong to $\mathcal{N}(4,\gamma)$ with $\gamma = 0.1$.

$M = 5$: For $\gamma = 0.1$ the set $\mathcal{N}(5,\gamma)$ contains all values of N in the range $100 \leq N \leq 100\,000$, except for integral multiples of 3 and 37 additional values.

We are now in the position to formulate our main result. We assume that initially only the first pair of modes ± 1 is excited, that is, the initial data satisfy (10) for $j_0 = 1$.

Theorem 1. *Fix $\gamma > 0$ and $\rho \geq 1$ arbitrarily, and let M and K be positive integers satisfying $M < K$ and $K + M = 10$. Then, there exist $\delta_0 > 0$ and $c > 0$, $C > 0$ such that the following holds: if*

(1) the dimension of the system satisfies $N \geq 41$ *and* $N \in \mathcal{N}(M,\gamma)$,
(2) the total energy E *is bounded such that* $\delta := \sqrt{E\,N^3} \leq \delta_0$,
(3) the initial normal mode energies satisfy $E_j(0) = 0$ *for* $j \neq \pm 1$,

then, over long times

$$t \leq c\,N^2\,\delta^{-M-1},$$

the normal mode energies satisfy the estimates

$$|E_1(t) - E_1(0)| \leq C\,E\,\delta^2, \tag{16}$$

$$E_j(t) \leq C\,E\,\delta^{2(j-1)}, \qquad j = 1,\ldots,K, \tag{17}$$

$$\sum_{j=K}^{N} \rho^{2j} E_j(t) \leq C\,E\,\delta^{2(K-1)}. \tag{18}$$

This theorem is proved in Sections 5 through 8. Moreover, the proof shows that the theorem remains valid if the single-mode excitation condition (10) is replaced by the geometric-decay condition that (17) and (18) hold at $t = 0$, with a given constant C_0 in place of C and a $\rho_0 > \rho$ in place of ρ.

We note that the maximal time interval is obtained with $M = 4$, for which we have $t \leq c\,N^2\,\delta^{-5}$. Longer time intervals for slightly weaker estimates will be obtained in Section 9, where we account for the almost-resonance $\omega_5 - 2\omega_4 + 3\omega_1 = \mathcal{O}(N^{-5})$ that leads to the restriction $K + M \leq 10$ in the above theorem.

The decay of the mode energies with powers of δ was first observed numerically by Flach, Ivanchenko and Kanakov [12].

To our knowledge, so far the only rigorous long-time bounds of the mode energies in FPU systems with large particle numbers N have been given by Bambusi and Ponno [1]. There it is shown for $E = N^{-3}$ and large N that the mode energies satisfy

$$E_j(t) \leq E(C_1 \mathrm{e}^{-\sigma j} + C_2 N^{-1}) \qquad \text{for} \quad t \leq T N^3,$$

where the factor T can be fixed arbitrarily, and the constants C_1, C_2 and σ depend on T.

In comparison, our result is concerned with the case $E \ll N^{-3}$. In this situation we obtain exponential decay for all j and estimates over much longer time intervals. For $E = \delta_0^2 N^{-3}$, however, our estimates are proved only on time intervals of length $\mathcal{O}(N^2)$. The proof of [1] exploits the relationship of the FPU α-model with the KdV equation for the scaling $E = N^{-3}$. The proof of Theorem 1 is completely different and is based on the technique of modulated Fourier expansions.

The proof of Theorem 1 gives explicit formulas for the dominant terms in $E_j(t)$:

$$E_j(t) = 2\,E\,\delta^{2(j-1)}\big(|c_-|^{2j} + |c_+|^{2j}\big)\big(|a_j(t)|^2 + \mathcal{O}(\delta^2 + N^{-2})\big), \tag{19}$$

where $c_\pm = \frac{1}{\sqrt{2\epsilon}}\big(\omega_1 u_1(0) \pm \mathrm{i}\,\dot u_1(0)\big)$ with $\epsilon = E/(2N)$, so that $|c_-|^2 + |c_+|^2 = 1$, and the first functions $a_j(t)$ are given by

$$
\begin{aligned}
a_1(t) &= \frac{1}{2}\\
a_2(t) &= \frac{1}{\pi^2}\Big(\mathrm{e}^{\mathrm{i}(2\omega_1-\omega_2)t} - 1\Big)\\
a_3(t) &= \frac{1}{\pi^4}\Big(\frac{3}{2}\Big(\mathrm{e}^{\mathrm{i}(3\omega_1-\omega_3)t} - 1\Big) - 2\Big(\mathrm{e}^{\mathrm{i}(\omega_1+\omega_2-\omega_3)t} - 1\Big)\Big).
\end{aligned}
$$

All the linear combinations of frequencies appearing in these formulas are of size $\mathcal{O}(N^{-3})$. In particular, we have

$$
2\omega_1 - \omega_2 = \frac{\pi^3}{4\,N^3} + \mathcal{O}(N^{-5}).
$$

The function $a_2(t)$ is thus periodic with period $T \approx 8\pi^{-2}N^3$. This agrees extremely well with Figure 2, where the choice (11) of the initial values corresponds to $c_+ = 0$ and $c_- = -\mathrm{i}$.

We remark that the mode energy values $E_j \approx E\,j^2(\delta/\pi^2)^{2(j-1)}$ stated in [12] are qualitatively similar, though not identical in quantity to the expressions given here.

5. Weak non-resonance inequalities

The frequencies $\omega_j = 2\sin\big(\frac{j\pi}{2N}\big)$ are almost in resonance for large N:

$$
\omega_j + \omega_\ell - \omega_{j+\ell} = 8\,\sin\Big(\frac{j\pi}{4N}\Big)\,\sin\Big(\frac{\ell\pi}{4N}\Big)\,\sin\Big(\frac{(j+\ell)\pi}{4N}\Big) = \mathcal{O}(N^{-3})
$$

for small j and ℓ. Such a near-resonant situation leads to small denominators in the construction of the modulated Fourier expansion that is given in the next section. We therefore present a series of technical lemmas that deal with the almost-resonances among the FPU frequencies.

We consider multi-indices $\mathbf{k} = (k_1, \ldots, k_N)$ with integers k_ℓ, we define

$$
\mu(\mathbf{k}) = \sum_{\ell=1}^{N} \ell |k_\ell|,
$$

and we denote $\langle j\rangle = (0, \ldots, 0, 1, 0, \ldots, 0)$ the $|j|$-th unit vector. Furthermore, we write $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)$ and $\mathbf{k}\cdot\boldsymbol{\omega} = k_1\omega_1 + \ldots + k_N\omega_N$. Our approach with modulated Fourier expansions leads to small denominators $\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2$ which have to be bounded from below. We first consider $(j,\mathbf{k})$ with small $|j|$ and small $\mu(\mathbf{k})$.

Lemma 1. *For pairs $(j,\mathbf{k})$ satisfying $\max\big(|j|, \mu(\mathbf{k})\big) \le 10$ and $\mathbf{k} \ne \pm\langle j\rangle$, we have for $N \ge 27$*

$$
\big|\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2\big| \ge \begin{cases} \dfrac{\pi}{4\,N}\Big(|\omega_j| + |\mathbf{k}\cdot\boldsymbol{\omega}|\Big) & \text{if} \quad \sum_{\ell=1}^N \ell\,k_\ell \ne \pm j \\[2mm] \dfrac{\pi^3}{8\,N^3}\Big(|\omega_j| + |\mathbf{k}\cdot\boldsymbol{\omega}|\Big) & \text{else.} \end{cases}
$$

The above estimate also holds for $\mu(\mathbf{k}) = 11$ *except for* $(j, \mathbf{k}) = (\pm 5, \pm(3\langle 1\rangle - 2\langle 4\rangle))$*, for which* $\omega_5^2 - (3\omega_1 - 2\omega_4)^2 = \mathcal{O}(N^{-6})$.

Proof. For large N we expand the expression $|\omega_j| - |\mathbf{k}\cdot\boldsymbol{\omega}|$ into a Taylor series, use the remainder bounds $|\sin x - x| \le x^3/6$ and $|\sin x - x + x^3/6| \le x^5/120$, and thus obtain the estimates for $N \ge 100$. We employ the fact that for $\mu(\mathbf{k}) \le 10$ the relation $j + \sum \ell\, k_\ell = 0$ implies $|j^3 + \sum \ell^3\, k_\ell| \ge 6$ (checked numerically). Also the remaining finitely many cases are checked numerically. □

We next consider the expression $\left|\omega_j^2 - (\omega_l + \mathbf{k}\cdot\boldsymbol{\omega})^2\right|$, where l is large and j and $\mu(\mathbf{k})$ are small.

Lemma 2. *For pairs* $(j, \langle l\rangle + \mathbf{k})$ *satisfying* $|j| \le 11$, $l \ge 7$, $\mu(\mathbf{k}) \le 5$, *and* $\langle l\rangle + \mathbf{k} \ne \pm\langle j\rangle$, *we have for* $N \ge 41$

$$\left|\omega_j^2 - (\omega_l + \mathbf{k}\cdot\boldsymbol{\omega})^2\right| \ge \begin{cases} \dfrac{\pi}{8\,N}\Big(|\omega_j| + |\omega_l + \mathbf{k}\cdot\boldsymbol{\omega}|\Big) & \textit{if} \quad l + \sum_{\ell=1}^N \ell\, k_\ell \ne \pm j \\ \dfrac{\pi^3}{8\,N^3}\Big(|\omega_j| + |\omega_l + \mathbf{k}\cdot\boldsymbol{\omega}|\Big) & \textit{else.} \end{cases}$$

Proof. For large $l \ge l_0 = 20$, for $0 \le j \le 11$, $\mu(\mathbf{k}) \le 5$, and $N \ge 37$, we prove

$$\omega_l + \mathbf{k}\cdot\boldsymbol{\omega} - \omega_j \ge \omega_{l_0} + \mathbf{k}\cdot\boldsymbol{\omega} - \omega_j \ge \frac{\pi}{N}.$$

The second inequality is obtained from $l_0 + \sum_{\ell=1}^N \ell\, k_\ell - j \ge 4 > 1$ by estimating the remainder in the Taylor series expansion. The finitely many remaining cases are verified numerically. □

We finally consider the expression $\left|\omega_j^2 - (\omega_{j-r} + \mathbf{k}\cdot\boldsymbol{\omega})^2\right|$ for arbitrarily large j, but with small $|r|$ and small $\mu(\mathbf{k})$. The factor $\left|\omega_j + (\omega_{j-r} + \mathbf{k}\cdot\boldsymbol{\omega})\right|$ will be bounded from below by ω_j and the expression $\omega_j - (\omega_{j-r} + \mathbf{k}\cdot\boldsymbol{\omega})$ is given by

$$4\,\sin\Big(\frac{r\,\pi}{4N}\Big)\,\cos\Big(\frac{(2j-r)\pi}{4N}\Big) \;-\; \frac{\pi}{N}\sum_{\ell=1}^N \ell\, k_\ell \;+\; \frac{\pi^3}{24\,N^3}\sum_{\ell=1}^N \ell^3\, k_\ell \;+\; \mathcal{O}\Big(\frac{1}{N^5}\Big). \tag{20}$$

Lemma 3. *Let* M, r *be integers such that* $M \le 15$ *and* $|r| \le M$, *and consider pairs* $(j, \langle j - r\rangle + \mathbf{k})$ *satisfying* $2M < j \le \min(N, N + r)$, $\mu(\mathbf{k}) \le M$, *and* $\langle j - r\rangle + \mathbf{k} \ne \pm\langle j\rangle$. *With* $s = \sum_\ell \ell\, k_\ell$ *we then have for all dimensions satisfying* $N \ge \max\big(\pi\sqrt{7/2}\,M,\ \pi\,(M^3 + 6)/12\big)$ *that*

$$\left|\omega_j^2 - (\omega_{j-r} + \mathbf{k}\cdot\boldsymbol{\omega})^2\right| \ge \begin{cases} \dfrac{\pi\,\omega_j}{2\,N} & \textit{if} \quad s < \min(r,0) \quad \textit{or} \quad s > \max(r,0) \\ \dfrac{\pi^3\,\omega_j}{8\,N^3} & \textit{if} \quad s = r = 0 \\ \dfrac{r^2\,\pi^2\,\omega_j}{8\,N^2} & \textit{if} \quad s = 0 \quad \textit{and} \quad r \ne 0 \\ \dfrac{|r|\,\pi\,\omega_j^3}{32\,N} & \textit{if} \quad s = r \quad \textit{and} \quad r \ne 0. \end{cases} \tag{21}$$

A counter-example for the second estimate is $\mathbf{k}\cdot\boldsymbol{\omega} = \omega_5 - 2\omega_4 + 3\omega_1 = \mathcal{O}(N^{-5})$, *for which* $\mu(\mathbf{k}) = 16$.

Proof. The assumptions on the indices imply $2j - r \le 2N - |r|$. The first term in (20) is thus a monotonic function of j with asymptotic values ranging from

$$\frac{r\,\pi}{N} \qquad \longrightarrow \qquad \frac{r\,|r|\,\pi^2}{4\,N^2}$$

when j goes from $j = M$ to $j = \max(N, N + r)$. This observation implies the first inequality of (21), because the second term in (20) is dominant in this case. Rigourous estimates prove the inequality for $N^2 \ge \pi^2 M^3/6$.

The second inequality follows as in Lemma 1 for $N^2 \ge \pi^2 M^5/3840$ from the fact that for $0 < \mu(\mathbf{k}) \le 15$ the condition $\sum_\ell \ell\, k_\ell = 0$ implies $|\sum_\ell \ell^3 k_\ell| \ge 6$.

For $s = 0$ and $r \ne 0$, the first term in (20) is dominant and therefore we have the third inequality of (21) for $N \ge \pi(M^3 + 6)/12$.

For the proof of the last inequality we note that

$$\begin{aligned}|\omega_j - \omega_{j-r} - \omega_r| &= 8\,\sin\Big(\frac{j\pi}{4N}\Big)\,\sin\Big(\frac{|r|\pi}{4N}\Big)\,\sin\Big(\frac{(j-r)\pi}{4N}\Big)\\ &\ge \frac{\omega_j}{2}\,\sin\Big(\frac{|r|\pi}{4N}\Big)\Big(\omega_j\,\cos\Big(\frac{|r|\pi}{4N}\Big) - 4\,\sin\Big(\frac{|r|\pi}{4N}\Big)\Big)\end{aligned}$$

where we have used the addition theorem for $\sin(\alpha - \beta)$ and the inequality $2\sin(\alpha/2) \ge \sin\alpha$. We therefore obtain

$$|\omega_j - \omega_{j-r} - \mathbf{k}\cdot\boldsymbol{\omega}| \ge \frac{|r|\,\pi\,\omega_j^2}{32\,N} + \chi_j$$

with

$$\chi_j = \frac{\omega_j^2}{4}\,\Big(\sin\Big(\frac{|r|\pi}{2N}\Big) - \frac{|r|\,\pi}{8\,N}\Big) - \frac{\omega_j\,|r|^2\pi^2}{8\,N^2} - |\omega_r - \mathbf{k}\cdot\boldsymbol{\omega}|.$$

To prove $\chi_j \ge 0$, we notice that it is an increasing function of j (for $j \ge 2M$), and bounded from below by its value at $j = 2M$. For $r = \sum_\ell \ell\, k_\ell$ and $\mu(\mathbf{k}) \le M$ we have $\sum_\ell \ell(|k_\ell| \pm k_\ell) \le M \pm r$, so that $k_\ell = 0$ for $2\ell > M + |r|$, and therefore $\sum_\ell \ell^3 |k_\ell| \le M(M + |r|)^2/4$. Using $x - x^3/6 \le \sin x \le x$ we obtain the lower bound

$$\chi_{2M} \ge \frac{|r|^3\pi^3}{24\,N^3}\,A\Big(\frac{M}{|r|}\Big) - \frac{|r|^3 M^2\pi^5}{48\,N^5}\,B\Big(\frac{M}{|r|}\Big)$$

with

$$A(x) = 9\,x^2 - 6\,x - 1 - x(x+1)^2/4, \qquad B(x) = 1 + 6\,x^2.$$

An inspection of these functions shows that the quotient satisfies $0 < B(x)/A(x) \le B(1)/A(1) = 7$ on an interval including $1 \le x \le 15$. Hence, $\chi_{2M} \ge 0$ for $N^2 \ge \pi^2 M^2 7/2$, which completes the proof of the lemma. □

It remains to consider the situation where $r \ne 0$ and $0 < s/r < 1$. We have a near-cancellation of the first two terms in (20) if the index j minimizes (15). We denote such an integer by $j^*(s, r, N)$.

Lemma 4. *Let $\mu(\mathbf{k}) \le M$, $2M < j \le \min(N, N + r)$, $0 < |r| \le M$ and $0 < s/r < 1$, where $s = \sum_\ell \ell\, k_\ell$. For $N \ge 0.64\,M^3$ and $j \ne j^*(s, r, N)$, we have the lower bound*

$$\big|\omega_j^2 - (\omega_{j-r} + \mathbf{k}\cdot\boldsymbol{\omega})^2\big| \;\ge\; \frac{\pi\,\omega_j}{8\,N^2}\,\sqrt{r}.$$

Proof. We define α by $\cos\alpha = \frac{s}{r}$, and we use $\cos\beta - \cos\alpha = 2\,\sin(\frac{\alpha+\beta}{2})\,\sin(\frac{\alpha-\beta}{2})$ with $\beta = \frac{(2j-r)\pi}{4N}$. The estimate then follows from $|\alpha-\beta| \geq \frac{\pi}{4N}$ and from $2\,\sin(\frac{\alpha+\beta}{2}) \geq 2\,\sin(\frac{\alpha}{2}) = \sqrt{2(1-\cos\alpha)} = \sqrt{\frac{2(r-s)}{r}} \geq \sqrt{\frac{2}{r}}$. $\square$

The most critical case for a lower bound of $\left|\omega_j^2 - (\omega_{j-r} + \mathbf{k}\cdot\boldsymbol{\omega})^2\right|$ is when $j = j^*(s,r,N)$. This is why we restrict the dimension N of the FPU-system to values satisfying the non-resonance condition of Definition 1. Because of property (42) below we consider only even values of $r-s$ in Definition 1.

6. Modulated Fourier expansion

Our principal tool for studying the long-time behaviour of mode energies is a modulated Fourier expansion, which was originally introduced for the study of numerical energy conservation in Hamiltonian ordinary differential equations in the presence of high oscillations [17,8]. This technique was also successfully applied to the long-time analysis of weakly nonlinear Hamiltonian partial differential equations [9,15,16]. The idea is to separate rapid oscillations from slow variations by a two-scale ansatz of the form

$$\omega_j u_j(t) \approx \sum_{\mathbf{k}\in\mathcal{K}_j} z_j^{\mathbf{k}}(\tau)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega})t} \qquad \text{with} \quad \tau = N^{-3}t, \tag{22}$$

where $\mathcal{K}_j$ is a finite set of multi-indices $\mathbf{k} = (k_1,\ldots,k_N)$ with integers k_ℓ. The products of complex exponentials $\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega})t} = \prod_{\ell=1}^N \mathrm{e}^{\mathrm{i}k_\ell\omega_\ell t}$ account for the nonlinear interaction of different modes. The slowly varying modulation functions $z_j^{\mathbf{k}}(\tau)$ are yet to be determined from a system of modulation equations.

6.1. *Choice of the interaction set.*

We will choose the sets $\mathcal{K}_j$ such that the interaction of low modes with low modes and high modes with low modes is incorporated, but the interaction of high modes with high modes is discarded. It turns out that for our purposes an appropriate choice of the multi-index set $\mathcal{K}_j$ is obtained as follows: fix positive integers K and M with

$$K + M \leq 10 \quad \text{and} \quad M < K.$$

We define, with the notation introduced at the beginning of Section 5,

$$\begin{aligned}\mathcal{K} = \{(j,\mathbf{k})\,:\,&\max\big(|j|,\mu(\mathbf{k})\big) \leq K+M\}\\ &\cup\ \{(j,\pm\langle l\rangle + \mathbf{k})\,:\,|j| \leq K+M,\ \mu(\mathbf{k}) \leq M,\ l \geq K+1\}\\ &\cup\ \{(j,\pm\langle j-r\rangle + \mathbf{k})\,:\,|j| \geq K+M,\ |r| \leq M,\ \mu(\mathbf{k}) \leq M\}.\end{aligned}$$

This set consists of pairs $(j,\mathbf{k})$ for which we have obtained lower bounds for $\left|\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2\right|$ in Section 5. We let

$$\mathcal{K}_j = \{\mathbf{k}\,:\,(j,\mathbf{k}) \in \mathcal{K}\}.$$

For convenience we set $z_j^{\mathbf{k}}(\tau) = 0$ for multi-indices $\mathbf{k}$ that are not in $\mathcal{K}_j$.

6.2. Choice of norm. We work with a weighted ℓ^2 norm for $2N$-periodic sequences $\mathbf{u} = (u_j)_{j=-N}^{N-1}$,

$$\|\mathbf{u}\|^2 = \sum_{j=-N}^{N-1} \sigma_j \, |u_j|^2, \qquad \sigma_j = |N\omega_j|^{2s} \rho^{2|N\omega_j|} \tag{23}$$

with $s > 1/2$ and $\rho \geq 1$. This choice is motivated by two facts. On the one hand, the extended norm

$$\|(\mathbf{u}, \dot{\mathbf{u}})\|^2 = \|\boldsymbol{\Omega}\mathbf{u}\|^2 + \|\dot{\mathbf{u}}\|^2 \tag{24}$$

with $\boldsymbol{\Omega} = \operatorname{diag}(\omega_j)$ can be written in terms of the mode energies as

$$\|(\mathbf{u}(t), \dot{\mathbf{u}}(t))\|^2 = \frac{1}{2N} \sum_{j=1}^{N}{}' \, \sigma_j E_j(t),$$

where the prime indicates that the last term in the sum is taken with the factor 1/2. On the other hand, with the given choice of σ_j, the norm behaves well with convolutions.

Lemma 5. *For the norm (23), we have*

$$\|\mathbf{u} * \mathbf{v}\| \leq c \, \|\mathbf{u}\| \cdot \|\mathbf{v}\| \qquad \textit{where} \qquad (\mathbf{u} * \mathbf{v})_j = \sum_{j_1 + j_2 = j \bmod 2N} u_{j_1} \, v_{j_2}, \tag{25}$$

with c depending on $s > 1/2$, but not on N and $\rho \geq 1$.

Proof. We note the bound $|\omega_j| \leq |\omega_{j_1}| + |\omega_{j_2}|$ for $j_1 + j_2 = j \bmod 2N$, which follows from the addition theorem for $\sin(\alpha + \beta)$. Together with the inequality $2|j| \leq N|\omega_j| \leq \pi|j|$, this yields

$$\sum_{j_1 + j_2 = j \bmod 2N} \sigma_{j_1}^{-1} \sigma_{j_2}^{-1} \leq c \, \sigma_j^{-1}.$$

The result then follows with the Cauchy-Schwarz inequality:

$$\sum_j \sigma_j |(\mathbf{u} * \mathbf{v})_j|^2 \leq \sum_j \Big(\sigma_j \sum_{j_1 + j_2 \equiv j} \sigma_{j_1}^{-1} \sigma_{j_2}^{-1} \Big) \sum_{j_1 + j_2 \equiv j} \sigma_{j_1} |u_{j_1}|^2 \, \sigma_{j_2} |v_{j_2}|^2,$$

where $\equiv$ means congruence modulo $2N$. $\square$

6.3. Statement of result. The approximability of the solution of the FPU system by modulated Fourier expansions is described in the following theorem.

Theorem 2. *Fix $\gamma > 0$, $\rho \geq 1$, $s > 1/2$, and let M and K be positive integers satisfying $M < K$ and $K + M = 10$. Then, there exist $\delta_0 > 0$ and $C_0 > 0$, $C_1 > 0$, $C > 0$ such that the following holds: if*

(1) the dimension of the system satisfies $N \geq 41$ and $N \in \mathcal{N}(M, \gamma)$,

(2) the energy is bounded such that $\delta := \sqrt{E \, N^3} \leq \delta_0$,

(3) the initial values satisfy $E_0(0) = 0$ *and*

$$
\begin{aligned}
E_j(0) &\le C_0\, E\, \delta^{2(j-1)} \qquad \textit{for} \quad 0 < j < K, \\
\sum_{K \le |j| \le N} \sigma_j\, E_j(0) &\le C_0\, E\, \delta^{2(K-1)}
\end{aligned}
\tag{26}
$$

with the weights $\sigma_j = |N\omega_j|^{2s} \rho^{2|N\omega_j|}$ *from (23);*

then, the solution $\mathbf{u}(t) = (u_j(t))_{j=-N}^{N-1}$ *admits an expansion*

$$
\omega_j u_j(t) = \sum_{\mathbf{k} \in \mathcal{K}_j} z_j^{\mathbf{k}}(\tau)\, \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega})t} + \omega_j r_j(t) \qquad \textit{with} \quad \tau = N^{-3} t,
\tag{27}
$$

where the remainder $\mathbf{r}(t) = (r_j(t))_{j=-N}^{N-1}$ *is bounded in the norm (24) by*

$$
\|(\mathbf{r}(t), \dot{\mathbf{r}}(t))\| \le C \sqrt{\epsilon}\, \delta^{K+M} N^{-2} t \qquad \textit{for} \quad 0 \le t \le \min(N^3, \epsilon^{-1/2})
\tag{28}
$$

with the specific energy $\epsilon = E/(2N)$*. The modulation functions* $z_j^{\mathbf{k}}(\tau)$ *are polynomials, identically zero for* $j = 0$*, and for* $0 \le \tau \le 1$ *bounded by*

$$
\begin{aligned}
\Big(\sum_{\mathbf{k} \in \mathcal{K}_j} |z_j^{\mathbf{k}}(\tau)| \Big)^2 &\le C_1\, \epsilon\, \delta^{2(|j|-1)} \qquad \textit{for} \quad 0 < |j| < K, \\
\sum_{K \le |j| \le N} \sigma_j \Big(\sum_{\mathbf{k} \in \mathcal{K}_j} |z_j^{\mathbf{k}}(\tau)| \Big)^2 &\le C_1\, \epsilon\, \delta^{2(K-1)}.
\end{aligned}
\tag{29}
$$

The same bounds hold for all derivatives of $z_j^{\mathbf{k}}$ *with respect to the slow time* $\tau = N^{-3} t$*. Moreover, the modulation functions satisfy* $z_{-j}^{-\mathbf{k}} = \overline{z_j^{\mathbf{k}}}$*.*

The above estimates imply, in particular, a solution bound in the weighted energy norm: for $t \le \min(N^3, \epsilon^{-1/2})$,

$$
\begin{aligned}
E_j(t) &\le C_1\, E\, \delta^{2(|j|-1)} \qquad \text{for} \quad 0 < |j| < K, \\
\sum_{K \le |j| \le N} \sigma_j\, E_j(t) &\le C_1\, E\, \delta^{2(K-1)}.
\end{aligned}
\tag{30}
$$

We shall see later that such estimates actually hold on much longer time intervals. The rest of this section is devoted to the proof of Theorem 2.

6.4. Formal modulation equations. Formally inserting the ansatz (22) into (3) and equating terms with the same exponential $\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega})t}$ lead to a relation

$$
\big(\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2 \big)\, z_j^{\mathbf{k}} + 2\mathrm{i}\, (\mathbf{k}\cdot\boldsymbol{\omega})\, N^{-3}\, \frac{dz_j^{\mathbf{k}}}{d\tau} + N^{-6}\, \frac{d^2 z_j^{\mathbf{k}}}{d\tau^2} = \dots\ ,
\tag{31}
$$

where the dots, coming from the non-linearity, represent terms that will be considered later. Since we aim at constructing functions without high oscillations, we have to look at the dominating terms in (31).

For $\mathbf{k} = \pm\langle j\rangle$, where $\langle j\rangle = (0,\ldots,0,1,0,\ldots,0)$ with the non-zero entry in the $|j|$-th position, the first term on the left-hand side of (31) vanishes and the second term with the time derivative $dz_j^{\mathbf{k}}/d\tau$ can be viewed as the dominant term. Recall that we only have to consider $j \neq 0$. For $\mathbf{k} \in \mathcal{K}_j$ and $\mathbf{k} \neq \pm\langle j\rangle$, the first term is dominant according to the non-resonance estimates of Section 5.

To derive the equations defining the coefficient functions $z_j^{\mathbf{k}}$, we have to study the non-linearity when (22) is inserted into (3). We get for $\mathbf{k} = \pm\langle j\rangle$ that

$$\begin{aligned}
&\pm 2\mathrm{i}\,\omega_j\,N^{-3}\frac{dz_j^{\pm\langle j\rangle}}{d\tau} + N^{-6}\,\frac{d^2 z_j^{\pm\langle j\rangle}}{d\tau^2} \\
&\qquad = -\mathrm{i}\,\omega_j^2 \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)} \sum_{\mathbf{k}^1+\mathbf{k}^2=\pm\langle j\rangle} z_{j_1}^{\mathbf{k}^1} z_{j_2}^{\mathbf{k}^2},
\end{aligned} \tag{32}$$

and for $\mathbf{k} \neq \pm\langle j\rangle$

$$\begin{aligned}
&\bigl(\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2\bigr)\, z_j^{\mathbf{k}} + 2\mathrm{i}\,(\mathbf{k}\cdot\boldsymbol{\omega})N^{-3}\,\frac{dz_j^{\mathbf{k}}}{d\tau} + N^{-6}\,\frac{d^2 z_j^{\mathbf{k}}}{d\tau^2} \\
&\qquad = -\mathrm{i}\,\omega_j^2 \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)} \sum_{\mathbf{k}^1+\mathbf{k}^2=\mathbf{k}} z_{j_1}^{\mathbf{k}^1} z_{j_2}^{\mathbf{k}^2}.
\end{aligned} \tag{33}$$

In addition, the initial conditions for u_j need to be taken care of. They will yield the initial conditions for the functions $z_j^{\pm\langle j\rangle}$ for $j \neq 0$:

$$\sum_{\mathbf{k}\in\mathcal{K}_j} z_j^{\mathbf{k}}(0) = \omega_j u_j(0)\,, \qquad \sum_{\mathbf{k}\in\mathcal{K}_j} \Bigl(\mathrm{i}\,(\mathbf{k}\cdot\boldsymbol{\omega})\, z_j^{\mathbf{k}}(0) + N^{-3}\,\frac{dz_j^{\mathbf{k}}}{d\tau}(0)\Bigr) = \omega_j \dot u_j(0). \tag{34}$$

6.5. Construction of coefficient functions for the modulated Fourier expansion. We construct the functions $z_j^{\mathbf{k}}(\tau)$ such that they are solutions of the system (32)–(33) up to a small defect and satisfy the initial conditions (34) for all j. We write the functions as a formal series in $\delta = \sqrt{EN^3}$,

$$z_j^{\mathbf{k}}(\tau) = \sqrt{2\epsilon}\sum_{m\ge 1} \delta^{m-1}\, z_{j,m}^{\mathbf{k}}(\tau) = N^{-2}\sum_{m\ge 1} \delta^{m}\, z_{j,m}^{\mathbf{k}}(\tau), \tag{35}$$

insert them into the relations (32)–(34), and compare like powers of δ:

$$\begin{aligned}
&\pm 2\mathrm{i}\,\omega_j\,N^{-3}\frac{dz_{j,m}^{\pm\langle j\rangle}}{d\tau} + N^{-6}\,\frac{d^2 z_{j,m}^{\pm\langle j\rangle}}{d\tau^2} \\
&= -\mathrm{i}\,\omega_j^2 N^{-2} \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)} \sum_{\mathbf{k}^1+\mathbf{k}^2=\pm\langle j\rangle}\ \sum_{m_1+m_2=m} z_{j_1,m_1}^{\mathbf{k}^1} z_{j_2,m_2}^{\mathbf{k}^2},
\end{aligned} \tag{36}$$

and for $\mathbf{k} \neq \pm\langle j\rangle$,

$$\begin{aligned}
&\bigl(\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2\bigr)\, z_{j,m}^{\mathbf{k}} + 2\mathrm{i}\,(\mathbf{k}\cdot\boldsymbol{\omega})N^{-3}\,\frac{dz_{j,m}^{\mathbf{k}}}{d\tau} + N^{-6}\,\frac{d^2 z_{j,m}^{\mathbf{k}}}{d\tau^2} \\
&= -\mathrm{i}\,\omega_j^2 N^{-2} \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)} \sum_{\mathbf{k}^1+\mathbf{k}^2=\mathbf{k}}\ \sum_{m_1+m_2=m} z_{j_1,m_1}^{\mathbf{k}^1} z_{j_2,m_2}^{\mathbf{k}^2}.
\end{aligned} \tag{37}$$

This yields an equation for $z_{j,m}^{\mathbf{k}}$ ($\mathbf{k} \neq \pm\langle j\rangle$) and for the derivative of $z_{j,m}^{\pm\langle j\rangle}$ with a right-hand side that depends only on the product $z_{j_1,m_1}^{\mathbf{k}^1} z_{j_2,m_2}^{\mathbf{k}^2}$ with $m_1+m_2=m$ so that both m_1 and m_2 are strictly smaller than m. Initial values for $z_{j,m}^{\pm\langle j\rangle}$ are obtained from (34). These differential equations possess a unique polynomial solution, since the only polynomial solution to

$$z - \alpha\,\frac{dz}{d\tau} - \beta\,\frac{d^2 z}{d\tau^2} = p \tag{38}$$

with a polynomial p of degree d is given by

$$z = \sum_{\ell=0}^{d}\Big(\alpha\,\frac{d}{d\tau} + \beta\,\frac{d^2}{d\tau^2}\Big)^{\ell} p. \tag{39}$$

We now show in detail how the coefficient functions $z_{j,m}^{\mathbf{k}}(\tau)$ for the first few values of $m = 1, 2, \dots$ are constructed. According to the bound (26) we assume for $1 \le j \le K$ that

$$\omega_j u_j(0) = \sqrt{2\epsilon}\,\delta^{|j|-1} a_j, \quad \dot u_j(0) = \sqrt{2\epsilon}\,\delta^{|j|-1} b_j \tag{40}$$

(consequently $\omega_j u_{-j}(0) = \sqrt{2\epsilon}\,\delta^{|j|-1}\overline{a}_j$, $\dot u_{-j}(0) = \sqrt{2\epsilon}\,\delta^{|j|-1}\overline{b}_j$) with factors a_j and b_j whose absolute values are bounded independently of N, ϵ, and δ.

Case $m=1$*:* For $m=1$, the right-hand sides in (36) and (37) are zero. The relation (37) shows that $z_{j,1}^{\mathbf{k}}(\tau) = 0$ for $\mathbf{k} \neq \pm\langle j\rangle$. Equation (36) shows that $\frac{d}{d\tau} z_{j,1}^{\pm\langle j\rangle}(\tau) = 0$ and hence $z_{j,1}^{\pm\langle j\rangle}(\tau)$ is a constant function. Its value is obtained from (34). It vanishes for $j \neq \pm 1$, and is given by

$$z_{1,1}^{\pm\langle 1\rangle}(0) = \frac{1}{2}(a_1 \mp \mathrm{i} b_1), \qquad z_{-1,1}^{\pm\langle 1\rangle}(0) = \frac{1}{2}(\overline{a}_1 \mp \mathrm{i}\overline{b}_1)$$

for the only non-vanishing functions with $m=1$.

Case $m=2$*:* The products $z_{1,1}^{\pm\langle 1\rangle} z_{1,1}^{\pm\langle 1\rangle}$ and $z_{-1,1}^{\pm\langle 1\rangle} z_{-1,1}^{\pm\langle 1\rangle}$ with all combinations of signs give a non-zero contribution to the right-hand side of (37). Notice that $j=0$ need not be considered, because $u_0 = 0$. This gives the constant functions

$$\begin{aligned} z_{2,2}^{\mathbf{0}}(\tau) &= -2\mathrm{i}\, z_{1,1}^{\langle 1\rangle}(0) z_{1,1}^{-\langle 1\rangle}(0)\, N^{-2} \\ \big(\omega_2^2 - 4\omega_1^2\big)\, z_{2,2}^{2\langle 1\rangle}(\tau) &= -\mathrm{i}\,\omega_2^2\, z_{1,1}^{\langle 1\rangle}(0) z_{1,1}^{\langle 1\rangle}(0)\, N^{-2} \end{aligned}$$

and similar formulas for $z_{-2,2}^{\mathbf{0}}(\tau)$, $z_{2,2}^{-2\langle 1\rangle}(\tau)$, and $z_{-2,2}^{\pm 2\langle 1\rangle}(\tau)$.

From (36) we obtain $\frac{d}{d\tau} z_{2,2}^{\pm\langle 2\rangle}(\tau) = 0$, and hence $z_{2,2}^{\pm\langle 2\rangle}(\tau)$ is constant. The initial values are determined from (34), i.e.,

$$\begin{aligned} \big(z_{2,2}^{\langle 2\rangle}(0) + z_{2,2}^{-\langle 2\rangle}(0)\big) + \big(z_{2,2}^{2\langle 1\rangle}(0) + z_{2,2}^{-2\langle 1\rangle}(0)\big) &= a_2 \\ \mathrm{i}\,\big(z_{2,2}^{\langle 2\rangle}(0) - z_{2,2}^{-\langle 2\rangle}(0)\big) + 2\mathrm{i}\,\frac{\omega_1}{\omega_2}\big(z_{2,2}^{2\langle 1\rangle}(0) - z_{2,2}^{-2\langle 1\rangle}(0)\big) &= b_2, \end{aligned}$$

since the values $z_{2,2}^{\pm 2\langle 1\rangle}(0)$ are already known. In the same way we get $z_{-2,2}^{\pm\langle 2\rangle}(\tau)$. These functions are the only non-vanishing functions for $m=2$.

Case $m = 3$*:* The nonzero terms in the right-hand side of (37) are formed of products $z^{\mathbf{k}^1}_{j_1,m_1} z^{\mathbf{k}^2}_{j_2,m_2}$, where one index among m_1, m_2 is 1 and the other is 2. We thus obtain formulas for $z^{\pm 3\langle 1\rangle}_{1,3}(\tau)$, $z^{\pm\langle 1\rangle\pm\langle 2\rangle}_{1,3}(\tau)$, $z^{\pm 3\langle 1\rangle}_{3,3}(\tau)$, $z^{\pm\langle 1\rangle\pm\langle 2\rangle}_{3,3}(\tau)$, $z^{\pm\langle 1\rangle}_{3,3}(\tau)$, and similar formulas with negative index j. All these functions are constant. The relation (36) leads to differential equations with vanishing right-hand side for $z^{\pm\langle 3\rangle}_{3,3}(\tau)$ and with constant right-hand side for $z^{\pm\langle 1\rangle}_{1,3}$, which thus becomes a linear polynomial in τ.

Case $4 \le m < K$*:* Assuming that the functions $z^{\mathbf{k}^1}_{j_1,p}$ are known for $p \le m-1$ and all j_1 and $\mathbf{k}^1$, the relation (37) permits us to compute $z^{\mathbf{k}}_{j,m}$ for $\mathbf{k} \ne \pm\langle j\rangle$. We further obtain $z^{\pm\langle j\rangle}_{j,m}$ from (36) and (34). By this construction many of the functions are constant and some of them are polynomials in τ, of degree at most $m/2$.

Case $K \le m \le K + M$ *with* $M < K$*:* A new situtation arises for $m = K$, because for $|j| \ge K$ we have, according to the bound (26), that

$$\omega_j u_j(0) = \sqrt{2\epsilon}\,\delta^{K-1} a_j, \quad \dot u_j(0) = \sqrt{2\epsilon}\,\delta^{K-1} b_j \tag{41}$$

with $a_j, b_j = \mathcal{O}(1)$. Hence, for $m = K$, we get differential equations for all diagonal functions $z^{\pm\langle j\rangle}_{j,m}$ with initial values that in general are not zero. Otherwise, the construction can be continued as before.

Since all coefficient functions are created from diagonal quantities $z^{\langle j\rangle}_j(\tau)$, we have for all $m \ge 1$ that

$$z^{\mathbf{k}}_{j,m}(\tau) = 0 \qquad \text{if} \qquad j \ne \mu(\mathbf{k}) \bmod 2. \tag{42}$$

Furthermore, as long as $m \le K$, we have

$$z^{\mathbf{k}}_{j,m}(\tau) = 0 \qquad \text{if} \qquad \begin{cases} \quad m < |j| \quad \text{or} \quad m < \mu(\mathbf{k}) \quad \text{or} \\ m \ne j \bmod 2 \quad \text{or} \quad m \ne \mu(\mathbf{k}) \bmod 2. \end{cases}$$

6.6. Bounds of the modulation functions. The above construction yields functions $z^{\mathbf{k}}_{j,m}$ that are bounded by certain non-positive powers of N. From the explicit formulas of Section 6.5 we have

$$\begin{aligned}
&z^{\pm\langle 1\rangle}_{1,1}(\tau) = \mathcal{O}(1)\\
&z^{\pm\langle 2\rangle}_{2,2}(\tau) = \mathcal{O}(1), \quad z^{\pm 2\langle 1\rangle}_{2,2}(\tau) = \mathcal{O}(1), \quad z^{\mathbf{0}}_{2,2}(\tau) = \mathcal{O}(N^{-2})\\
&z^{\pm\langle 3\rangle}_{3,3}(\tau) = \mathcal{O}(1), \quad z^{\pm 3\langle 1\rangle}_{3,3}(\tau) = \mathcal{O}(1)\\
&z^{\pm\langle 1\rangle\pm\langle 2\rangle}_{3,3}(\tau) = \mathcal{O}(1), \quad z^{\pm\langle 1\rangle\mp\langle 2\rangle}_{3,3}(\tau) = \mathcal{O}(N^{-2})\\
&z^{\pm\langle 1\rangle}_{1,3}(\tau) = \mathcal{O}(N^{-2}) + \mathcal{O}(\tau), \quad z^{\pm\langle 1\rangle}_{3,3}(\tau) = \mathcal{O}(N^{-2})\\
&z^{\pm 3\langle 1\rangle}_{1,3}(\tau) = \mathcal{O}(N^{-2}), \quad z^{\pm\langle 1\rangle\pm\langle 2\rangle}_{1,3}(\tau) = \mathcal{O}(N^{-2}), \quad z^{\pm\langle 1\rangle\mp\langle 2\rangle}_{1,3}(\tau) = \mathcal{O}(1)
\end{aligned}$$

and similarly for negative index j. In the course of this subsection we show that for $m \le K + M$,

$$\sum_{j=-N}^{N-1} \sigma_j \Big(\sum_{\mathbf{k}\in\mathcal{K}_j} |z_{j,m}^{\mathbf{k}}(\tau)| \Big)^2 \le C \qquad \text{for} \quad 0 \le \tau \le 1$$

with a constant C that depends on K and M, but is independent of N.

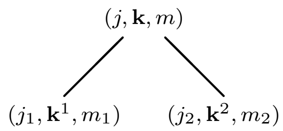

Fig. 3. Binary decomposition.

The behaviour of $z_{j,m}^{\mathbf{k}}$ can best be understood by using rooted binary trees. For multi-indices $\mathbf{k} \ne \pm\langle j\rangle$ we see from (33) that $z_{j,m}^{\mathbf{k}}$ is determined by the products $z_{j_1,m_1}^{\mathbf{k}^1} z_{j_2,m_2}^{\mathbf{k}^2}$, where $j = j_1 + j_2$, $\mathbf{k} = \mathbf{k}^1 + \mathbf{k}^2$, and $m = m_1 + m_2$, as illustrated in Figure 3. Recursively applying this reduction, we see that $z_{j,m}^{\mathbf{k}}$ can be bounded in terms of products of diagonal terms $z_{\ell,p}^{\pm\langle\ell\rangle}$ with $p < m$. For example, $z_{3,5}^{\langle 1\rangle}$ contains a term $\big(z_{1,1}^{\langle 1\rangle}\big)^2 z_{1,3}^{-\langle 1\rangle}$. This is illustrated by the binary tree of Figure 4.

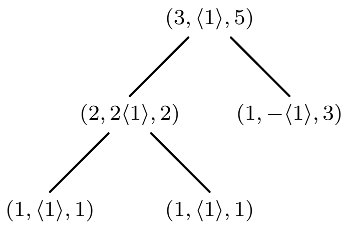

Fig. 4. Example of a binary tree with root $(3, \langle 1\rangle, 5)$.

In such rooted binary trees, there are two types of vertices: leaves of the form $(\ell, \pm\langle\ell\rangle, p)$ and inner vertices $(j, \mathbf{k}, m)$ with $\mathbf{k} \in \mathcal{K}_j$ and $\mathbf{k} \ne \pm\langle j\rangle$ which have two branches. For $\mathbf{k} \in \mathcal{K}_j$ and $m \le K + M$, we denote by $\mathcal{B}_{j,m}^{\mathbf{k}}$ the set of all rooted binary trees of this kind with root $(j, \mathbf{k}, m)$. We note that the number of trees in $\mathcal{B}_{j,m}^{\mathbf{k}}$ is independent of N.

We estimate the modulation functions in a time-dependent norm on the space of polynomials of degree not exceeding $K + M$,

$$|||z|||_\tau = \sum_{i\ge 0} \frac{1}{i!} \Big| \frac{d^i z}{d\tau^i}(\tau) \Big|. \tag{43}$$

By the weak non-resonance estimates of Section 5, we have for all $(j,\mathbf{k}) \in \mathcal{K}$ with $\mathbf{k} \neq \pm\langle j\rangle$ that

$$\alpha_j^{\mathbf{k}} = -2\mathrm{i}\frac{(\mathbf{k}\cdot\boldsymbol{\omega})N^{-3}}{\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2}, \quad \beta_j^{\mathbf{k}} = -\frac{N^{-6}}{\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2} \quad \text{satisfy} \quad |\alpha_j^{\mathbf{k}}| + |\beta_j^{\mathbf{k}}| \le C \tag{44}$$

with a constant C that does not depend on $(j,\mathbf{k}) \in \mathcal{K}$ and N. Since for a polynomial $p(\tau)$ of degree n, we have $|||dp/d\tau|||_\tau \le n\,|||p|||_\tau$, the solution (39) of (38) is bounded by $|||z|||_\tau \le C(n)|||p|||_\tau$ with a constant depending on n. Equation (37) thus yields the recursive bound, for $\mathbf{k} \neq \pm\langle j\rangle$ and for $m \le K+M$,

$$|||z_{j,m}^{\mathbf{k}}|||_\tau \le \gamma_j^{\mathbf{k}} \sum_{j_1+j_2=j \bmod 2N} \; \sum_{\mathbf{k}^1+\mathbf{k}^2=\mathbf{k}} \; \sum_{m_1+m_2=m} |||z_{j_1,m_1}^{\mathbf{k}^1}|||_\tau \cdot |||z_{j_2,m_2}^{\mathbf{k}^2}|||_\tau \tag{45}$$

with

$$\gamma_j^{\mathbf{k}} = \frac{C\,\omega_j^2\,N^{-2}}{|\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2|}, \tag{46}$$

where C depends on K and M, but not on N, j, and $\mathbf{k}$.

We resolve this recurrence relation via the binary trees. We denote by $\kappa_{\ell,p}^{+}(b)$ the number of appearances of $(\ell,\langle\ell\rangle,p)$ and its complex conjugate $(-\ell,-\langle\ell\rangle,p)$ among the leaves of a tree $b \in \mathcal{B}_{j,m}^{\mathbf{k}}$, and by $\kappa_{\ell,p}^{-}(b)$ the number of appearances of $(\ell,-\langle\ell\rangle,p)$ and $(-\ell,\langle\ell\rangle,p)$. We then obtain

$$|||z_{j,m}^{\mathbf{k}}|||_\tau \le \sum_{b\in\mathcal{B}_{j,m}^{\mathbf{k}}} \Gamma(b) \prod_{\ell=1}^{N} \prod_{p=1}^{m-1} |||z_{\ell,p}^{\langle\ell\rangle}|||_\tau^{\kappa_{\ell,p}^{+}(b)}\, |||z_{\ell,p}^{-\langle\ell\rangle}|||_\tau^{\kappa_{\ell,p}^{-}(b)} \tag{47}$$

with a factor $\Gamma(b)$ that is given recursively as follows: $\Gamma(b)=1$ for a tree with a single node $(j,\pm\langle j\rangle,m)$, and

$$\Gamma(b) = \gamma_j^{\mathbf{k}}\,\Gamma(b_1)\Gamma(b_2)$$

for a binary tree b having a root $(j,\mathbf{k},m)$ with $\mathbf{k} \neq \pm\langle j\rangle$ and composed of subtrees b_1 and b_2.

Lemma 6. *The estimates of Section 5 yield*

$$\Gamma(b) \le Const.$$

with a constant independent of $N \in \mathcal{N}(M,\gamma)$ (but depending on K, M and γ).

Proof. We have $\gamma_j^{\mathbf{k}} = \mathcal{O}(1)$ for all values of $(j,\mathbf{k})$ with the exception of the cases corresponding to the second inequality of (21) or to that of (14), where we only have $\gamma_j^{\mathbf{k}} = \mathcal{O}(\omega_j N)$ which is unbounded for large j. Nevertheless, we obtain the statement of the lemma, because the non-resonance estimates of Section 5 permit to prove

$$\gamma_j^{\mathbf{k}}\gamma_{j_1}^{\mathbf{k}^1}\gamma_{j_2}^{\mathbf{k}^2} = \mathcal{O}(1) \quad \text{for} \quad j_1+j_2 = j \bmod 2N \quad \text{and} \quad \mathbf{k}^1+\mathbf{k}^2=\mathbf{k}, \tag{48}$$

where we have set $\gamma_\ell^{\pm\langle\ell\rangle}=1$ and $\gamma_j^{\mathbf{k}}=0$ if $j \notin \mathcal{K}_j$.

Indeed, let $\mathbf{k} = \langle j - r\rangle + \mathbf{h}$ with $s = \sum_\ell \ell\, h_\ell$, and assume that $\mathbf{k}^1$ is close to $\langle j-r\rangle$, so that $\mu(\mathbf{k}^2)$ and $|j_2|$ are bounded by $2M \le K+M$. We distinguish two situations. If $s_2 = \sum_\ell \ell\, k_\ell^2$ satisfies $s_2 \ne \pm j_2$, it follows from the first inequality of Lemma 1 that $\gamma_{j_2}^{\mathbf{k}^2} = \mathcal{O}(N^{-2})$. This is sufficient to get (48). If $s_2 = j_2$ or $s_2 = -j_2$, so that $\gamma_{j_2}^{\mathbf{k}^2} = \mathcal{O}(1)$, an inspection of the possible values for $(j_1, \mathbf{k}^1)$ shows that $\gamma_{j_1}^{\mathbf{k}^1} = \mathcal{O}(N^{-1})$. Also in this case we have (48). □

By (45) and (47) the functions $z_{j,m}^{\mathbf{k}}(\tau)$ with $\mathbf{k} \ne \pm\langle j\rangle$ are estimated in terms of the diagonal functions $z_{\ell,p}^{\pm\langle \ell\rangle}(\tau)$ with $p < m$. With similar arguments we find that the derivative of $z_{j,m}^{\pm\langle j\rangle}(\tau)$, given by (36), is bounded by

$$|||\frac{d}{d\tau} z_{j,m}^{\pm\langle j\rangle}|||_\tau \le \omega_j N \sum_{j_1+j_2=j \bmod 2N} \sum_{\mathbf{k}^1+\mathbf{k}^2=\pm\langle j\rangle} \sum_{m_1+m_2=m} |||z_{j_1,m_1}^{\mathbf{k}^1}|||_\tau \cdot |||z_{j_2,m_2}^{\mathbf{k}^2}|||_\tau \,. \tag{49}$$

Here we note that the factor $\omega_j N$, which is large for large $|j|$, is compensated by the factors present in $z_{j_1,m_1}^{\mathbf{k}^1}$ and $z_{j_2,m_2}^{\mathbf{k}^2}$. This is a consequence of

$$\gamma_{j_1}^{\mathbf{k}^1} \gamma_{j_2}^{\mathbf{k}^2} = \mathcal{O}((\omega_j N)^{-1}) \quad \text{for} \quad j_1 + j_2 = j \bmod 2N \;\; \text{and} \;\; \mathbf{k}^1 + \mathbf{k}^2 = \pm\langle j\rangle, \tag{50}$$

which follows from the non-resonance estimates of Section 5 exactly as in the proof of Lemma 6. The estimate $|||z|||_\tau \le |z(0)| + 2\,|||\frac{dz}{d\tau}|||_\tau$ for $\tau \le 1$ together with an induction argument thus yields

$$|||z_{j,m}^{\pm\langle j\rangle}|||_\tau \le |z_{j,m}^{\pm\langle j\rangle}(0)| + C \sum_{b\in\mathcal{B}_{j,m}^{\pm\langle j\rangle}} \prod_{\ell=1}^{N} \prod_{p=1}^{m-1} |||z_{\ell,p}^{\langle \ell\rangle}|||_\tau^{\kappa_{\ell,p}^+(b)}\, |||z_{\ell,p}^{-\langle \ell\rangle}|||_\tau^{\kappa_{\ell,p}^-(b)}, \tag{51}$$

where $\mathcal{B}_{j,m}^{\pm\langle j\rangle}$ is the set of binary trees with root $(j, \pm\langle j\rangle, m)$, and $\kappa_{\ell,p}^{\pm}(b)$ have the same meaning as in (47).

We arrive at the point where we have to estimate the initial value $z_{j,m}^{\pm\langle j\rangle}(0)$, defined by (34). It is bounded in terms of the scaled quantities a_j, b_j, $z_{j,m}^{\mathbf{k}}(0)$ for $\mathbf{k} \ne \pm\langle j\rangle$, and of $\frac{d}{d\tau} z_{j,m}^{\mathbf{k}}(0)$ for all $\mathbf{k} \in \mathcal{K}_j$. Using (49) this implies the boundedness of $z_{j,m}^{\pm\langle j\rangle}(0)$ for $|j| \le K+M$ and $m \le K+M$. For large $|j| > K+M$ and $m = K$ the function $z_{j,m}^{\pm\langle j\rangle}(\tau)$ is constant, and we have

$$z_{j,m}^{\pm\langle j\rangle}(0) = \frac{1}{2}(a_j \mp \mathrm{i} b_j), \qquad z_{-j,m}^{\pm\langle j\rangle}(0) = \frac{1}{2}(\overline{a}_j \mp \mathrm{i}\overline{b}_j),$$

because in this case $z_{j,m}^{\mathbf{k}}(0) = 0$ for all $\mathbf{k} \ne \pm\langle j\rangle$. By induction we thus obtain for $|j| > K+M$ and $K \le m \le K+M$ that

$$|z_{j,m}^{\pm\langle j\rangle}(0)|^2 \le C \sum_{|\ell-j|\le M} \Big(|a_\ell|^2 + |b_\ell|^2 \Big).$$

By (51) we obtain a similar estimate for $|||z_{j,m}^{\pm\langle j\rangle}|||_\tau$, and by (47) for all $|||z_{j,m}^{\mathbf{k}}|||_\tau$. Using

$$\sum_{j=-N}^{N-1} \sigma_j \left(|a_j|^2 + |b_j|^2 \right) \le C,$$

which follows from the assumption (26) on the initial values, we obtain the following result.

Lemma 7. *For $m \le K+M$, we have for $\tau \le 1$*

$$\sum_{j=-N}^{N-1} \sigma_j \Big(\sum_{\mathbf{k}\in\mathcal{K}_j} |||z_{j,m}^{\mathbf{k}}|||_\tau \Big)^2 \le C. \tag{52}$$

Together with the observation that

$$z_{j,m}^{\mathbf{k}} = 0 \quad \text{ for } \quad m < \min(|j|, K), \tag{53}$$

Lemma 7 yields (29). Analogous bounds are obtained for the derivatives. As a further consequence of this reduction process by binary trees, we note that

$$z_{j,m}^{\mathbf{k}} = 0 \qquad \text{if} \qquad k_\ell \ne 0 \quad \text{and} \quad m < \min(\ell, K), \tag{54}$$

and the only non-vanishing coefficient functions for $m = \min(\ell, K)$ are the diagonal terms $z_{\ell,m}^{\pm\langle\ell\rangle}$.

6.7. Defect in the modulation equations. We consider the series (35) truncated after $K+M$ terms,

$$z_j^{\mathbf{k}}(\tau) = \frac{1}{N^2} \sum_{m\le K+M} \delta^m \, z_{j,m}^{\mathbf{k}}(\tau). \tag{55}$$

As an approximation to the solution of (3) we thus take

$$\omega_j \widetilde{u}_j(t) = \sum_{\mathbf{k}\in\mathcal{K}_j} z_j^{\mathbf{k}}(\tau)\, \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega})t}. \tag{56}$$

The construction is such that $\widetilde{u}_j(0) = u_j(0)$ and $\dot{\widetilde{u}}_j(0) = \dot{u}_j(0)$ for all j. The functions $z_j^{\mathbf{k}}(\tau)$ do not satisfy the modulation equations (32)-(33) exactly. We denote the defects, for $j = -N, \dots, N-1$ and arbitrary multi-indices $\mathbf{k} = (k_1, \dots, k_N)$, by

$$\begin{aligned} d_j^{\mathbf{k}} \;&=\; \big(\omega_j^2 - (\mathbf{k}\cdot\boldsymbol{\omega})^2\big)\, z_j^{\mathbf{k}} + 2\mathrm{i}\,(\mathbf{k}\cdot\boldsymbol{\omega}) N^{-3}\, \frac{dz_j^{\mathbf{k}}}{d\tau} + N^{-6}\, \frac{d^2 z_j^{\mathbf{k}}}{d\tau^2} \\ &\quad + \mathrm{i}\,\omega_j^2 \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)} \sum_{\mathbf{k}^1+\mathbf{k}^2=\mathbf{k}} z_{j_1}^{\mathbf{k}^1} z_{j_2}^{\mathbf{k}^2}. \end{aligned} \tag{57}$$

By construction of the coefficient functions $z_{j,m}^{\mathbf{k}}$, the coefficients of δ^m in a δ-expansion of the defect vanish for $m \le K+M$. We thus have

$$\omega_j^{-2}\, d_j^{\mathbf{k}} = \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)} \sum_{\mathbf{k}^1+\mathbf{k}^2=\mathbf{k}} \; \sum_{m=K+M+1}^{2(K+M)} \; \sum_{m_1+m_2=m} z_{j_1,m_1}^{\mathbf{k}^1} z_{j_2,m_2}^{\mathbf{k}^2}\, \delta^m N^{-4}. \tag{58}$$

Using the estimate (52) of the modulation functions and Lemma 5, the defect is thus bounded by

$$\sum_{j=-N}^{N-1} \sigma_j \Big(\omega_j^{-2} \sum_{\mathbf{k}} |||d_j^{\mathbf{k}}|||_\tau \Big)^2 \le C \left(\delta^{K+M+1} N^{-4} \right)^2 \qquad \text{for} \ \ 0 \le \tau \le 1. \qquad (59)$$

With (50), a stronger bound is obtained for the diagonal defects:

$$\sum_{j=-N}^{N-1} \sigma_j \Big(\omega_j^{-1} \, |||d_j^{\pm\langle j\rangle}|||_\tau \Big)^2 \le C \left(\delta^{K+M+1} N^{-5} \right)^2 \qquad \text{for} \ \ 0 \le \tau \le 1. \qquad (60)$$

The defect $d_j^{\mathbf{k}}$ vanishes for multi-indices $(j,\mathbf{k})$ that cannot be decomposed as $j = j_1 + j_2$ and $\mathbf{k} = \mathbf{k}^1 + \mathbf{k}^2$ with $\mathbf{k}^1 \in \mathcal{K}_{j_1}, \mathbf{k}^2 \in \mathcal{K}_{j_2}$.

6.8. Remainder term of the modulated Fourier expansion. We compare the approximation (56) with the exact solution $u_j(t)$ of (3) for initial values satisfying (10). With the notation of Section 2 this yields, for $j = -N, \dots, N-1$,

$$\begin{aligned} \ddot{u}_j + \omega_j^2 \, u_j + \nabla_{-j} U(\mathbf{u}) &= 0 \\ \ddot{\widetilde{u}}_j + \omega_j^2 \, \widetilde{u}_j + \nabla_{-j} U(\widetilde{\mathbf{u}}) &= \vartheta_j \end{aligned}$$

where the defect $\vartheta_j(t)$ is given by

$$\omega_j \vartheta_j(t) = \sum_{\mathbf{k}} d_j^{\mathbf{k}}(\tau) \, \mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega})t} \qquad \text{with} \quad \tau = N^{-3} t,$$

and, by (59), bounded by

$$\sum_{j=-N}^{N-1} \sigma_j \, |\omega_j^{-1} \vartheta_j(t)|^2 \le C \left(\delta^{K+M+1} N^{-4} \right)^2 \qquad \text{for} \quad t \le N^3. \qquad (61)$$

The initial values satisfy $\widetilde{u}_j(0) = u_j(0)$ and $\dot{\widetilde{u}}_j(0) = \dot{u}_j(0)$.

The nonlinearity $\overline{\nabla} U(\mathbf{u}) = \mathbf{b}(\mathbf{u},\mathbf{u})$ is a quadratic form corresponding to the symmetric bilinear form

$$\mathbf{b}(\mathbf{u},\mathbf{v}) = \Bigg(-\mathrm{i}\,\omega_j \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)} \, \omega_{j_1} \, \omega_{j_2} \; u_{j_1} \, v_{j_2} \Bigg)_{j=-N}^{N-1}.$$

Using Lemma 5 and $|\omega_j| \le 2$, we thus obtain the bound

$$\|\mathbf{b}(\mathbf{u},\mathbf{v})\| \le 2 \, \|\boldsymbol{\Omega}^{-1} \mathbf{b}(\mathbf{u},\mathbf{v})\| \le 2\,c\,\|\boldsymbol{\Omega}\mathbf{u}\| \cdot \|\boldsymbol{\Omega}\mathbf{v}\| \, .$$

We then estimate the error $r_j(t) = u_j(t) - \widetilde{u}_j(t)$ by standard arguments: we rewrite the second-order differential equation as a system of first-order differential equations

$$\begin{aligned} \omega_j \dot{\widetilde{u}}_j &= \omega_j \, \widetilde{v}_j \\ \dot{\widetilde{v}}_j &= -\omega_j \, (\omega_j \, \widetilde{u}_j) + b_j(\widetilde{\mathbf{u}},\widetilde{\mathbf{u}}) + \vartheta_j(t). \end{aligned}$$

We use the variation-of-constants formula and the Lipschitz bound

$$\|\mathbf{b}(\mathbf{u},\mathbf{u}) - \mathbf{b}(\widetilde{\mathbf{u}},\widetilde{\mathbf{u}})\| \le 2\,c\,\big(\|\boldsymbol{\Omega}\mathbf{u}\| + \|\boldsymbol{\Omega}\widetilde{\mathbf{u}}\|\big)\,\|\boldsymbol{\Omega}\mathbf{u} - \boldsymbol{\Omega}\widetilde{\mathbf{u}}\| \le 4\,c\,C\,\sqrt{\epsilon}\|\boldsymbol{\Omega}\mathbf{u} - \boldsymbol{\Omega}\widetilde{\mathbf{u}}\|$$

for $\|\boldsymbol{\Omega}\mathbf{u}\| \le C\sqrt{\epsilon}$, $\|\boldsymbol{\Omega}\widetilde{\mathbf{u}}\| \le C\sqrt{\epsilon}$. By (61) we have $\|\vartheta(t)\| \le C\delta^{K+M+1}N^{-4}$ for $t \le N^3$. The Gronwall inequality then shows that the error satisfies the bound (28) for $t \le \min(N^3, \epsilon^{-1/2})$. This completes the proof of Theorem 2.

7. Bounds in terms of the diagonal modulation functions

The following result bounds the norms (43) of the non-diagonal modulation functions $z_j^{\mathbf{k}}$ with $\mathbf{k} \ne \pm\langle j\rangle$ in terms of those of the diagonal functions $z_\ell^{\pm\langle\ell\rangle}$.

Lemma 8. *Under the assumptions of Theorem 2, we have the bound*

$$|||N^2 z_j^{\mathbf{k}}|||_\tau \le C \prod_{\ell=1}^{N}\Big(|||N^2 z_\ell^{\langle\ell\rangle}|||_\tau + |||N^2 z_\ell^{-\langle\ell\rangle}|||_\tau\Big)^{|k_\ell|} + N^2\vartheta_j^{\mathbf{k}} \quad \textit{for} \quad \tau \le 1, \tag{62}$$

for $|j| \le N$ *and* $\mathbf{k} = (k_1,\dots,k_N) \in \mathcal{K}_j$, *with*

$$\sum_{j=-N}^{N-1} \sigma_j \Big(\sum_{\mathbf{k}\in\mathcal{K}_j} |\vartheta_j^{\mathbf{k}}|\Big)^2 \le \widehat{C}\,\big(\delta^{K+M+1}N^{-3}\big)^2. \tag{63}$$

The constants C and $\widehat{C}$ are independent of E, N and τ.

We recall that $N^2 z_\ell^{\pm\langle\ell\rangle}(\tau) = \mathcal{O}(\delta^{\min(\ell,K)})$.

Proof. The proof uses arguments similar to those in Section 6.6. The defect equation (57) yields, for $\mathbf{k} \ne \langle j\rangle$,

$$|||N^2 z_j^{\mathbf{k}}|||_\tau \le \gamma_j^{\mathbf{k}} \sum_{j_1+j_2=j \bmod 2N}\ \sum_{\mathbf{k}^1+\mathbf{k}^2=\mathbf{k}} |||N^2 z_{j_1}^{\mathbf{k}^1}|||_\tau \cdot |||N^2 z_{j_2}^{\mathbf{k}^2}|||_\tau + \gamma_j^{\mathbf{k}}\omega_j^{-2}N^2|||d_j^{\mathbf{k}}|||_\tau$$

with $\gamma_j^{\mathbf{k}}$ of (46). We denote by $\mathcal{B}_j^{\mathbf{k}}$ the set of binary trees with root $(j,\mathbf{k})$ that are obtained by omitting the labels m' at all nodes $(j',\mathbf{k}',m')$ of trees in $\mathcal{B}_{j,m}^{\mathbf{k}}$ for all $m \le K+M$, and by $\kappa_\ell^{\pm}(b)$ the number of appearances of $(\ell,\pm\langle\ell\rangle)$ among the leaves of a tree b. We then obtain the bound

$$|||N^2 z_j^{\mathbf{k}}|||_\tau \le \sum_{b\in\mathcal{B}_j^{\mathbf{k}}} \Gamma(b) \prod_{\ell\ne 0} |||N^2 z_\ell^{\langle\ell\rangle}|||_\tau^{\kappa_\ell^+(b)}\, |||N^2 z_\ell^{-\langle\ell\rangle}|||_\tau^{\kappa_\ell^-(b)} + N^2\vartheta_j^{\mathbf{k}} \tag{64}$$

with the same factor $\Gamma(b)$ as in (47), and with $\vartheta_j^{\mathbf{k}}$ satisfying the estimate (63) because of (59) and the bound $\gamma_j^{\mathbf{k}} = \mathcal{O}(\omega_j N)$. Since every tree in $\mathcal{B}_j^{\mathbf{k}}$ contains the leaf $(\ell,\pm\langle\ell\rangle)$ at least $|k_{|\ell|}|$ times, this yields (together with $z_{-\ell}^{\mp\langle\ell\rangle} = \overline{z_\ell^{\pm\langle\ell\rangle}}$)

$$|||N^2 z_j^{\mathbf{k}}|||_\tau \le \sum_{b\in\mathcal{B}_j^{\mathbf{k}}} \Gamma(b) \prod_{\ell\ge 1}\Big(|||N^2 z_\ell^{\langle\ell\rangle}|||_\tau + |||N^2 z_\ell^{-\langle\ell\rangle}|||_\tau\Big)^{|k_\ell|} + N^2\vartheta_j^{\mathbf{k}}. \tag{65}$$

Since $\Gamma(b) \le Const.$ and the number of trees in $\mathcal{B}_j^{\mathbf{k}}$ is independent of N, this implies the stated result. $\square$

We also need another bound that follows from (64).

Lemma 9. *Under the assumptions of Theorem 2, we have the bound*

$$|||z_j^{\mathbf{k}}|||_\tau \le C \sum_{\ell=1}^{N} \delta^{|j-\ell|} \Big(|||z_\ell^{\langle \ell\rangle}|||_\tau + |||z_\ell^{-\langle \ell\rangle}|||_\tau \Big) + \vartheta_j^{\mathbf{k}} \quad \textit{for} \quad \tau \le 1, \tag{66}$$

for $0 < j \le N$ *and* $\mathbf{k} \in \mathcal{K}_j$, *where* $\vartheta_j^{\mathbf{k}}$ *is bounded by* (63). *The sum in* (66) *is actually only over* ℓ *with* $|j-\ell| < K+M$. *We recall* $z_{-j}^{-\mathbf{k}} = \overline{z_j^{\mathbf{k}}}$.

Proof. For a tree $b \in \mathcal{B}_j^{\mathbf{k}}$, let $(\ell, \pm\langle \ell\rangle)$ be a leaf such that $|\ell|$ is largest among the leaves of b. For the other leaves $(\ell_i, \pm\langle \ell_i\rangle)$ we have $\ell + \sum_i \ell_i = j$ and therefore $|j-\ell| \le \sum_i |\ell_i| < K+M$. Since $z_{\ell_i}^{\pm\langle \ell_i\rangle}$ is bounded by $\mathcal{O}(N^{-2}\delta^{|\ell_i|})$, we obtain the result from (64). $\square$

The next lemma bounds the derivatives of the diagonal functions in terms of their function values.

Lemma 10. *Under the assumptions of Theorem 2, we have the bounds*

$$|||z_j^{\langle j\rangle}|||_\tau + |||z_j^{-\langle j\rangle}|||_\tau \le 2\Big(|z_j^{\langle j\rangle}(\tau)| + |z_j^{-\langle j\rangle}(\tau)| \Big) + \vartheta_j \quad \textit{for} \quad \tau \le 1, \tag{67}$$

for $j = -N, \dots, N-1$, *with*

$$\sum_{j=-N}^{N-1} \sigma_j \, \vartheta_j^2 \le C \left(\delta^{K+M+1} N^{-2} \right)^2. \tag{68}$$

The constant C *is independent of* E, N *and* τ.

Proof. From the defect equation (57) with $\mathbf{k} = \pm\langle j\rangle$ we obtain

$$\begin{aligned} |||N^2 \frac{dz_j^{\pm\langle j\rangle}}{d\tau}|||_\tau \le 2\omega_j N \sum_{j_1+j_2=j \bmod 2N} \; \sum_{\mathbf{k}^1+\mathbf{k}^2=\pm\langle j\rangle} |||N^2 z_{j_1}^{\mathbf{k}^1}|||_\tau \cdot |||N^2 z_{j_2}^{\mathbf{k}^2}|||_\tau \\ + \omega_j^{-1} N^5 |||d_j^{\pm\langle j\rangle}|||_\tau. \end{aligned}$$

The last term is bounded by (60). We now use (65) for $z_{j_1}^{\mathbf{k}^1}$ and $z_{j_2}^{\mathbf{k}^2}$ and note that by (50) we have $\Gamma(b_1)\Gamma(b_2) = \mathcal{O}((\omega_j N)^{-1})$ for all $b_1 \in \mathcal{B}_{j_1}^{\mathbf{k}^1}$ and $b_2 \in \mathcal{B}_{j_2}^{\mathbf{k}^2}$. This gives us

$$\begin{aligned} & |||N^2 \frac{dz_j^{\pm\langle j\rangle}}{d\tau}|||_\tau \le \\ & C_1 \sum_{j_1+j_2=j \bmod 2N} \; \sum_{\mathbf{k}^1+\mathbf{k}^2=\pm\langle j\rangle} \; \prod_{\ell\ge 1} \Big(|||N^2 z_\ell^{\langle \ell\rangle}|||_\tau + |||N^2 z_\ell^{-\langle \ell\rangle}|||_\tau \Big)^{|k_\ell^1|+|k_\ell^2|} + N^2 \vartheta_j \\ & \le C_2 \delta^2 \Big(|||N^2 z_j^{\langle j\rangle}|||_\tau + |||N^2 z_j^{-\langle j\rangle}|||_\tau \Big) + N^2 \vartheta_j, \end{aligned}$$

where ϑ_j is bounded by (68) because of (60) and (63). For the second inequality we note that the number of terms in the sums is independent of N, and that $(\pm j, \pm\langle j\rangle)$ (for some combination of signs) must appear among the leaves, in addition to at least two further leaves that account for the presence of the factor δ^2. On the other hand, we have

$$|||z_j^{\pm\langle j\rangle}|||_\tau \leq |z_j^{\pm\langle j\rangle}(\tau)| + |||\frac{dz_j^{\pm\langle j\rangle}}{d\tau}|||_\tau .$$

Hence we obtain

$$(1 - C_2\delta^2)\big(|||z_j^{\langle j\rangle}|||_\tau + |||z_j^{-\langle j\rangle}|||_\tau\big) \leq |z_j^{\langle j\rangle}(\tau)| + |z_j^{-\langle j\rangle}(\tau)| + \vartheta_j ,$$

which yields the result if $C_2\delta^2 \leq \frac{1}{2}$. $\square$

8. Almost-invariant energies of the modulation system

We now show that the system of equations determining the modulation functions has almost-invariants that bound the mode energies $E_j(t)$ from above. This fact will lead to the long-time energy bounds of Theorem 1. The construction of the almost-invariants takes up a line of arguments from [18, Chapter XIII] and [9, Section 4].

8.1. The extended potential. We introduce the functions

$$\mathbf{y}(t) = \big(y_j^{\mathbf{k}}(t)\big)_{(j,\mathbf{k})\in\mathcal{K}} \qquad \text{with} \quad \omega_j y_j^{\mathbf{k}}(t) = z_j^{\mathbf{k}}(\tau)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega})t} \tag{69}$$

with $\tau = N^{-3}t$, so that $u_j(t) \approx \widetilde{u}_j(t) = \sum_{\mathbf{k}} y_j^{\mathbf{k}}(t)$. By the construction of the functions $z_j^{\mathbf{k}}$ from the modulation equations, the functions $y_j^{\mathbf{k}}$ satisfy

$$\ddot{y}_j^{\mathbf{k}} + \omega_j^2 y_j^{\mathbf{k}} + \mathrm{i}\,\omega_j \sum_{j_1+j_2=j \bmod 2N} (-1)^{(j_1+j_2-j)/(2N)}\,\omega_{j_1}\,\omega_{j_2} \sum_{\mathbf{k}^1+\mathbf{k}^2=\mathbf{k}} y_{j_1}^{\mathbf{k}^1}\, y_{j_2}^{\mathbf{k}^2} = e_j^{\mathbf{k}}, \tag{70}$$

where the defects $e_j^{\mathbf{k}}(t) = \omega_j^{-1} d_j^{\mathbf{k}}(\tau)\,\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega})t}$ are bounded by (59). The nonlinearity is recognised as the partial derivative with respect to $y_{-j}^{-\mathbf{k}}$ of the *extended potential* $\mathcal{U}(\mathbf{y})$ given by

$$\mathcal{U}(\mathbf{y}) = -\frac{\mathrm{i}}{3} \sum_{j_1+j_2+j_3=0 \bmod 2N} (-1)^{(j_1+j_2+j_3)/(2N)}\,\omega_{j_1}\,\omega_{j_2}\,\omega_{j_3} \sum_{\mathbf{k}^1+\mathbf{k}^2+\mathbf{k}^3=\mathbf{0}} y_{j_1}^{\mathbf{k}^1}\, y_{j_2}^{\mathbf{k}^2}\, y_{j_3}^{\mathbf{k}^3}. \tag{71}$$

Hence, the modulation system can be rewritten as

$$\ddot{y}_j^{\mathbf{k}} + \omega_j^2 y_j^{\mathbf{k}} + \nabla_{-j}^{-\mathbf{k}}\,\mathcal{U}(\mathbf{y}) = e_j^{\mathbf{k}}, \tag{72}$$

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

8.2. Invariance under group actions. The extended potential turns out to be invariant under the continuous group action that, for an arbitrary real vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N) \in \mathbb{R}^N$ and for $\theta \in \mathbb{R}$, is given as

$$S_{\boldsymbol{\lambda}}(\theta)\mathbf{y} = \big(\mathrm{e}^{\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\lambda})\theta} y_j^{\mathbf{k}}\big)_{j,\mathbf{k}}.$$

Since the sum in the definition of $\mathcal{U}$ is over $\mathbf{k}^1 + \mathbf{k}^2 + \mathbf{k}^3 = \mathbf{0}$, we have

$$\mathcal{U}(S_{\boldsymbol{\lambda}}(\theta)\mathbf{y}) = \mathcal{U}(\mathbf{y}) \quad \text{for} \quad \theta \in \mathbb{R}. \tag{73}$$

Differentiating this relation wih respect to θ yields

$$0 = \frac{d}{d\theta}\Big|_{\theta=0} \mathcal{U}(S_{\boldsymbol{\lambda}}(\theta)\mathbf{y}) = \sum_{j=-N}^{N-1} \sum_{\mathbf{k}} \mathrm{i}(\mathbf{k}\cdot\boldsymbol{\lambda})\, y_j^{\mathbf{k}} \nabla_j^{\mathbf{k}} \mathcal{U}(\mathbf{y}). \tag{74}$$

In fact, the full Lagrangian of the system (72) without the perturbations $e_j^{\mathbf{k}}$,

$$\mathcal{L}(\mathbf{y}, \dot{\mathbf{y}}) = \frac{1}{2} \sum_{j=-N}^{N-1} \sum_{\mathbf{k}} \Big(\dot{y}_{-j}^{-\mathbf{k}} \dot{y}_j^{\mathbf{k}} - \omega_j^2 y_{-j}^{-\mathbf{k}} y_j^{\mathbf{k}} \Big) - \mathcal{U}(\mathbf{y}),$$

is invariant under the action of the one-parameter groups $S_{\boldsymbol{\lambda}}(\theta)$. By Noether's theorem, the corresponding Lagrangian system has a set of invariants $\mathcal{I}_{\boldsymbol{\lambda}}(\mathbf{y}, \dot{\mathbf{y}})$, which we now study as almost-invariants of the perturbed system (72).

8.3. Almost-invariant energies of the modulation system. We multiply (72) with $\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\lambda}) y_{-j}^{-\mathbf{k}}$ and sum over j and $\mathbf{k}$. Using (74), we obtain

$$\sum_j \sum_{\mathbf{k}} \mathrm{i}(\mathbf{k}\cdot\boldsymbol{\lambda}) \big(y_{-j}^{-\mathbf{k}} \ddot{y}_j^{\mathbf{k}} + \omega_j^2 y_{-j}^{-\mathbf{k}} y_j^{\mathbf{k}} \big) = \sum_j \sum_{\mathbf{k}} \mathrm{i}(\mathbf{k}\cdot\boldsymbol{\lambda}) y_{-j}^{-\mathbf{k}} e_j^{\mathbf{k}}\,,$$

where we notice that the second terms in the sum on the left-hand side cancel. The left-hand side then equals $-\frac{d}{dt}\mathcal{I}_{\boldsymbol{\lambda}}(\mathbf{y}, \dot{\mathbf{y}})$ with

$$\mathcal{I}_{\boldsymbol{\lambda}}(\mathbf{y}, \dot{\mathbf{y}}) = -\sum_{j=-N}^{N-1} \sum_{\mathbf{k}} \mathrm{i}(\mathbf{k}\cdot\boldsymbol{\lambda}) y_{-j}^{-\mathbf{k}} \dot{y}_j^{\mathbf{k}}\,. \tag{75}$$

Hence we obtain

$$\frac{d}{dt}\mathcal{I}_{\boldsymbol{\lambda}}(\mathbf{y}, \dot{\mathbf{y}}) = -\sum_{j=-N}^{N-1} \sum_{\mathbf{k}} \mathrm{i}(\mathbf{k}\cdot\boldsymbol{\lambda}) y_{-j}^{-\mathbf{k}} e_j^{\mathbf{k}}\,. \tag{76}$$

In the following we consider the almost-invariants as functions of the modulation sequence $\mathbf{z}(\tau) = (z_j^{\mathbf{k}}(\tau))$ and its derivative $(d\mathbf{z}/d\tau)(\tau)$ with respect to the slow time variable $\tau = N^{-3}t$, rather than of $\mathbf{y}(t)$ defined by (69) and $\dot{\mathbf{y}} = d\mathbf{y}/dt$. For $\boldsymbol{\lambda} = \omega_\ell \langle \ell \rangle$, a multiple of the ℓth unit vector, we write

$$\mathcal{E}_\ell\big(\mathbf{z}, \tfrac{d\mathbf{z}}{d\tau}\big) = 2N\,\mathcal{I}_{\omega_\ell\langle\ell\rangle}(\mathbf{y}, \dot{\mathbf{y}}) = 2N\,\omega_\ell\,\mathcal{I}_{\langle\ell\rangle}(\mathbf{y}, \dot{\mathbf{y}})$$

so that

$$\mathcal{E}_\ell\big(\mathbf{z}, \tfrac{d\mathbf{z}}{d\tau}\big) = -2N\,\omega_\ell \sum_{j=-N}^{N-1} \sum_{\mathbf{k}} \mathrm{i}k_\ell\,\omega_j^{-2} z_{-j}^{-\mathbf{k}} \Big(\mathrm{i}(\mathbf{k}\cdot\boldsymbol{\omega}) z_j^{\mathbf{k}} + N^{-3}\,\frac{dz_j^{\mathbf{k}}}{d\tau}\Big). \tag{77}$$

By the estimates of the modulation functions we have at the initial time

$$\begin{aligned} \Big|\mathcal{E}_\ell\big(\mathbf{z}(0), \tfrac{d\mathbf{z}}{d\tau}(0)\big)\Big| &\le \mathcal{C}_0\,E\,\delta^{2(\ell-1)} \qquad \text{for} \quad \ell = 1,\dots,K,\\ \sum_{\ell=K}^{N} \sigma_\ell \Big|\mathcal{E}_\ell\big(\mathbf{z}(0), \tfrac{d\mathbf{z}}{d\tau}(0)\big)\Big| &\le \mathcal{C}_0\,E\,\delta^{2(K-1)}, \end{aligned} \tag{78}$$

where $\mathcal{C}_0$ only depends on the initial values $\big(\mathbf{u}(0), \dot{\mathbf{u}}(0)\big)$. From (76) we have

$$N^{-3}\,\frac{d}{d\tau}\mathcal{E}_\ell\big(\mathbf{z}, \tfrac{d\mathbf{z}}{d\tau}\big) = -\mathrm{i}\,2N\,\omega_\ell \sum_{j=-N}^{N-1} \sum_{\mathbf{k}} k_\ell\, z_{-j}^{-\mathbf{k}}\,\omega_j^{-2} d_j^{\mathbf{k}}\,. \tag{79}$$

Theorem 3. *Under the conditions of Theorem 2 we have for $\tau \le 1$*

$$\Big|\frac{d}{d\tau}\mathcal{E}_\ell\big(\mathbf{z}(\tau), \tfrac{d\mathbf{z}}{d\tau}(\tau)\big)\Big| \le \vartheta_\ell,$$

where

$$\begin{aligned} \vartheta_\ell \;&\le\; C\,E\,\delta^{\ell+K+M-1} \qquad \textit{for} \quad \ell = 1,\dots,K\\ \sum_{\ell=K}^{N} \sigma_\ell\,\vartheta_\ell \;&\le\; C\,E\,\delta^{2K+M-1}. \end{aligned} \tag{80}$$

Proof. We insert (58) into (79) to obtain, with $L = \min(\ell, K)$,

$$N^{-3}\,\Big|\frac{d}{d\tau}\mathcal{E}_\ell\big(\mathbf{z}, \tfrac{d\mathbf{z}}{d\tau}\big)\Big| \le 2N\omega_\ell \sum |k_\ell\, z_{-j,m}^{-\mathbf{k}}\, z_{j_1,m_1}^{\mathbf{k}^1}\, z_{j_2,m_2}^{\mathbf{k}^2}|\,\delta^{L+K+M+1} N^{-6},$$

where the sum is over all indices j, j_1, j_2 with $j_1 + j_2 = j$, multi-indices $\mathbf{k}, \mathbf{k}^1, \mathbf{k}^2$ with $\mathbf{k}^1 + \mathbf{k}^2 = \mathbf{k}$ and indices $m, m_1, m_2 \le K + M$ with $m \ge L$ (note (54)) and $m_1 + m_2 \ge K + M + 1$. This sum contains a number of terms that is independent of N. Estimating the modulation functions by (45) yields

$$\Big|\frac{d}{d\tau}\mathcal{E}_\ell\big(\mathbf{z}, \tfrac{d\mathbf{z}}{d\tau}\big)\Big| \le 2N\omega_\ell \sum |k_\ell|\,\gamma_{-j}^{-\mathbf{k}}\gamma_{j_1}^{\mathbf{k}^1}\gamma_{j_2}^{\mathbf{k}^2}\Big(\prod_{i=3}^{8} |||z_{j_i,m_i}^{\mathbf{k}^i}|||\Big)\,\delta^{L+K+M+1} N^{-3},$$

where the number of terms in the sum is still independent of N. As in (48), the non-resonance estimates of Section 5 yield

$$\omega_\ell\,k_\ell\,\gamma_{-j}^{-\mathbf{k}}\gamma_{j_1}^{\mathbf{k}^1}\gamma_{j_2}^{\mathbf{k}^2} = \mathcal{O}(N^{-1})$$

for $j_1 + j_2 = j \bmod 2N$ and $\mathbf{k}^1 + \mathbf{k}^2 = \mathbf{k}$. Among the terms in the product over i, two terms satisfy $|j_i - \ell| \le 2M$, and all others $|j_i| \le K + M$. For the latter, we estimate $|||z_{j_i,m_i}^{\mathbf{k}^i}|||$ by a constant, for the two others we use the Cauchy-Schwarz inequality together with the estimate (52) to arrive at the stated result. □

8.4. Almost-invariant energies and diagonal modulation functions. The following lemma shows in particular that $|z_\ell^{\pm\langle\ell\rangle}(\tau)|^2$ is bounded in terms of the almost-invariant $\mathcal{E}_\ell$.

Theorem 4. *Under the conditions of Theorem 2, there exists $c > 0$ independent of E and N such that*

$$\mathcal{E}_\ell\big(\mathbf{z}(\tau), \frac{d\mathbf{z}}{d\tau}(\tau)\big) \geq (1 - c\delta^2)\, 4N\Big(|z_\ell^{\langle\ell\rangle}(\tau)|^2 + |z_\ell^{-\langle\ell\rangle}(\tau)|^2\Big) - \vartheta_\ell \quad \textit{and}$$
$$\mathcal{E}_\ell\big(\mathbf{z}(\tau), \frac{d\mathbf{z}}{d\tau}(\tau)\big) \leq (1 + c\delta^2)\, 4N\Big(|z_\ell^{\langle\ell\rangle}(\tau)|^2 + |z_\ell^{-\langle\ell\rangle}(\tau)|^2\Big) + \vartheta_\ell$$

for $\tau \leq 1$, with ϑ_ℓ bounded as in (80).

Proof. $\mathcal{E}_\ell$ has the four terms $|z_{\pm\ell}^{\pm\langle\ell\rangle}|^2$ in the sum (77) and further terms containing $z_j^{\mathbf{k}}$ for $(j, \mathbf{k}) \in \mathcal{K}$ with $\mathbf{k} \neq \pm\langle j\rangle$ and $k_\ell \neq 0$. For such $(j, \mathbf{k})$ we note that $\omega_j^{-2}(k_\ell\omega_\ell)(\mathbf{k}\cdot\boldsymbol{\omega})$ is bounded independently of N, and from Lemmas 8 and 10 we have the bound

$$\big|z_j^{\mathbf{k}}(\tau)\big| + \Big|\frac{d}{d\tau} z_j^{\mathbf{k}}(\tau)\Big| \leq C\delta\Big(|z_\ell^{\langle\ell\rangle}(\tau)| + |z_\ell^{-\langle\ell\rangle}(\tau)|\Big) + \vartheta_j^{\mathbf{k}} + \widetilde{\vartheta}_\ell$$

with $\vartheta_j^{\mathbf{k}}$ bounded as in (63), and with $\widetilde{\vartheta}_\ell$ bounded as in (68). This yields the result. □

8.5. Bounding the mode energies by the almost-invariant energies. We are now in the position to bound the mode energies $E_j(t) = E_j(\mathbf{u}(t), \dot{\mathbf{u}}(t))$ of (8) in terms of the almost-invariants

$$\mathcal{E}_\ell(t) = \mathcal{E}_\ell\big(\mathbf{z}(\tau), \frac{d\mathbf{z}}{d\tau}(\tau)\big) \quad \text{for} \quad \tau = N^{-3}t.$$

Theorem 5. *Under the conditions of Theorem 2, we have*

$$E_j(t) \leq (\mathcal{C} + c\delta^2) \sum_{\ell=1}^{N} \delta^{2|j-\ell|}\, \mathcal{E}_\ell(t) + \vartheta_j \quad \textit{for} \quad t \leq \min(N^3, \epsilon^{-1/2}),$$

where ϑ_j is bounded as in (80) and $\mathcal{C}, c$ are independent of E and N. The constant $\mathcal{C}$ only depends on C_0 of (78), and c depends on C_0 of (26). The sum is actually only over ℓ with $|j - \ell| < K + M$.

Proof. We insert in $\epsilon_j(t)$ the modulated Fourier expansion (27) for $\omega_j u_j(t)$ and $\dot{u}_j(t)$ and use the remainder bound (28). We use Lemmas 9 and 10 to bound the modulation functions $z_j^{\mathbf{k}}$ in terms of the diagonal functions $z_\ell^{\pm\langle\ell\rangle}$, and Theorem 4 to bound the diagonal functions in terms of the almost-invariants $\mathcal{E}_\ell$. This yields the stated result. □

Theorem 6. *Under the conditions of Theorem 2, we have*

$$|E_1(t) - \mathcal{E}_1(t)| \leq C\, E\, \delta^2 \quad \textit{for} \quad t \leq \min(N^3, \epsilon^{-1/2}).$$

Proof. We insert in $\epsilon_1(t)$ the modulated Fourier expansion (27). Using the estimate (52) we obtain

$$E_1(t) = 4N\Big(|z_1^{\langle 1\rangle}(\tau)|^2 + |z_1^{-\langle 1\rangle}(\tau)|^2\Big) + \mathcal{O}(E\,\delta^2).$$

Together with Theorem 4 this gives the result. □

8.6. Dependence of the almost-invariant energies on the initial values.

Lemma 11. *In the situation of Theorem 2, consider perturbed initial values* $(\widetilde{\mathbf{u}}(0), \dot{\widetilde{\mathbf{u}}}(0))$ *whose difference to* $(\mathbf{u}(0), \dot{\mathbf{u}}(0))$ *is bounded in the norm (24) by*

$$\big\|\big(\mathbf{u}(0)-\widetilde{\mathbf{u}}(0), \dot{\mathbf{u}}(0)-\dot{\widetilde{\mathbf{u}}}(0)\big)\big\| \le \vartheta \qquad \textit{with} \quad \vartheta \le \widetilde{C}\,\delta^{K+M+1}N^{-2}.$$

Then, the difference of the almost-invariant energies of the associated modulation functions $(z_j^{\mathbf{k}})$ *and* $(\widetilde{z}_j^{\mathbf{k}})$ *is bounded by*

$$\Big|\mathcal{E}_\ell\big(\mathbf{z}(\tau), \frac{d\mathbf{z}}{d\tau}(\tau)\big) - \mathcal{E}_\ell\big(\widetilde{\mathbf{z}}(\tau), \frac{d\widetilde{\mathbf{z}}}{d\tau}(\tau)\big)\Big| \le C\,\vartheta\,\delta^\ell N^{-1} \qquad \textit{for} \quad \ell = 1,\dots,K$$

$$\sum_{\ell=K}^{N} \sigma_\ell \Big|\mathcal{E}_\ell\big(\mathbf{z}(\tau), \frac{d\mathbf{z}}{d\tau}(\tau)\big) - \mathcal{E}_\ell\big(\widetilde{\mathbf{z}}(\tau), \frac{d\widetilde{\mathbf{z}}}{d\tau}(\tau)\big)\Big| \le C\,\vartheta\,\delta^K N^{-1}$$

for $\tau \le 1$*, with a constant* C *that is independent of* E *and* N*.*

Proof. We follow the lines of the proof of (52), taking differences in the recursions instead of direct bounds. Omitting the details, we obtain

$$\sum_{j=-N}^{N-1} \sigma_j \Big(\sum_{\mathbf{k}\in\mathcal{K}_j} |||z_j^{\mathbf{k}} - \widetilde{z}_j^{\mathbf{k}}|||_\tau \Big)^2 \le C\,\vartheta^2$$

for $\tau \le 1$, with a constant C that is independent of E and N. Together with the definition of $\mathcal{E}_\ell$ and the bounds (52), this yields the result. □

8.7. From short to long time intervals. By the estimates of the modulation functions we have for the almost-invariants at the initial time the estimates (78), where $\mathcal{C}_0$ can be chosen to depend only on the constant C_0 of (26). We apply Theorem 3 repeatedly on intervals of length 1. As long as the solution $\mathbf{u}(t)$ of (3) satisfies the smallness condition (26) with a larger constant $\widehat{C}_0$ in place of C_0, Theorem 2 gives a modulated Fourier expansion corresponding to starting values $(\mathbf{u}(t_n), \dot{\mathbf{u}}(t_n))$ at $t_n = n$. We denote the sequence of modulation functions of this expansion by $\mathbf{z}_n(\tau)$. The estimate (28) of Theorem 2 for $t = 1$ allows us to apply Lemma 11 with $\vartheta \le \widehat{C}_1\delta^{K+M+1}N^{-4}$ (with $\widehat{C}_1$ depending on $\widehat{C}_0$) to obtain, for $\tau = N^{-3}$,

$$\Big|\mathcal{E}_\ell\big(\mathbf{z}_n(\tau), \frac{d\mathbf{z}_n}{d\tau}(\tau)\big) - \mathcal{E}_\ell\big(\mathbf{z}_{n+1}(0), \frac{d\mathbf{z}_{n+1}}{d\tau}(0)\big)\Big| \le \widehat{C}\,\delta^{\ell+K+M+1}N^{-5} \qquad \text{for} \quad \ell = 1,\dots,K,$$

$$\sum_{\ell=K}^{N} \sigma_\ell \Big|\mathcal{E}_\ell\big(\mathbf{z}_n(\tau), \frac{d\mathbf{z}_n}{d\tau}(\tau)\big) - \mathcal{E}_\ell\big(\mathbf{z}_{n+1}(0), \frac{d\mathbf{z}_{n+1}}{d\tau}(0)\big)\Big| \le \widehat{C}\,\delta^{2K+M+1}N^{-5}$$

with $\widehat{C}$ depending on $\widehat{C}_0$. Theorem 3 now yields the same estimates with $\tau = 0$, possibly with a different constant $\widehat{C}$ depending on $\widehat{C}_0$. Summing up we obtain

$$\Big|\mathcal{E}_\ell\big(\mathbf{z}_n(0), \frac{d\mathbf{z}_n}{d\tau}(0)\big) - \mathcal{E}_\ell\big(\mathbf{z}_0(0), \frac{d\mathbf{z}_0}{d\tau}(0)\big)\Big| \le \widehat{C}\,\delta^{\ell+K+M+1}N^{-5}t_n \qquad \text{for} \quad \ell = 1,\dots,K,$$

$$\sum_{\ell=K}^{N} \sigma_\ell \Big|\mathcal{E}_\ell\big(\mathbf{z}_n(0), \frac{d\mathbf{z}_n}{d\tau}(0)\big) - \mathcal{E}_\ell\big(\mathbf{z}_0(0), \frac{d\mathbf{z}_0}{d\tau}(0)\big)\Big| \le \widehat{C}\,\delta^{2K+M+1}N^{-5}t_n,$$

and the same estimates hold when the argument 0 of $\mathbf{z}_n$ is replaced by $\tau \le N^{-3}$. Again $\widehat{C}$ may be different and depends on $\widehat{C}_0$. For $t_n \le c_0\, N^2\delta^{-M-1}$ with $c_0 = \mathcal{C}_0/\widehat{C}$, the first expression is smaller than $\mathcal{C}_0\delta^{\ell+K}N^{-3}$, and the second one is smaller than $\mathcal{C}_0\delta^{2K}N^{-3}$. Hence we obtain, for $n \le c_0\, N^2\delta^{-M-1}$ and $\tau \le N^{-3}$,

$$\Big|\mathcal{E}_\ell\big(\mathbf{z}_n(\tau), \tfrac{d\mathbf{z}_n}{d\tau}(\tau)\big)\Big| \le 2\,\mathcal{C}_0\,\delta^{2\ell}N^{-3} \quad \text{for} \quad \ell = 1,\dots,K,$$
$$\sum_{\ell=K}^{N} \sigma_\ell \Big|\mathcal{E}_\ell\big(\mathbf{z}_n(\tau), \tfrac{d\mathbf{z}_n}{d\tau}(\tau)\big)\Big| \le 2\,\mathcal{C}_0\,\delta^{2K}N^{-3}.$$

By Theorem 5 we therefore obtain, for $t \le c_0\, N^2\delta^{-M-1}$,

$$E_j(t) \le (\mathcal{C} + c\delta^2)\,\delta^{2j}N^{-3} \le 2\,\mathcal{C}\,\delta^{2j}N^{-3} \quad \text{for} \quad \ell = 1,\dots,K,$$
$$\sum_{j=K}^{N} \sigma_j\, E_j(t) \le (\mathcal{C} + c\delta^2)\,\delta^{2K}N^{-3} \le 2\,\mathcal{C}\,\delta^{2K}N^{-3},$$

where $\mathcal{C}$ only depends on $\mathcal{C}_0$ and hence on C_0, but not on $\widehat{C}_0$. Provided that $\widehat{C}_0$ has been chosen such that $\widehat{C}_0 \ge 2\,\mathcal{C}$, we see that the solution satisfies the smallness condition (26) up to times $t \le c_0\, N^2\delta^{-M-1}$, so that the construction of the modulated Fourier expansions on each of the subintervals of length 1 is indeed feasible with bounds that hold uniformly in n. The proof of Theorem 1 is thus complete.

9. Including the first near-resonance

The non-resonance estimates of Section 5 are crucial for the construction of the modulated Fourier expansion. The restriction $\max(|j|, \mu(\mathbf{k})) \le 10$ in Lemma 1 leads to the rather severe restriction $K + M \le 10$ in Theorem 1, which together with the condition $K < M$ yields $M \le 4$ and hence limits the result to a time scale $t \le N^2\delta^{-5}$. We discuss the case $K = 6$, $M = 5$ (so that $K + M = 11$) in order to stretch the time interval by a further factor δ^{-1}.

A difficulty now arises in the construction of $z_j^{\mathbf{k}}$ with $j = 5$ and the multi-index $\mathbf{k} = (-3, 0, 0, 2, 0, 0, \dots, 0) \in \mathcal{K}_j$, because here $\omega_j - \mathbf{k}\cdot\boldsymbol{\omega} = \omega_5 - 2\omega_4 + 3\omega_1 = \mathcal{O}(N^{-5})$; see Lemma 1. We introduce the lowest-order resonance module

$$\mathcal{M} = \{ n \cdot (-3, 0, 0, 2, -1, 0, \dots, 0) \,|\, n \in \mathbb{Z} \},$$

which is included in the larger resonance module

$$\mathcal{M} \subset \bigg\{ \mathbf{k} \in \mathbb{Z}^N \;\bigg|\; \sum_{\ell=1}^{N} \ell\, k_\ell = 0,\; \sum_{\ell=1}^{N} \ell^3\, k_\ell = 0 \bigg\}.$$

We remove from the modulated Fourier expansion (22) pairs $(j, \mathbf{k})$ for which $\mathbf{k} - \langle j \rangle \in \mathcal{M}$ or $\mathbf{k} + \langle j \rangle \in \mathcal{M}$. To keep the defect small, we have to modify the definition of the diagonal coefficient functions $z_j^{\pm\langle j \rangle}$.

We consider the pair $(j, \mathbf{k})$ with $j = 5$ and $\mathbf{k} = (-3, 0, 0, 2, 0, \dots)$, for which $\mathbf{k} - \langle j \rangle \in \mathcal{M}$. The use of (33) would lead to a small denominator of size $\mathcal{O}(N^{-6})$, which then makes the construction of the modulated Fourier expansion impossible. Instead of using (33) we set $z_j^{\mathbf{k}} = 0$ and include the terms $z_{j_1}^{\mathbf{k}^1} z_{j_2}^{\mathbf{k}^2}$ with $j_1 + j_2 = j = 5$ and $\mathbf{k}^1 + \mathbf{k}^2 = \mathbf{k} = (-3, 0, 0, 2, 0, \ldots)$ in the defining formula (32) for $z_j^{\langle j\rangle}$. This leads to

$$2\mathrm{i}\,\omega_5\, N^{-3} \frac{dz_5^{\langle 5\rangle}}{d\tau} + N^{-6}\, \frac{d^2 z_5^{\langle 5\rangle}}{d\tau^2} = -\mathrm{i}\,\omega_5^2\, \mathrm{e}^{\mathrm{i}\alpha\tau} z_3^{\langle 4\rangle - \langle 1\rangle} z_2^{\langle 4\rangle - 2\langle 1\rangle} + \ldots$$

with $\alpha = N^3(2\omega_4 - 3\omega_1 - \omega_5)$, which is small of size $\mathcal{O}(N^{-2})$. Since, by construction, each of the coefficient functions $z_{j_i}^{\mathbf{k}_i}(\tau)$ contains a factor N^{-2}, we are concerned with a differential equation of the form

$$\frac{dz}{d\tau} + \beta\, \frac{d^2 z}{d\tau^2} = p(\tau)\, \mathrm{e}^{\mathrm{i}\alpha\tau}, \qquad z(0) = z_0, \tag{81}$$

where α and β are of size $\mathcal{O}(N^{-2})$, and $p(\tau)$ is a polynomial with coefficients that are uniformly bounded in N. For large N, this differential equation is of singular perturbation type, and the general solution will have oscillations with a high frequency of size $\mathcal{O}(N^2)$. We are interested in a particular solution that does not have such high oscillations. The ansatz $\frac{dz}{d\tau}(\tau) = A(\tau)\, \mathrm{e}^{\mathrm{i}\alpha\tau}$ transforms (81) into a linear differential equation for $A(\tau)$ with polynomial right-hand side, for which a polynomial solution can be found as in (39). Integration of $A(\tau)\, \mathrm{e}^{\mathrm{i}\alpha\tau}$ then yields a smooth solution of (81) of the form $z_0 + C(\tau)\, \mathrm{e}^{\mathrm{i}\alpha\tau} - C(0)$, where $C(\tau)$ is a polynomial with uniformly bounded coefficients.

The only difference to the computations of Section 6 is that the coefficient functions $z_j^{\mathbf{k}}(\tau)$ are no longer polynomials, but linear combinations of polynomials multiplied with exponentials $\mathrm{e}^{\mathrm{i}\alpha\tau}$, where $\alpha = \mathcal{O}(N^{-2})$. Such functions also satisfy $|||z|||_\tau \le C$ for $\tau \le 1$.

In the following we write $\mathbf{h} \sim \mathbf{k}$ if $\mathbf{h} - \mathbf{k} \in \mathcal{M}$. For a given multi-index $\mathbf{k}$ we collect all modulation functions $z_j^{\mathbf{h}}(\tau)$ with $\mathbf{h} \sim \mathbf{k}$ and instead of (69) we consider the functions

$$\omega_j y_j^{\mathbf{k}}(t) = \sum_{\mathbf{h}\sim\mathbf{k}} \mathrm{e}^{\mathrm{i}(\mathbf{h}\cdot\boldsymbol{\omega})t} z_j^{\mathbf{h}}(\tau)$$

with $\tau = N^{-3}t$, so that $y_j^{\mathbf{k}}$ depends only on the equivalence class $[\mathbf{k}] = \mathbf{k} + \mathcal{M}$ of $\mathbf{k}$. The construction of the modulation functions $z_j^{\mathbf{k}}$ has been modified in such a way that in the modulation equations (70) for the above-defined $y_j^{\mathbf{k}}$, the sum is now over equivalence classes of multi-indices $[\mathbf{k}^1]$, $[\mathbf{k}^2]$ with $\mathbf{k}^1 + \mathbf{k}^2 \sim \mathbf{k}$ instead of multi-indices $\mathbf{k}^1$, $\mathbf{k}^2$ with $\mathbf{k}^1 + \mathbf{k}^2 = \mathbf{k}$ (cf. also [8] for the construction of resonant modulated Fourier expansions). Consequently, in the extended potential $\mathcal{U}$ of (71) the sum is now over $\mathbf{k}^1 + \mathbf{k}^2 + \mathbf{k}^3 \sim \mathbf{0}$. Therefore, the group invariance property (74) is no longer true for all $\boldsymbol{\lambda} \in \mathbb{R}^N$, but only for $\boldsymbol{\lambda} \perp \mathcal{M}$, i.e., if $3\lambda_1 - 2\lambda_4 + \lambda_5 = 0$. Only then $\mathcal{I}_{\boldsymbol{\lambda}}$ is an almost-invariant. This does not affect the almost-invariant energies $\mathcal{E}_j$ with $j \ne 1, 4, 5$, but instead of $\mathcal{E}_1, \mathcal{E}_4, \mathcal{E}_5$ we have only two independent almost-invariants

$$2\,\mathcal{I}_1 + 3\,\mathcal{I}_4 \quad \text{and} \quad \mathcal{I}_4 + 2\,\mathcal{I}_5$$

(where $\mathcal{I}_j = \mathcal{E}_j/\omega_j$ is the corresponding action). Though Theorem 5 is still valid in this modified setting, it does not bound the normal mode energies E_j

any longer in terms of almost-invariants. A way out is to bound the critical $\mathcal{E}_1, \mathcal{E}_4, \mathcal{E}_5$ in terms of the remaining almost-invariants, using that $\omega_\ell \mathcal{I}_\ell = \mathcal{E}_\ell \geq 0$ by Theorem 4 (which remains valid in the present modified setting):

$$\begin{aligned}
\mathcal{E}_1 &\leq N\,\omega_1(2\mathcal{I}_1 + 3\mathcal{I}_4)\\
\mathcal{E}_4 &\leq 2N\,\omega_4(\mathcal{I}_4 + 2\mathcal{I}_5)\\
\mathcal{E}_5 &\leq N\,\omega_5(\mathcal{I}_4 + 2\mathcal{I}_5).
\end{aligned}$$

However, we then obtain essentially the same bound for $\mathcal{E}_4$ and $\mathcal{E}_5$ and consequently for E_4 and E_5. These considerations lead to the following extension of Theorem 1 to the case $K = 6$, $M = 5$.

Theorem 7. *Fix $\gamma > 0$ and $\rho \geq 1$. Then, there exist $\delta_0 > 0$ and $c > 0$, $C > 0$ such that the following holds: if*

(1) the dimension of the system satisfies $N \geq 32$ and $N \in \mathcal{N}(5, \gamma)$,
(2) the total energy E is bounded such that $\delta := \sqrt{E\,N^3} \leq \delta_0$,
(3) the initial normal mode energies satisfy $E_j(0) = 0$ for $j \neq \pm 1$,

then, over long times

$$t \leq c\,N^2\,\delta^{-6},$$

the normal mode energies satisfy the estimates

$$|E_1(t) - E_1(0)| \leq C\,E\,\delta^2, \tag{82}$$

$$E_j(t) \leq C\,E\,\delta^{2(j-1)}, \qquad j = 1, \ldots, 4, \tag{83}$$

$$E_5(t) \leq C\,E\,\delta^6, \tag{84}$$

$$\sum_{j=6}^{N} \rho^{2j} E_j(t) \leq C\,E\,\delta^8. \tag{85}$$

The proof uses a variant of Theorem 2 where $K = 6$, $M = 5$, and the initial energies in the modes satisfy (83)–(85). The partly resonant modulated Fourier expansion is constructed as described above. One verifies that Theorems 4 and 5 remain valid also in this modified situation, and an analogue of Theorem 3 holds for the remaining almost-invariants.

The inclusion of further near-resonances to arrive at even longer time-scales appears feasible in principle, but is beyond the scope of this paper.

10. Quartic and higher-order potentials

We consider the Fermi–Pasta–Ulam β-model, i.e., the lattice (1) with potential $V(x) = x^4/4$ (this corresponds to $\alpha = 0$ and $\beta = 1$ in the problem of Section 2). For the discrete Fourier coefficients we obtain the system

$$\ddot{u}_j + \omega_j^2\,u_j \;=\; \omega_j \sum_{j_1+j_2+j_3=j \bmod 2N} (-1)^{(j_1+j_2+j_3-j)/(2N)}\,\omega_{j_1}\,\omega_{j_2}\,\omega_{j_3}\;u_{j_1}\,u_{j_2}\,u_{j_3}$$

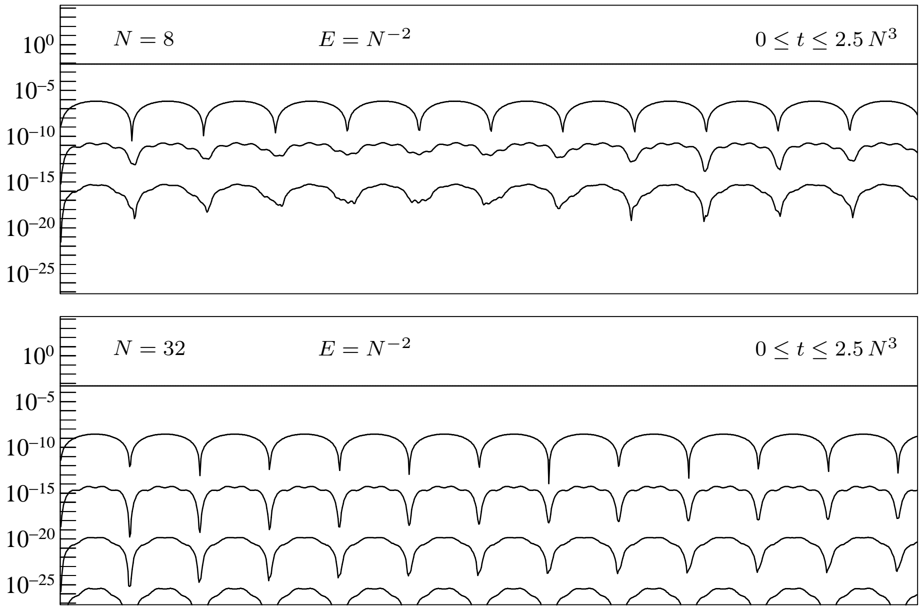

Fig. 5. Normal mode energies $E_j(t)$ as functions of time for the FPU β-problem with initial values (11); increasing j corresponds to decreasing values of $E_j(t)$. For even j we have $E_j(t) = 0$.

with frequencies ω_j as in Section 2. This is a complex Hamiltonian system with potential

$$U(\mathbf{u}) \;=\; \frac{1}{4} \sum_{j_1+j_2+j_3+j_4=0 \bmod 2N} (-1)^{(j_1+j_2+j_3+j_4)/(2N)}\, \omega_{j_1}\, \omega_{j_2}\, \omega_{j_3}\, \omega_{j_4}\, u_{j_1}\, u_{j_2}\, u_{j_3}\, u_{j_4}.$$

The normal mode energies are defined as in Section 2.

We performed numerical experiments with the FPU β-model. Figure 5 shows the mode energies $E_j(t)$ as functions of time for $E = N^{-2}$. It is observed that time intervals proportional to N^3 are again a natural time scale for the slow motion of the energies. Further numerical experiments indicate that the mode energies behave like

$$E_{2j-1}(t) \approx E\, \delta^{2(j-1)} f_j(N^{-3}t) \quad \text{for } j \ge 2, \quad \text{with } \; \delta = EN. \tag{86}$$

Theorem 8. *Consider the potential $V(x) = x^{q+1}$ with an integer $q \ge 2$, which gives a nonlinearity of degree q in (3). Fix $\gamma > 0$, $\rho \ge 1$, and let M and K be positive integers satisfying $M < K$ and $K + M = 10$. Then, there exist $\delta_0 > 0$ and $c > 0$, $C > 0$ such that the following holds: if*

(1) the dimension of the system satisfies $N \ge 41$, $N \in \mathcal{N}(M, \gamma)$, and N is an integral multiple of $q - 1$,

(2) the total energy E is bounded such that $\delta := \sqrt{E^{q-1}\, N^{5-q}} \le \delta_0$,

(3) the initial normal mode energies satisfy $E_j(0) = 0$ for $j \ne \pm 1$,

then, over long times

$$t \le c\, N^2\, \delta^{-M-1},$$

the normal mode energies satisfy the estimates

$$|E_1(t) - E_1(0)| \le C\, E\, \delta^2, \tag{87}$$

$$E_{1+(j-1)(q-1)}(t) \le C\, E\, \delta^{2(j-1)}, \qquad j = 1, \dots, K, \tag{88}$$

$$\sum_{j=K}^{N} \rho^{2j} E_{1+(j-1)(q-1)}(t) \le C\, E\, \delta^{2(K-1)}. \tag{89}$$

The other mode energies are identically zero.

To our knowledge, this is the first rigorous result on the long-time behaviour of the mode energies in the FPU β-model for large particle numbers N. The proof by modulated Fourier expansions is very similar to that of Theorem 1 and is therefore not presented in detail. We have an analogue of Theorem 2, now with $\delta = \sqrt{E^{q-1}\, N^{5-q}}$, and valid on time intervals $t \le \min(N^3, \epsilon^{-(q-1)/2})$ with $\epsilon = E/(2N)$. Theorems 3 to 6 remain valid with subscript ℓ replaced by $1 + (\ell - 1)(q - 1)$.

The proof again gives explicit formulas for the dominant terms for $\delta \ll 1$:

$$E_{1+(j-1)(q-1)}(t) \; = 2\, E\, \delta^{2(j-1)} C\big(|a_j(t)|^2 + \mathcal{O}(\delta^2 + N^{-2})\big),$$

where $C = |c_-|^{2+2(j-1)(q-1)} + |c_+|^{2+2(j-1)(q-1)}$ with $c_\pm = \frac{1}{\sqrt{2\epsilon}}\big(\omega_1 u_1(0) \pm \mathrm{i}\, \dot u_1(0)\big)$ as in (19), and the first functions $a_j(t)$ are given by

$$a_1(t) = \frac{1}{2}, \qquad a_2(t) = \frac{3}{\pi^2\, 2^{q-2}(q+1)(q-1)} \Big(\mathrm{e}^{\mathrm{i}(q\omega_1 - \omega_q)t} - 1\Big).$$

All the linear combinations of frequencies appearing in these formulas are of size $\mathcal{O}(N^{-3})$. In particular, we have

$$q\omega_1 - \omega_q = \frac{(q+1)q(q-1)\,\pi^3}{24\, N^3} + \mathcal{O}(N^{-5}).$$

For the β-model ($q = 3$), the period of $a_2(t)$ is approximately 4 times smaller than that for the α-model ($q = 2$). In Figure 5 we have therefore chosen an interval that is 4 times smaller than in Figure 2, so that for large N the same number of periods is covered.

The decay of the mode energies with powers of $\delta = EN \ll 1$ in the β-model was observed numerically by Flach, Ivanchenko and Kanakov [12]. The mode energy values $E_{2j-1} \approx E(3\delta/8\pi^2)^{2(j-1)}$ stated in [12] are qualitatively similar, though not identical to the expressions obtained with our derivation.

11. Conclusion and perspectives

We have given a rigorous analysis of the FPU α-model for large particle numbers N with initial excitation in the first mode and total energy E distinctly smaller than the inverse of the third power of the particle number, that is, $\delta = \sqrt{EN^3} \ll 1$. Our results show the presence of several time-scales in the problem: On the fast time scale $t \sim N$, there is almost-harmonic oscillation. On the time scale $t \sim N^3$, there is energy flow into the higher modes and the *formation of a packet*: the mode energies have bounds that decay geometrically with the mode number (with a rate proportional to δ at least for the first modes) and they behave quasi-periodically in time. These energy bounds persevere over a longer time scale $t \sim N^2\delta^{-5}$ (*metastability*). Analogous results are obtained for the FPU β-model (for $EN \ll 1$) and for higher-order nonlinearities.

It is instructive to compare and contrast our results and techniques with those of Bambusi and Ponno [1]. That paper considers the scaling $\delta \sim 1$ as opposed to our $\delta \ll 1$. It obtains results on metastability for time scales $t \sim N^3$. The proofs in [1] use a *resonant* normal form, which turns out to be integrable as a pair of uncoupled KdV equations for the FPU α-model. Integrability is lost, however, in the resonant normal forms of higher-order nonlinearities [2]. In contrast, we work with a *non-resonant* modulated Fourier expansion in time, for which the associated modulation system possesses a full set of almost-invariant energies that bound the mode energies from above. Integrability plays no role here. Both [1] and the present paper work far from the thermodynamic limit $E/N \sim 1$ as $N \to \infty$.

The techniques developed in this paper are not restricted to the particular case of the FPU α- or β-model with excitation of only the first mode. They allow for extensions to:

- other single excited modes (of low or high frequency), several excited modes, packets of excited modes
- other potentials, other nonlinearities (that are combinations of convolutions and multipliers in mode coordinates)
- multidimensional lattices
- Hamiltonian partial differential equations such as nonlinear wave equations and Schrödinger equations
- numerical discretizations.

The availability of (weak) non-resonance estimates for the frequencies is a key issue, which influences the actual time scales and has to be studied from case to case. In each case, however, there appear the phenomena of *formation of a packet* of modes with geometrically decaying energies and *metastability* of the packet over longer time scales.

For numerical experiments relating to some of the above issues in the FPU problem we refer to [4,5,19] and further papers cited therein. Formation of a packet and metastability in Hamiltonian partial differential equations have been studied, analytically using modulated Fourier expansions and also numerically, in the doctoral thesis of Gauckler [15].

Acknowledgements. We thank Giancarlo Benettin and Antonio Ponno for interesting discussions on the FPU problem and for helpful hints to the FPU literature, and Daniel Weiss and Ludwig Gauckler for helpful comments on an earlier version. This work has been supported by Fonds National Suisse, Project No. 200020-126638, and by DFG, Project LU 532/4-1.

References

1. D. Bambusi and A. Ponno, *On metastability in FPU*, Comm. Math. Phys. **264** (2006), 539–561.
2. D. Bambusi and A. Ponno, *Resonance, Metastability and Blow up in FPU*, Lecture Notes in Physics **728** (2008), 191.
3. G. Benettin, A. Carati, L. Galgani, and A. Giorgilli, *The Fermi-Pasta-Ulam problem and the metastability perspective*, The Fermi-Pasta-Ulam problem, Lecture Notes in Phys., vol. 728, Springer, Berlin, 2008, pp. 152–189.
4. G. Benettin and G. Gradenigo, *A study of the Fermi–Pasta–Ulam problem in dimension two*, Chaos: An Interdisciplinary Journal of Nonlinear Science **18** (2008), 013112.
5. G. Benettin, R. Livi, and A. Ponno, *The Fermi-Pasta-Ulam Problem: Scaling Laws vs. Initial Conditions*, Journal of Statistical Physics **135** (2009), no. 5, 873–893.
6. L. Berchialla, L. Galgani, and A. Giorgilli, *Localization of energy in FPU chains*, Discrete and Continuous Dynamical Systems (DCDS-A) **11** (2004), no. 4, 855–866.
7. G.P. Berman and F.M. Izrailev, *The Fermi–Pasta–Ulam problem: fifty years of progress*, Chaos: An Interdisciplinary Journal of Nonlinear Science **15** (2005), 015104.
8. D. Cohen, E. Hairer, and C. Lubich, *Numerical energy conservation for multi-frequency oscillatory differential equations*, BIT **45** (2005), 287–305.
9. ———, *Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions*, Arch. Ration. Mech. Anal. **187** (2008), 341–368.
10. J. DeLuca, A. J. Lichtenberg, and S. Ruffo, *Energy transitions and time scales to equipartition in the Fermi-Pasta-Ulam oscillator chain*, Phys. Rev. E **51** (1995), no. 4, 2877–2885.
11. E. Fermi, J. Pasta, and S. Ulam, *Studies of non linear problems*, Tech. Report LA-1940, Los Alamos, 1955, Later published in E. Fermi: Collected Papers Chicago 1965 and Lect. Appl. Math. 15 143 1974.
12. S. Flach, M. V. Ivanchenko, and O. I. Kanakov, *q-breathers in Fermi-Pasta-Ulam chains: Existence, localization, and stability*, Phys. Rev. E **73** (2006), no. 036618.
13. J. Ford, *The Fermi-Pasta-Ulam problem: paradox turns discovery*, Physics Reports **213** (1992), 271–310.
14. G. Gallavotti (ed.), *The Fermi-Pasta-Ulam problem*, Lecture Notes in Physics, vol. 728, Springer, Berlin, 2008, A status report.
15. L. Gauckler, *Long-time analysis of Hamiltonian partial differential equations and their discretizations*, Ph.D. thesis, Univ. Tübingen, 2010.
16. L. Gauckler and C. Lubich, *Nonlinear Schrödinger equations and their spectral semi-discretizations over long times*, Found. Comput. Math. **10** (2010), 141–169.
17. E. Hairer and C. Lubich, *Long-time energy conservation of numerical methods for oscillatory differential equations*, SIAM J. Numer. Anal. **38** (2001), 414–441.
18. E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations*, 2nd ed., Springer Series in Computational Mathematics 31, Springer-Verlag, Berlin, 2006.
19. S. Paleari and T. Penati, *Numerical methods and results in the FPU problem*, The Fermi-Pasta-Ulam problem, Lecture Notes in Phys., vol. 728, Springer, Berlin, 2008, pp. 239–282.
20. T.P. Weissert, *The Genesis of Simulation in Dynamics: Pursuing the Fermi–Pasta–Ulam Problem*, Springer-Verlag, New York, 1997.