

# Energy separation in oscillatory Hamiltonian systems without any non-resonance condition

Ludwig Gauckler<sup>1</sup>, Ernst Hairer<sup>2</sup>, Christian Lubich<sup>3</sup>

<sup>1</sup> Institut für Mathematik, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany. E-mail: gauckler@math.tu-berlin.de

<sup>2</sup> Section de mathématiques, 2-4 rue du Lièvre, Université de Genève, CH-1211 Genève 4, Switzerland. E-mail: Ernst.Hairer@unige.ch

<sup>3</sup> Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle, D-72076 Tübingen, Germany. E-mail: Lubich@na.uni-tuebingen.de

**Abstract:** We consider multiscale Hamiltonian systems in which harmonic oscillators with several high frequencies are coupled to a slow system. It is shown that the oscillatory energy is nearly preserved over long times  $\varepsilon^{-N}$  for arbitrary  $N > 1$ , where  $\varepsilon^{-1}$  is the size of the smallest high frequency. The result is uniform in the frequencies and does not require non-resonance conditions.

## 1. Introduction and statement of the main result

As has been discussed, e.g., in [2–5,15], the problem of slow rates of thermalization in statistical mechanics can be understood by studying multiscale problems of the following type: For momenta  $\mathbf{p} = (p_0, p_1, \dots, p_n)$  and positions  $\mathbf{q} = (q_0, q_1, \dots, q_n)$  with  $p_j, q_j \in \mathbb{R}^{d_j}$  consider a Hamiltonian that couples high-frequency harmonic oscillators with a Hamiltonian of slow motion,

$$H(\mathbf{p}, \mathbf{q}) = H_\omega(\mathbf{p}, \mathbf{q}) + H_{\text{slow}}(\mathbf{p}, \mathbf{q}), \quad (1)$$

where the oscillatory and slow-motion energies are given by

$$H_\omega(\mathbf{p}, \mathbf{q}) = \sum_{j=1}^n \frac{1}{2} (|p_j|^2 + \omega_j^2 |q_j|^2), \quad H_{\text{slow}}(\mathbf{p}, \mathbf{q}) = \frac{1}{2} |p_0|^2 + U(\mathbf{q}) \quad (2)$$

with high angular frequencies

$$\omega_j \geq \frac{1}{\varepsilon}, \quad 0 < \varepsilon \ll 1. \quad (3)$$

The coupling potential  $U(\mathbf{q})$  is assumed smooth with derivatives bounded independently of the small parameter  $\varepsilon$ .

Weakly anharmonic systems similar to (1)–(3) have been studied in numerous publications, from different viewpoints and with different mathematical techniques, ever since the seminal paper by Fermi, Pasta & Ulam ([9], reprinted in [10]). For some recent work, see [1, 8, 10, 12, 14] and references therein. The system (1)–(3) has been considered in [4] with the double motivation of explaining slow rates of thermalization and the realization of holonomic constraints.

The equations of motion are

$$\ddot{q}_j + \omega_j^2 q_j = -\nabla_j U(\mathbf{q}), \quad j = 0, \dots, n, \quad (4)$$

where  $\nabla_j$  denotes the partial derivative with respect to  $q_j$  and  $\omega_0 = 0$ . We are interested in the energy exchange (or lack thereof) between the fast and the slow system, where in contrast to previous work we assume no condition on the frequencies other than the lower bound (3).

**Theorem 1.** *Fix an energy bound  $E > 0$ , an integer  $N \geq 1$ , and further a radius  $\rho > 0$  and a set  $K \subset \mathbb{R}^{d_0}$  such that the potential  $U$  has bounded derivatives of all orders in a  $\rho$ -neighbourhood of  $K \times 0 \times \dots \times 0$ . Then there exist  $C > 0$  and  $\varepsilon^* > 0$  such that the following holds for  $0 < \varepsilon \leq \varepsilon^*$ : Whenever the frequencies satisfy  $\omega_j \geq \varepsilon^{-1}$  for  $j = 1, \dots, n$  and the initial values  $(\mathbf{p}(0), \mathbf{q}(0))$  are such that*

$$H_\omega(\mathbf{p}(0), \mathbf{q}(0)) \leq E, \quad (5)$$

*then, along the solution of (4) to these initial values, the oscillatory energy deviates from its starting value by no more than*

$$|H_\omega(\mathbf{p}(t), \mathbf{q}(t)) - H_\omega(\mathbf{p}(0), \mathbf{q}(0))| \leq C \varepsilon^{3/4} \quad \text{for } 0 \leq t \leq \varepsilon^{-N}, \quad (6)$$

*provided that  $q_0(t)$  stays in the set  $K$  for such long times. The threshold  $\varepsilon^*$  and the constant  $C$  depend on  $n$  and  $N$ , on the energy bound  $E$  and on bounds of derivatives of the potential  $U$ .*

We emphasize that no non-resonance condition is required in this result. The estimate is uniform in the frequencies  $\omega_j \geq \varepsilon^{-1}$ . This sets Theorem 1 apart from existing related results in Hamiltonian perturbation theory.

For fixed strongly non-resonant frequencies as well as for exactly resonant frequencies that satisfy a diophantine non-resonance condition outside a resonance module, the long-time estimate (6) is a particular case of the energy separation results shown by Benettin, Galgani and Giorgilli [4]. The proof of Theorem 1 is based on modulated Fourier expansions with suitably modified frequencies and does not use the canonical transformation techniques of Hamiltonian perturbation theory on which the proofs in [4] are based. The idea of modifying frequencies to avoid almost-resonant situations has also been used in [2] in a different context where all frequencies are sufficiently close to a single high frequency.

The results in [4] are proved over times that are exponentially long in negative fractional powers of  $\varepsilon$  in the case of a real-analytic potential. We expect that then also Theorem 1 holds over exponentially long times, but we will not investigate this question in this paper. See, however, [6] for a corresponding result over exponentially long times for a single high frequency proved via modulated Fourier expansions.

We remark that the order  $\varepsilon^{3/4}$  can be improved, with the same proof, to  $\varepsilon^{1-\delta}$  for arbitrary  $\delta > 0$ , with  $\varepsilon^*$  and  $C$  depending additionally on  $\delta$  and deteriorating as  $\delta \rightarrow 0$ .

The energy bound  $H_\omega(\mathbf{p}, \mathbf{q}) \leq \text{Const.}$  implies  $q_j = \mathcal{O}(\varepsilon)$  for  $j = 1, \dots, n$ . Theorem 1 and conservation of the total energy therefore yield that also the energy  $\frac{1}{2}|p_0|^2 + U(q_0, 0, \dots, 0)$  of the isolated slow system is nearly preserved over time  $\varepsilon^{-N}$  along solutions of (4).

We present a numerical example in the next section and prove Theorem 1 in the remainder of the paper.

## 2. Numerical illustration

We consider the problem with  $n = 7$  frequencies

$$\varepsilon\omega_j = 1, 1 + \varepsilon^2, 1 + \varepsilon, 1 + \varepsilon^{3/4}, 1 + \varepsilon^{2/3}, 1 + \varepsilon^{1/2}, 2$$

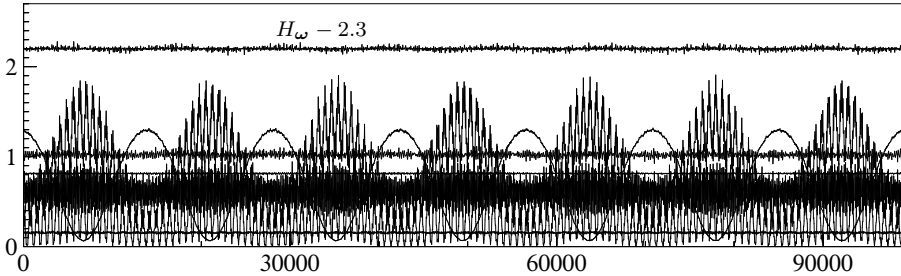
for  $j = 1, \dots, n$ , and with the potential

$$U(\mathbf{q}) = \frac{1}{2}q_0^2 + (aq_0 + q_1 + q_2 + 2q_3 + 3q_4 + q_5 + q_6 + 3q_7)^3$$

depending on a parameter  $a$ . The initial values are chosen as

$$\begin{aligned} \mathbf{q}(0) &= (1, 0.3\varepsilon, 0.4\varepsilon, 0.7\varepsilon, -1.1\varepsilon, 0.4\varepsilon, -0.6\varepsilon, -0.7\varepsilon) \\ \mathbf{p}(0) &= (-0.2, 0.6, 0.7, -0.9, -0.9, 0.4, -1.1, 0.8). \end{aligned}$$

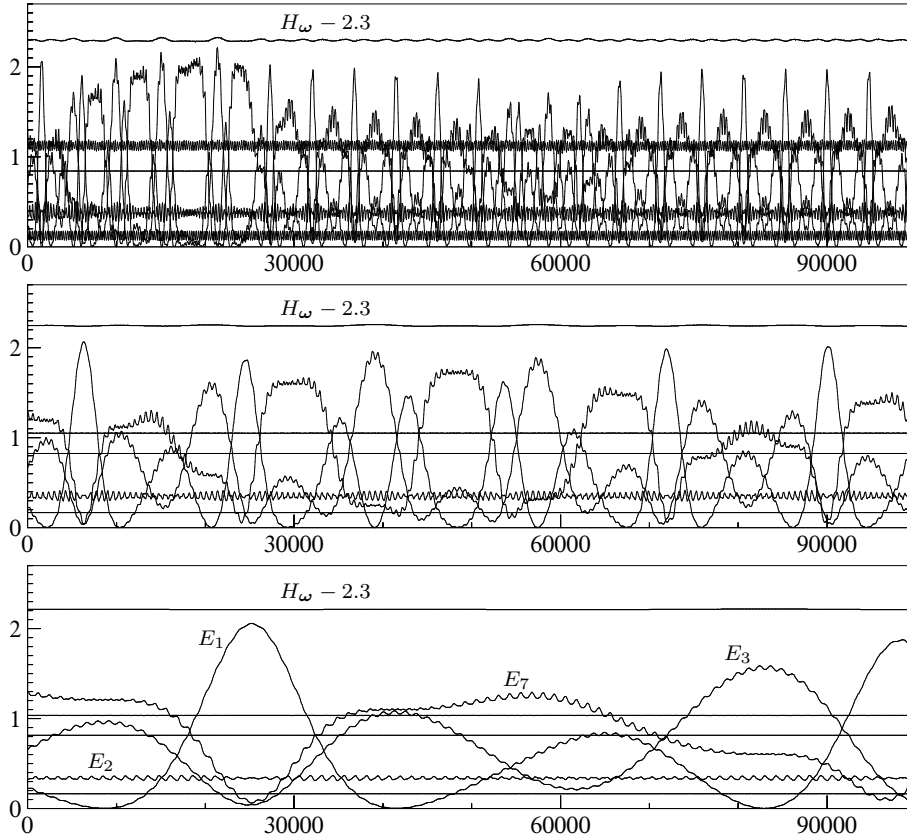
The numerical solution is computed with a trigonometric integrator (symplectic method of Deuffhard, see [13, Chap. XIII]), which is applied with constant step size  $\Delta t = 0.01\varepsilon$ . Taking the step size twice as large gives nearly identical figures.



**Fig. 1.** Individual oscillatory energies and shifted total oscillatory energy with parameter  $a = 0.5$  and  $\varepsilon = 0.005$ .

We consider two different values for the parameter  $a$ . Figure 1 shows the individual oscillatory energies  $E_j = \frac{1}{2}(|p_j|^2 + \omega_j^2|q_j|^2)$  as well as the total oscillatory energy  $H_\omega$  on an interval of length 100 000 for  $a = 0.5$  (for  $a = 1$  the energies explode) and for  $\varepsilon = 0.005$ . The initial oscillatory energies are close to 0.22, 0.32, 0.65, 1.05, 0.17, 0.82, 1.3. We observe that the oscillatory energies for the modes  $j = 4, 5, 6$  remain nearly constant, whereas for the other modes, which are closer to a 1:1 and 1:2 resonance, there is a significant energy exchange. In

the figure the constant value 2.3 is subtracted from the total oscillatory energy, so that it fits better into the picture. We repeated the experiment with other values of  $\varepsilon$  and found that the maximal deviation (6) of the total oscillatory energy is  $4.56 \cdot 10^{-1}$  for  $\varepsilon = 0.02$ ,  $1.85 \cdot 10^{-1}$  for  $\varepsilon = 0.01$ , and  $9.54 \cdot 10^{-2}$  for  $\varepsilon = 0.005$ , which indicates an  $\mathcal{O}(\varepsilon)$ -behaviour.



**Fig. 2.** Individual oscillatory energies and total oscillatory energy with parameter  $a = \varepsilon$ . The upper picture is for  $\varepsilon = 0.02$ , the middle picture for  $\varepsilon = 0.01$ , and the lower picture for  $\varepsilon = 0.005$ .

Figure 2 shows the same experiment for the parameter  $a = \varepsilon$  and for three different values of  $\varepsilon$ . Here, the maximal deviation (6) of the total oscillatory energy is  $3.95 \cdot 10^{-2}$  for  $\varepsilon = 0.02$ ,  $1.41 \cdot 10^{-2}$  for  $\varepsilon = 0.01$ , and  $4.81 \cdot 10^{-3}$  for  $\varepsilon = 0.005$ .

### 3. Modulated Fourier expansion with modified frequencies

*3.1. Changing almost-resonant frequencies to exactly resonant frequencies.* We deal with almost-resonant frequencies by introducing modified frequencies close

to the original frequencies such that almost-resonances among the original frequencies become exact resonances among the modified frequencies. For a multi-index  $\mathbf{k} = (k_1, \dots, k_n) \in \mathbb{Z}^n$ , we denote  $\|\mathbf{k}\| = |k_1| + \dots + |k_n|$ . The vector of frequencies is  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)$ , and we write  $\mathbf{k} \cdot \boldsymbol{\omega} = \sum_{j=1}^n k_j \omega_j$ .

In order to decide whether a linear combination  $\mathbf{k} \cdot \boldsymbol{\omega}$  of frequencies is considered as almost-resonant or as non-resonant we use the fact that there is a *gap* in the linear combinations of frequencies.

**Lemma 1.** *There exist  $\mu > 0$  depending only on  $n$  and  $N$  and  $\alpha \in [\mu, \frac{1}{4}]$  depending on the frequencies such that there is no multi-index  $\mathbf{k} \in \mathbb{Z}^n$  with  $\|\mathbf{k}\| \leq N+1$  and  $\varepsilon^{-\alpha} \leq |\mathbf{k} \cdot \boldsymbol{\omega}| \leq \varepsilon^{-\alpha-\mu}$ .*

*Proof.* Let us denote the number of all multi-indices  $\mathbf{k} \in \mathbb{Z}^n$  with  $\|\mathbf{k}\| \leq N+1$  by  $M$ , which obviously depends only on  $n$  and  $N$ . Consider the logarithms to base  $\varepsilon^{-1}$  of the  $|\mathbf{k} \cdot \boldsymbol{\omega}|$  with  $\|\mathbf{k}\| \leq N+1$ , which gives a set  $\{a_1, \dots, a_M\}$ . With  $\mu = 1/(4M+4)$ , there is an interval of width  $\mu$  in  $[\mu, \frac{1}{4}]$  that does not contain any of the  $a_i$ . If we denote this interval by  $[\alpha, \alpha + \mu]$ , then there is no  $\mathbf{k} \in \mathbb{Z}^n$  with  $\|\mathbf{k}\| \leq N+1$  and  $\varepsilon^{-\alpha} \leq |\mathbf{k} \cdot \boldsymbol{\omega}| \leq \varepsilon^{-\alpha-\mu}$ .  $\square$

Let  $\alpha$  and  $\mu$  be as in the previous lemma. We consider multi-indices  $\mathbf{k} \in \mathbb{Z}^n$  with  $\|\mathbf{k}\| \leq N+1$  and  $|\mathbf{k} \cdot \boldsymbol{\omega}| \geq \varepsilon^{-\alpha-\mu}$  as non-resonant, whereas those with  $|\mathbf{k} \cdot \boldsymbol{\omega}| \leq \varepsilon^{-\alpha}$  are considered as almost-resonant and collected in the set

$$\mathcal{R} = \{\mathbf{k} \in \mathbb{Z}^n : |\mathbf{k} \cdot \boldsymbol{\omega}| \leq \varepsilon^{-\alpha}, \|\mathbf{k}\| \leq N+1\}.$$

This set generates the module  $\mathcal{M}$  that contains all finite linear combinations of multi-indices in  $\mathcal{R}$  with integer coefficients. We select a maximum number  $d \leq n$  of  $\mathbb{R}$ -linearly independent vectors  $\mathbf{k}^1, \dots, \mathbf{k}^d \in \mathcal{R}$ . Their real-linear span contains the module  $\mathcal{M}$ .

We modify the almost-resonant frequencies  $\omega_j$  to exactly resonant frequencies

$$\varpi_j = \omega_j + \vartheta_j, \quad j = 1, \dots, n,$$

by determining  $\boldsymbol{\vartheta} = (\vartheta_1, \dots, \vartheta_n)$  as a solution of minimal norm of the (possibly underdetermined) system of linear equations

$$\mathbf{k}^i \cdot \boldsymbol{\omega} + \mathbf{k}^i \cdot \boldsymbol{\vartheta} = 0, \quad i = 1, \dots, d.$$

We then have

$$\|\boldsymbol{\vartheta}\| \leq \gamma \varepsilon^{-\alpha}, \quad (7)$$

where  $\gamma$  depends only on  $n$  and  $N$ , and

$$\mathbf{k} \cdot \boldsymbol{\varpi} = 0 \quad \text{for all } \mathbf{k} \in \mathcal{M}, \quad (8)$$

i.e., almost-resonances in the original frequencies become exact resonances in the modified frequencies. We further have the lower bound

$$|\mathbf{k} \cdot \boldsymbol{\varpi}| \geq \frac{1}{2} \varepsilon^{-\alpha-\mu} \quad \text{for } \mathbf{k} \notin \mathcal{M}, \|\mathbf{k}\| \leq N+1, \quad (9)$$

provided that  $0 < \varepsilon \leq \varepsilon^*$ , where  $\varepsilon^*$  depends only on  $n$  and  $N$ , i.e., non-resonant multi-indices for the original frequencies are still non-resonant for the modified frequencies. It will be convenient to let  $\varpi_0 = 0$  and  $\vartheta_0 = 0$ .

*3.2. Short-time approximation by a modulated Fourier expansion.* We let  $\mathcal{N}$  be a set of representatives of the equivalence classes in  $\mathbb{Z}^n/\mathcal{M}$ , which are chosen such that for each  $\mathbf{k} \in \mathcal{N}$ , the norm of  $\mathbf{k}$  is minimal in the equivalence class  $[\mathbf{k}] = \mathbf{k} + \mathcal{M}$ , and with  $\mathbf{k} \in \mathcal{N}$ , also  $-\mathbf{k} \in \mathcal{N}$ . We denote  $\mathcal{K} = \{\mathbf{k} \in \mathcal{N} : \|\mathbf{k}\| \leq N\}$ .

We make the approximation ansatz (*modulated Fourier expansion*)

$$q_j(t) \approx \sum_{\mathbf{k} \in \mathcal{K}} z_j^{\mathbf{k}}(\varepsilon^{-\alpha}t) e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t}, \quad j = 0, 1, \dots, n. \quad (10)$$

The modulation functions  $z_j^{\mathbf{k}}(\tau)$  for  $\tau = \varepsilon^{-\alpha}t$  with  $\alpha$  from Lemma 1 are determined such that all their derivatives with respect to  $\tau$  are bounded independently of  $\varepsilon$  and that they yield a small defect  $\delta_j^{\mathbf{k}}$  in the modulation equations, which are derived by inserting the expansion (10) into the differential equation (4) and comparing the coefficients of  $e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t}$ :

$$\begin{aligned} (\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2) z_j^{\mathbf{k}} + 2i(\mathbf{k} \cdot \boldsymbol{\varpi}) \varepsilon^{-\alpha} \frac{dz_j^{\mathbf{k}}}{d\tau} + \varepsilon^{-2\alpha} \frac{d^2 z_j^{\mathbf{k}}}{d\tau^2} \\ = (2\varpi_j \vartheta_j - \vartheta_j^2) z_j^{\mathbf{k}} - \nabla_j^{-\mathbf{k}} \mathcal{U}(\mathbf{z}) + \delta_j^{\mathbf{k}}. \end{aligned} \quad (11)$$

Here  $\nabla_j^{-\mathbf{k}}$  denotes the gradient with respect to  $z_j^{-\mathbf{k}}$  and the modulation potential  $\mathcal{U}(\mathbf{z})$  for  $\mathbf{z} = (z_j^{\mathbf{k}})$  is given as

$$\mathcal{U}(\mathbf{z}) = U(\mathbf{z}_0^{\mathbf{0}}) + \sum_{m=1}^N \sum_{j_1, \dots, j_m=0}^n \sum_{\mathbf{k}^1 + \dots + \mathbf{k}^m \in \mathcal{M}} \frac{1}{m!} \partial_{j_1} \dots \partial_{j_m} U(\mathbf{z}_0^{\mathbf{0}})(z_{j_1}^{\mathbf{k}^1}, \dots, z_{j_m}^{\mathbf{k}^m}),$$

where  $\mathbf{z}_0^{\mathbf{0}} = (z_0^{\mathbf{0}}, 0, \dots, 0)$  and the indices  $(j_l, \mathbf{k}^l) = (0, \mathbf{0})$  are excluded from the sum. Note that  $z_j^{\mathbf{k}}$  are vectors in  $\mathbb{R}^{d_j}$  and correspondingly  $\partial_{j_1} \dots \partial_{j_m} U(\mathbf{z}_0^{\mathbf{0}})$  is a multilinear form on  $\mathbb{R}^{d_{j_1}} \times \dots \times \mathbb{R}^{d_{j_m}}$ ; in contrast  $\nabla_j U$  denotes the gradient, a column vector in  $\mathbb{R}^{d_j}$ . In the following we let  $\langle j \rangle = (0, \dots, 1, \dots, 0)$  denote the  $j$ th unit vector in  $\mathbb{Z}^n$ .

**Proposition 1.** *In the situation of Theorem 1, the solution  $q_j(t)$  ( $j = 0, \dots, n$ ) of (4) admits an expansion*

$$q_j(t) = \sum_{\mathbf{k} \in \mathcal{K}} z_j^{\mathbf{k}}(\varepsilon^{-\alpha}t) e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t} + r_j(t), \quad 0 \leq t \leq \varepsilon^\alpha, \quad (12)$$

where the coefficient functions  $z_j^{\mathbf{k}}(\tau)$  satisfy  $z_j^{-\mathbf{k}} = \overline{z_j^{\mathbf{k}}}$  and are bounded for  $0 \leq \tau \leq 1$ , together with all their derivatives with respect to  $\tau$  up to any fixed order, by

$$|z_0^{\mathbf{0}}(\tau)| \leq C_1, \quad |z_j^{\pm \langle j \rangle}(\tau)| \leq C_1 \omega_j^{-1} \leq C_1 \varepsilon,$$

and for all other  $(j, \mathbf{k})$ ,

$$|z_j^{\mathbf{k}}(\tau)| \leq C_1 |\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2|^{-1} \varepsilon^{\|\mathbf{k}\|}.$$

The functions  $z_j^{\mathbf{k}}$  satisfy the modulation equations (11) with a defect bounded by

$$|\delta_j^{\mathbf{k}}(\tau)| \leq C_2 \varepsilon^{N+1}, \quad 0 \leq \tau \leq 1. \quad (13)$$

The remainder term is bounded by

$$(\omega_j^2 |r_j(t)|^2 + |\dot{r}_j(t)|^2)^{1/2} \leq C_3 \varepsilon^{N+1}, \quad 0 \leq t \leq \varepsilon^\alpha.$$

The constants  $C_1, C_2, C_3$  are independent of  $\varepsilon$  and the frequencies  $\omega_j \geq \varepsilon^{-1}$ , but depend on  $n$  and  $N$ , on the energy bound  $E$  and on bounds of derivatives of the potential  $U$ .

This will be proved in Section 4. The proof follows the lines of previous proofs for modulated Fourier expansions; compare for example [13, Chap. XIII] and, in technically more complicated situations than here, [7, 11, 12].

*3.3. The almost-invariant.* We introduce the functions  $y_j^{\mathbf{k}}(t) = z_j^{\mathbf{k}}(\varepsilon^{-\alpha}t) e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t}$ , which we collect in the vector  $\mathbf{y} = (y_j^{\mathbf{k}})$  for  $j = 0, \dots, n$  and  $\mathbf{k} \in \mathcal{K}$ . Since  $\mathbf{k} \cdot \boldsymbol{\varpi} = 0$  for  $\mathbf{k} \in \mathcal{M}$ , the modulation potential satisfies  $\mathcal{U}(\mathbf{y}) = \mathcal{U}(\mathbf{z})$ . By (11), we have with rotated defects  $d_j^{\mathbf{k}}(t) = \delta_j^{\mathbf{k}}(\varepsilon^{-\alpha}t) e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t}$  that

$$\ddot{y}_j^{\mathbf{k}} + \omega_j^2 y_j^{\mathbf{k}} = -\nabla_j^{-\mathbf{k}} \mathcal{U}(\mathbf{y}) + d_j^{\mathbf{k}}, \quad (14)$$

where  $\nabla_j^{-\mathbf{k}}$  denotes the partial derivative with respect to  $y_j^{-\mathbf{k}}$ .

Without the defects  $d_j^{\mathbf{k}}$ , equation (14) is again a Hamiltonian system. Since the sum in  $\mathcal{U}(\mathbf{y})$  is over multi-indices  $\mathbf{k}^1, \dots, \mathbf{k}^m$  with  $\mathbf{k}^1 + \dots + \mathbf{k}^m \in \mathcal{M}$ , the extended potential  $\mathcal{U}$  is invariant under the action of the one-parameter group  $S(\theta)\mathbf{y} = (e^{i\theta(\mathbf{k} \cdot \boldsymbol{\varpi})} y_j^{\mathbf{k}})$ :

$$\mathcal{U}(S(\theta)\mathbf{y}) = \mathcal{U}(\mathbf{y}) \quad \text{for all } \theta \in \mathbb{R},$$

because  $\mathbf{k} \cdot \boldsymbol{\varpi} = 0$  for  $\mathbf{k} \in \mathcal{M}$ . By Noether's theorem this leads to a conserved quantity  $\mathcal{E}$  of the system (14) considered without defects. As we now show, this quantity is almost conserved in our situation including the defects  $d_j^{\mathbf{k}}$ : We have

$$-\sum_{j=0}^n \sum_{\mathbf{k} \in \mathcal{K}} i(\mathbf{k} \cdot \boldsymbol{\varpi}) y_j^{-\mathbf{k}} \nabla_j^{-\mathbf{k}} \mathcal{U}(\mathbf{y}) = \frac{d}{d\theta} \mathcal{U}(S(\theta)\mathbf{y}) \Big|_{\theta=0} = 0,$$

and for

$$\mathcal{E}(\mathbf{y}, \dot{\mathbf{y}}) = -i \sum_{j=0}^n \sum_{\mathbf{k} \in \mathcal{K}} (\mathbf{k} \cdot \boldsymbol{\varpi}) y_j^{-\mathbf{k}} \dot{y}_j^{\mathbf{k}} \quad (15)$$

we thus obtain by multiplying (14) with  $-i(\mathbf{k} \cdot \boldsymbol{\varpi}) y_j^{-\mathbf{k}}$  and summing over  $j$  and  $\mathbf{k}$  that

$$\frac{d}{dt} \mathcal{E}(\mathbf{y}(t), \dot{\mathbf{y}}(t)) = -i \sum_{j=0}^n \sum_{\mathbf{k} \in \mathcal{K}} (\mathbf{k} \cdot \boldsymbol{\varpi}) y_j^{-\mathbf{k}}(t) d_j^{\mathbf{k}}(t).$$

Since the right-hand side is  $\mathcal{O}(\varepsilon^{N+1})$  by the estimates of Proposition 1,  $\mathcal{E}$  is an almost-invariant. Moreover,  $\mathcal{E}$  turns out to be close to the oscillatory energy  $H_\omega$  along the solution of (4).

**Proposition 2.** *In the situation of Proposition 1 we have for  $0 \leq t \leq \varepsilon^\alpha$*

$$\begin{aligned} |\mathcal{E}(\mathbf{y}(t), \dot{\mathbf{y}}(t)) - \mathcal{E}(\mathbf{y}(0), \dot{\mathbf{y}}(0))| &\leq C\varepsilon^{N+1} \\ |\mathcal{E}(\mathbf{y}(t), \dot{\mathbf{y}}(t)) - H_\omega(\mathbf{p}(t), \mathbf{q}(t))| &\leq C'\varepsilon^{3/4}. \end{aligned}$$

*The constants  $C$  and  $C'$  are independent of  $\varepsilon$  and the frequencies  $\omega_j \geq \varepsilon^{-1}$ .*

*Proof.* The first inequality has already been derived above. Concerning the second inequality, we note that with  $\tau = \varepsilon^{-\alpha}t$  we have for  $\mathbf{k} \neq \mathbf{0}$

$$\dot{y}_j^{\mathbf{k}}(t) = e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t} \left( \varepsilon^{-\alpha} \frac{dz_j^{\mathbf{k}}}{d\tau}(\tau) + i(\mathbf{k} \cdot \boldsymbol{\varpi})z_j^{\mathbf{k}}(\tau) \right) = i(\mathbf{k} \cdot \boldsymbol{\varpi})y_j^{\mathbf{k}}(t) + \mathcal{O}(\varepsilon^{1-\alpha}).$$

With the bounds of Proposition 1 it follows that (omitting the argument  $t$ )

$$\begin{aligned} \mathcal{E}(\mathbf{y}, \dot{\mathbf{y}}) &= 2 \sum_{j=1}^n \varpi_j^2 |y_j^{(j)}|^2 + \mathcal{O}(\varepsilon^{1-\alpha}) \\ &= 2 \sum_{j=1}^n \omega_j^2 |y_j^{(j)}|^2 + \mathcal{O}(\varepsilon^{1-\alpha}). \end{aligned}$$

On the other hand, with Proposition 1 we also obtain

$$\begin{aligned} H_\omega(\mathbf{p}, \mathbf{q}) &= \frac{1}{2} \sum_{j=1}^n \left( |\dot{y}_j^{(j)} + \dot{y}_j^{-(j)}|^2 + \omega_j^2 |y_j^{(j)} + y_j^{-(j)}|^2 \right) + \mathcal{O}(\varepsilon) \\ &= 2 \sum_{j=1}^n \omega_j^2 |y_j^{(j)}|^2 + \mathcal{O}(\varepsilon^{1-\alpha}). \end{aligned}$$

Since  $\alpha \leq \frac{1}{4}$ , the result follows.  $\square$

*3.4. Transition from one interval to the next.* Propositions 1 and 2 are only valid on a short time interval  $0 \leq t \leq \varepsilon^\alpha$ . We now consider a second short time interval  $\varepsilon^\alpha \leq t \leq 2\varepsilon^\alpha$  and control in the following proposition the transition in the almost-invariant from the first to the second interval.

**Proposition 3.** *Let the conditions of Proposition 1 be fulfilled. Let  $z_j^{\mathbf{k}}(\varepsilon^{-\alpha}t)$  for  $0 \leq t \leq \varepsilon^\alpha$  be the coefficient functions as in Proposition 1 for initial data  $(\mathbf{p}(0), \mathbf{q}(0))$ , and let  $y_j^{\mathbf{k}}(t) = z_j^{\mathbf{k}}(\varepsilon^{-\alpha}t)e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t}$  and  $\mathbf{y}(t) = (y_j^{\mathbf{k}}(t))$ . Let further  $\tilde{\mathbf{y}}(t) = (\tilde{y}_j^{\mathbf{k}}(t))$  be the corresponding functions of the modulated Fourier expansion for  $0 \leq t \leq \varepsilon^\alpha$  to the initial data  $(\mathbf{p}(\varepsilon^\alpha), \mathbf{q}(\varepsilon^\alpha))$ , constructed as in Proposition 1. Then,*

$$|\mathcal{E}(\mathbf{y}(\varepsilon^\alpha), \dot{\mathbf{y}}(\varepsilon^\alpha)) - \mathcal{E}(\tilde{\mathbf{y}}(0), \dot{\tilde{\mathbf{y}}}(0))| \leq C\varepsilon^{N+1},$$

*where  $C$  is independent of  $\varepsilon$  and the frequencies  $\omega_j \geq \varepsilon^{-1}$ .*

The proof will be given in Section 4.



*3.5. From short to long time intervals.* Now we put many short time intervals together to get the long-time result of Theorem 1. For  $m = 0, 1, 2, \dots$ , let  $\mathbf{y}_m(t)$  contain the summands of the modulated Fourier expansion starting from  $(\mathbf{p}(m\varepsilon^\alpha), \mathbf{q}(m\varepsilon^\alpha))$ . As long as  $H_\omega(\mathbf{p}(m\varepsilon^\alpha), \mathbf{q}(m\varepsilon^\alpha)) \leq 2E$ , Proposition 2 yields for  $0 \leq t \leq \varepsilon^\alpha$

$$|\mathcal{E}(\mathbf{y}_m(t), \dot{\mathbf{y}}_m(t)) - \mathcal{E}(\mathbf{y}_m(0), \dot{\mathbf{y}}_m(0))| \leq C\varepsilon^{N+1}.$$

By Proposition 3,

$$|\mathcal{E}(\mathbf{y}_m(\varepsilon^\alpha), \dot{\mathbf{y}}_m(\varepsilon^\alpha)) - \mathcal{E}(\mathbf{y}_{m+1}(0), \dot{\mathbf{y}}_{m+1}(0))| \leq C\varepsilon^{N+1}.$$

Summing up these estimates over  $m$  and applying the triangle inequality yields, for  $0 \leq \theta \leq \varepsilon^\alpha$ ,

$$|\mathcal{E}(\mathbf{y}_m(\theta), \dot{\mathbf{y}}_m(\theta)) - \mathcal{E}(\mathbf{y}_0(0), \dot{\mathbf{y}}_0(0))| \leq 2(m+1)C\varepsilon^{N+1}.$$

By Proposition 2, we have at  $t = m\varepsilon^\alpha + \theta$

$$|\mathcal{E}(\mathbf{y}_m(\theta), \dot{\mathbf{y}}_m(\theta)) - H_\omega(\mathbf{p}(t), \mathbf{q}(t))| \leq C'\varepsilon^{3/4}.$$

Combining these bounds at  $t$  and at 0 we obtain for  $t \leq \varepsilon^{-N}$

$$|H_\omega(\mathbf{p}(t), \mathbf{q}(t)) - H_\omega(\mathbf{p}(0), \mathbf{q}(0))| \leq 2Ct\varepsilon^{N+1-\alpha} + 2C'\varepsilon^{3/4}.$$

This yields the bound of Theorem 1.

#### 4. Proof of Propositions 1 and 3

*4.1. Construction of the modulation functions.* We give an iterative construction of the functions  $z_j^{\mathbf{k}}$  ( $j = 0, \dots, n$  and  $\mathbf{k} \in \mathcal{K}$ ) such that after  $M = \lceil (N+1)/\mu \rceil$  iteration steps, the defect in equations (11) is of size  $\mathcal{O}(\varepsilon^{N+1})$ . We denote by  $\mathbf{z}^m = ([z_j^{\mathbf{k}}]^m)$  the  $m$ th iterate and distinguish between the following cases:

1. For  $j = 0$  and  $\mathbf{k} = \mathbf{0}$  the first two terms in (11) disappear and we iterate with a second order differential equation for  $[z_0^{\mathbf{0}}]^{m+1}$ :

$$\varepsilon^{-2\alpha} \frac{d^2[z_0^{\mathbf{0}}]^{m+1}}{d\tau^2} = -\nabla_0^{-\mathbf{0}} \mathcal{U}(\mathbf{z}^m).$$

2. For  $j \neq 0$  and  $\mathbf{k} = \pm\langle j \rangle$  with the  $j$ th unit vector  $\langle j \rangle = (0, \dots, 0, 1, 0, \dots, 0)$ , the first term in (11) disappears and we iterate using a first order differential equation for  $[z_j^{\pm\langle j \rangle}]^{m+1}$ :

$$\begin{aligned} \pm 2i\varpi_j \varepsilon^{-\alpha} \frac{d[z_j^{\pm\langle j \rangle}]^{m+1}}{d\tau} + \varepsilon^{-2\alpha} \frac{d^2[z_j^{\pm\langle j \rangle}]^m}{d\tau^2} \\ = (2\varpi_j\vartheta_j - \vartheta_j^2)[z_j^{\pm\langle j \rangle}]^{m+1} - \nabla_j^{\mp\langle j \rangle} \mathcal{U}(\mathbf{z}^m). \end{aligned} \quad (16)$$

We note that by (7) we have for sufficiently small  $\varepsilon$

$$\left| \frac{2\varpi_j\vartheta_j - \vartheta_j^2}{2\varpi_j\varepsilon^{-\alpha}} \right| \leq 2\gamma.$$

3. In all other cases we iterate with an explicit equation for  $[z_j^{\mathbf{k}}]^{m+1}$ :

$$\begin{aligned} (\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2) [z_j^{\mathbf{k}}]^{m+1} + 2i(\mathbf{k} \cdot \boldsymbol{\varpi}) \varepsilon^{-\alpha} \frac{d[z_j^{\mathbf{k}}]^m}{d\tau} + \varepsilon^{-2\alpha} \frac{d^2[z_j^{\mathbf{k}}]^m}{d\tau^2} \\ = (2\varpi_j \vartheta_j - \vartheta_j^2) [z_j^{\mathbf{k}}]^m - \nabla_j^{-\mathbf{k}} \mathcal{U}(\mathbf{z}^m). \end{aligned} \quad (17)$$

For  $j \neq 0$  we note that  $\pm\langle j \rangle - \mathbf{k} \notin \mathcal{M}$  for  $\mathbf{k} \in \mathcal{K}$  and  $\mathbf{k} \neq \pm\langle j \rangle$ , and therefore we have by (9)

$$\left| \frac{(\mathbf{k} \cdot \boldsymbol{\varpi}) \varepsilon^{-\alpha}}{\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2} \right| \leq 2\varepsilon^\mu, \quad \left| \frac{\varepsilon^{-2\alpha}}{\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2} \right| \leq 2\varepsilon^{1-\alpha+\mu}$$

and

$$\left| \frac{2\varpi_j \vartheta_j - \vartheta_j^2}{\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2} \right| \leq 8\gamma\varepsilon^\mu.$$

For  $j = 0$  and  $\mathbf{k} \notin \mathcal{M}$  we have similar estimates directly by (9).

We need initial values  $[z_0^{\mathbf{0}}]^{m+1}(0)$ ,  $\frac{d}{d\tau}[z_0^{\mathbf{0}}]^{m+1}(0)$  and  $[z_j^{\pm\langle j \rangle}]^{m+1}(0)$  for  $j \neq 0$ . They are uniquely determined by the equations (for  $j = 0, \dots, n$ )

$$q_j(0) = \sum_{\mathbf{k} \in \mathcal{K}} [z_j^{\mathbf{k}}]^{m+1}(0), \quad \dot{q}_j(0) = \sum_{\mathbf{k} \in \mathcal{K}} \left( \varepsilon^{-\alpha} \frac{d[z_j^{\mathbf{k}}]^{m+1}}{d\tau}(0) + i(\mathbf{k} \cdot \boldsymbol{\varpi}) [z_j^{\mathbf{k}}]^{m+1}(0) \right) \quad (18)$$

after inserting (17) for  $\mathbf{k} \neq \pm\langle j \rangle$  and replacing for  $j \neq 0$  the derivative of  $[z_j^{\pm\langle j \rangle}]^{m+1}$  with (16). The starting iterates are chosen as  $[z_0^{\mathbf{0}}]^0(\tau)$  equal to the solution at  $t = \varepsilon^\alpha \tau$  of  $\ddot{y} = -\nabla_0 U(y, 0, \dots, 0)$  with  $y(0) = q_0(0)$  and  $\dot{y}(0) = \dot{q}_0(0)$ , and for all other  $(j, \mathbf{k})$  with  $\mathbf{k} \neq \pm\langle j \rangle$ , we set  $[z_j^{\mathbf{k}}]^0(\tau) = 0$ . The diagonal functions  $[z_j^{\pm\langle j \rangle}]^0$  are then chosen as constant functions such that (18) is satisfied.

*4.2. Bounds of the modulation functions.* We estimate the modulation functions on intervals of length  $\tau = \mathcal{O}(1)$ , or equivalently  $t = \mathcal{O}(\varepsilon^\alpha)$ . With the above inequalities and the energy bound (5) we obtain by induction on  $m$ , for  $m = 0, \dots, M$ ,

$$[z_0^{\mathbf{0}}]^m = \mathcal{O}(1), \quad [z_j^{\pm\langle j \rangle}]^m = \mathcal{O}(\omega_j^{-1})$$

and for all other  $(j, \mathbf{k})$ ,

$$[z_j^{\mathbf{k}}]^m = \mathcal{O}(|\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2|^{-1} \varepsilon^{\|\mathbf{k}\|}),$$

and the same bounds hold for all their derivatives up to any prescribed order.

4.3. *Bounds of the defect.* We prove the estimate (13) of the defect  $\delta_j^{\mathbf{k}}$  in Proposition 1. The defect  $[\delta_j^{\mathbf{k}}]^m(\tau)$  in inserting the  $m$ th iterate  $\mathbf{z}^m = ([z_j^{\mathbf{k}}]^m)$  in the modulation equations (11) satisfies

$$[\delta_0^{\mathbf{0}}]^m = -\varepsilon^{-2\alpha} \left( \frac{d^2[z_0^{\mathbf{0}}]^{m+1}}{d\tau^2} - \frac{d^2[z_0^{\mathbf{0}}]^m}{d\tau^2} \right)$$

and for  $j = 1, \dots, n$

$$\begin{aligned} [\delta_j^{\pm\langle j \rangle}]^m &= \mp 2i \varpi_j \varepsilon^{-\alpha} \left( \frac{d[z_j^{\pm\langle j \rangle}]^{m+1}}{d\tau} - \frac{d[z_j^{\pm\langle j \rangle}]^m}{d\tau} \right) \\ &\quad + (2\varpi_j \vartheta_j - \vartheta_j^2) \left( [z_j^{\pm\langle j \rangle}]^{m+1} - [z_j^{\pm\langle j \rangle}]^m \right) \end{aligned}$$

and for all other  $(j, \mathbf{k})$

$$[\delta_j^{\mathbf{k}}]^m = -(\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2) \left( [z_j^{\mathbf{k}}]^{m+1} - [z_j^{\mathbf{k}}]^m \right).$$

In the following we work with the  $C^r([0, 1])$  norm, defined for  $\mathbf{v} = (v_j^{\mathbf{k}})$  by

$$\|\mathbf{v}\|_{C^r} = \max_{0 \leq \tau \leq 1} \max_{0 \leq l \leq r} \sum_{j=0}^n \sum_{\mathbf{k} \in \mathcal{K}} \left| \frac{d^l}{d\tau^l} v_j^{\mathbf{k}}(\tau) \right|,$$

and use the notation

$$\mathbf{Az} = \mathbf{v} = (v_j^{\mathbf{k}}) \quad \text{with} \quad \begin{cases} v_0^{\mathbf{0}} = \varepsilon^{-2\alpha} z_0^{\mathbf{0}} \\ v_j^{\pm\langle j \rangle} = 2i \varpi_j \varepsilon^{-\alpha} z_j^{\pm\langle j \rangle}, & j = 1, \dots, n \\ v_j^{\mathbf{k}} = (\varpi_j^2 - (\mathbf{k} \cdot \boldsymbol{\varpi})^2) z_j^{\mathbf{k}} & \text{else.} \end{cases}$$

With this notation and the above formulas for the defect we have

$$|[\delta_j^{\mathbf{k}}]^m| \leq C \|\mathbf{v}^{m+1} - \mathbf{v}^m\|_{C^2}.$$

We therefore study  $\mathbf{v}^{m+1} - \mathbf{v}^m$  and show by induction on  $m$  that for  $m \leq M$  and  $r = 2M - 2m + 2$ ,

$$\|\mathbf{v}^{m+1} - \mathbf{v}^m\|_{C^r} = \mathcal{O}(\varepsilon^{m\mu}), \quad (19)$$

which implies the estimate (13) of the defect in Proposition 1 if we use  $M \geq (N+1)/\mu$  iterations for the construction of the modulation functions. Estimate (19) is shown by analysing the iterative construction of modulation functions, where the  $(m+1)$ th iterates are the dominant terms and the  $m$ th iterates the non-dominant ones.

We first consider the diagonal elements  $[v_j^{\pm\langle j \rangle}]^m$ . From (18) we obtain after inserting (17) for  $\mathbf{k} \neq \pm\langle j \rangle$  that

$$|[v_j^{\pm\langle j \rangle}]^{m+1}(0) - [v_j^{\pm\langle j \rangle}]^m(0)| \leq C\varepsilon^\mu \|\mathbf{v}^m - \mathbf{v}^{m-1}\|_{C^2}.$$

Solving the differential equation (16) by the variation-of-constants formula then yields that

$$|[v_j^{\pm\langle j \rangle}]^{m+1}(\tau) - [v_j^{\pm\langle j \rangle}]^m(\tau)| \leq C\varepsilon^\mu \|\mathbf{v}^m - \mathbf{v}^{m-1}\|_{C^2}, \quad 0 \leq \tau \leq 1.$$

Repeated differentiation in (16) further shows that for  $\ell \leq 2M - 2m + 2$  and for  $0 \leq \tau \leq 1$

$$\left| \frac{d^\ell}{d\tau^\ell} [v_j^{\pm(j)}]^{m+1}(\tau) - \frac{d^\ell}{d\tau^\ell} [v_j^{\pm(j)}]^m(\tau) \right| \leq C\varepsilon^\mu \|\mathbf{v}^m - \mathbf{v}^{m-1}\|_{C^{\ell+2}}.$$

Using  $2\alpha \geq \mu$  we also bound, for  $r = 2M - 2m + 2$ ,

$$\|[v_0^{\mathbf{0}}]^{m+1} - [v_0^{\mathbf{0}}]^m\|_{C^r} \leq C\varepsilon^\mu \|\mathbf{v}^m - \mathbf{v}^{m-1}\|_{C^{r+2}}.$$

Similarly we have from the iteration formulas and the above inequalities that

$$\|\mathbf{v}^{m+1} - \mathbf{v}^m\|_{C^r} \leq C\varepsilon^\mu \|\mathbf{v}^m - \mathbf{v}^{m-1}\|_{C^{r+2}}.$$

This proves (19) and the estimate of the defect of Proposition 1.

*4.4. Solution approximation.* We consider the  $M$ th iterates of the modulation functions (with  $M = \lceil (N+1)/\mu \rceil$ ) and omit the superscript  $M$  on the modulation functions and their defects. The truncated modulated Fourier expansion

$$\tilde{q}_j(t) = \sum_{\mathbf{k} \in \mathcal{K}} z_j^{\mathbf{k}}(\varepsilon^{-\alpha}t) e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t}$$

satisfies the perturbed second order differential equation

$$\ddot{\tilde{q}}_j + \omega_j^2 \tilde{q}_j = -\nabla_j U(\tilde{\mathbf{q}}) + d_j, \quad j = 0, \dots, n,$$

with

$$d_j(t) = \sum_{\mathbf{k} \in \mathcal{K}} \delta_j^{\mathbf{k}}(\varepsilon^{-\alpha}t) e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t} - \sum_{\mathbf{k} \in \mathcal{N} \setminus \mathcal{K}} \nabla_j^{-\mathbf{k}} \mathcal{U}(\mathbf{z}(\varepsilon^{-\alpha}t)) e^{i(\mathbf{k} \cdot \boldsymbol{\varpi})t} + \rho_j(t) = \mathcal{O}(\varepsilon^{N+1}),$$

where  $\rho_j(t)$  denotes the gradient of the remainder term in the truncated Taylor expansion of the potential  $U$  about  $\mathbf{z}_0^{\mathbf{0}}(\varepsilon^{-\alpha}t)$ : omitting the arguments  $t$  and  $\tau = \varepsilon^{-\alpha}t$ , and writing  $\tilde{\mathbf{q}} = \mathbf{z}_0^{\mathbf{0}} + \hat{\mathbf{q}}$ ,

$$\rho_j = \nabla_j U(\tilde{\mathbf{q}}) - \nabla_j U(\mathbf{z}_0^{\mathbf{0}}) - \sum_{m=1}^N \sum_{j_1, \dots, j_m=0}^n \frac{1}{m!} \partial_{j_1} \dots \partial_{j_m} \nabla_j U(\mathbf{z}_0^{\mathbf{0}})(\hat{q}_{j_1}, \dots, \hat{q}_{j_m}).$$

After subtracting the differential equation (4), standard estimates using the variation of constants formula and the Gronwall inequality then yield the bound of  $r_j(t) = \tilde{q}_j(t) - q_j(t)$  for  $0 \leq t \leq \varepsilon^\alpha$  as stated in Proposition 1.

4.5. *Proof of Proposition 3.* Let  $\mathbf{z}^m = ([z_j^{\mathbf{k}}]^m)$  and  $\tilde{\mathbf{z}}^m = ([\tilde{z}_j^{\mathbf{k}}]^m)$  be the  $m$ th iterates of the construction of the modulation functions  $\mathbf{z}$  and  $\tilde{\mathbf{z}}$  for initial data  $(\mathbf{p}(0), \mathbf{q}(0))$  and  $(\mathbf{p}(\varepsilon^\alpha), \mathbf{q}(\varepsilon^\alpha))$ , respectively. With the iteration for  $\tilde{\mathbf{z}}^m$  and the defect formula for  $\mathbf{z}^M$  the same arguments as in the estimates of the defect yield for  $\mathbf{w}^m(\tau) = \Lambda \mathbf{z}^M(1 + \tau) - \Lambda \tilde{\mathbf{z}}^m(\tau)$  that

$$\|\mathbf{w}^M\|_{C^2} = \mathcal{O}(\varepsilon^{N+1}). \quad (20)$$

The only difference is that the estimate of  $[w_j^{\pm(j)}]^{m+1}(0)$  now reads

$$|[w_j^{\pm(j)}]^{m+1}(0)| \leq C\varepsilon^\mu \|\mathbf{w}^m\|_{C^2} + C\varepsilon^{N+1},$$

a bound that is obtained by inserting (12) in the defining relation (18) for  $[\tilde{z}_j^{\pm(j)}]^{m+1}(0)$  and by using the estimate of the remainder term from Proposition 1. Similarly the estimate of  $w_0^0$  has to be modified. The bound of Proposition 3 follows from (20).

*Acknowledgement.* This work has been supported by Fonds National Suisse, Project No. 200020-126638.

## References

1. K. Aoki, J. Lukkarinen, H. Spohn, *Energy transport in weakly anharmonic chains*, J. Stat. Phys. **124** (2006), 1105–1129.
2. D. Bambusi and A. Giorgilli, *Exponential stability of states close to resonance in infinite-dimensional Hamiltonian systems*, J. Statist. Phys. **71** (1993), 569–606.
3. G. Benettin, L. Galgani, and A. Giorgilli, *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part I*, Comm. Math. Phys. **113** (1987), 87–103.
4. G. Benettin, L. Galgani, and A. Giorgilli, *Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part II*, Comm. Math. Phys. **121** (1989), 557–601.
5. A. Carati and L. Galgani, *Theory of dynamical systems and the relations between classical and quantum mechanics*, Found. Phys. **31** (2001), 69–87.
6. D. Cohen, E. Hairer, and C. Lubich, *Modulated Fourier expansions of highly oscillatory differential equations*, Found. Comput. Math. **3** (2003), 327–345.
7. ———, *Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions*, Arch. Ration. Mech. Anal. **187** (2008), 341–368.
8. D. Dolgopyat and C. Liverani, *Energy transfer in a fast-slow Hamiltonian system*, Comm. Math. Phys. **308** (2011), 201–225.
9. E. Fermi, J. Pasta, and S. Ulam. *Studies of non linear problems*. Technical Report LA-1940, Los Alamos, 1955.
10. G. Gallavotti, editor. *The Fermi-Pasta-Ulam problem. A status report. Lecture Notes in Physics*, volume 728. Springer, Berlin, 2008.
11. L. Gauckler, E. Hairer, C. Lubich, and D. Weiss, *Metastable energy strata in weakly nonlinear wave equations*, Comm. Partial Differential Equations **37** (2012), 1391–1413.
12. E. Hairer and C. Lubich, *On the energy distribution in Fermi–Pasta–Ulam lattices*, Arch. Ration. Mech. Anal. **205** (2012), 993–1029.
13. 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.
14. F. Huveneers, *Thermal conductivity of disordered lattices in the weak coupling regime*, preprint (2012), arXiv:1203.3587.
15. A. Shudo, K. Ichiki, and S. Saito, *Origin of slow relaxation in liquid-water dynamics: A possible scenario for the presence of bottleneck in phase space*, Europhys. Lett. **73** (2006), 826–832.