DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS–SERIES B
Volume **6**, Number **3**, May **2006**

Website: http://AIMsciences.org

pp. **493–523**

DISPERSIVE EVOLUTION OF PULSES IN OSCILLATOR CHAINS WITH GENERAL INTERACTION POTENTIALS

Johannes Giannoulis

Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstraße 39, D–10117 Berlin, Germany

Alexander Mielke

Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstraße 39, D–10117 Berlin, Germany
and
Institut für Mathematik, Humboldt-Universität zu Berlin
Rudower Chaussee 25, D–12489 Berlin, Germany

(Communicated by Serge Aubry)

Abstract. We study the dispersive evolution of modulated pulses in a nonlinear oscillator chain embedded in a background field. The atoms of the chain interact pairwise with an arbitrary but finite number of neighbors. The pulses are modeled as macroscopic modulations of the exact spatiotemporally periodic solutions of the linearized model. The scaling of amplitude, space and time is chosen in such a way that we can describe how the envelope changes in time due to dispersive effects. By this multiscale ansatz we find that the macroscopic evolution of the amplitude is given by the nonlinear Schrödinger equation. The main part of the work is focused on the justification of the formally derived equation: We show that solutions which have initially the form of the assumed ansatz preserve this form over time-intervals with a positive macroscopic length. The proof is based on a normal-form transformation constructed in Fourier space, and the results depend on the validity of suitable nonresonance conditions.

1. **Introduction.** A major topic in the area of multiscale problems is the derivation of macroscopic, continuum models from microscopic, discrete ones. The prototype of a discrete many-particle system is a periodic lattice for modeling a crystal. Starting from the seminal work of Fermi, Pasta and Ulam [10], a lot of interest and work (cf. e.g. the article collection in [5]) was attracted to the simplest, one-dimensional model, namely the monoatomic, infinite oscillator chain

$$\ddot{x}_j = \sum_{m=1}^{M} [V_m'(x_{j+m}-x_j) - V_m'(x_j-x_{j-m})] - W'(x_j), \quad j \in \mathbb{Z}. \tag{1}$$

Here $j \in \mathbb{Z}$ is the particle index and $x_j(t) \in \mathbb{R}$ is the deviation from the equilibrium position $j \in \mathbb{Z}$ at time $t \geq 0$. The motion is driven by the pair potentials V_m with its m-th neighbors (up to the M-th ones) and by the on-site potential W coupling the atoms to a background field.

2000 *Mathematics Subject Classification.* Primary: 34E13, 34C20, 37K60; Secondary: 70F45.
Key words and phrases. nonlinear oscillator chain, multiscale theory, modulational theory, nonlinear Schrödinger equation, normal-form transformation, nonresonance conditions.

We are interested in the macroscopic limit which is obtained by choosing well-prepared initial conditions: We choose the initial data in a specified class of functions and want to obtain an evolution equation within this function class, which we call the macroscopic limit problem. This approach is motivated by the theory of modulation equations which evolved in the late 1960's for problems in fluid mechanics (see [31] for a survey on this subject). If the linearized model has a space-time periodic solution one asks how initial modulations of this pattern evolve in time. The modulations occur on much larger spatial and temporal scales, such that the modulation equation is a macroscopic equation.

This is only one among a huge variety of possible approaches for investigating the oscillator chain and deriving macroscopic limits, which reflect different viewpoints and aims. Apart from methods and results in the framework of nonequilibrium statistical mechanics (cf. for a survey e.g. [4, 37]), in a more deterministic setting we would like to mention the following groups of questions: First, one can focus on completely integrable systems like the Toda lattices (with $M = 1$, $V(y) = \mathrm{e}^y$ and $W \equiv 0$, see, e.g., [6, 9]). Second, a big body of work is concentrated on the dynamics of special types of solutions like solitons, breathers or wave trains [13]–[18], [22]–[25], [28, 29]. Third, one can be interested in the response of the oscillator chain to a simple initial disturbance [1] or to Riemann initial data [6, 9].

Our paper is embedded in that body of work which is focused on the derivation and rigorous justification of partial differential equations as macroscopic limits describing the dynamics of a discrete lattice. In the framework of harmonic lattices [30] considers general polyatomic crystals in any dimension. It is shown that the weak continuum limit describing the macroscopic evolution of displacements and velocities is the equation of linear elastodynamics, and that the weak limit of the local energy density can be described by Wigner-Husimi measures, satisfying a transport equation parametrized by the microscopic wave vector. Here, the macroscopic space and time variables are modeled as $y = \varepsilon j$ and $\tau = \varepsilon t$, respectively. In the nonlinear, anharmonic setting the same hyperbolic scaling is used in [7, 11, 21], where for $W \equiv 0$ the modulations of large-amplitude traveling waves are considered, and the derived macroscopic limit is the so-called Whitham modulation equation. A similar modulation ansatz has been used in [20] for the discrete nonlinear Schrödinger equation $\mathrm{i}\dot{A}_j + c_1(A_{j-1} - 2A_j + A_{j+1}) + c_2|A_j|^2 A_j = 0$ with $A_j(t) \in \mathbb{C}$.

Solutions without internal microscopic structure are obtained in the long wavelength limit which leads to the Korteweg-de Vries equation, see [14, 26, 36]. There, for $W \equiv 0$ small-amplitude solutions of the long-wave form $x_j(t) = \varepsilon^2 U(\varepsilon^3 t, \varepsilon(x - ct)) + \mathcal{O}(\varepsilon^4)$ are studied, and it is justified that U satisfies the KdV equation $\partial_\tau U + \kappa_1 U \partial_\xi U + \kappa_2 \partial_\xi^3 U = 0$. Similar investigations have been made in [3, 33], where the KdV is derived for a long periodic FPU chain to analyze long-time averages of the energy distribution in the Fourier modes. For solutions with internal microstructure the modulational approach is used in the physical literature for a long time, see e.g., [12, 34], where kinks, breathers, bright and dark solitons for FPU and Klein-Gordon systems are constructed.

The purpose of our work is to make these construction mathematically rigorous and give quantitative estimates of the involved errors. Like [2, 19] we are concerned with modulations of the form

$$x_j(t) = \varepsilon A(\varepsilon^2 t, \varepsilon(j - c_{\mathrm{gr}} t))\mathrm{e}^{\mathrm{i}(\omega t + \vartheta_0 j)} + \text{c.c.} + \mathcal{O}(\varepsilon^2),$$

where (c.c. abbreviates "complex conjugate" and) A satisfies the nonlinear Schrödinger equation $\mathrm{i}\partial_\tau A = \gamma_1 \partial_\xi^2 A + \gamma_2 |A|^2 A$. These solutions are microscopically periodic in space and time via $\mathrm{e}^{\mathrm{i}(\omega t + \vartheta_0 j)}$ and are modulated by the complex-valued envelope A which depends on the macroscopic time $\tau = \varepsilon^2 t$ and the macroscopic spatial variable $\xi = \varepsilon(j - c_{\mathrm{gr}} t)$.

Our aim is to generalize [2, 19] in two directions. First, we allow for general interaction potentials leading to quadratic terms in the nonlinearities. Second, we allow for pair interaction potentials between 1 to M neighbors. To be more specific, we consider potentials of the form

$$\begin{aligned} V_m(d) &:= \frac{\alpha_{m,1}}{2} d^2 + \frac{\alpha_{m,2}}{3} d^3 + \frac{\alpha_{m,3}}{4} d^4 + \mathcal{O}(d^5), \\ W(x) &:= \frac{\beta_1}{2} x^2 + \frac{\beta_2}{3} x^3 + \frac{\beta_3}{4} x^4 + \mathcal{O}(x^5) \end{aligned} \tag{2}$$

for $m = 1, \ldots, M$. (In particular, [2, 19] relates to the case $M = 1$ and $\alpha_{1,2} = 0 = \beta_2$, which leads to a much simpler analysis.) The linearized model is given by

$$\ddot{x}_j = L_j(x) := \sum_{m=1}^{M} \alpha_{m,1}(x_{j+m} - 2x_j + x_{j-m}) - \beta_1 x_j. \tag{3}$$

It has the basic solutions $x_j(t) = \mathrm{e}^{\mathrm{i}(\widetilde{\omega} t + \vartheta j)}$, where the wave number ϑ and the frequency $\widetilde{\omega}$ have to satisfy the dispersion relation $\widetilde{\omega}^2 = \omega^2(\vartheta)$ with

$$\omega^2(\vartheta) := \beta_1 + 2 \sum_{m=1}^{M} \alpha_{m,1}[1 - \cos(m\vartheta)], \quad \vartheta \in (-\pi, \pi]. \tag{DR}$$

Throughout, we require that a stability condition

$$\omega^2(\vartheta) > 0 \quad \text{for all } \vartheta \in (-\pi, \pi] \tag{SC}$$

holds, and we take $\omega(\vartheta) > 0$. In the case of interactions only between nearest neighbors ($M = 1$, $V := V_1$, $\alpha_k := \alpha_{1,k}$) (SC) is equivalent to $\min\{\beta_1, 4\alpha_1 + \beta_1\} > 0$ (cf. (21)). In the following, we consider always a fixed wave number $\vartheta_0 \in (-\pi, \pi]$, and write shortly ω, ω', ω'' to denote $\omega(\vartheta_0)$, $\omega'(\vartheta_0)$, $\omega''(\vartheta_0)$, respectively. The associated basic mode $\mathbf{E}(t, j) := \mathrm{e}^{\mathrm{i}(\omega t + \vartheta_0 j)}$ is considered to be the microscopic pattern of reference. The function $A : [0, \infty) \times \mathbb{R} \to \mathbb{C}$ modulates the microscopic pattern $\mathbf{E}$ such that the solutions under consideration take the form

$$x_j(t) = (X_\varepsilon^A)_j(t) + \mathcal{O}(\varepsilon^2) \quad \text{with} \quad (X_\varepsilon^A)_j(t) := \varepsilon A(\varepsilon^2 t, \varepsilon(j - c_{\mathrm{gr}} t)) \mathbf{E}(t, j) + \text{c.c.}. \tag{4}$$

Since these solutions are small, they lead to dynamics which are close to the linear one. Only the extremely long time scale enables us to see how the amplitude A changes due to dispersive and nonlinear effects. In the hyperbolic scaling $\tau = \varepsilon t$ with $\xi = \varepsilon j$ one only sees hyperbolic transport effects, but neither dispersion nor nonlinearities.

Inserting the ansatz (4) into (1), it turns out that this provides a useful approximation for solutions of (1) only if the group velocity c_{gr} equals $-\omega'$, and A satisfies the associated nonlinear Schrdinger equation (NLSE)

$$2\mathrm{i}\omega \partial_\tau A = \omega \omega'' \partial_\xi^2 A + \rho |A|^2 A, \tag{5}$$

where ρ can be calculated explicitly (cf. (19)). A formal derivation of (5) is obtained by assuming that solutions in the form (4) exist (cf. Section 2).

The mathematical justification of the NLSE (5) is carried out in Section 4: We show that solutions $t \mapsto (x_j(t))_{j \in \mathbb{Z}}$ which start at $t = 0$ in the form of the ansatz

(4) stay in this form over intervals $[0, \tau_0/\varepsilon^2]$ of positive macroscopic length $\tau_0 > 0$. More precisely, Theorem 2 states the following: Given a sufficiently smooth solution A of NLSE (5), $\tau_0 > 0$ and $d > 0$, there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \le \varepsilon_0$ any solution x of (1) with

$$\|(x(0), \dot{x}(0)) - (X_\varepsilon^A(0), \dot{X}_\varepsilon^A(0))\|_{\ell^2\times\ell^2} \le d\varepsilon^{3/2}$$

satisfies the estimate

$$\|(x(t), \dot{x}(t)) - (X_\varepsilon^A(t), \dot{X}_\varepsilon^A(t))\|_{\ell^2\times\ell^2} \le C\varepsilon^{3/2} \text{ for } t \in [0, \tau_0/\varepsilon^2].$$

We prove this result in principle by the same approach we used in our previous paper [19] on this subject. There, we considered the situation of only nearest-neighbor interactions and restricted the justification of the NLSE on the case of cubic leading terms of the nonlinearity in (1) (i.e. $V'''(0) = 0 = W'''(0)$ or, equivalently, $\alpha_2 = 0 = \beta_2$), since exactly this assumption enabled us to use the method developed in [27], relying on a Gronwall-type argument. For the general case including quadratic leading terms this Gronwall-type argument can not be used directly and we circumvent this difficulty by a method which was developed in [35] for hyperbolic PDEs.

The idea is to apply to the system $\dot{\widetilde{x}} = \widetilde{L}\widetilde{x} + \widetilde{Q}(\widetilde{x}, \widetilde{x}) + \widetilde{M}(\widetilde{x})$ (corresponding to our microscopic model (1)) a suitable normal-form transformation (near-identity transformation) $F : \widetilde{x} \mapsto \widetilde{y} = F(\widetilde{x})$, such that the transformed system $\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + N(\widetilde{y})$ has a nonlinearity N with cubic leading terms. Then, we prove for the transformed system a result equivalent to Theorem 2, by using the Gronwall-type argument mentioned above. Transforming this result back into the variable $\widetilde{x}$, we obtain Theorem 2.

The construction of the normal-form transformation F is carried out in Fourier space (Section 3). An essential condition in normal-form theory is a *nonresonance condition* of third order on our fixed $\vartheta_0 \in (-\pi, \pi]$:

$$\exists\, C_{\vartheta_0}^{\mathrm{NR}} > 0 : \quad \inf_{\substack{s,\sigma=1,2;\\ \theta\in(-\pi,\pi]}} |\omega(\vartheta_0) + (-1)^s\omega(\theta) + (-1)^\sigma\omega(\vartheta_0 - \theta)| \ge C_{\vartheta_0}^{\mathrm{NR}} > 0. \quad (\mathrm{NR3})_{\vartheta_0}$$

Our first result is Theorem 2 which is proved under a strengthened version of $(\mathrm{NR3})_{\vartheta_0}$, which we call *uniform nonresonance condition*

$$\exists\, C_{\mathrm{unif}}^{\mathrm{NR}} > 0 : \quad \inf_{\substack{s,\sigma=1,2;\\ \vartheta,\theta\in(-\pi,\pi]}} |\omega(\vartheta) + (-1)^s\omega(\theta) + (-1)^\sigma\omega(\vartheta - \theta)| \ge C_{\mathrm{unif}}^{\mathrm{NR}} > 0. \quad (\mathrm{NR3})_{\mathrm{unif}}$$

(In the case of nearest-neighbor interactions $(\mathrm{NR3})_{\mathrm{unif}}$ holds if and only if the coefficients α_1, β_1 of the harmonic parts of the potentials V, W satisfy the condition $\min\{\beta_1, (16/3)\alpha_1 + \beta_1\} > 0$ (cf. (24)), which is slightly sharper than the stability condition (SC) $\min\{\beta_1, 4\alpha_1 + \beta_1\} > 0$. However, for $\alpha_1 \ge 0$, both conditions reduce to $\beta_1 = W''(0) > 0$.)

Under the more general condition $(\mathrm{NR3})_{\vartheta_0}$ the analysis is more subtle. We obtain an analogous justification result by using the higher-order approximation

$$X_\varepsilon^{A,2} := \varepsilon A\mathbf{E} + \varepsilon^2 \left(\frac{\beta_2}{\delta_0}|A|^2 + A_{2,1}\mathbf{E} + \frac{a}{\delta_2}A^2\mathbf{E}^2 \right) + \text{c.c.}, \tag{6}$$

with $\delta_n := n^2\omega^2(\vartheta_0) - \omega^2(n\vartheta_0)$ and $a := 4\mathrm{i}\sum_{m=1}^M \alpha_{m,2}\sin(m\vartheta_0)[1 - \cos(m\vartheta_0)] + \beta_2$, where A solves the NLSE (5) for $\tau \in [0, \tau_0]$ and $A_{2,1} : [0, \tau_0] \times \mathbb{R} \to \mathbb{C}$ solves the equation

$$\begin{aligned} 2\mathrm{i}\omega\partial_\tau A_{2,1} &= \omega\omega''\partial_\xi^2 A_{2,1} + \rho(2|A|^2 A_{2,1} + A^2\overline{A}_{2,1}) \\ &\quad -2\omega'\partial_\tau\partial_\xi A - \tfrac{\mathrm{i}(\omega^2)'''}{6}\partial_\xi^3 A + 2e|A|^2\partial_\xi A, \end{aligned} \tag{7}$$

where again e can be given explicitly (cf. (20)). This equation is obtained formally in the course of the formal derivation of the NLSE by increasing the order of considered scales ε^k to $k = 4$ (cf. Section 2). Clearly, by increasing the order of our approximation we consider estimates for the error with respect to an original solution which are also of higher order, namely ε^α with $\alpha \in (2, 5/2]$. The precise result is proven in Section 4.3. Note, however, that also the approximation $X_\varepsilon^{A,2}$ depends only on the solution A of the NLSE (5), since assuming initially $A_{2,1}(0,\cdot) \equiv 0$, the solution $A_{2,1}$ of (7) depends only on A.

We believe that the results obtained in this paper can be extended in principle also to the multidimensional case. However, for a precise investigation and a description of probable limitations (at least at a technical level) in the adaptation of the argument used in the one-dimensional case presented here, we refer the reader to future work.

2. The formal derivation of the NLSE.

The formal derivation of the NLSE as a modulation equation for the oscillator chain model (1) with only nearest-neighbor interactions ($M = 1$) has been presented in full detail in [19, Section 2]. There, sort of a step-by-step method was used, which was restricted to the concrete situation. More general situations are treated in [28].

Here, since we want to derive the NLSE in the case of generalized interaction potentials ($M > 1$) and especially since we need to consider also additional modulation equations (cf. (7)), we take the opportunity to present the formal derivation in a more general way, culminating in the equation system (18), which can be used in some sense algorithmically in order to determine the functions $A_{k,n}$ of an approximation $X_\varepsilon^{A,p}$ (cf. (8)) for arbitrary $p \in \mathbb{N}$.

Since we want to study the macroscopic evolution of modulated solutions of the form (4), it is natural to insert such an ansatz into our microscopic model (1) in order to derive an evolution equation for the macroscopic modulation function $A : [0,\infty)\times\mathbb{R} \to \mathbb{C}$. But, inserting such an ansatz into the nonlinear problem (1) will generate higher harmonic terms (with factors $\mathbf{E}^n$) having scaling parameters ε^k, $k \in \mathbb{N}$. Hence, anticipating this situation, we insert into (1) the multiple scale ansatz

$$X_\varepsilon^{A,p} := \sum_{k=1}^{p} \varepsilon^k \sum_{n=-k}^{k} A_{k,n}\mathbf{E}^n \tag{8}$$

with $A_{k,n} = A_{k,n}(\tau,\xi) \in \mathbb{C}$ and $A_{k,-n} = \overline{A}_{k,n}$ where $\tau = \varepsilon^2 t,\ \xi = \varepsilon(j - c_{\mathrm{gr}}t)$ for $j \in \mathbb{Z}$, $t \geq 0$. Obviously, $A_{1,1} = A$.

The idea is now to expand the left- and right- hand side of the equation

$$(\ddot{X}_\varepsilon^{A,p})_j = \sum_{m=1}^{M}\{V_m'[(X_\varepsilon^{A,p})_{j+m} - (X_\varepsilon^{A,p})_j] - V_m'[(X_\varepsilon^{A,p})_j - (X_\varepsilon^{A,p})_{j-m}]\} - W'[(X_\varepsilon^{A,p})_j] \tag{9}$$

in terms of $\varepsilon^k\mathbf{E}^n$. Then, by equating the left- and right-hand side coefficients of each of these terms for $k = 1,\dots,p,\ n = 0,\dots,k$ separately, we will obtain an hierarchy of equations for the functions $A_{k,n}$.

Since

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(A_{k,n}\mathbf{E}^n) = [\mathrm{i}n\omega(\vartheta_0) + \varepsilon(-c_{\mathrm{gr}})\partial_\xi + \varepsilon^2\partial_\tau]^2 A_{k,n}\mathbf{E}^n,$$

we obtain from (8) for the left-hand side by summing according to orders of ε

$$\ddot{X}_\varepsilon^{A,p} = \sum_{k=1}^{p+4}\varepsilon^k \sum_{q=1}^{\min\{k,p\}} \sum_{n=-q}^{q} \sum_{\mu+2\nu=k-q} c_{n\mu\nu}\partial_\tau^\nu\partial_\xi^\mu A_{q,n}\mathbf{E}^n,$$

and thus

$$\ddot{X}_\varepsilon^{A,p} = \sum_{k=1}^{p}\varepsilon^k \sum_{q=1}^{k} \sum_{n=-q}^{q} \sum_{\mu+2\nu=k-q} c_{n\mu\nu}\partial_\tau^\nu\partial_\xi^\mu A_{q,n}\mathbf{E}^n + \varepsilon^{p+1} r_\varepsilon^{D,p} \tag{10}$$

with

$$\varepsilon^{p+1} r_\varepsilon^{D,p} := \sum_{k=p+1}^{p+4}\varepsilon^k \sum_{q=1}^{p} \sum_{n=-q}^{q} \sum_{\mu+2\nu=k-q} c_{n\mu\nu}\partial_\tau^\nu\partial_\xi^\mu A_{q,n}\mathbf{E}^n, \tag{11}$$

where $\mu, \nu \in \mathbb{N}_0$ and

$$c_{n\mu\nu} = \gamma_{\mu\nu}[\mathrm{i}n\omega(\vartheta_0)]^{2-\mu-\nu}(-c_{\mathrm{gr}})^\mu, \quad \gamma_{\mu\nu} = \begin{cases} 0 & \text{for } \mu+\nu > 2, \\ 2 & \text{for } \mu+\nu \le 2 \text{ and } \mu = 1 \text{ or } \nu = 1, \\ 1 & \text{else.} \end{cases}$$

In order to treat the terms on the right-hand side of (9) we introduce for convenience the expression

$$\begin{aligned} \partial_j^{\pm m} X_\varepsilon^{A,p} &:= \pm\,[(X_\varepsilon^{A,p})_{j\pm m} - (X_\varepsilon^{A,p})_j] \\ &= \pm\sum_{k=1}^{p}\varepsilon^k \sum_{n=-k}^{k} [A_{k,n}(\tau, \xi \pm \varepsilon m)\mathrm{e}^{\pm \mathrm{i}mn\vartheta_0} - A_{k,n}]\mathbf{E}^n. \end{aligned}$$

By Taylor expansion we obtain

$$\partial_j^{\pm m} X_\varepsilon^{A,p} = \pm\sum_{k=1}^{2p}\varepsilon^k \sum_{q=\max\{1,k-p\}}^{\min\{k,p\}} \sum_{n=-q}^{q} f_{q(k-q)n}^{\pm m}\mathbf{E}^n \tag{12}$$

with

$$f_{q0n}^{\pm m} = (\mathrm{e}^{\pm \mathrm{i}mn\vartheta_0} - 1)A_{q,n}, \qquad f_{qrn}^{\pm m} = \mathrm{e}^{\pm \mathrm{i}mn\vartheta_0}\frac{(\pm m)^r}{r!}\partial_\xi^r A_{q,n} \quad \text{for} \quad r = 1, \dots, p-1,$$

$$f_{qpn}^{\pm m} = \mathrm{e}^{\pm \mathrm{i}mn\vartheta_0}\frac{(\pm m)^p}{p!}\partial_\xi^p A_{q,n}(\tau, \xi \pm \theta_{pqn}^{\pm\varepsilon m}\varepsilon m), \quad \text{where} \quad \theta_{pqn}^{\pm\varepsilon m} \in (0,1).$$

Using

$$\sum_{m=1}^{M}\alpha_{m,1}(f_{k0n}^{+m} + f_{k0n}^{-m}) - \beta_1 A_{k,n} = -\omega^2(n\vartheta_0)A_{k,n}$$

and

$$\sum_{m=1}^{M}\alpha_{m,1}(f_{qrn}^{+m} + f_{qrn}^{-m}) = \frac{(-\mathrm{i})^{r+2}}{r!}\left.\frac{\mathrm{d}^r\omega^2(\vartheta)}{\mathrm{d}\vartheta^r}\right|_{\vartheta=n\vartheta_0}\partial_\xi^r A_{q,n} \quad \text{for } r = 1, \dots, p-1,$$

we obtain for the linear part of the right-hand side

$$\begin{aligned} L_j X_\varepsilon^{A,p} &= \sum_{m=1}^{M} \alpha_{m,1}(\partial_j^{+m} X_\varepsilon^{A,p} - \partial_j^{-m} X_\varepsilon^{A,p}) - \beta_1 (X_\varepsilon^{A,p})_j \\ &= \sum_{k=1}^{p} \varepsilon^k \sum_{q=1}^{k} \sum_{n=-q}^{q} \frac{(-\mathrm{i})^{k-q+2}}{(k-q)!} \left. \frac{\mathrm{d}^{k-q}\omega^2(\vartheta)}{\mathrm{d}\vartheta^{k-q}} \right|_{\vartheta=n\vartheta_0} \partial_\xi^{k-q} A_{q,n} \mathbf{E}^n + \varepsilon^{p+1} r_\varepsilon^{L,p} \end{aligned} \tag{13}$$

with

$$\varepsilon^{p+1} r_\varepsilon^{L,p} := \sum_{k=p+1}^{2p} \varepsilon^k \sum_{q=k-p}^{p} \sum_{n=-q}^{q} \sum_{m=1}^{M} \alpha_{m,1} (f_{q(k-q)n}^{+m} + f_{q(k-q)n}^{-m}) \mathbf{E}^n. \tag{14}$$

In order to treat the nonlinear terms of (9) we split (12) into two parts with respect to orders of ε: $\partial_j^{\pm m} X_\varepsilon^{A,p} = \partial_j^{\pm m} X_\varepsilon^{A,\le p} + \partial_j^{\pm m} X_\varepsilon^{A,>p}$ with

$$\partial_j^{\pm m} X_\varepsilon^{A,\le p} := \pm \sum_{k=1}^{p} \varepsilon^k \sum_{q=1}^{k} \sum_{n=-q}^{q} f_{q(k-q)n}^{\pm m} \mathbf{E}^n.$$

Then, the nonlinear part of the right-hand side of equation (9) reads

$$\begin{aligned} N_j(X_\varepsilon^{A,p}) = &\sum_{s=2}^{p} \{ \sum_{m=1}^{M} \alpha_{m,s} [(\partial_j^{+m} X_\varepsilon^{A,\le p})^s - (\partial_j^{-m} X_\varepsilon^{A,\le p})^s] - \beta_s (X_\varepsilon^{A,p})_j^s \} \\ &+ \sum_{s=2}^{p} \sum_{m=1}^{M} \alpha_{m,s} \sum_{\sigma=1}^{s} \binom{s}{\sigma} [(\partial_j^{+m} X_\varepsilon^{A,\le p})^{s-\sigma} (\partial_j^{+m} X_\varepsilon^{A,>p})^\sigma \\ &\qquad - (\partial_j^{-m} X_\varepsilon^{A,\le p})^{s-\sigma} (\partial_j^{-m} X_\varepsilon^{A,>p})^\sigma] \\ &+ \sum_{m=1}^{M} [v_{m,p}(\partial_j^{+m} X_\varepsilon^{A,p}) - v_{m,p}(\partial_j^{-m} X_\varepsilon^{A,p})] - w_p[(X_\varepsilon^{A,p})_j] \end{aligned}$$

with

$$v_{m,p}(d) := V_m'(d) - \sum_{s=1}^{p} \alpha_{m,s} d^s, \quad w_p(x) := W'(x) - \sum_{s=1}^{p} \beta_s x^s. \tag{15}$$

In the following we use the general formula

$$\sum_{s=2}^{p} (\sum_{k=1}^{p} \varepsilon^k a_k)^s = \sum_{k=2}^{p} \varepsilon^k \sum_{s=2}^{k} \sum_{|(i)_s|=k} a_{(i)_s} + \sum_{k=p+1}^{p^2} \varepsilon^k \sum_{s=[(k-1)/p]+1}^{k} \sum_{|(i)_s|=k} a_{(i)_s},$$

where $(i)_s := (i_1, \dots, i_s)$ with $i_t \in \{1, \dots, p\}$, $|(i)_s| := \sum_{t=1}^{s} i_t$ and $a_{(i)_s} := \prod_{t=1}^{s} a_{i_t}$. Applying this formula on $a_k := \sum_{n=-k}^{k} A_{k,n} \mathbf{E}^n$, we obtain

$$\begin{aligned} \sum_{s=2}^{p} \beta_s (X_\varepsilon^{A,p})^s = &\sum_{k=2}^{p} \varepsilon^k \sum_{n=-k}^{k} \sum_{s=2}^{k} \beta_s \sum_{\substack{|(i,\nu)_s|=(k,n), \\ (|\nu|)_s \le (i)_s}} A_{(i,\nu)_s} \mathbf{E}^n \\ &+ \sum_{k=p+1}^{p^2} \varepsilon^k \sum_{s=[(k-1)/p]+1}^{k} \beta_s \sum_{|(i)_s|=k} a_{(i)_s} \end{aligned}$$

with $|(i,\nu)_s| = (k,n) :\Leftrightarrow (\sum_{t=1}^s i_t{=}k \text{ and} \sum_{t=1}^s \nu_t{=}n)$, $(|\nu|)_s \le (i)_s :\Leftrightarrow |\nu_t| \le i_t$ and $A_{(i,\nu)_s} := \prod_{t=1}^s A_{i_t,\nu_t}$. Analogously, since

$$\partial_j^{\pm m} X_\varepsilon^{A,\le p} = \sum_{k=1}^p \varepsilon^k \sum_{n=-k}^k \sum_{q=|n|}^k d_{(k-q)n}^{\pm m} \partial_\xi^{k-q} A_{q,n} \mathbf{E}^n$$

(by $\sum_{q=1}^k \sum_{n=-q}^q a_{qn} = \sum_{n=-k}^k \sum_{q=\max\{1,|n|\}}^k a_{qn}$) with

$$d_{0n}^{\pm m} := \pm(\mathrm{e}^{\pm \mathrm{i} mn\vartheta_0}-1), \qquad d_{rn}^{\pm m} := \pm \mathrm{e}^{\pm \mathrm{i} mn\vartheta_0} \frac{(\pm m)^r}{r!} \quad \text{for } r = 1,\dots,p{-}1,$$

by setting $b_k^{\pm m} := \sum_{n=-k}^k \sum_{q=\max\{1,|n|\}}^k d_{(k-q)n}^{\pm m} \partial_\xi^{k-q} A_{q,n} \mathbf{E}^n$, we obtain

$$\begin{aligned}\sum_{s=2}^p \alpha_{m,s} (\partial_j^{\pm m} X_\varepsilon^{A,\le p})^s &= \\ &= \sum_{k=2}^p \varepsilon^k \sum_{n=-k}^k \sum_{s=2}^k \alpha_{m,s} \sum_{\substack{|(i,\nu)_s|=(k,n),\\ (|\nu|)_s \le (i)_s}} \sum_{\substack{(\max\{1,|\nu|\})_s\\ \le (q)_s \le (i)_s}} d_{(i-q,\nu)_s}^{\pm m} \partial_\xi^{(i-q)_s} A_{(q,\nu)_s} \mathbf{E}^n \\ &\quad + \sum_{k=p+1}^{p^2} \varepsilon^k \sum_{s=[(k-1)/p]+1}^k \alpha_{m,s} \sum_{|(i)_s|=k} b_{(i)_s}^{\pm m}\end{aligned}$$

with $(\max\{1,|\nu|\})_s := (\max\{1,|\nu_1|\},\dots,\max\{1,|\nu_s|\})$, $(i{-}q)_s := (i_1{-}q_1,\dots,i_s{-}q_s)$ and

$$d_{(r,\nu)_s}^{\pm m} := \prod_{t=1}^s d_{r_t\nu_t}^{\pm m}, \qquad \partial_\xi^{(r)_s} A_{(q,\nu)_s} := \prod_{t=1}^s \partial_\xi^{r_t} A_{q_t,\nu_t}.$$

Thus, we can write the nonlinear terms of (9) as a sum over orders of ε:

$$\begin{aligned}N_j(X_\varepsilon^{A,p}) = &\sum_{k=2}^p \varepsilon^k \sum_{n=-k}^k \sum_{s=2}^k \sum_{\substack{|(i,\nu)_s|=(k,n),\\ (|\nu|)_s \le (i)_s}} \sum_{\substack{(\max\{1,|\nu|\})_s\\ \le (q)_s \le (i)_s}} (-1) e_{(i-q,\nu)_s}^{M,s} \partial_\xi^{(i-q)_s} A_{(q,\nu)_s} \mathbf{E}^n \\ &+ \varepsilon^{p+1} r_\varepsilon^{N,p} \qquad (16)\end{aligned}$$

with

$$e_{(r,\nu)_s}^{M,s} := \begin{cases} \sum_{m=1}^M \alpha_{m,s} (\prod_{t=1}^s d_{r_t\nu_t}^{-m} - \prod_{t=1}^s d_{r_t\nu_t}^{+m}) & \text{for } (r)_s \ne (0)_s, \\ \sum_{m=1}^M \alpha_{m,s} (\prod_{t=1}^s d_{0\nu_t}^{-m} - \prod_{t=1}^s d_{0\nu_t}^{+m}) + \beta_s & \text{for } (r)_s = (0)_s \end{cases}$$

and

$$\begin{aligned}\varepsilon^{p+1} r_\varepsilon^{N,p} := &\sum_{k=p+1}^{p^2} \varepsilon^k \sum_{s=[(k-1)/p]+1}^k \sum_{|(i)_s|=k} [\sum_{m=1}^M \alpha_{m,s} (b_{(i)_s}^{+m} - b_{(i)_s}^{-m}) - \beta_s a_{(i)_s}] \\ &+ \sum_{s=2}^p \sum_{m=1}^M \alpha_{m,s} \sum_{\sigma=1}^s \binom{s}{\sigma} [(\partial_j^{+m} X_\varepsilon^{A,\le p})^{s-\sigma} (\partial_j^{+m} X_\varepsilon^{A,>p})^\sigma \\ &\qquad - (\partial_j^{-m} X_\varepsilon^{A,\le p})^{s-\sigma} (\partial_j^{-m} X_\varepsilon^{A,>p})^\sigma]\end{aligned}$$

$$+\sum_{m=1}^{M}[v_{m,p}(\partial_j^{+m}X_\varepsilon^{A,p})-v_{m,p}(\partial_j^{-m}X_\varepsilon^{A,p})]-w_p[(X_\varepsilon^{A,p})_j]. \tag{17}$$

Hence, equating the coefficients of the left- and right-hand side for each term $\varepsilon^k\mathbf{E}^n$ with $k=1,\dots,p$ and $n=0,\dots,k$ (the terms for $n=-k,\dots,-1$ can be omitted since they are just the complex conjugates of the terms for $n=1,\dots,k$), we obtain the equations that determine the functions $A_{k,n}$

$$\begin{aligned}
&\delta_n(\vartheta_0)A_{k,n}\\
&=\sum_{q=\max\{1,n\}}^{k-1}\{\sum_{\mu+2\nu=k-q}c_{n\mu\nu}\partial_\tau^\nu\partial_\xi^\mu A_{q,n}+\frac{(-\mathrm{i})^{k-q}}{(k-q)!}\,\frac{\mathrm{d}^{k-q}\omega^2(\vartheta)}{\mathrm{d}\vartheta^{k-q}}\bigg|_{\vartheta=n\vartheta_0}\partial_\xi^{k-q}A_{q,n}\}\\
&\quad+\sum_{s=2}^{k}\sum_{\substack{|(i,\nu)_s|=(k,n),\\(|\nu|)_s\le(i)_s}}\sum_{\substack{(\max\{1,|\nu|\})_s\\ \le(q)_s\le(i)_s}}e^{M,s}_{(i-q,\nu)_s}\partial_\xi^{(i-q)_s}A_{(q,\nu)_s}
\end{aligned}\tag{18}$$

with $\delta_n(\vartheta_0):=n^2\omega^2(\vartheta_0)-\omega^2(n\vartheta_0)$.

By this formalism we can calculate hierarchically the determining equations for the functions $A_{k,n}$ of the approximation (8) with $p=3$ and $p=4$ in which we are interested here. Note, that it is $A=A_{1,1}$, $A_{k,-n}=\overline{A}_{k,n}$. Thus, for $k=1$, $n=0,1$ we obtain only the equation $-\omega^2(0)A_{1,0}=0$ which yields $A_{1,0}=0$, since $\omega^2(0)=\beta_1>0$ by (SC). The function $A=A_{1,1}$ remains undetermined. For $k=2$, $n=0,1,2$ we obtain (with $A_{1,0}=0$)

$$\begin{aligned}
-\omega^2(0)A_{2,0}&=2\beta_2|A|^2,\\
0&=2\mathrm{i}\omega(\vartheta_0)[c_{\mathrm{gr}}+\omega'(\vartheta_0)]\partial_\xi A,\\
\delta_2(\vartheta_0)A_{2,2}&=aA^2
\end{aligned}$$

with $a:=4\mathrm{i}\sum_{m=1}^M\alpha_{m,2}\sin(m\vartheta_0)[1-\cos(m\vartheta_0)]+\beta_2$. The second equation yields $c_{\mathrm{gr}}=-\omega'(\vartheta_0)$, since $\omega(\vartheta_0)\neq0$ by (SC). The function $A_{2,1}$ remains undetermined. In the following we use the abbreviation

$$\gamma_{\kappa\lambda\mu\nu}:=\sum_{m=1}^{M}\alpha_{m,\kappa}m^\lambda\{2\mathrm{i}\sin(m\vartheta_0)\}^\mu\{2[1-\cos(m\vartheta_0)]\}^\nu.$$

Hence, $a=\gamma_{2011}+\beta_2$. Using the results we obtained until now, the equations for $k=3$ read

$$\begin{aligned}
-\omega^2(0)A_{3,0}&=2\beta_2(A\overline{A}_{2,1}+\text{c.c.})-2\gamma_{2101}(A\partial_\xi\overline{A}+\text{c.c.}),\\
0&=2\mathrm{i}\omega(\vartheta_0)\partial_\tau A-\omega(\vartheta_0)\omega''(\vartheta_0)\partial_\xi^2A-\rho A|A|^2,\\
\delta_2(\vartheta_0)A_{3,2}&=2bA\partial_\xi A+2aAA_{2,1},\\
\delta_3(\vartheta_0)A_{3,3}&=cA^3
\end{aligned}$$

with

$$\begin{aligned}
\rho&:=2(\gamma_{2011}^2-\beta_2^2)/\delta_2+4\beta_2^2/\beta_1-3(\gamma_{3002}+\beta_3), \qquad (19)\\
b&:=\gamma_{1111}a/\delta_2+3\gamma_{2101}-\gamma_{2102},\\
c&:=2(3\gamma_{2011}-\gamma_{2012}+\beta_2)a/\delta_2-3\gamma_{3002}+\gamma_{3003}+\beta_3.
\end{aligned}$$

The function $A_{3,1}$ remains undetermined. Note, that the equation for $k=3$, $n=1$ is the nonlinear Schrödinger equation (5) which determines the evolution of A. Thus, if we are interested only in the formal derivation of this equation we can insert in

(1) the improved approximation (8) for $p = 3$ and stop here (and set $A_{3,1} \equiv 0$), since at this stage all the functions $A_{k,n}$ of our approximation $X_\varepsilon^A = X_\varepsilon^{A,1}$, namely $A_{1,0}$ and A, are determined.

However, as we will see later on, we need also the approximation $X_\varepsilon^{A,2}$. In order to determine $A_{2,1}$ we have to insert the improved approximation $X_\varepsilon^{A,4}$ into (1) and calculate by the formalism (18) the functions $A_{4,n}$: By using the previous results, we obtain

$$\begin{aligned}
-\omega^2(0)A_{4,0} &= d_1\partial_\xi^2|A|^2 + [\gamma_{2210}(\partial_\xi^2 A)\overline{A}\text{+c.c.}] - 2\gamma_{2101}[\partial_\xi(A_{2,1}\overline{A})\text{+c.c.}] \\
&\quad + 2\beta_2|A_{2,1}|^2 + 2\beta_2(A_{3,1}\overline{A}\text{+c.c.}) + d_2|A|^4, \\
0 &= 2\mathrm{i}\omega(\vartheta_0)\partial_\tau A_{2,1} - \omega(\vartheta_0)\omega''(\vartheta_0)\partial_\xi^2 A_{2,1} - \rho(2|A|^2A_{2,1}+A^2\overline{A}_{2,1}) \\
&\quad + 2\omega'(\vartheta_0)\partial_\tau\partial_\xi A + (\mathrm{i}/6)[\omega^2(\vartheta_0)]'''\partial_\xi^3 A - 2e|A|^2\partial_\xi A, \\
\delta_2(\vartheta_0)A_{4,2} &= 8\mathrm{i}\omega(\vartheta_0)(a/\delta_2)A\partial_\tau A + f_1\partial_\xi(A\partial_\xi A) + \gamma_{2210}A\partial_\xi^2 A + 2b\partial_\xi(AA_{2,1}) \\
&\quad + a(A_{2,1}^2+2AA_{3,1}) + f_2A^2|A|^2, \\
\delta_3(\vartheta_0)A_{4,3} &= gA^2\partial_\xi A + 3cA^2A_{2,1}, \\
\delta_4(\vartheta_0)A_{4,4} &= hA^4
\end{aligned}$$

with

$$\begin{aligned}
d_1 &:= [\gamma_{1201}-\omega(\vartheta_0)\omega''(\vartheta_0)]\beta_2/\beta_1, \\
d_2 &:= 2\beta_2|a/\delta_2|^2 + 4\beta_2^3/\beta_1^2 + 6\beta_3\beta_2(1/\delta_2-2/\beta_1) + 6\beta_4, \\
e &:= 2[2(3\gamma_{2101}-\gamma_{2102})\gamma_{2011}+\gamma_{1111}(\gamma_{2011}^2-\beta_2^2)/\delta_2]/\delta_2 + 3\gamma_{3111}, \qquad (20) \\
f_1 &:= [3\gamma_{1201}-\gamma_{1202}-2\omega(\vartheta_0)\omega''(\vartheta_0)]a/\delta_2 + 2\gamma_{1111}b/\delta_2 + \gamma_{2211} - 2\gamma_{2210}, \\
f_2 &:= -4\beta_2^2a/\beta_1\delta_2 + 2(\beta_2-3\gamma_{2011}+\gamma_{2012})c/\delta_3 + 6(\beta_3-\gamma_{3021})a/\delta_2 - 6\beta_3\beta_2/\beta_1 \\
&\quad + 4(\gamma_{4012}+\beta_4), \\
g &:= -3\gamma_{1130}c/\delta_3 + 2(3\gamma_{2103}-16\gamma_{2102}+18\gamma_{2101})a/\delta_2 + 4(3\gamma_{2011}-\gamma_{2012}+\beta_2)b/\delta_2 \\
&\quad + 3(2\gamma_{3111}-\gamma_{3112}), \\
h &:= (\gamma_{2031}-2\gamma_{2030}+\beta_2)a^2/\delta_2^2 + 2(\gamma_{2031}-\gamma_{2012}+6\gamma_{2011}+\beta_2)c/\delta_3 \\
&\quad + 3(2\gamma_{3021}-\gamma_{3022}+\beta_3)a/\delta_2 + \gamma_{4013}-2\gamma_{4012}+\beta_4.
\end{aligned}$$

The function $A_{4,1}$ remains undetermined. Since the equation for $k = 4$, $n = 1$ determines $A_{2,1}$, we now have determined all the functions $A_{k,n}$ of the improved approximation $X_\varepsilon^{A,2}$ in which we are interested and can stop here, setting $A_{3,1} \equiv A_{4,1} \equiv 0$.

Thus, we have established the following result:

Theorem 1. *If the microscopic oscillator chain (1) has for all $\varepsilon \in (0, \varepsilon_0)$ solutions of the form*

$$x_j(t) = (X_\varepsilon^A)_j(t) + \mathcal{O}(\varepsilon^2) \quad \text{with} \quad (X_\varepsilon^A)_j(t) = \varepsilon A(\tau,\xi)\mathbf{E}(t,j) + \text{c.c.},$$

where $\tau = \varepsilon^2 t$, $\xi = \varepsilon(j+\omega' t)$ and $A : [0,\tau_0]\times\mathbb{R}\to\mathbb{C}$ is a smooth function, then A necessarily has to satisfy the NLSE (5). Analogously, if (1) has for all $\varepsilon \in (0,\varepsilon_0)$ solutions of the form

$$x = X_\varepsilon^{A,2} + \mathcal{O}(\varepsilon^3) \quad \text{with} \quad X_\varepsilon^{A,2} = \varepsilon A\mathbf{E} + \varepsilon^2\left(\frac{\beta_2}{\delta_0}|A|^2 + A_{2,1}\mathbf{E} + \frac{a}{\delta_2}A^2\mathbf{E}^2\right) + \text{c.c.},$$

where $\delta_n(\vartheta_0) := n^2\omega^2(\vartheta_0) - \omega^2(n\vartheta_0)$*,* $a := 4\mathrm{i}\sum_{m=1}^{M}\alpha_{m,2}\sin(m\vartheta_0)[1-\cos(m\vartheta_0)]+\beta_2$*, then* A *and* $A_{2,1} : [0,\tau_0]\times\mathbb{R}\to\mathbb{C}$*,* $(\tau,\xi)\mapsto A_{2,1}(\tau,\xi)$*, necessarily have to satisfy the NLSE (5) and equation (7), respectively.*

We call this result a formal derivation, since the existence of solutions satisfying such expansions is not clear at all. Theorem 1 constitutes merely a necessary condition on the amplitude A (and $A_{2,1}$) of the ansatz (4) (and (6), respectively), in the sense that if solutions of these forms exist, their amplitudes have to satisfy the NLSE (5) (and (7)). The purpose of the justification of the NLSE (and (7) in the second case) is to show that solutions starting in modulation form at time $t = 0$ will indeed maintain this modulation form on suitably long time scales.

Note that in order to determine the functions $A_{k,n}$ of the improved approximation $X_\varepsilon^{A,p}$ (cf. (8)) by the formalism (18), it has to hold $\delta_0(\vartheta_0) = -\omega^2(0) \neq 0$ and $\delta_n(\vartheta_0) = n^2\omega^2(\vartheta_0)-\omega^2(n\vartheta_0) \neq 0$ for $n = 2,\ldots,p$. Under the stability condition (SC): $\omega^2(\vartheta) > 0$ for all $\vartheta \in (-\pi,\pi]$, this is satisfied if the nonresonance conditions of second order

$$n\omega(\vartheta_0)-\omega(n\vartheta_0) \neq 0 \quad \text{for } n = 2,\ldots,p \qquad (\mathrm{NR2})^p_{\vartheta_0}$$

hold. In the case of only nearest-neighbor interactions ($M = 1$ in (1)) one can formulate these conditions as conditions on the coefficients of the potentials V and W, and relate them to the nonresonance conditions $(\mathrm{NR3})_{\mathrm{unif}}$ and $(\mathrm{NR3})_{\vartheta_0}$ which are necessary for the justification of the NLSE.

Proposition 1. *In the case of nearest-neighbor interactions (*$M = 1$ *with* $\alpha_1 := \alpha_{1,1}$*) the following relations hold true:*

$$(\mathrm{SC}) \Longleftrightarrow \min\{\beta_1, 4\alpha_1+\beta_1\} > 0, \qquad (21)$$

For $n = 2,3,4$*:* $\forall\, \vartheta \in (-\pi,\pi]$*:* $\delta_n(\vartheta) > 0 \Longleftrightarrow \min\{\beta_1, \frac{2^{n+2}}{n^2-1}\alpha_1+\beta_1\} > 0, \qquad (22)$

$$(\mathrm{SC}) \text{ and } (\mathrm{NR3})_{\vartheta_0} \Longrightarrow (\mathrm{NR2})^4_{\vartheta_0}, \qquad (23)$$

$$(\mathrm{SC}) \text{ and } (\mathrm{NR3})_{\mathrm{unif}} \Longleftrightarrow \min\{\beta_1, (16/3)\alpha_1+\beta_1\} > 0. \qquad (24)$$

Remark 1. First, the stability condition (SC) restricts us by (21) to the harmonic coefficients $\beta_1 > 0$ and $\alpha_1/\beta_1 > -1/4$. Second, for $\alpha_1/\beta_1 > -3/16$ we obtain by (22) and (24) that $(\mathrm{NR3})_{\mathrm{unif}}$ implies $(\mathrm{NR2})^3_{\vartheta_0}$. Finally, from (22) it follows that in order to guarantee $(\mathrm{NR2})^4_{\vartheta_0}$ we have to require $2\omega(\vartheta_0) \neq \omega(2\vartheta_0)$ in the case $\alpha_1/\beta_1 \in (-15/64, -3/16]$ and $2\omega(\vartheta_0) \neq \omega(2\vartheta_0)$, $4\omega(\vartheta_0) \neq \omega(4\vartheta_0)$ in the case $\alpha_1/\beta_1 \in (-1/4, -15/64]$. By (23), both conditions follow from $(\mathrm{NR3})_{\vartheta_0}$.

Proof of Proposition 1. Ad (21): The equivalence follows immediately from (DR) and (SC).

Ad (22): For $n = 2,3,4$ we have

$$\delta_n(\vartheta) = n^2\omega^2(\vartheta)-\omega^2(n\vartheta) = 2^n\alpha_1 g_n(\cos\vartheta) + (n^2-1)\beta_1 \qquad (25)$$

with

$$g_2(c) = (1-c)^2, \quad g_3(c) = c^3-3c+2, \quad g_4(c) = (1-c)^2[(1+c)^2+1] \quad \text{for } c \in [-1,1]$$

and $\min g_n = 0$, $\max g_n = 4$. This yields (22).

Note, that since $(\mathrm{NR2})^4_{\vartheta_0}$ is equivalent to $\delta_n(\vartheta_0) \neq 0$ for $n = 2,3,4$ (if (SC): $\omega^2(\vartheta) > 0$ for all $\vartheta \in (-\pi,\pi]$ holds), it is, by (25), equivalent to

$$\frac{\alpha_1}{\beta_1} \neq \frac{-(n^2-1)}{2^n g_n(\cos\vartheta_0)} =: f_n(\cos\vartheta_0). \qquad (26)$$

In Figure 1 we plotted $f_n(c)$ for $n = 2$ (black), $n = 3$ (dashed light grey) and $n = 4$ (dotted dark grey) over $c = \cos\vartheta_0 \in [-1, 0.25]$ (left) and $c \in [-1, -0.4]$ (right). In relation to Remark 1, the plots illustrate that f_3 approximates $-1/4$ from below, while $f_1(c)$ and $f_2(c)$ take also values in $[-1/4, -3/16)$ for $c = \cos\vartheta_0$ close to -1. Hence, for such ϑ_0, the coefficients α_1/β_1 for which (26) (and thus $(\mathrm{NR2})^4_{\vartheta_0}$) holds, depend on ϑ_0. For $\alpha_1/\beta_1 > -3/16$ (26) is fulfilled for all ϑ_0.

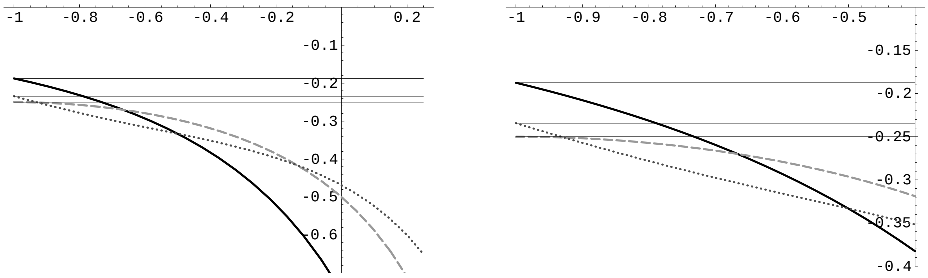

FIGURE 1. Plots over $c \in [-1, 0.25]$ (left) and $c \in [-1, -0.4]$ (right) of $f_n(c)$ defined by (26) for $n = 2$ (black), $n = 3$ (dashed light grey), $n = 4$ (dotted dark grey).

Ad (23): We have to show $n\omega(\vartheta_0) - \omega(n\vartheta_0) \neq 0$ for $n = 2, 3, 4$. For $n = 2$, setting $s = 2$, $\sigma = 1$, $\theta = -\vartheta_0$ into $(\mathrm{NR3})_{\vartheta_0}$ and using $\omega(\vartheta) = \omega(-\vartheta)$ we obtain $|2\omega(\vartheta_0) - \omega(2\vartheta_0)| \geq C^{\mathrm{NR}}_{\vartheta_0} > 0$. For $n = 3$, we obtain from (21) and (22) $\delta_3(\vartheta_0) = 9\omega^2(\vartheta_0) - \omega^2(3\vartheta_0) > 0$. Since $\omega(\vartheta) > 0$ for all ϑ, it follows $3\omega(\vartheta_0) - \omega(3\vartheta_0) > 0$. Finally, for $n = 4$, let us assume that $4\omega(\vartheta_0) = \omega(4\vartheta_0)$, which can be written as $f(-3\vartheta_0) = \omega(3\vartheta_0) - 3\omega(\vartheta_0)$ with $f(\theta) := \omega(\vartheta_0) + \omega(\theta) - \omega(\vartheta_0 - \theta)$. By what we showed in the previous case $n = 3$, it is $f(-3\vartheta_0) < 0$. Hence, since $f(0) = \omega(0) > 0$ and f is continuous, there exists a $\tilde{\theta}$ with $|\tilde{\theta}| \in (0, 3|\vartheta_0|)$ and $f(\tilde{\theta}) = 0$. But this contradicts $(\mathrm{NR3})_{\vartheta_0}$ with $s = 2$ and $\sigma = 1$.

Ad (24): By (SC) and $\omega(\vartheta) = \omega(-\vartheta) > 0$, $(\mathrm{NR3})_{\mathrm{unif}}$ can be reduced to

$$\exists\, C^{\mathrm{NR}}_{\mathrm{unif}} > 0 : \quad \inf_{\vartheta,\theta \in (-\pi,\pi]} [\omega(\vartheta) + \omega(\vartheta - \theta) - \omega(\theta)] \geq C^{\mathrm{NR}}_{\mathrm{unif}} > 0.$$

For $\alpha_1 = 0$ we take $C^{\mathrm{NR}}_{\mathrm{unif}} = \beta_1^{1/2} > 0$. For $\alpha_1 < 0$ we have $\mu_- := \min\{\omega(\vartheta) : \vartheta \in [-\pi, \pi]\} = \omega(\pm\pi) = (4\alpha_1 + \beta_1)^{1/2} > 0$ and $\mu_+ := \max\{\omega(\vartheta) : \vartheta \in [-\pi, \pi]\} = \omega(0) = \beta_1^{1/2} > 0$. Thus,

$$\inf_{\vartheta,\theta \in (-\pi,\pi]} [\omega(\vartheta) + \omega(\vartheta - \theta) - \omega(\theta)] \geq 2\mu_- - \mu_+ = \frac{16\alpha_1 + 3\beta_1}{2\mu_- + \mu_+}.$$

Equality is attained for $\vartheta = \pi$, $\theta = 0$. Hence, $(\mathrm{NR3})_{\mathrm{unif}}$ is satisfied in the case $\alpha_1 < 0$, if and only if $16\alpha_1 + 3\beta_1 > 0$. For $\alpha_1 > 0$ equality is not attained, since it is $\mu_- = \omega(0)$ and $\mu_+ = \omega(\pm\pi)$ but there exists no $\vartheta \in \mathbb{T}$ with $\vartheta = \vartheta \mp \pi = 0$. In this case, we show $\omega(\vartheta) + \omega(\vartheta - \theta) - \omega(\theta) \geq C^{\mathrm{NR}}_{\mathrm{unif}} > 0$ for all $\vartheta, \theta \in (-\pi, \pi]$, and (24) is proved:

Since $\omega(\vartheta) = 2\alpha_1^{1/2} \left[\sin^2(\vartheta/2) + \gamma\right]^{1/2}$ with $\gamma := \beta_1/4\alpha_1 > 0$ is 2π-periodic and continuous, it suffices to show

$$f(\vartheta, \theta) := \left[\sin^2[(\vartheta - \theta)/2] + \gamma\right]^{1/2} + \left[\sin^2(\vartheta/2) + \gamma\right]^{1/2} - \left[\sin^2(\theta/2) + \gamma\right]^{1/2} > 0$$

for all $\vartheta, \theta \in (-\pi, \pi]$. First, we prove

$$\widetilde{f}(\vartheta,\theta) := |\sin[(\vartheta-\theta)/2]| + \left[\sin^2(\vartheta/2)+\gamma\right]^{1/2} - \left[\sin^2(\theta/2)+\gamma\right]^{1/2} \geq 0.$$

This estimate is sharp, since $\widetilde{f}(\theta,\theta) = 0$. With $\beta := \sin^2(\vartheta/2)$, $\beta_0 := \sin^2(\theta/2)$ it is equivalent to

$$\sin^2[(\vartheta-\theta)/2] \geq \beta + \beta_0 + 2\gamma - 2\left[(\beta+\gamma)(\beta_0+\gamma)\right]^{1/2}.$$

From $\sin^2(x-y) = \sin^2 x + \sin^2 y - 2\sin x \sin y \cos(x-y)$, we obtain $\sin^2[(\vartheta-\theta)/2] = \beta + \beta_0 \pm 2\eta$ with $\eta := (\beta\beta_0)^{1/2}\,|\cos[(\vartheta-\theta)/2]| \geq 0$. Thus, we want to prove

$$\left[(\beta+\gamma)(\beta_0+\gamma)\right]^{1/2} \pm \eta \geq \gamma \quad \forall\, \beta, \beta_0 \in [0,1],$$

which surely holds if $\left[(\beta+\gamma)(\beta_0+\gamma)\right]^{1/2} - \eta \geq \gamma$, i.e., if

$$(\beta+\gamma)(\beta_0+\gamma) - (\eta+\gamma)^2 = \beta\beta_0 - \eta^2 + \gamma(\beta+\beta_0 - 2\eta) \geq 0.$$

But this holds true, since $\eta \in [0, (\beta\beta_0)^{1/2}]$. Hence, since

$$f(\vartheta,\theta) = \widetilde{f}(\vartheta,\theta) + \left[\sin^2[(\vartheta-\theta)/2]+\gamma\right]^{1/2} - |\sin[(\vartheta-\theta)/2]|$$

and $\min\limits_{\alpha\in[0,1]}\{(\alpha+\gamma)^{1/2} - \alpha^{1/2}\} = (1+\gamma)^{1/2}-1$, we obtain $f(\vartheta,\theta) \geq \gamma/[(1+\gamma)^{1/2}+1]$, i.e.,

$$\inf_{\vartheta,\theta\in(-\pi,\pi]} [\omega(\vartheta)+\omega(\vartheta-\theta)-\omega(\theta)] \geq \frac{\beta_1}{(4\alpha_1+\beta_1)^{1/2} + 2\alpha_1^{1/2}} = \frac{\mu_-^2}{\mu_+ + (\mu_+^2-\mu_-^2)^{1/2}} > 0.$$

Note, that the minimum of $\widetilde{f}(\vartheta,\theta)$ is attained for $\vartheta = \theta$, which yields $f(\theta,\theta) = \gamma^{1/2}$, whereas the minimum of $f(\vartheta,\theta)-\widetilde{f}(\vartheta,\theta)$ is attained for $\vartheta = \theta \pm \pi$, which yields $f(\theta\pm\pi,\theta) \geq \gamma^{1/2} > (1+\gamma)^{1/2}-1$. Hence, in the last two estimates above we can actually replace $\geq$ by $>$. Note also, that the bound of the last estimate tends to μ_- for $\mu_+-\mu_- \to 0$ ($\alpha_1 \to 0$) and to 0 for $\mu_+-\mu_- \to \infty$ ($\alpha_1 \to \infty$). □

3. **The normal-form transformation.** For the justification of the NLSE (5) we need to rewrite our oscillator chain model (1) as a first-order ordinary system

$$\dot{\widetilde{x}} = \widetilde{L}\widetilde{x} + \widetilde{Q}(\widetilde{x},\widetilde{x}) + \widetilde{M}(\widetilde{x}) \tag{27}$$

in the Banach space $Y := \ell^2\times\ell^2$, where

$$\widetilde{x} = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}, \quad \widetilde{L} = \begin{pmatrix} 0 & 1 \\ L & 0 \end{pmatrix}, \quad \widetilde{Q}(\widetilde{x},\widetilde{y}) = \begin{pmatrix} 0 \\ Q(x,y) \end{pmatrix}, \quad \widetilde{M}(\widetilde{x}) = \begin{pmatrix} 0 \\ M(x) \end{pmatrix}$$

with L defined by (3), and

$$[Q(x,y)]_j := \sum_{m=1}^{M} \alpha_{m,2}[(x_{j+m}-x_j)(y_{j+m}-y_j)-(x_j-x_{j-m})(y_j-y_{j-m})] - \beta_2 x_j y_j, \tag{28}$$

$$[M(x)]_j := \sum_{m=1}^{M} [v_{m,2}(x_{j+m}-x_j) - v_{m,2}(x_j-x_{j-m})] - w_2(x_j) \tag{29}$$

with $v_{m,2}$ and w_2 defined in (15).

On the Banach space Y we use the energy norm

$$\|(x,y)\|_Y^2 := \|x\|_E^2 + \|y\|_{\ell^2}^2 \quad \text{with} \quad \|x\|_E^2 := \sum_{m=1}^{M} \alpha_{m,1} \sum_{j\in\mathbb{Z}} |x_{j+m}-x_j|^2 + \beta_1 \|x\|_{\ell^2}^2 \tag{30}$$

and $\|y\|_{\ell^2}^2 = \sum_{j\in\mathbb{Z}} |y_j|^2$. The norms $\|\cdot\|_{\ell^2}$ and $\|\cdot\|_E$ are equivalent by our stability assumption (SC):

$$\mu_-^2 \|x\|_{\ell^2}^2 \le \|x\|_E^2 \le \mu_+^2 \|x\|_{\ell^2}^2 \tag{31}$$

with $\mu_-^2 := \min\{\omega^2(\vartheta) : \vartheta \in [-\pi,\pi]\}$ and $\mu_+^2 := \max\{\omega^2(\vartheta) : \vartheta \in [-\pi,\pi]\}$, which follows easily by Fourier transformation.

The full oscillator chain is a Hamiltonian system whose solutions make the sum H of kinetic and potential energy

$$H(x,\dot{x}) = \frac{1}{2}\|\dot{x}\|_{\ell^2}^2 + \sum_{j\in\mathbb{Z}} \Big[\sum_{m=1}^{M} V_m(x_{j+m}-x_j) + W(x_j) \Big]$$

constant with respect to time. The norm $\|\cdot\|_Y$ is defined in such a way that its square is twice the quadratic part of H. The flow of the linearized system (3) preserves this norm: The solutions $\widetilde{x} : t \mapsto \widetilde{x}(t) = \mathrm{e}^{t\widetilde{L}}\widetilde{x}(0)$ of (3) satisfy $\|\widetilde{x}(t)\|_Y = \|\widetilde{x}(0)\|_Y$ for all $t \in \mathbb{R}$ (cf. [19, Proposition 3.1]).

The dispersive scaling $\tau = \varepsilon^2 t$ means that a Gronwall estimate for the errors has to be made on the microscopic time interval $[0, \tau_0/\varepsilon^2]$. If $\widetilde{Q}$ vanishes, then the remaining nonlinearity $\widetilde{M}$ in (27) has Lipschitz constant $\mathcal{O}(\varepsilon^2)$ in a ball containing X_ε^A. However, for $\widetilde{Q} \not\equiv 0$ the Lipschitz constant of $\widetilde{Q}+\widetilde{M}$ is $\mathcal{O}(\varepsilon)$. Hence, the Gronwall-type argument for the error control works in a natural way for systems (27) with $\widetilde{Q} \equiv 0$ only, see [19]. To apply the Gronwall argument here we follow the approach devised in [35] and apply a normal-form transformation on the system (27), with the aim to transform it into a system with only cubic and higher order nonlinear terms. We introduce the normal-form transformation $F : Y \to Y$ with

$$\widetilde{y} = F(\widetilde{x}) := \widetilde{x} + B(\widetilde{x},\widetilde{x}), \tag{32}$$

where the bilinear form $B : Y\times Y \to Y$ remains to be determined. Applying this transformation on (27) we obtain

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + \overline{Q}(\widetilde{x},\widetilde{x}) + \overline{M}(\widetilde{x}) \tag{33}$$

with

$$\overline{Q}(\widetilde{x},\widetilde{x}) := -\widetilde{L}B(\widetilde{x},\widetilde{x}) + B(\widetilde{L}\widetilde{x},\widetilde{x}) + B(\widetilde{x},\widetilde{L}\widetilde{x}) + \widetilde{Q}(\widetilde{x},\widetilde{x}), \tag{34}$$

$$\overline{M}(\widetilde{x}) := B(\widetilde{Q}(\widetilde{x},\widetilde{x})+\widetilde{M}(\widetilde{x}),\widetilde{x}) + B(\widetilde{x},\widetilde{Q}(\widetilde{x},\widetilde{x})+\widetilde{M}(\widetilde{x})) + \widetilde{M}(\widetilde{x}). \tag{35}$$

The terms of quadratic order with respect to $\widetilde{x}$ are given by $\overline{Q}$. The terms of cubic and higher order of $\widetilde{x}$ are subsumed by $\overline{M}$.

Now we require for $B = (B_1, B_2)$ to satisfy $\overline{Q}(\widetilde{x},\widetilde{x}) = 0$ for all $\widetilde{x} \in Y$. This is equivalent to

$$\begin{cases} B_2(\widetilde{x},\widetilde{x}) = B_1(\widetilde{L}\widetilde{x},\widetilde{x}) + B_1(\widetilde{x},\widetilde{L}\widetilde{x}), \\ LB_1(\widetilde{x},\widetilde{x}) - B_2(\widetilde{L}\widetilde{x},\widetilde{x}) - B_2(\widetilde{x},\widetilde{L}\widetilde{x}) = Q(x,x). \end{cases}$$

Setting

$$B_2(\widetilde{x},\widetilde{y}) := B_1(\widetilde{L}\widetilde{x},\widetilde{y}) + B_1(\widetilde{x},\widetilde{L}\widetilde{y}), \tag{36}$$

the first equation is fulfilled, and the second reads

$$LB_1(\widetilde{x},\widetilde{x}) - B_1(\widetilde{L}^2\widetilde{x},\widetilde{x}) - 2B_1(\widetilde{L}\widetilde{x},\widetilde{L}\widetilde{x}) - B_1(\widetilde{x},\widetilde{L}^2\widetilde{x}) = Q(x,x),$$

i.e., by $B_1(\widetilde{x},\widetilde{x}) = B_1(x,\dot{x};x,\dot{x})$,

$$LB_1(x,\dot{x};x,\dot{x}) - B_1(Lx,L\dot{x};x,\dot{x}) - 2B_1(\dot{x},Lx;\dot{x},Lx) - B_1(x,\dot{x};Lx,L\dot{x}) = Q(x,x). \tag{37}$$

We determine $B_1 : Y\times Y \to \ell^2$ via its Fourier transform. We denote the Fourier transform of $x \in \ell^2$ by $\widehat{x} \in \mathrm{L}^2(\mathbb{T})$ with $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, where

$$\widehat{x}(\vartheta) = \sum_{j\in\mathbb{Z}} x_j\, \mathrm{e}^{-\mathrm{i}\vartheta j} \quad \text{for } \vartheta \in \mathbb{T}.$$

The inverse of the Fourier transform is given by

$$x_j = \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\vartheta)\, \mathrm{e}^{\mathrm{i}\vartheta j}\, \mathrm{d}\vartheta \quad \text{for } j \in \mathbb{Z}.$$

For the linear operator L defined by (3) we have $\widehat{L}\widehat{x} := \widehat{Lx} :\ \vartheta \mapsto -\omega^2(\vartheta)\widehat{x}(\vartheta)$. Using the convolution

$$\widehat{xy}(\vartheta) = \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\vartheta-\theta)\widehat{y}(\theta)\, \mathrm{d}\theta := (\widehat{x} * \widehat{y})(\vartheta) \quad \text{for } x,y \in \ell^2,\ \vartheta \in \mathbb{T},$$

we obtain for Q defined by (28) the Fourier transform

$$[\widehat{Q}(\widehat{x},\widehat{y})](\vartheta) := \widehat{Q(x,y)}(\vartheta) = \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\vartheta-\theta) q(\vartheta,\theta)\widehat{y}(\theta)\, \mathrm{d}\theta \tag{38}$$

with

$$q(\vartheta,\theta) := 2\mathrm{i}\sum_{m=1}^{M} \alpha_{m,2}\, [\sin(m\vartheta) - \sin(m(\vartheta-\theta)) - \sin(m\theta)] - \beta_2. \tag{39}$$

The Fourier transform of $B_1(\widetilde{x},\widetilde{y})$ with $\widetilde{x}=(x_1,x_2)$, $\widetilde{y}=(y_1,y_2) \in Y=\ell^2\times\ell^2$ has the general form

$$\begin{aligned}[\widehat{B}_1(\widehat{\widetilde{x}},\widehat{\widetilde{y}})](\vartheta) :=& \widehat{B_1(\widetilde{x},\widetilde{y})}(\vartheta)\\ =& \frac{1}{2\pi}\int_{\mathbb{T}} \begin{pmatrix}\widehat{x}_1(\vartheta-\theta), & \widehat{x}_2(\vartheta-\theta)\end{pmatrix} \begin{pmatrix} b_{11}(\vartheta,\theta) & b_{12}(\vartheta,\theta) \\ b_{21}(\vartheta,\theta) & b_{22}(\vartheta,\theta)\end{pmatrix} \begin{pmatrix}\widehat{y}_1(\theta)\\ \widehat{y}_2(\theta)\end{pmatrix} \mathrm{d}\theta\end{aligned}$$

for $\vartheta \in \mathbb{T}$. Thus, the Fourier transform of equation (37)

$$\begin{aligned}&-\omega^2\widehat{B}_1(\widehat{x},\dot{\widehat{x}};\widehat{x},\dot{\widehat{x}}) + \widehat{B}_1(\omega^2\widehat{x},\omega^2\dot{\widehat{x}};\widehat{x},\dot{\widehat{x}}) - 2\widehat{B}_1(\dot{\widehat{x}},-\omega^2\widehat{x};\dot{\widehat{x}},-\omega^2\widehat{x}) + \\ &\qquad + \widehat{B}_1(\widehat{x},\dot{\widehat{x}};\omega^2\widehat{x},\omega^2\dot{\widehat{x}}) = \widehat{Q}(\widehat{x},\widehat{x})\end{aligned}$$

holds for all $(\widehat{x},\dot{\widehat{x}})$ if and only if

$$\alpha(\vartheta,\theta)\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22}\end{pmatrix} - 2\begin{pmatrix}\omega^2(\vartheta-\theta)\omega^2(\theta)b_{22} & -\omega^2(\vartheta-\theta)b_{21} \\ -\omega^2(\theta)b_{12} & b_{11}\end{pmatrix} = \begin{pmatrix} q & 0 \\ 0 & 0\end{pmatrix}$$

with $\alpha(\vartheta,\theta) := \omega^2(\vartheta-\theta) + \omega^2(\theta) - \omega^2(\vartheta)$ holds for all $(\vartheta,\theta) \in (-\pi,\pi]^2$. This yields

$$b_{11} = \frac{\alpha q}{\alpha^2-\beta^2}, \qquad b_{22} = \frac{2q}{\alpha^2-\beta^2}, \qquad b_{12} = b_{21} = 0 \tag{40}$$

with $\beta(\vartheta,\theta) := 2\omega(\vartheta-\theta)\omega(\theta)$, provided that $(\alpha^2-\beta^2)(\vartheta,\theta) \neq 0$ for all $(\vartheta,\theta) \in (-\pi,\pi]^2$. Hence, in this case $B_1 : Y\times Y \to \ell^2$ is given by

$$B_1(\widetilde{x},\widetilde{y}) = b_1(x_1,y_1) + b_2(x_2,y_2) \quad \text{for } \widetilde{x} = (x_1,x_2),\ \widetilde{y} = (y_1,y_2), \tag{41}$$

where $b_i : \ell^2 \times \ell^2 \to \ell^2$, $i = 1, 2$, are defined by

$$[\widehat{b}_i(\widehat{x}_i, \widehat{y}_i)](\vartheta) = \frac{1}{2\pi} \int_{\mathbb{T}} \widehat{x}_i(\vartheta - \theta) b_{ii}(\vartheta, \theta) \widehat{y}_i(\theta) \, \mathrm{d}\theta \quad \text{for } \vartheta \in \mathbb{T} \tag{42}$$

with the b_{ii} determined by (40). From (36) we obtain $B_2 : Y \times Y \to \ell^2$:

$$B_2(\widetilde{x}, \widetilde{y}) = b_1(x_2, y_1) + b_1(x_1, y_2) + b_2(Lx_1, y_2) + b_2(x_2, Ly_1) \tag{43}$$

for $\widetilde{x} = (x_1, x_2)$, $\widetilde{y} = (y_1, y_2)$. This determines $B : Y \times Y \to Y$ with $B = (B_1, B_2)$.

Since

$$\begin{aligned}(\alpha^2 - \beta^2)(\vartheta, \theta) = {} & [\omega(\vartheta - \theta) - \omega(\theta) - \omega(\vartheta)] \, [\omega(\vartheta - \theta) - \omega(\theta) + \omega(\vartheta)] \\ & \times [\omega(\vartheta - \theta) + \omega(\theta) - \omega(\vartheta)] \, [\omega(\vartheta - \theta) + \omega(\theta) + \omega(\vartheta)] \, ,\end{aligned} \tag{44}$$

the condition $(\alpha^2 - \beta^2)(\vartheta, \theta) \neq 0$ is fulfilled for all $(\vartheta, \theta) \in (-\pi, \pi]^2$ if and only if our uniform nonresonance condition is satisfied:

$$\exists \, C^{\mathrm{NR}}_{\mathrm{unif}} > 0 : \quad \inf_{\substack{s, \sigma = 1, 2; \\ \vartheta, \theta \in (-\pi, \pi]}} |\omega(\vartheta) + (-1)^s \omega(\theta) + (-1)^\sigma \omega(\vartheta - \theta)| \geq C^{\mathrm{NR}}_{\mathrm{unif}} > 0. \tag{NR3$_{\mathrm{unif}}$}$$

(In the case of only nearest-neighbor interactions $M = 1$ (NR3)$_{\mathrm{unif}}$ is equivalent to $\min\{\beta_1, (16/3)\alpha_{1,1} + \beta_1\} > 0$, cf. (24) in Proposition 1.)

Since $B = (B_1, B_2)$ is given via (41), (43), in order to obtain an estimate for B, we use the following estimate for a (general) bilinear form $b : \ell^2 \times \ell^2 \to \ell^2$.

Proposition 2. *For a bilinear form $b : \ell^2 \times \ell^2 \to \ell^2$ with the Fourier transform*

$$[\widehat{b}(\widehat{x}, \widehat{y})](\vartheta) = \frac{1}{2\pi} \int_{\mathbb{T}} \widehat{x}(\vartheta - \theta) \beta(\vartheta, \theta) \widehat{y}(\theta) \, \mathrm{d}\theta \quad \textit{for } x, y \in \ell^2 \textit{ and } \vartheta \in \mathbb{T},$$

where $\beta \in \mathrm{H}^3(\mathbb{T} \times \mathbb{T})$, there exists a $c_b > 0$ depending on β, such that the following estimate holds:

$$\|b(x, y)\|_{\ell^2} \leq c_b \|x\|_{\ell^2} \|y\|_{\ell^\infty} \quad \textit{for } x, y \in \ell^2.$$

Proof. The general form of b is given by

$$[b(x, y)]_j = \sum_{k, l \in \mathbb{Z}} b^j_{k,l} x_k y_l \quad \text{for } j \in \mathbb{Z}.$$

Using the translation operator $T : \ell^2 \to \ell^2$ defined by $(Tx)_j := x_{j+1}$, we obtain $Tb(x, y) = b(Tx, Ty)$, since

$$\widehat{Tb(x, y)}(\vartheta) = \mathrm{e}^{\mathrm{i}\vartheta} [\widehat{b}(\widehat{x}, \widehat{y})](\vartheta) = \frac{1}{2\pi} \int_{\mathbb{T}} \mathrm{e}^{\mathrm{i}(\vartheta - \theta)} \widehat{x}(\vartheta - \theta) \beta(\vartheta, \theta) \mathrm{e}^{\mathrm{i}\theta} \widehat{y}(\theta) \, \mathrm{d}\theta = [\widehat{b}(\widehat{Tx}, \widehat{Ty})](\vartheta).$$

Thus, from

$$\sum_{k, l \in \mathbb{Z}} b^{j+1}_{k,l} x_k y_l = [Tb(x, y)]_j = [b(Tx, Ty)]_j = \sum_{k, l \in \mathbb{Z}} b^j_{k,l} x_{k+1} y_{l+1} = \sum_{k, l \in \mathbb{Z}} b^j_{k-1, l-1} x_k y_l$$

we obtain $b^j_{k,l} = b^{j-1}_{k-1,l-1}$ and, hence, iteratively $b^j_{k,l} = b^0_{k-j,l-j}$ for all $j, k, l \in \mathbb{Z}$.

Since $b^j_{k,l} = [b(e^k, e^l)]_j$, where $\{e^k : k \in \mathbb{Z}\}$ with $(e^k)_i := \delta_{ki}$ (δ the Kronecker-symbol) and $\widehat{e^k}(\vartheta) = \mathrm{e}^{-\mathrm{i}\vartheta k}$ is the orthonormal system of the Hilbert space ℓ^2, we have

$$b^0_{k,l} = [b(e^k, e^l)]_0 = \frac{1}{2\pi} \int_{\mathbb{T}} [\widehat{b}(\widehat{e^k}, \widehat{e^l})](\vartheta) \, \mathrm{d}\vartheta = \frac{1}{4\pi^2} \int_{\mathbb{T}} \int_{\mathbb{T}} \mathrm{e}^{-\mathrm{i}(\vartheta - \theta)k} \, \beta(\vartheta, \theta) \, \mathrm{e}^{-\mathrm{i}\theta l} \, \mathrm{d}\theta \, \mathrm{d}\vartheta.$$

Using the Fourier representation of β

$$\beta(\vartheta,\theta)=\sum_{m,n\in\mathbb{Z}} b_{m,n}\,\mathrm{e}^{-\mathrm{i}(\vartheta m+\theta n)} \quad \text{with} \quad b_{m,n}=\frac{1}{4\pi^2}\int_{\mathbb{T}}\int_{\mathbb{T}}\beta(\vartheta,\theta)\,\mathrm{e}^{\mathrm{i}(\vartheta m+\theta n)}\,\mathrm{d}\vartheta\,\mathrm{d}\theta,$$

we obtain $b^0_{k,l}=b_{-k,k-l}$. From $\beta\in \mathrm{H}^s(\mathbb{T}\times\mathbb{T})$ it follows

$$|b_{m,n}|\le C\frac{\|\beta\|_s}{(1+m^2+n^2)^{s/2}} \quad \text{for all } m,n\in\mathbb{Z}.$$

Hence, we obtain

$$|b^j_{k,l}|=|b^0_{k-j,l-j}|=|b_{j-k,k-l}|\le C\frac{\|\beta\|_s}{[1+(j-k)^2+(k-l)^2]^{s/2}} \quad \text{for all } j,k,l\in\mathbb{Z}.$$

By this, we have

$$\begin{aligned}
\|b(x,y)\|^2_{\ell^2} &\le \sum_{j\in\mathbb{Z}}\left(\sum_{k,l\in\mathbb{Z}}|b^j_{k,l}||x_k||y_l|\right)^2\\
&\le C^2\|\beta\|_s^2\|y\|^2_{\ell^\infty}\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}}|x_k|\sum_{l\in\mathbb{Z}}\frac{1}{[1+(j-k)^2+(k-l)^2]^{s/2}}\right)^2\\
&\le C^2\|\beta\|_s^2\|y\|^2_{\ell^\infty}\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}}|x_k|\frac{1}{[1+(j-k)^2]^{s/4}}\sum_{l\in\mathbb{Z}}\frac{1}{[1+(k-l)^2]^{s/4}}\right)^2\\
&\le C^2\|\beta\|_s^2\|y\|^2_{\ell^\infty}(1+2\zeta(s/2))^2\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}}|x_k|\frac{1}{[1+(j-k)^2]^{s/4}}\right)^2\\
&\le C^2\|\beta\|_s^2\|y\|^2_{\ell^\infty}(1+2\zeta(s/2))^4\|x\|^2_{\ell^2},
\end{aligned}$$

where we used

$$\|x*y\|^2_{\ell^2}\le\|x\|^2_{\ell^2}\|y\|^2_{\ell^1} \quad \text{with} \quad (x*y)_j:=\sum_{k\in\mathbb{Z}}x_ky_{j-k},$$

$$\zeta(p):=\sum_{k=1}^\infty\frac{1}{k^p} \quad \text{for} \quad p>1.$$

Choosing $s=3$, our proposition holds with $c_b:=C^2\|\beta\|_3^2(1+2\zeta(3/2))^4$. □

We can use the previous proposition in order to obtain an estimate for $B=(B_1,B_2)$ in the case where the uniform nonresonance condition $(\mathrm{NR3})_{\mathrm{unif}}$ holds. By (44), and the analyticity of ω and q given by (39), the $b_{ii}(\vartheta,\theta)$, $i=1,2$, defined by (40) are analytic with respect to $\vartheta,\theta\in(-\pi,\pi]$. Hence, there exist $c_i>0$, such that

$$\|b_i(x,y)\|_{\ell^2}\le c_i\|x\|_{\ell^2}\|y\|_{\ell^\infty} \quad \text{for } x,y\in\ell^2,\ i=1,2.$$

By the definitions of B_1 (41) and B_2 (43), and $\|\widetilde{y}\|_\infty:=\max\{\|y_1\|_{\ell^\infty},\|y_2\|_{\ell^\infty}\}$, we obtain

$$\begin{aligned}
\|B_1(\widetilde{x},\widetilde{y})\|^2_{\ell^2} &\le (\|b_1(x_1,y_1)\|_{\ell^2}+\|b_2(x_2,y_2)\|_{\ell^2})^2\\
&\le (c_1\|x_1\|_{\ell^2}+c_2\|x_2\|_{\ell^2})^2\|\widetilde{y}\|^2_\infty,\\
\|B_2(\widetilde{x},\widetilde{y})\|^2_{\ell^2} &\le (\|b_1(x_2,y_1)\|_{\ell^2}+\|b_1(x_1,y_2)\|_{\ell^2}+\|b_2(Lx_1,y_2)\|_{\ell^2}+\|b_2(x_2,Ly_1)\|_{\ell^2})^2\\
&\le (c_1+c_2C)^2(\|x_1\|_{\ell^2}+\|x_2\|_{\ell^2})^2\|\widetilde{y}\|^2_\infty,
\end{aligned}$$

where we used $\|Lx\|_{\ell^2} \leq C\|x\|_{\ell^2}$ and $\|Lx\|_{\ell^\infty} \leq C\|x\|_{\ell^\infty}$. Thus, setting $c := \max\{c_1, c_2\}$, and using (31) and $\|x_1\|_{\ell^2}+\|x_2\|_{\ell^2} \leq \mu\|\widetilde{x}\|_Y$ with $\mu := (\mu_- + 1)/\mu_-$, we obtain

$$\|B(\widetilde{x},\widetilde{y})\|_Y^2 \leq \mu_+^2\|B_1(\widetilde{x},\widetilde{y})\|_{\ell^2}^2+\|B_2(\widetilde{x},\widetilde{y})\|_{\ell^2}^2 \leq C_B^2\|\widetilde{x}\|_Y^2\|\widetilde{y}\|_\infty^2 \quad \text{for } \widetilde{x},\widetilde{y} \in Y \tag{45}$$

with $C_B^2 := \mu^2 c^2[\mu_+^2+(1+C)^2]$. By $\|\widetilde{y}\|_\infty \leq \mu\|\widetilde{y}\|_Y$, this yields also

$$\|B(\widetilde{x},\widetilde{y})\|_Y \leq \mu C_B\|\widetilde{x}\|_Y\|\widetilde{y}\|_Y \quad \text{for } \widetilde{x},\widetilde{y} \in Y. \tag{46}$$

Moreover, $B : Y{\times}Y \to Y$ is symmetric. Indeed, from $q(\vartheta,\vartheta-\theta) = q(\vartheta,\theta)$, $\alpha(\vartheta,\vartheta-\theta) = \alpha(\vartheta,\theta)$ and $\beta(\vartheta,\vartheta-\theta) = \beta(\vartheta,\theta)$, we obtain $b_{ii}(\vartheta,\vartheta-\theta) = b_{ii}(\vartheta,\theta)$ for $i = 1,2$ (cf. (40)). By

$$\begin{aligned}[\widehat{b}_i(\widehat{x},\widehat{y})](\vartheta) &= \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\vartheta-\theta)b_{ii}(\vartheta,\theta)\widehat{y}(\theta)\,\mathrm{d}\theta \\ &= \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\theta)b_{ii}(\vartheta,\theta)\widehat{y}(\vartheta-\theta)\,\mathrm{d}\theta = [\widehat{b}_i(\widehat{y},\widehat{x})](\vartheta),\end{aligned}$$

it follows $b_i(x,y) = b_i(y,x)$ for $x,y \in \ell^2$. By (41), (43), we obtain $B_i(\widetilde{x},\widetilde{y}) = B_i(\widetilde{y},\widetilde{x})$, i.e., $B(\widetilde{x},\widetilde{y}) = B(\widetilde{y},\widetilde{x})$ for $\widetilde{x},\widetilde{y} \in Y$.

Hence, in the case where the uniform nonresonance condition $(\text{NR3})_{\text{unif}}$ holds, we obtain by the normal-form transformation $F : Y \to Y$ with $F(\widetilde{x}) = \widetilde{x} + B(\widetilde{x},\widetilde{x})$, where the bilinear form $B = (B_1, B_2) : Y{\times}Y \to Y$ is defined via (41), (43), the system

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + \overline{M}(\widetilde{x}) \quad \text{with} \quad \overline{M}(\widetilde{x}) = 2B(\widetilde{Q}(\widetilde{x},\widetilde{x})+\widetilde{M}(\widetilde{x}),\widetilde{x}) + \widetilde{M}(\widetilde{x}). \tag{47}$$

By the Implicit Function Theorem, the inverse mapping $\rho := F^{-1}$ of the transformation F exists on a ball $B_{\varepsilon_\rho}(0) \subset Y$ of radius $\varepsilon_\rho > 0$ and center $F(0) = 0$, and it holds $\rho \in \mathrm{C}^1(B_{\varepsilon_\rho}(0), Y)$ and $D\rho(0) = I$. Hence, it exists a $C_\rho > 0$, such that

$$\|\rho(\widetilde{y}_1)-\rho(\widetilde{y}_2)\|_Y \leq C_\rho\|\widetilde{y}_1-\widetilde{y}_2\|_Y \quad \text{for } \widetilde{y}_1,\widetilde{y}_2 \in Y \text{ with } \|\widetilde{y}_1\|_Y, \|\widetilde{y}_2\|_Y < \varepsilon_\rho. \tag{48}$$

Indeed, by the properties of B, the Fréchet derivative of F is given by $DF : Y \to \mathcal{L}(Y,Y)$ with $DF(\widetilde{x}) = I + 2B(\widetilde{x},\cdot)$ and, thus, $DF(0) = I$. Moreover, (46) gives $F \in \mathrm{C}^1(Y,Y)$, since

$$\|DF(\widetilde{x})-DF(\widetilde{x}_0)\|_{\mathcal{L}(Y,Y)} = 2\|B(\widetilde{x}-\widetilde{x}_0,\cdot)\|_{\mathcal{L}(Y,Y)} \leq 2\mu C_B\|\widetilde{x}-\widetilde{x}_0\|_Y \quad \text{for } \widetilde{x},\widetilde{x}_0 \in Y.$$

Thus, for sufficiently small $\widetilde{x} \in Y$ the system (47) reads

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + N(\widetilde{y}) \quad \text{with} \quad N(\widetilde{y}) := \overline{M}(\rho(\widetilde{y})). \tag{49}$$

By the definition of $\overline{M}$ and $\rho(\widetilde{y}) = \widetilde{y} + \mathcal{O}(\widetilde{y}^2)$, the nonlinearity N of the transformed system (49) has only cubic and higher order terms, but no quadratic ones. As already mentioned, this is the crucial motivation in applying the normal-form transformation F on the system (27), since it enables us to apply the Gronwall type argument already used in [19] for the justification of the NLSE derived for systems with cubic leading terms in their nonlinearity (cf. Section 4.2).

In order to illustrate the normal-form transformation better, let us consider the simplest example of an oscillator chain with only nearest-neighbor interactions and a quadratic nonlinearity

$$\ddot{x}_j = (Lx)_j - \beta_2 x_j^2, \quad \text{where} \quad (Lx)_j = \alpha_1(x_{j+1}-2x_j+x_{j-1}) - \beta_1 x_j$$

with $\alpha_1, \beta_1 > 0$ and dispersion relation $\omega^2(\vartheta) = 2\alpha_1(1-\cos\vartheta) + \beta_1$. According to the general framework we presented above, the normal-form transformation has the form

$$y = x - \beta_2[b(x,x)+c(\dot{x},\dot{x})],$$

where the bilinear forms $b, c : \ell^2\times\ell^2 \to \ell^2$ are given by

$$[b(x,x)]_j = \sum_{k,l\in\mathbb{Z}} b_{j-k,k-l}x_k x_l \quad \text{with} \quad b_{m,n} = \frac{1}{4\pi^2}\int_{\mathbb{T}}\int_{\mathbb{T}} \frac{\alpha(\vartheta,\theta)}{\gamma(\vartheta,\theta)}\, \mathrm{e}^{\mathrm{i}(\vartheta m+\theta n)}\, \mathrm{d}\vartheta\mathrm{d}\theta$$

and

$$[c(x,x)]_j = \sum_{k,l\in\mathbb{Z}} c_{j-k,k-l}x_k x_l \quad \text{with} \quad c_{m,n} = \frac{1}{4\pi^2}\int_{\mathbb{T}}\int_{\mathbb{T}} \frac{2}{\gamma(\vartheta,\theta)}\, \mathrm{e}^{\mathrm{i}(\vartheta m+\theta n)}\, \mathrm{d}\vartheta\mathrm{d}\theta$$

and with $\alpha(\vartheta,\theta) := \omega^2(\vartheta-\theta)+\omega^2(\theta)-\omega^2(\vartheta)$ and $\gamma(\vartheta,\theta) := \alpha^2(\vartheta,\theta)-4\omega^2(\vartheta-\theta)\omega^2(\theta)$ (cf. (32), (39), (40) and the proof of Proposition 2). Since $\alpha_1, \beta_1 > 0$, it is $\gamma(\vartheta,\theta) \neq 0$ for all $(\vartheta,\theta) \in \mathbb{T}^2$ (cf. (24) in Proposition 1). The transformed system has then the form (cf. (47))

$$\ddot{y} = Ly - 2\beta_2[b(x^2,x)+2c(x\dot{x},\dot{x})+c(x^2,2Lx-\beta_2x^2)],$$

which obviously has only cubic and biquadratic nonlinear terms in $x, \dot{x}$. By using the inverse normal-form transformation as described above we obtain cubic and higher order nonlinear terms in $y, \dot{y}$.

4. **The justification of the NLSE.**

4.1. **Estimate of the residual.**

The procedure of the formal derivation of NLSE consisted in equating the left- and right- hand side coefficients of each term $\varepsilon^k\mathbf{E}^n$ of the expansion in such terms of equation (9) for $k = 1,\dots,p$ (and $n = 0,\dots,k$). Hence, for the improved approximation $X_\varepsilon^{A,p}$ with the $A_{k,n}$ calculated in Section 2 the residual terms have the form

$$\mathrm{res}(X_\varepsilon^{A,p}) := \ddot{X}_\varepsilon^{A,p} - LX_\varepsilon^{A,p} - N(X_\varepsilon^{A,p}) = \varepsilon^{p+1}(r_\varepsilon^{D,p}-r_\varepsilon^{L,p}-r_\varepsilon^{N,p}) \tag{50}$$

(cf. (10), (13) and (16)). From (11) we obtain that there exists a $C_D > 0$, depending on $\omega : \vartheta \mapsto \omega(\vartheta)$, ϑ_0, and ε_0 such that for $\varepsilon \le \varepsilon_0$

$$|(r_\varepsilon^{D,p})_j(t)| \le C_D \sum_{k=p+1}^{p+4}\sum_{q=1}^{p}\sum_{n=-q}^{q}\sum_{\substack{\mu+2\nu=k-q,\\ \mu+\nu\le 2}} |\partial_\tau^\nu\partial_\xi^\mu A_{q,n}(\tau,\xi)| \tag{51}$$

with $\tau = \varepsilon^2 t$, $\xi = \varepsilon(j+\omega'(\vartheta_0)t)$ holds. Analogously, from (14) and (17) we obtain that there exists a $C > 0$, depending on ω, ϑ_0, $V_m, W \in \mathrm{C}^{p+2}(\mathbb{R})$ $(m = 1,\dots,M)$ and ε_0, such that for $\varepsilon \le \varepsilon_0$ and $\varepsilon^2 t \in [0,\tau_0]$

$$\begin{aligned} &|(r_\varepsilon^{L,p})_j(t)| + |(r_\varepsilon^{N,p})_j(t)| \\ \le\;& C\sum_{s,r=0}^{p}\sum_{l=1}^{p}\sum_{n=0}^{l}\|\partial_\xi^r A_{l,n}(\tau,\cdot)\|_{\mathrm{L}^\infty(\mathbb{R})}^s \Bigg[\sum_{r=0}^{p-1}\sum_{l=1}^{p}\sum_{n=0}^{l}|\partial_\xi^r A_{l,n}(\tau,\xi)| \\ &+ \sum_{m=1}^{M}\sum_{l=1}^{p}\sum_{n=0}^{l}\left(|\partial_\xi^p A_{l,n}(\tau,\xi+\theta_{pqn}^{+\varepsilon m}\varepsilon m)| + |\partial_\xi^p A_{l,n}(\tau,\xi-\theta_{pqn}^{-\varepsilon m}\varepsilon m)|\right)\Bigg] \end{aligned}$$

with $\theta_{pqn}^{\pm\varepsilon m} \in (0,1)$ (cf. (12)) holds. For the estimation of $v_{m,p}$ and w_p in $N(X_\varepsilon^{A,p})$ we used the mean value theorem. Hence, the above estimate holds as long as $\|\partial_\xi^r A_{l,n}(\tau,\cdot)\|_{\mathrm{L}^\infty(\mathbb{R})} \le d$ is satisfied for all $r = 0,\dots,p$, $l = 1,\dots,p$, $n = 0,\dots,l$ and all $\tau \in [0,\tau_0]$ (and C depends also on d). By Sobolev's embedding theorem, this is fulfilled if $\|A_{l,n}(\tau,\cdot)\|_{\mathrm{H}^{p+1}(\mathbb{R})} \le \tilde{d}$ for $\tau \in [0,\tau_0]$.

Thus, applying Proposition 3.3 of [19]

$$\sum_{j\in\mathbb{Z}} \sup_{|s|\le 1} |\phi\,(\varepsilon(j{+}c{+}s))|^2 \le \frac{8}{\varepsilon}\,\|\phi\|^2_{\mathrm{H}^1(\mathbb{R})} \quad \text{for } \phi \in \mathrm{H}^1(\mathbb{R}),\ \varepsilon \in (0,1),\ c \in \mathbb{R},$$

we obtain

$$\|r_\varepsilon^{L,p}(t)\|_{\ell^2} + \|r_\varepsilon^{N,p}(t)\|_{\ell^2} \le \varepsilon^{-1/2}\tilde{C}\sum_{s=1}^{p+1}\sum_{r=0}^{p}\sum_{l=1}^{p}\sum_{n=0}^{l} \|\partial_\xi^r A_{l,n}(\tau,\cdot)\|^s_{H^1(\mathbb{R})}.$$

The same argument yields $\|r_\varepsilon^{D,p}(t)\|_{\ell^2} = \mathcal{O}(\varepsilon^{-1/2})$ for $\varepsilon \le \varepsilon_0$ and $\varepsilon^2 t \le \tau_0$, if and only if the derivatives appearing in (51) satisfy $\|\partial_\tau^\nu \partial_\xi^\mu A_{q,n}(\tau,\cdot)\|_{\mathrm{H}^1(\mathbb{R})} \le c$ for $\tau \in [0,\tau_0]$ and some $c > 0$. If this is the case, by (50) and $\mathrm{res}(\widetilde{X}_\varepsilon^{A,p}) := (0, \mathrm{res}(X_\varepsilon^{A,p}))$ we finally obtain

$$\|\mathrm{res}(\widetilde{X}_\varepsilon^{A,p})(t)\|_Y = \|\mathrm{res}(X_\varepsilon^{A,p})(t)\|_{\ell^2} \le \tilde{C}_r \varepsilon^{p+1/2} \quad \text{for } \varepsilon \le \varepsilon_0 \text{ and } \varepsilon^2 t \le \tau_0. \tag{52}$$

Looking at our systems of determining equations for the functions $A_{l,n}$ and taking into account that $\partial_\tau A,\ \partial_\tau A_{2,1}$ is equivalent to $\partial_\xi^2 A,\ \partial_\xi^2 A_{2,1}$, respectively, we see that the needed regularity conditions on $A_{l,n}$ are satisfied for $A(\tau,\cdot) \in \mathrm{H}^6(\mathbb{R})$ if $p = 3$ (where $A_{2,1}$ remained undetermined and thus can be assumed as equivalently vanishing) and for $A(\tau,\cdot) \in \mathrm{H}^7(\mathbb{R})$ and $A_{2,1}(\tau,\cdot) \in \mathrm{H}^6(\mathbb{R})$ if $p = 4$. From the determining equations of A and $A_{2,1}$ it follows by standard arguments of the theory of semilinear wave equations (cf. e.g., [32, 38]) that there exists some $\tau_0 > 0$ such that the required regularity on A and $A_{2,1}$ is preserved for $\tau \in [0,\tau_0]$ if we assume initially $A(0,\cdot) \in \mathrm{H}^6(\mathbb{R})$ in the case $p = 3$ and $A(0,\cdot) \in \mathrm{H}^7(\mathbb{R})$, $A_{2,1}(0,\cdot) \in \mathrm{H}^6(\mathbb{R})$ in the case $p = 4$.

Obviously, under this regularity conditions we obtain by the same reasoning as above that for $X_\varepsilon^{A,p}$ with $p = 3, 4$ and the calculated coefficients $A_{k,n}$ (with $A_{2,1} = A_{3,1} = 0$ if $p = 3$ and $A_{3,1} = A_{4,1} = 0$ if $p = 4$) there exists a $C > 0$, depending on $V_m, W, \omega, \vartheta_0, A, \tau_0$ (and $A_{2,1}$ if $p = 4$), such that

$$\|X_\varepsilon^{A,p}(t)\|_{\ell^\infty},\ \|\dot{X}_\varepsilon^{A,p}(t)\|_{\ell^\infty},\ \|X_\varepsilon^{A,p}(t)\|_{\ell^2},\ \|\dot{X}_\varepsilon^{A,p}(t)\|_{\ell^2} \le \varepsilon^{1/2} C$$

for $\varepsilon \le \varepsilon_0 \le 1$ and $\varepsilon^2 t \le \tau_0$, which leads by the definitions (30) and $\|(x,y)\|_\infty := \max\{\|x\|_{\ell^\infty}, \|y\|_{\ell^\infty}\}$ to

$$\|\widetilde{X}_\varepsilon^{A,p}(t)\|_\infty \le \varepsilon C, \quad \|\widetilde{X}_\varepsilon^{A,p}(t)\|_Y \le \varepsilon^{1/2} C_1 \quad \text{for } \varepsilon \le \varepsilon_0 \le 1 \text{ and } \varepsilon^2 t \le \tau_0 \tag{53}$$

with $C_1 := C(\mu_+^2{+}1)^{1/2}$ (cf. (31)). Analogously, there exist $C_2, C_2', C_3 > 0$ such that

$$\begin{cases} \|\widetilde{X}_\varepsilon^{A,p}(t) {-} \widetilde{X}_\varepsilon^{A}(t)\|_Y \le C_2 \varepsilon^{3/2}, & \|\widetilde{X}_\varepsilon^{A,p}(t) {-} \widetilde{X}_\varepsilon^{A}(t)\|_\infty \le C_2' \varepsilon^2, \\ \|\widetilde{X}_\varepsilon^{A,p}(t) {-} \widetilde{X}_\varepsilon^{A,2}(t)\|_Y \le C_3 \varepsilon^{5/2} & \qquad \text{for } \varepsilon \le \varepsilon_0 \le 1,\ \varepsilon^2 t \le \tau_0. \end{cases} \tag{54}$$

After these preliminary remarks we start the justification of the NLSE (5) with the case where the uniformly nonresonant condition $(\mathrm{NR3})_{\mathrm{unif}}$ holds. As we have seen in Section 3 in this case we are able to transform the system (27) into a system (49) with only cubic and higher order nonlinear terms, which allows us to use our Gronwall type argument. Then, we will examine the situation where only the weaker nonresonance condition $(\mathrm{NR3})_{\vartheta_0}$ holds.

4.2. **Justification under the uniform nonresonance condition (NR3)$_{\text{unif}}$.** We consider the transformed system (49) $\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + N(\widetilde{y})$ and the associated transformed approximation

$$\widetilde{Y}_\varepsilon^{A,3} := F(\widetilde{X}_\varepsilon^{A,3}) = \widetilde{X}_\varepsilon^{A,3} + B(\widetilde{X}_\varepsilon^{A,3}, \widetilde{X}_\varepsilon^{A,3}).$$

The residual term of the approximation $\widetilde{Y}_\varepsilon^{A,3}$ is given by

$$\begin{aligned} \operatorname{res}(\widetilde{Y}_\varepsilon^{A,3}) &:= \dot{\widetilde{Y}}_\varepsilon^{A,3} - \widetilde{L}\widetilde{Y}_\varepsilon^{A,3} - N(\widetilde{Y}_\varepsilon^{A,3}) \\ &= \dot{\widetilde{X}}_\varepsilon^{A,3} + 2B(\dot{\widetilde{X}}_\varepsilon^{A,3}, \widetilde{X}_\varepsilon^{A,3}) - \widetilde{L}\widetilde{X}_\varepsilon^{A,3} - \widetilde{L}B(\widetilde{X}_\varepsilon^{A,3}, \widetilde{X}_\varepsilon^{A,3}) - \overline{M}(\widetilde{X}_\varepsilon^{A,3}) \\ &= \dot{\widetilde{X}}_\varepsilon^{A,3} + 2B(\operatorname{res}(\widetilde{X}_\varepsilon^{A,3}), \widetilde{X}_\varepsilon^{A,3}) - \widetilde{L}\widetilde{X}_\varepsilon^{A,3} - \widetilde{Q}(\widetilde{X}_\varepsilon^{A,3}, \widetilde{X}_\varepsilon^{A,3}) - \widetilde{M}(\widetilde{X}_\varepsilon^{A,3}) \\ &= \operatorname{res}(\widetilde{X}_\varepsilon^{A,3}) + 2B(\operatorname{res}(\widetilde{X}_\varepsilon^{A,3}), \widetilde{X}_\varepsilon^{A,3}), \end{aligned}$$

where we used (34) (with $\overline{Q} = 0$) and (35). From (52) and (53), it follows by (45)

$$\|\operatorname{res}(\widetilde{Y}_\varepsilon^{A,3})(t)\|_Y \le C_r\varepsilon^{7/2} \quad \text{for } \varepsilon \le \varepsilon_0,\ \varepsilon^2 t \le \tau_0 \quad \text{with } C_r := (1{+}2C_BC\varepsilon_0)\widetilde{C}_r. \tag{55}$$

Inserting the error $\varepsilon^{3/2}\widetilde{R} := \widetilde{y} - \widetilde{Y}_\varepsilon^{A,3}$ between a solution $\widetilde{y}$ of the transformed system (49) and the transformed approximation $\widetilde{Y}_\varepsilon^{A,3}$ into (49), we obtain by the definition of the residual term $\operatorname{res}(\widetilde{Y}_\varepsilon^{A,3})$ the differential equation for the error

$$\dot{\widetilde{R}} = \widetilde{L}\widetilde{R} + \varepsilon^{-3/2}[N(\widetilde{Y}_\varepsilon^{A,3}{+}\varepsilon^{3/2}\widetilde{R}){-}N(\widetilde{Y}_\varepsilon^{A,3}){-}\operatorname{res}(\widetilde{Y}_\varepsilon^{A,3})].$$

The semigroup associated to the linear problem $\dot{\widetilde{R}} = \widetilde{L}\widetilde{R}$ is given by $G(t) = \mathrm{e}^{t\widetilde{L}}$. Hence, the differential equation for the error can be transformed by the variation of constants formula into

$$\begin{aligned} \widetilde{R}(t) =& \mathrm{e}^{t\widetilde{L}}\widetilde{R}(0) \\ &+ \varepsilon^{-3/2}\int_0^t \mathrm{e}^{(t-s)\widetilde{L}}\left[N(\widetilde{Y}_\varepsilon^{A,3}(s){+}\varepsilon^{3/2}\widetilde{R}(s)) - N(\widetilde{Y}_\varepsilon^{A,3}(s)) - \operatorname{res}(\widetilde{Y}_\varepsilon^{A,3}(s))\right]\mathrm{d}s. \end{aligned} \tag{56}$$

Assuming $\|\widetilde{x}(0){-}\widetilde{X}_\varepsilon^{A}(0)\|_Y \le d\varepsilon^{3/2}$, (54) yields $\|\widetilde{x}(0){-}\widetilde{X}_\varepsilon^{A,3}(0)\|_Y \le (d{+}C_2)\varepsilon^{3/2}$, and thus by (53): $\|\widetilde{x}(0){+}\widetilde{X}_\varepsilon^{A,3}(0)\|_Y \le [(d{+}C_2)\varepsilon_0{+}2C_1]\varepsilon^{1/2}$ for $\varepsilon \le \varepsilon_0 < 1$. This yields, by $\widetilde{y} - \widetilde{Y}_\varepsilon^{A,3} = \widetilde{x} - \widetilde{X}_\varepsilon^{A,3} + B(\widetilde{x}{-}\widetilde{X}_\varepsilon^{A,3}, \widetilde{x}{+}\widetilde{X}_\varepsilon^{A,3})$ and (46),

$$\|\widetilde{R}(0)\|_Y = \varepsilon^{-3/2}\|\widetilde{y}(0){-}\widetilde{Y}_\varepsilon^{A,3}(0)\|_Y \le \widetilde{d} \quad \text{for } \varepsilon \le \varepsilon_0 < 1 \tag{57}$$

with $\widetilde{d} := (d{+}C_2)\{1{+}\mu C_B[(d{+}C_2)\varepsilon_0{+}2C_1]\varepsilon_0^{1/2}\}$.

Now, let us assume for the moment that we can show, that there exist a constant $C_N > 0$ independent of a given $D > 0$, and an $\varepsilon_0 > 0$ depending on D, such that

$$\|N(\widetilde{Y}_\varepsilon^{A,3}(t){+}\varepsilon^{3/2}\widetilde{R}(t)) - N(\widetilde{Y}_\varepsilon^{A,3}(t))\|_Y \le C_N\varepsilon^2\|\varepsilon^{3/2}\widetilde{R}(t)\|_Y \tag{58}$$

holds for $\varepsilon \le \varepsilon_0$, $\varepsilon^2 t \le \tau_0$, $\|\widetilde{R}(t)\|_Y \le D$. Then, by $\|\mathrm{e}^{t\widetilde{L}}\|_{Y\to Y} = 1$, (55), (57) and (58), equation (56) yields

$$\|\widetilde{R}(t)\|_Y \le \widetilde{d} + \varepsilon^2\left(C_N\int_0^t \|\widetilde{R}(s)\|_Y\,\mathrm{d}s + tC_r\right) \quad \text{for } \varepsilon \le \varepsilon_0,\ \varepsilon^2 t \le \tau_0$$

as long as $\|\widetilde{R}(s)\|_Y \le D$ holds for $s \in [0,t]$. By Gronwall's inequality, it follows

$$\|\widetilde{R}(t)\|_Y \le (\widetilde{d}{+}\varepsilon^2 tC_r)\mathrm{e}^{\varepsilon^2 tC_N} \quad \text{for } \varepsilon \le \varepsilon_0,\ \varepsilon^2 t \le \tau_0$$

as long as $\|\widetilde{R}(s)\|_Y \leq D$ holds for $s \in [0,t]$. This is fulfilled if we choose $D := (\widetilde{d}+\tau_0 C_r)\mathrm{e}^{\tau_0 C_N}$. Hence, for an $\varepsilon_0 > 0$ associated to this D by (58), we have obtained

$$\|\widetilde{y}(t)-\widetilde{Y}_\varepsilon^{A,3}(t)\|_Y = \|\widetilde{R}(t)\|_Y \varepsilon^{3/2} \leq D\varepsilon^{3/2} \quad \text{for } \varepsilon \leq \varepsilon_0,\, \varepsilon^2 t \leq \tau_0.$$

For $\|\widetilde{y}(t)-\widetilde{Y}_\varepsilon^{A,3}(t)\|_Y$, $\|\widetilde{Y}_\varepsilon^{A,3}(t)\|_Y < \varepsilon_\rho/2$ it is $\|\widetilde{y}(t)\|_Y < \varepsilon_\rho$. Thus, since $\widetilde{x}(t) = \rho(\widetilde{y}(t))$ and $\widetilde{X}_\varepsilon^{A,3}(t) = \rho(\widetilde{Y}_\varepsilon^{A,3}(t))$, it follows from (48)

$$\|\widetilde{x}(t)-\widetilde{X}_\varepsilon^{A,3}(t)\|_Y \leq C_\rho \|\widetilde{y}(t)-\widetilde{Y}_\varepsilon^{A,3}(t)\|_Y \leq C_\rho D\varepsilon^{3/2} \quad \text{for } \varepsilon \leq \varepsilon_0,\, \varepsilon^2 t \leq \tau_0.$$

Hence, we finally obtain by (54)

$$\|\widetilde{x}(t)-\widetilde{X}_\varepsilon^{A}(t)\|_Y \leq C\varepsilon^{3/2} \quad \text{for } \varepsilon^2 t \leq \tau_0,\, \varepsilon \leq \varepsilon_0 < 1 \quad \text{with } C := C_\rho D + C_2.$$

Thus, except for the proof of (58) that is presented below, we have established the following theorem, which constitutes under the uniform nonresonance condition $(\mathrm{NR3})_{\mathrm{unif}}$ our justification of the validity of the NLSE (5) as a macroscopic limit for the oscillator chain model (1).

Theorem 2. *Assume that V_m, $W \in \mathrm{C}^5(\mathbb{R})$ in (1) have the form (2) and that the stability condition (SC) and the nonresonance conditions $(NR2)^3_{\vartheta_0}$ and $(NR3)_{\mathrm{unif}}$ hold. Let $A : [0,\tau_0]\times\mathbb{R} \to \mathbb{C}$, $\tau_0 > 0$, be a solution of the NLSE (5) with $A(0,\cdot) \in \mathrm{H}^6(\mathbb{R})$ and let X_ε^A be the formal approximation given in (4) with $\mathrm{c}_{\mathrm{gr}} = -\omega'$. Then, for each $d > 0$ there exist $\varepsilon_0, C > 0$ such that for all $\varepsilon \in (0,\varepsilon_0)$ the following statement holds:*

Any solution $\widetilde{x}$ of (27) with an initial condition $\widetilde{x}(0)$ satisfying

$$\|\widetilde{x}(0)-\widetilde{X}_\varepsilon^{A}(0)\|_Y \leq d\varepsilon^{3/2}$$

fulfills the estimate

$$\|\widetilde{x}(t)-\widetilde{X}_\varepsilon^{A}(t)\|_Y \leq C\varepsilon^{3/2} \quad \textit{for} \quad t \in [0,\tau_0/\varepsilon^2].$$

(In the case of nearest-neighbor interactions $(\mathrm{NR2})^3_{\vartheta_0}$ is implied by (SC) and $(\mathrm{NR3})_{\mathrm{unif}}$, cf. Remark 1.)

Proof of (58). For $y, \Delta \in Y$ with $\|y\|_Y + \|\Delta\|_Y < \varepsilon_\rho$ and $\widetilde{x}_1 := \rho(y+\Delta)$, $\widetilde{x}_2 := \rho(y) \in Y$ we have by (49) and (35)

$$\begin{aligned} N(y+\Delta) - N(y) =& 2B(\widetilde{Q}(\widetilde{x}_1,\widetilde{x}_1)-\widetilde{Q}(\widetilde{x}_2,\widetilde{x}_2),\widetilde{x}_1) + 2B(\widetilde{Q}(\widetilde{x}_2,\widetilde{x}_2),\widetilde{x}_1-\widetilde{x}_2) \\ &+ 2B(\widetilde{M}(\widetilde{x}_1)-\widetilde{M}(\widetilde{x}_2),\widetilde{x}_1) + 2B(\widetilde{M}(\widetilde{x}_2),\widetilde{x}_1-\widetilde{x}_2) \\ &+ \widetilde{M}(\widetilde{x}_1) - \widetilde{M}(\widetilde{x}_2). \end{aligned}$$

From (45) $\|B(\widetilde{x},\widetilde{y})\|_Y \leq C_B \|\widetilde{x}\|_\infty \|\widetilde{y}\|_Y$ it follows

$$\begin{aligned} \|N(y+\Delta)-N(y)\|_Y \leq& 2C_B\|\widetilde{x}_1\|_\infty(\|\widetilde{Q}(\widetilde{x}_1,\widetilde{x}_1)-\widetilde{Q}(\widetilde{x}_2,\widetilde{x}_2)\|_Y+\|\widetilde{M}(\widetilde{x}_1)-\widetilde{M}(\widetilde{x}_2)\|_Y) \\ &+ 2C_B(\|\widetilde{Q}(\widetilde{x}_2,\widetilde{x}_2)\|_\infty+\|\widetilde{M}(\widetilde{x}_2)\|_\infty)\|\widetilde{x}_1-\widetilde{x}_2\|_Y \\ &+ \|\widetilde{M}(\widetilde{x}_1)-\widetilde{M}(\widetilde{x}_2)\|_Y. \end{aligned} \quad (59)$$

By the definition of M (29), we have

$$\|M(x_1)-M(x_2)\|_{\ell^2}^2 = \sum_{j\in\mathbb{Z}}\Bigg(\sum_{m=1}^{M}[v_{m,2}(x_{j+m}^{(1)}-x_j^{(1)})-v_{m,2}(x_j^{(1)}-x_{j-m}^{(1)}) \\ -v_{m,2}(x_{j+m}^{(2)}-x_j^{(2)})+v_{m,2}(x_j^{(2)}-x_{j-m}^{(2)})]-w_2(x_j^{(1)})+w_2(x_j^{(2)})\Bigg)^2$$

with $x^{(i)} := x_i \in \ell^2$ for $i=1,2$. Since $v_{m,2}(d)=\mathcal{O}(d^3)$ and $w_2(x)=\mathcal{O}(x^3)$ (cf. (15)), it follows by the mean value theorem that for arbitrary $\delta>0$ there exists a constant $C>0$ depending on $v_{m,2}, w_2, \delta$ such that for $\|x_1\|_{\ell^\infty}, \|x_2\|_{\ell^\infty} \le \delta/\mu_-$

$$|v_{m,2}(x_{j\pm m}^{(1)}-x_j^{(1)})-v_{m,2}(x_{j\pm m}^{(2)}-x_j^{(2)})| \le C(\|x_1\|_{\ell^\infty}^2+\|x_2\|_{\ell^\infty}^2)|x_{j\pm m}^{(1)}-x_j^{(1)}-x_{j\pm m}^{(2)}+x_j^{(2)}|,$$
$$|w_2(x_j^{(1)})-w_2(x_j^{(2)})| \le C(\|x_1\|_{\ell^\infty}^2+\|x_2\|_{\ell^\infty}^2)|x_j^{(1)}-x_j^{(2)}|$$

holds. Hence, we obtain

$$\|M(x_1)-M(x_2)\|_{\ell^2} \le \tilde{C}(\|x_1\|_{\ell^\infty}^2+\|x_2\|_{\ell^\infty}^2)\|x_1-x_2\|_{\ell^2} \quad \text{for } \|x_1\|_{\ell^\infty}, \|x_2\|_{\ell^\infty} \le \delta/\mu_-,$$

and thus, by $\widetilde{M}(\widetilde{x}) = (0, M(x))$, the definitions of $\|\cdot\|_Y$ and $\|\cdot\|_\infty$ and (31),

$$\|\widetilde{M}(\widetilde{x}_1)-\widetilde{M}(\widetilde{x}_2)\|_Y \le C_M(\|\widetilde{x}_1\|_\infty^2+\|\widetilde{x}_2\|_\infty^2)\|\widetilde{x}_1-\widetilde{x}_2\|_Y \quad \text{for } \|\widetilde{x}_1\|_Y, \|\widetilde{x}_2\|_Y \le \delta \tag{60}$$

with $C_M := \tilde{C}/\mu_-$. By $\widetilde{M}(0)=0$ (since $v_{m,2}(0)=w_2(0)=0$), this yields also

$$\|\widetilde{M}(\widetilde{x})\|_\infty \le \|\widetilde{M}(\widetilde{x})\|_Y \le C_M\|\widetilde{x}\|_\infty^2\|\widetilde{x}\|_Y \quad \text{for } \|\widetilde{x}\|_Y \le \delta. \tag{61}$$

Analogously, since

$$|(x_{j+1}^{(1)}-x_j^{(1)})^2-(x_{j+1}^{(2)}-x_j^{(2)})^2| \le 2\|x_1+x_2\|_{\ell^\infty}(|x_{j+1}^{(1)}-x_{j+1}^{(2)}|+|x_j^{(1)}-x_j^{(2)}|),$$
$$|(x_j^{(1)})^2-(x_j^{(2)})^2| \le \|x_1+x_2\|_{\ell^\infty}|x_j^{(1)}-x_j^{(2)}|,$$

we obtain, by $\widetilde{Q}(\widetilde{x},\widetilde{x}) = (0, Q(x,x))$, the definition of Q (28), and (31)

$$\|\widetilde{Q}(\widetilde{x}_1,\widetilde{x}_1)-\widetilde{Q}(\widetilde{x}_2,\widetilde{x}_2)\|_Y \le C_Q\|\widetilde{x}_1+\widetilde{x}_2\|_\infty\|\widetilde{x}_1-\widetilde{x}_2\|_Y \tag{62}$$

with $C_Q := C/\mu_-$, $C^2>0$ being a polynomial of $|\alpha_m|$, $(m=1,\dots,M)$ and $|\beta_2|$. Moreover, it is

$$\|\widetilde{Q}(\widetilde{x},\widetilde{x})\|_\infty = \|Q(x,x)\|_{\ell^\infty} \le (6\sum_{m=1}^{M}|\alpha_{m,2}|+|\beta_2|)\|x\|_{\ell^\infty}^2 \le (6\sum_{m=1}^{M}|\alpha_{m,2}|+|\beta_2|)\|\widetilde{x}\|_\infty^2. \tag{63}$$

Inserting (60)–(63) in (59), we obtain

$$\|N(y+\Delta)-N(y)\|_Y \le C_n(\|\widetilde{x}_1\|_\infty^2+\|\widetilde{x}_2\|_\infty^2)\|\widetilde{x}_1-\widetilde{x}_2\|_Y \quad \text{for } \|\widetilde{x}_1\|_Y, \|\widetilde{x}_2\|_Y \le \delta,$$

where $C_n>0$ depends on C_B, V_m, W and δ, with arbitrary $\delta>0$. For $\widetilde{x}_1 = \rho(y+\Delta)$, $\widetilde{x}_2=\rho(y)$ with $\|y\|_Y+\|\Delta\|_Y < \varepsilon_\rho$, there exists a $\delta>0$, such that $\|\widetilde{x}_1\|_Y, \|\widetilde{x}_2\|_Y \le \delta$. Hence, we obtain by (48)

$$\|N(y+\Delta)-N(y)\|_Y \le C_nC_\rho(\|\widetilde{x}_1\|_\infty^2+\|\widetilde{x}_2\|_\infty^2)\|\Delta\|_Y \quad \text{for } \|y\|_Y+\|\Delta\|_Y < \varepsilon_\rho \tag{64}$$

with $C_n>0$ depending on C_B, V_m, W and ε_ρ.

For $y := \widetilde{Y}_\varepsilon^{A,3}(t) = \widetilde{X}_\varepsilon^{A,3}(t) + B(\widetilde{X}_\varepsilon^{A,3}(t), \widetilde{X}_\varepsilon^{A,3}(t))$, we obtain from (53) by (45)

$$\|y\|_Y \le C_1\varepsilon^{1/2}(1+C_BC\varepsilon) \quad \text{for } \varepsilon^2 t \le \tau_0,\ \varepsilon \le \varepsilon_0.$$

Let $\Delta := \varepsilon^{3/2} R$ with $\|R\|_Y \le D$, $D > 0$, and choose ε_0 sufficiently small, such that $\|y\|_Y + \|\Delta\|_Y < \varepsilon_\rho$. Since $\widetilde{x}_2 = \rho(y) = \widetilde{X}^{A,3}_\varepsilon(t)$ and $\widetilde{x}_1 = \rho(y+\Delta)$, it follows by (53), (48) and $\|\widetilde{x}\|_\infty \le \mu \|\widetilde{x}\|_Y$ with $\mu := \frac{m+1}{m}$

$$\begin{aligned}\|\widetilde{x}_1\|_\infty &\le \|\widetilde{x}_1 - \widetilde{x}_2\|_\infty + \|\widetilde{x}_2\|_\infty \le \mu \|\rho(y+\Delta) - \rho(y)\|_Y + C\varepsilon \\ &\le \mu C_\rho \varepsilon^{3/2} D + C\varepsilon \le (\mu C_\rho \varepsilon_0^{1/2} D + C)\varepsilon \quad \text{for } \varepsilon^2 t \le \tau_0,\ \varepsilon \le \varepsilon_0,\ \|R\|_Y \le D.\end{aligned}$$

Decreasing $\varepsilon_0 > 0$ further if needed, we obtain for given $D > 0$ e.g.

$$\|\widetilde{x}_1\|_\infty \le 2C\varepsilon \quad \text{for } \varepsilon^2 t \le \tau_0,\ \varepsilon \le \varepsilon_0,\ \|R\|_Y \le D.$$

Inserting these estimates in (64), we obtain (58) with $C_N := 5 C_n C_\rho C^2$, which is independent of D. $\square$

4.3. **Justification under the nonresonance condition $(\mathrm{NR3})_{\vartheta_0}$.** In the previous section we proved Theorem 2 under the uniform nonresonance condition $(\mathrm{NR3})_{\mathrm{unif}}$. Now we only assume the weaker local nonresonance condition $(\mathrm{NR3})_{\vartheta_0}$ for the given, fixed $\vartheta_0 \in \mathbb{T}$. Our approach follows closely ideas in [35].

By the representation (44) of $\gamma := \alpha^2 - \beta^2$, $(\mathrm{NR3})_{\vartheta_0}$ implies that there exist $c, \delta > 0$ such that

$$|\gamma(\vartheta,\theta)| > c \quad \text{for } (\vartheta,\theta) \in \mathbb{T}\times\mathbb{T} \text{ with } |\vartheta - \vartheta_0| < 2\delta \tag{65}$$

(here and in the following the differences $\vartheta - \vartheta_0$ are taken $\bmod 2\pi$), which certainly means that it is also

$$|\gamma(\vartheta,-\theta)| > c \quad \text{for } (\vartheta,\theta) \in \mathbb{T}\times\mathbb{T} \text{ with } |\vartheta - \vartheta_0| < 2\delta.$$

By the structure of γ, and since ω is an even function, we have the symmetries

$$\gamma(\vartheta,-\theta) = \gamma(-\theta,\vartheta), \quad \gamma(\vartheta,\theta) = \gamma(\theta,\vartheta), \quad \gamma(\vartheta,\theta) = \gamma(\vartheta,\vartheta-\theta) \quad \text{for } (\vartheta,\theta) \in \mathbb{T}\times\mathbb{T}.$$

This yields

$$|\gamma(\vartheta,\theta)| > c \quad \text{for } (\vartheta,\theta) \in \Gamma(\vartheta_0, 2\delta)$$

with

$$\Gamma(\vartheta_0,\delta) := \{(\vartheta,\theta) \in \mathbb{T}\times\mathbb{T} :\ |\vartheta \pm \vartheta_0| < \delta \ \vee\ |\theta \pm \vartheta_0| < \delta \ \vee\ |\vartheta - \theta \pm \vartheta_0| < \delta\}$$

(cf. Figure 2).

Hence, the nonresonance condition $(\mathrm{NR3})_{\vartheta_0}$ only guarantees that the denominator $\gamma = \alpha^2 - \beta^2$ of b_{11}, b_{22} in (40) is nonvanishing for $(\vartheta,\theta) \in \Gamma(\vartheta_0,\delta)$. Thus, we are able to apply the normal-form transformation of Section 3 only in a neighborhood of ϑ_0. We define $B' = (B'_1, B'_2) : Y\times Y \to Y$ by (40)–(43) with the b_i and b_{ii}, $i = 1,2$, replaced by b'_i and $b'_{ii} := \chi_\delta b_{ii}$, respectively, where $\chi_\delta : \mathbb{T}\times\mathbb{T} \to [0,1]$ is such that $\chi_\delta \in \mathrm{H}^3(\mathbb{T}\times\mathbb{T})$, $\chi_\delta(\vartheta,\theta) = \chi_\delta(\vartheta,\vartheta-\theta)$, and

$$\chi_\delta(\vartheta,\theta) = 1 \quad \text{for } (\vartheta,\theta) \in \Gamma(\vartheta_0,\delta), \quad \chi_\delta(\vartheta,\theta) = 0 \quad \text{for } (\vartheta,\theta) \in \mathbb{T}\times\mathbb{T} \setminus \Gamma(\vartheta_0, 2\delta).$$

Analogously, we define $\widetilde{Q}' = (0, Q') : Y\times Y \to Y$ by (38), replacing Q and q by Q' and $q' := \chi_\delta q$, respectively. The bilinear forms B', $\widetilde{Q}'$ inherit the symmetry of B, $\widetilde{Q}$ (i.e., $q(\vartheta,\theta) = q(\vartheta,\vartheta-\theta)$, $b_{ii}(\vartheta,\theta) = b_{ii}(\vartheta,\vartheta-\theta)$), since $(\vartheta,\theta) \in \Gamma(\vartheta_0,\delta) \Leftrightarrow (\vartheta,\vartheta-\theta) \in \Gamma(\vartheta_0,\delta)$. Moreover, in analogy to the uniformly nonresonant case $(\mathrm{NR3})_{\mathrm{unif}}$, we have

$$2B'(\widetilde{L}\widetilde{x},\widetilde{x}) - \widetilde{L}B'(\widetilde{x},\widetilde{x}) + \widetilde{Q}'(\widetilde{x},\widetilde{x}) = 0 \tag{66}$$

(cf. (34)). Thus, applying to (27) the normal-form transformation $F' : Y \to Y$ with

$$\widetilde{y} = F'(\widetilde{x}) := \widetilde{x} + B'(\widetilde{x},\widetilde{x}),$$

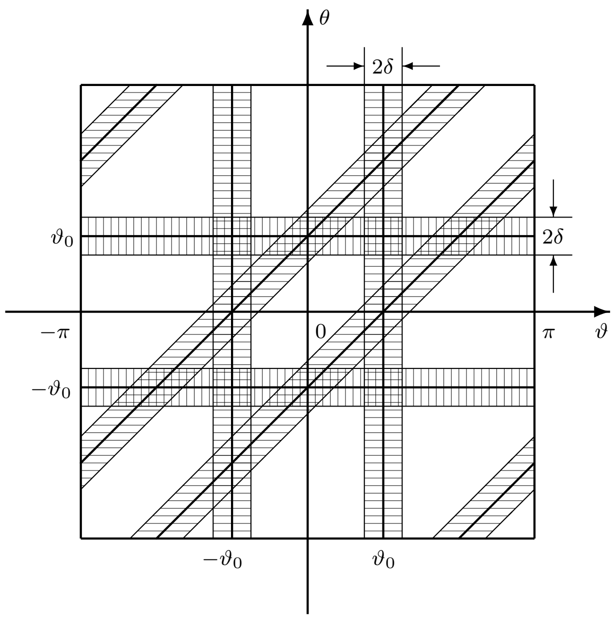

FIGURE 2. Sketch of $\Gamma(\vartheta_0,\delta)$ for $\vartheta_0 \in (-\pi,\pi]$, $\delta > 0$

we obtain

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + \widetilde{S}(\widetilde{x},\widetilde{x}) + \overline{M}'(\widetilde{x}) \tag{67}$$

with

$$\begin{aligned} \widetilde{S}(\widetilde{x},\widetilde{y}) &:= \widetilde{Q}(\widetilde{x},\widetilde{y}) - \widetilde{Q}'(\widetilde{x},\widetilde{y}), \\ \overline{M}'(\widetilde{x}) &:= 2B'(\widetilde{Q}(\widetilde{x},\widetilde{x}) + \widetilde{M}(\widetilde{x}),\widetilde{x}) + \widetilde{M}(\widetilde{x}). \end{aligned} \tag{68}$$

By definition, $\widetilde{S}(\widetilde{x},\widetilde{y}) = (0, S(x,y))$ with $S(x,y) := Q(x,y) - Q'(x,y)$ is bilinear and symmetric, and has the Fourier transform

$$[\widehat{S}(\widehat{x},\widehat{y})](\vartheta) := \widehat{S(x,y)}(\vartheta) = \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\vartheta-\theta)s(\vartheta,\theta)\widehat{y}(\theta)\,\mathrm{d}\theta$$

with $s := (1-\chi_\delta)q$. Since $b'_{ii}, s \in \mathrm{H}^3(\mathbb{T}\times\mathbb{T})$, we obtain by Proposition 2

$$\|b'_i(x,y)\|_{\ell^2} \le c'_i\|x\|_{\ell^2}\|y\|_{\ell^\infty} \quad (i=1,2), \qquad \|S(x,y)\|_{\ell^2} \le c_s\|x\|_{\ell^2}\|y\|_{\ell^\infty} \tag{69}$$

for $x,y \in \ell^2$, and thus

$$\|B'(\widetilde{x},\widetilde{y})\|_Y \le C'_B\|\widetilde{x}\|_Y\|\widetilde{y}\|_\infty, \quad \|\widetilde{S}(\widetilde{x},\widetilde{y})\|_Y \le C_S\|\widetilde{x}\|_Y\|\widetilde{y}\|_\infty \qquad \text{for } \widetilde{x},\widetilde{y} \in Y \tag{70}$$

with $C_S = c_s/\mu_-$ and C'_B defined by C_B with c, c_i replaced by c', c'_i (cf. (45)), which yields also

$$\|B'(\widetilde{x},\widetilde{y})\|_Y \le \mu C'_B\|\widetilde{x}\|_Y\|\widetilde{y}\|_Y, \quad \|\widetilde{S}(\widetilde{x},\widetilde{y})\|_Y \le \mu C_S\|\widetilde{x}\|_Y\|\widetilde{y}\|_Y \qquad \text{for } \widetilde{x},\widetilde{y} \in Y \tag{71}$$

(cf. (46)). Thus, like for $\rho = F^{-1}$, it follows from the Implicit Function Theorem also for $\rho' := (F')^{-1}$ the existence of constants $C'_\rho, \varepsilon'_\rho > 0$, such that

$$\|\rho'(\widetilde{y}_1)-\rho'(\widetilde{y}_2)\|_Y \le C'_\rho\|\widetilde{y}_1-\widetilde{y}_2\|_Y \quad \text{for } \widetilde{y}_1, \widetilde{y}_2 \in Y \text{ with } \|\widetilde{y}_1\|_Y, \|\widetilde{y}_2\|_Y < \varepsilon'_\rho \tag{72}$$

(cf. (48)). Hence, for sufficiently small $\widetilde{x} \in Y$ the system (67) reads in analogy to (49)

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + S'(\widetilde{y},\widetilde{y}) + N'(\widetilde{y}) \quad \text{with } S'(\widetilde{y},\widetilde{y}) := \widetilde{S}(\rho'(\widetilde{y}),\rho'(\widetilde{y})),\ N'(\widetilde{y}) := \overline{M}'(\rho'(\widetilde{y})). \tag{73}$$

Since we were able to apply a normal-form transformation only on a neighborhood of ϑ_0, it is obvious that the transformed system (73) contains still a quadratic term $S'(\widetilde{y},\widetilde{y})$ related to $(\vartheta,\theta) \in \mathbb{T}^2 \setminus \Gamma(\vartheta_0,\delta)$. Even if this term can be controlled (cf. (82)), we will see that its existence means that for a justification result in the sense of Theorem 2 we can only obtain estimates for the error between original solution and approximation of order ε^α with $\alpha > 2$ in comparison to the order $\varepsilon^{3/2}$ appearing in Theorem 2, where the stronger uniform nonresonance condition $(\mathrm{NR3})_{\mathrm{unif}}$ was assumed. In order to show that this increase of the order α is necessary, we present the method of justification in a more abstract way then in Section 4.2, and deduce the needed order α.

In Section 2 we carried out the formal derivation of the NLSE for the multiple scale ansatz $X_\varepsilon^{A,p}$ given in (8) for $p = 3$ and $p = 4$. (In the uniformly nonresonant case $p = 3$ was sufficient; in the present case it will turn out that we need $p = 4$.) This means that we inserted (8) in our original microscopic model (1) and equated the left- and right-hand side coefficients of each term $\varepsilon^k\mathbf{E}^n$ for all $k = 1,\dots,p$, $n = -k,\dots,k$. Thus, we obtained determining equations for all macroscopic functions $A_{k,n}$ (or set $A_{k,n} = 0$), and moreover, as showed in Section 4.2, the estimate

$$\|\mathrm{res}(\widetilde{X}_\varepsilon^{A,p})(t)\|_Y \le \tilde{C}_r\varepsilon^{p+1/2} \quad \text{for } \varepsilon \le \varepsilon_0 \text{ and } \varepsilon^2 t \le \tau_0 \tag{52}$$

under certain regularity conditions on $A_{k,n}$ and for potentials $V_m, W \in \mathrm{C}^{p+2}(\mathbb{R})$. For $p = 4$ it suffices to require $V_m, W \in \mathrm{C}^6(\mathbb{R})$, and $A(0,\cdot) \in \mathrm{H}^7(\mathbb{R})$, $A_{2,1}(0,\cdot) \in \mathrm{H}^6(\mathbb{R})$ for the solutions A of the NLSE (5) and $A_{2,1}$ of (7), which guarantees that there exist constants $C_A, \tau_0 > 0$ such that

$$\max_{k+2l\le 7}\|\partial_\tau^l\partial_\xi^k A(\tau,\cdot)\|_{\mathrm{L}^2(\mathbb{R})} + \max_{k+2l\le 6}\|\partial_\tau^l\partial_\xi^k A_{2,1}(\tau,\cdot)\|_{\mathrm{L}^2(\mathbb{R})} \le C_A \quad \text{for } \tau \le \tau_0 \tag{74}$$

(cf. Section 4.1).

Then, considering the transformed system (73) and the associated transformed approximation $\widetilde{Y}_\varepsilon^{A,p} := F'(\widetilde{X}_\varepsilon^{A,p}) = \widetilde{X}_\varepsilon^{A,p} + B'(\widetilde{X}_\varepsilon^{A,p},\widetilde{X}_\varepsilon^{A,p})$, the residual terms are given by

$$\begin{aligned}
\mathrm{res}(\widetilde{Y}_\varepsilon^{A,p}) :=& \dot{\widetilde{Y}}_\varepsilon^{A,p} - \widetilde{L}\widetilde{Y}_\varepsilon^{A,p} - S'(\widetilde{Y}_\varepsilon^{A,p},\widetilde{Y}_\varepsilon^{A,p}) - N'(\widetilde{Y}_\varepsilon^{A,p})\\
=& \dot{\widetilde{Y}}_\varepsilon^{A,p} + 2B'(\dot{\widetilde{X}}_\varepsilon^{A,p},\widetilde{X}_\varepsilon^{A,p}) - \widetilde{L}\widetilde{X}_\varepsilon^{A,p} - \widetilde{L}B'(\widetilde{X}_\varepsilon^{A,p},\widetilde{X}_\varepsilon^{A,p})\\
&- \widetilde{S}(\widetilde{X}_\varepsilon^{A,p},\widetilde{X}_\varepsilon^{A,p}) - \overline{M}'(\widetilde{X}_\varepsilon^{A,p})\\
=& \dot{\widetilde{X}}_\varepsilon^{A,p} + 2B'(\mathrm{res}(\widetilde{X}_\varepsilon^{A,p}),\widetilde{X}_\varepsilon^{A,p}) - \widetilde{L}\widetilde{X}_\varepsilon^{A,p} - \widetilde{Q}(\widetilde{X}_\varepsilon^{A,p},\widetilde{X}_\varepsilon^{A,p}) - \widetilde{M}(\widetilde{X}_\varepsilon^{A,p})\\
=& \mathrm{res}(\widetilde{X}_\varepsilon^{A,p}) + 2B'(\mathrm{res}(\widetilde{X}_\varepsilon^{A,p}),\widetilde{X}_\varepsilon^{A,p}),
\end{aligned}$$

where we used (66) and (68) with $\widetilde{x} = \widetilde{X}_\varepsilon^{A,p}$. From (52) and (53), it follows by $(70)_1$

$$\|\mathrm{res}(\widetilde{Y}_\varepsilon^{A,p})(t)\|_Y \le \widetilde{C}'_r\varepsilon^{p+1/2} \quad \text{for } \varepsilon \le \varepsilon_0,\ \varepsilon^2 t \le \tau_0 \quad \text{with } \widetilde{C}'_r := (1+2C'_B C\varepsilon_0)\widetilde{C}_r. \tag{75}$$

Inserting into (73) the error $\varepsilon^\alpha \widetilde{R}' := \widetilde{y} - \widetilde{Y}_\varepsilon^{A,p}$ between a solution $\widetilde{y}$ of (73) and the transformed approximation $\widetilde{Y}_\varepsilon^{A,p}$, we obtain the differential equation for the error

$$\begin{aligned}\dot{\widetilde{R}}' = \widetilde{L}\widetilde{R}' &+ \varepsilon^{-\alpha}[S'(\widetilde{Y}_\varepsilon^{A,p}+\varepsilon^\alpha \widetilde{R}', \widetilde{Y}_\varepsilon^{A,p}+\varepsilon^\alpha \widetilde{R}') - S'(\widetilde{Y}_\varepsilon^{A,p}, \widetilde{Y}_\varepsilon^{A,p})] + \\ &+ \varepsilon^{-\alpha}[N'(\widetilde{Y}_\varepsilon^{A,p}+\varepsilon^\alpha \widetilde{R}') - N'(\widetilde{Y}_\varepsilon^{A,p})] - \varepsilon^{-\alpha}\mathrm{res}(\widetilde{Y}_\varepsilon^{A,p}),\end{aligned}$$

which, using the semigroup $G(t) = \mathrm{e}^{t\widetilde{L}}$ associated to the linear problem $\dot{\widetilde{R}}' = \widetilde{L}\widetilde{R}'$ and the variation of constants formula, yields

$$\begin{aligned}&\|\widetilde{R}'(t)\|_Y \le \|\widetilde{R}'(0)\|_Y + \\ &+ \varepsilon^{-\alpha}\int_0^t \Big(\|S'(\widetilde{Y}_\varepsilon^{A,p}(s)+\varepsilon^\alpha \widetilde{R}'(s), \widetilde{Y}_\varepsilon^{A,p}(s)+\varepsilon^\alpha \widetilde{R}'(s)) - S'(\widetilde{Y}_\varepsilon^{A,p}(s), \widetilde{Y}_\varepsilon^{A,p}(s))\|_Y + \\ &\qquad + \|N'(\widetilde{Y}_\varepsilon^{A,p}(s)+\varepsilon^\alpha \widetilde{R}'(s)) - N'(\widetilde{Y}_\varepsilon^{A,p}(s))\|_Y + \|\mathrm{res}(\widetilde{Y}_\varepsilon^{A,p}(s))\|_Y\Big)\,\mathrm{d}s. \quad (76)\end{aligned}$$

Now, let us assume that there exist constants $d', C_r', C_N', C_S' > 0$ independent of a given $D' > 0$, and an $\varepsilon_0 > 0$ depending on D', such that the estimates

$$\|\widetilde{R}'(0)\|_Y \le d', \quad (77)$$

$$\|\mathrm{res}(\widetilde{Y}_\varepsilon^{A,p})(t)\|_Y \le C_r' \varepsilon^{\alpha+2}, \quad (78)$$

$$\|N'(\widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}'(t)) - N'(\widetilde{Y}_\varepsilon^{A,p}(t))\|_Y \le C_N' \varepsilon^{\alpha+2}\|\widetilde{R}'(t)\|_Y, \quad (79)$$

$$\begin{aligned}\|S'(\widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}'(t), \widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}(t)') - S'(\widetilde{Y}_\varepsilon^{A,p}(t), \widetilde{Y}_\varepsilon^{A,p}(t))\|_Y \le \\ \le C_S' \varepsilon^{\alpha+2}\|\widetilde{R}'(t)\|_Y \quad (80)\end{aligned}$$

hold for $\varepsilon \le \varepsilon_0$, $\varepsilon^2 t \le \tau_0$, $\|\widetilde{R}'(t)\|_Y \le D'$. Inserting these estimates in (76), we obtain by Gronwall's Lemma for $D' := (d' + \tau_0 C_r')\mathrm{e}^{\tau_0(C_N' + C_S')}$ and its associated $\varepsilon_0 > 0$ the estimate $\|\widetilde{R}'(t)\|_Y \le D'$ for $\varepsilon \le \varepsilon_0$, $\varepsilon^2 t \le \tau_0$, i.e.,

$$\|\widetilde{y}(t) - \widetilde{Y}_\varepsilon^{A,p}(t)\|_Y \le D'\varepsilon^\alpha \quad \text{for } \varepsilon \le \varepsilon_0,\ \varepsilon^2 t \le \tau_0$$

(cf. the argument in Section 4.2). Thus, for $\varepsilon_0 > 0$ such that $\|\widetilde{y}(t) - \widetilde{Y}_\varepsilon^{A,p}(t)\|_Y < \varepsilon_\rho'/2$, $\|\widetilde{Y}_\varepsilon^{A,p}(t)\|_Y < \varepsilon_\rho'/2$ the inequality (72) yields

$$\|\widetilde{x}(t) - \widetilde{X}_\varepsilon^{A,p}(t)\|_Y \le C_\rho' D' \varepsilon^\alpha \quad \text{for } \varepsilon \le \varepsilon_0,\ \varepsilon^2 t \le \tau_0,$$

and for $\alpha \in (2, 5/2]$ we obtain by $(54)_3$

$$\|\widetilde{x}(t) - \widetilde{X}_\varepsilon^{A,2}(t)\|_Y \le C'\varepsilon^\alpha \quad \text{for } \varepsilon^2 t \le \tau_0,\ \varepsilon \le \varepsilon_0 < 1 \quad \text{with } C' := C_\rho' D' + \varepsilon_0^{5/2-\alpha} C_3.$$

Let us now verify the estimates (77)–(80), and thereby deduce the required values of α and p: Since N' consists of the cubic and higher order nonlinear terms, the proof of (79) follows along the same lines as that of (58) in Section 4.2: By (68), $(70)_1$, (72), (73), we obtain for $\widetilde{x}_1' := \rho'(\widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}'(t))$, $\widetilde{x}_2' := \rho'(\widetilde{Y}_\varepsilon^{A,p}(t)) = \widetilde{X}_\varepsilon^{A,p}(t)$ under the condition

$$\|\widetilde{Y}_\varepsilon^{A,p}(t)\|_Y + \|\varepsilon^\alpha \widetilde{R}'(t)\|_Y \le \varepsilon_\rho' \quad (81)$$

the estimate

$$\|N'(\widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}'(t)) - N'(\widetilde{Y}_\varepsilon^{A,p}(t))\|_Y \le C_n' C_\rho'(\|\widetilde{x}_1'\|_\infty^2 + \|\widetilde{x}_2'\|_\infty^2)\|\varepsilon^\alpha \widetilde{R}'(t)\|_Y$$

with $C_n' > 0$ depending only on C_B', V_m, W and ε_ρ' (cf. (64)). Moreover, for $\alpha > 1$ we can show, that for given $D' > 0$ there exists an $\varepsilon_0 > 0$ such that (81) and

$$\|\widetilde{x}_1'\|_\infty^2 + \|\widetilde{x}_2'\|_\infty^2 \leq 5C^2\varepsilon^2 \quad \text{for } \varepsilon \leq \varepsilon_0,\ \varepsilon^2 t \leq \tau_0,\ \|R'(t)\|_Y \leq D'$$

are satisfied, with $C > 0$ independent of D'. Setting $C_N' := 5C_n'C_\rho'C^2$ we obtain (79).

In order to prove (80), we decompose S' using its bilinearity and symmetry (with $\widetilde{x}_1', \widetilde{x}_2'$ as before):

$$\begin{aligned} &S'(\widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}'(t), \widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}'(t)) - S'(\widetilde{Y}_\varepsilon^{A,p}(t), \widetilde{Y}_\varepsilon^{A,p}(t)) \\ =&\widetilde{S}(\widetilde{x}_1', \widetilde{x}_1') - \widetilde{S}(\widetilde{x}_2', \widetilde{x}_2') = \widetilde{S}(\widetilde{x}_1'-\widetilde{x}_2', \widetilde{x}_1'-\widetilde{x}_2') + 2\widetilde{S}(\widetilde{x}_2', \widetilde{x}_1'-\widetilde{x}_2') \\ =&\widetilde{S}(\widetilde{x}_1'-\widetilde{x}_2', \widetilde{x}_1'-\widetilde{x}_2') + 2\widetilde{S}(\widetilde{X}_\varepsilon^{A,p}(t)-\widetilde{X}_\varepsilon^{A}(t), \widetilde{x}_1'-\widetilde{x}_2') + 2\widetilde{S}(\widetilde{X}_\varepsilon^{A}(t), \widetilde{x}_1'-\widetilde{x}_2'). \end{aligned}$$

By (72) we obtain under condition (81) $\|x_1'-\widetilde{x}_2'\|_Y \leq \varepsilon^\alpha C_\rho' \|\widetilde{R}'(t)\|_Y$. This yields by $(71)_2$

$$\|\widetilde{S}(\widetilde{x}_1'-\widetilde{x}_2', \widetilde{x}_1'-\widetilde{x}_2')\|_Y \leq \varepsilon^{2\alpha}\mu C_S (C_\rho')^2 \|\widetilde{R}'(t)\|_Y^2,$$

and by $(54)_2$ and $(70)_2$

$$\|\widetilde{S}(\widetilde{X}_\varepsilon^{A,4}(t)-\widetilde{X}_\varepsilon^{A}(t), x_1'-\widetilde{x}_2')\|_Y \leq \varepsilon^{2+\alpha} C_S C_2' C_\rho' \|\widetilde{R}'(t)\|_Y$$

for $\varepsilon \leq \varepsilon_0 < 1$, $\varepsilon^2 t \leq \tau_0$. Finally, let us assume for the moment that we can show that there exists a $C_P > 0$ such that

$$\|\widetilde{S}(\widetilde{X}_\varepsilon^{A}(t), \widetilde{z})\|_Y \leq C_S C_P \varepsilon^2 \|\widetilde{z}\|_Y \quad \text{for } \widetilde{z} \in Y \text{ and } \varepsilon \leq \varepsilon_0,\ \varepsilon^2 t \leq \tau_0. \tag{82}$$

Then, for given $D' > 0$ there exists an $\varepsilon_0 > 0$ satisfying (81) and

$$\begin{aligned} &\|S'(\widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}'(t), \widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha \widetilde{R}'(t)) - S'(\widetilde{Y}_\varepsilon^{A,p}(t), \widetilde{Y}_\varepsilon^{A,p}(t))\|_Y \\ \leq\ & \varepsilon^{\alpha+2} C_S C_\rho' (\varepsilon_0^{\alpha-2}\mu C_\rho' D' + 2C_2' + 2C_P) \|\widetilde{R}'(t)\|_Y \end{aligned}$$

for $\varepsilon \leq \varepsilon_0$, $\varepsilon^2 t \leq \tau_0$, $\|\widetilde{R}'(t)\|_Y \leq D'$. Thus, in order to obtain (80) with a constant C_S' independent of D', e.g., $C_S' = 2C_S C_\rho'(\mu C_\rho' + C_2' + C_P)$, which for given D' can be achieved by controlling ε_0, we have to require $\alpha > 2$.

The estimate (78) for the residual $\mathrm{res}(\widetilde{Y}_\varepsilon^{A,p})$ follows from (75) if $p + 1/2 \geq \alpha + 2$, with $C_r' := \widetilde{C}_r' \varepsilon_0^{p+1/2-\alpha-2}$. Since we need $\alpha > 2$, we require necessarily $p > 3/2$, i.e. at least $p = 4$, which is also sufficient for $\alpha \leq 5/2$, and optimal for $\alpha = 5/2$.

Finally, estimate (77) is equivalent to $\|\widetilde{y}(0) - \widetilde{Y}_\varepsilon^{A,p}(0)\|_Y \leq \varepsilon^\alpha d'$ for $\varepsilon \leq \varepsilon_0$. By $\widetilde{y} - \widetilde{Y}_\varepsilon^{A,p} = \widetilde{x} - \widetilde{X}_\varepsilon^{A,p} + B'(\widetilde{x} - \widetilde{X}_\varepsilon^{A,p}, \widetilde{x} - \widetilde{X}_\varepsilon^{A,p}) + 2B'(\widetilde{x} - \widetilde{X}_\varepsilon^{A,p}, \widetilde{X}_\varepsilon^{A,p})$, and $(71)_1$, $(53)_2$, it has to hold $\|\widetilde{x}(0) - \widetilde{X}_\varepsilon^{A,p}(0)\|_Y \leq \varepsilon^\alpha \widetilde{d}$ for $\varepsilon \leq \varepsilon_0$ and a $\widetilde{d} > 0$. For $\alpha \in (2, 5/2]$ this is by $(54)_3$ equivalent to the assumption $\|\widetilde{x}(0) - \widetilde{X}_\varepsilon^{A,2}(0)\|_Y \leq \varepsilon^\alpha c'$ for $\varepsilon \leq \varepsilon_0$ and a $c' > 0$.

Hence, except for the estimate (82) which is proven below, we proved the following result for the case were only the nonresonance condition $(\mathrm{NR3})_{\vartheta_0}$ holds:

Theorem 3. *Assume that V_m, $W \in \mathrm{C}^6(\mathbb{R})$ in (1) have the form (2) and that the stability condition (SC) and the nonresonance conditions $(NR2)_{\vartheta_0}^4$ and $(NR3)_{\vartheta_0}$ hold. Let A, $A_{2,1} : [0, \tau_0] \times \mathbb{R} \to \mathbb{C}$, $\tau_0 > 0$, be the solutions of the NLSE (5) with $A(0, \cdot) \in \mathrm{H}^7(\mathbb{R})$ and of (7) with $A_{2,1}(0, \cdot) \in \mathrm{H}^6(\mathbb{R})$, respectively, and let $X_\varepsilon^{A,2}$ be the formal approximation (6). Then, for each $c' > 0$ there exist $\varepsilon_0, C' > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the following statement holds:*

Any solution $\widetilde{x}$ of (27) with an initial condition $\widetilde{x}(0)$ satisfying

$$\|\widetilde{x}(0)-\widetilde{X}_\varepsilon^{A,2}(0)\|_Y \leq c'\varepsilon^\alpha \quad \textit{with } \alpha \in (2, 5/2]$$

fulfills the estimate

$$\|\widetilde{x}(t)-\widetilde{X}_\varepsilon^{A,2}(t)\|_Y \leq C'\varepsilon^\alpha \quad \textit{for} \quad t \in [0, \tau_0/\varepsilon^2].$$

(In the case of nearest-neighbor interactions $(\mathrm{NR2})^4_{\vartheta_0}$ is implied by (SC) and $(\mathrm{NR3})_{\vartheta_0}$, cf. (23) in Proposition 1.)

In particular, the theorem holds for $A_{2,1}(0,\cdot) \equiv 0$. Then, the approximation $X_\varepsilon^{A,2}$ depends only on the solution A of the initial value problem for the NLSE (5) with $A(0,\cdot) \in \mathrm{H}^7(\mathbb{R})$, since $A_{2,1}$ can be calculated by (7) with $A_{2,1}(0,\cdot) \equiv 0$.

Proof of (82). It suffices to prove the existence of a $C_p > 0$ such that

$$\|S(X_\varepsilon^A(t), z)\|_{\ell^2} \leq c_s C_p \varepsilon^2 \|z\|_{\ell^2} \quad \text{for } z \in \ell^2 \text{ and } \varepsilon \leq \varepsilon_0,\ \varepsilon^2 t \leq \tau_0$$

with the $c_s > 0$ given by $(69)_2$. We define $\mathcal{P} : \ell^2 \to \ell^2$ by $\widehat{\mathcal{P}}\widehat{x} := \widehat{\mathcal{P}x} := \gamma\widehat{x}$ with

$$\gamma(\vartheta) := \begin{cases} 0 & \text{for } \vartheta \in \mathbb{T} \text{ with } |\vartheta \pm \vartheta_0| < \delta, \\ 1 & \text{else} \end{cases}$$

with the $\delta > 0$ given in (65). Thus, by the definitions of S and $\Gamma(\vartheta_0, \delta)$, we obtain $\gamma(\vartheta-\theta)s(\vartheta,\theta) = s(\vartheta,\theta)\gamma(\theta) = s(\vartheta,\theta)$ for all $(\vartheta,\theta) \in \mathbb{T}\times\mathbb{T}$, which implies $\widehat{S}(\widehat{\mathcal{P}}\widehat{x}, \widehat{y}) = \widehat{S}(\widehat{x}, \widehat{\mathcal{P}}\widehat{y}) = \widehat{S}(\widehat{x}, \widehat{y})$, and hence $S(\mathcal{P}x, y) = S(x, \mathcal{P}y) = S(x, y)$. This yields by (69)

$$\|S(x,z)\|_{\ell^2} \leq c_s \|\mathcal{P}x\|_{\ell^\infty} \|z\|_{\ell^2}.$$

Since $\|x\|_{\ell^\infty}^2 \leq \|x\|_{\ell^2}^2 = \|\widehat{x}\|_{\mathrm{L}^2(\mathbb{T})}^2 := \frac{1}{2\pi}\int_{\mathbb{T}} |\widehat{x}(\vartheta)|^2 \,\mathrm{d}\vartheta$, it remains to show

$$\|\widehat{\mathcal{P}}\widehat{X}_\varepsilon^A(t)\|_{\mathrm{L}^2(\mathbb{T})} \leq C_p\varepsilon^2 \quad \text{for } \varepsilon \leq \varepsilon_0,\ \varepsilon^2 t \leq \tau_0. \tag{83}$$

Setting $a_j := a(\varepsilon j) := A(\varepsilon^2 t, \varepsilon(j+\omega' t))$ for $j \in \mathbb{Z}$, we obtain

$$[\widehat{X}_\varepsilon^A(t)](\vartheta) = \varepsilon \mathrm{e}^{\mathrm{i}\omega t}\widehat{a}(\vartheta-\vartheta_0) + \varepsilon \mathrm{e}^{-\mathrm{i}\omega t}\overline{\widehat{a}}(-\vartheta-\vartheta_0),$$

and hence

$$|[\widehat{X}_\varepsilon^A(t)](\vartheta)|^2 \leq \varepsilon^2 2(|\widehat{a}(\vartheta-\vartheta_0)|^2 + |\widehat{a}(-\vartheta-\vartheta_0)|^2).$$

By the definition of $\widehat{\mathcal{P}}$, this yields

$$\|\widehat{\mathcal{P}}\widehat{X}_\varepsilon^A(t)\|_{\mathrm{L}^2(\mathbb{T})}^2 \leq \varepsilon^2 4 \frac{1}{2\pi}\int_{\delta \leq |\eta| \leq \pi} |\widehat{a}(\eta)|^2 \,\mathrm{d}\eta.$$

Considering $\phi \in \ell^2$ with $\phi_0 = 1$, $\phi_1 = \phi_{-1} = -1/2$ and $\phi_k = 0$ for $k \in \mathbb{Z}$, $|k| \geq 2$, we obtain $\widehat{\phi}(\eta) = 1 - \cos\eta$ and

$$\int_{\delta \leq |\eta| \leq \pi} |\widehat{a}(\eta)|^2 \,\mathrm{d}\eta \leq \frac{1}{(1-\cos\delta)^2}\int_{-\pi}^{\pi} |\widehat{\phi}(\eta)\widehat{a}(\eta)|^2 \,\mathrm{d}\eta = \frac{2\pi}{(1-\cos\delta)^2}\|\phi * a\|_{\ell^2}^2.$$

Since

$$\begin{aligned} |(\phi * a)_j|^2 &= |\sum_{k\in\mathbb{Z}} \phi_k a_{j-k}|^2 = |a_j - \frac{1}{2}(a_{j+1}+a_{j-1})|^2 = \frac{1}{4}|(a_{j+1}-a_j)-(a_j-a_{j-1})|^2 \\ &= \frac{\varepsilon^2}{4}|a'(\varepsilon x_j^+)-a'(\varepsilon x_j^-)|^2 = \frac{\varepsilon^4}{4}\Big|\int_{x_j^-}^{x_j^+} a''(\varepsilon x)\,\mathrm{d}x\Big|^2 \leq \frac{\varepsilon^4}{4}\Big(\int_{j-1}^{j+1} |a''(\varepsilon x)|\,\mathrm{d}x\Big)^2 \\ &\leq \frac{\varepsilon^4}{2}\int_{j-1}^{j+1} |a''(\varepsilon x)|^2 \,\mathrm{d}x \end{aligned}$$

with $x_j^- \in (j-1,j)$, $x_j^+ \in (j,j+1)$, we obtain

$$\|\phi * a\|_{\ell^2}^2 \le \varepsilon^4 \int_{\mathbb{R}} |a''(\varepsilon x)|^2 \,\mathrm{d}x = \varepsilon^3 \int_{\mathbb{R}} |a''(\xi)|^2 \,\mathrm{d}\xi = \varepsilon^3 \|\partial_\xi^2 A(\varepsilon^2 t,\cdot)\|_{\mathrm{L}^2(\mathbb{R})}^2.$$

Hence, by (74) we obtain (83), and thus (82), with $C_p := \varepsilon_0^{1/2} 2C_A/(1-\cos\delta)$. $\square$

Acknowledgments. This work has been supported by the DFG Priority Program 1095 "Analysis, Modeling and Simulation of Multiscale Problems" under Mi 459/3.

REFERENCES

[1] A. M. Balk, A. V. Cherkaev and L. I. Slepyan, *Dynamics of chains with non-monotone stress-strain relations. I. Model and numerical experiments. II. Nonlinear waves and waves of phase transition*, J. Mech. Phys Solids, **49** (2001) no. 1, 131–148, 149–171.

[2] D. Bambusi, A. Carati and A. Ponno, *The Nonlinear Schrödinger equation as a resonant normal form*, Discrete Cont. Dyn. Syst. Ser. B, **2** (2002), 109–128.

[3] D. Bambusi and A. Ponno, *On Metastability in FPU*, Commun. Math. Phys., to appear.

[4] C. Boldrighini, *Macroscopic limits of microscopic systems*, Rend. Mat. Appl. (7), **16** (1996), no. 1, 1–107.

[5] D. K. Campbell, P. Rosenau and G. M. Zaslavsky (eds.) *Focus issue: The "Fermi-Pasta-Ulam" problem – the first 50 years*, Chaos, **15** (2005), no. 1, 015101.

[6] P. Deift, S. Kamvissis, T. Kriecherbauer and X. Zhou, *The Toda rarefaction problem*, Comm. Pure Appl. Math., **49** (1996), no. 1, 35–83.

[7] W. Dreyer, M. Herrmann and A. Mielke, *Micro–macro transitions for the atomic chain via Whitham's modulation equation*, Nonlinearity, **19** (2006), 471–500.

[8] W. Dreyer and M. Kunik, *Cold, thermal and oscillator closure of the atomic chain*, J. Phys. A, **33** (2000), no. 10, 2097–2129, (Corrigendum: J. Phys. A, **33** (2000) 2458).

[9] P. Deift, T. Kriecherbauer, and S. Venakides. *Forced lattice vibrations. I, II*, Comm. Pure Appl. Math., **48** (1995), 1187–1249, 1251–1298.

[10] E. Fermi, J. Pasta, and S. Ulam, *Studies of nonlinear problems*, Los Alamos Scientific Laboratory of the University of California, Report LA-1940, May 1955.

[11] A.-M. Filip and S. Venakides, *Existence and modulation of traveling waves in particle chains*, Comm. Pure Appl. Math., **52** (1999), no. 6, 693–735.

[12] N. Flytzanis, S. Pnevmatikos, and M. Remoissenet, *Kink, breather and asymmetric envelope or dark solitons in nonlinear chains. I. Monatomic chain*, J. Phys. C, **18** (1985), 4603–4630.

[13] G. Friesecke and K. Matthies, *Atomic-scale localization of high-energy solitary waves on lattices*, Physica D, **171** (2002), no. 4, 211–220.

[14] G. Friesecke and R. L. Pego, *Solitary waves on FPU lattices: I. Qualitative properties, renormalization and continuum limit*, Nonlinearity, **12** (1999), 1601–1627.

[15] G. Friesecke and R. L. Pego, *Solitary waves on FPU lattices: II. Linear implies nonlinear stability*, Nonlinearity, **15** (2002), 1343–1359.

[16] G. Friesecke and R. L. Pego, *Solitary waves on FPU lattices: III. Howland-type Floquet theory*, Nonlinearity, **17** (2004), 207–227.

[17] G. Friesecke and R. L. Pego, *Solitary waves on FPU lattices: IV. Proof of stability at low energy*, Nonlinearity, **17** (2004), 229–251.

[18] G. Friesecke and J. A. D. Wattis, *Existence theorem for solitary waves on lattices*, Comm. Math. Phys., **161** (1994), no. 2, 391–418.

[19] J. Giannoulis and A. Mielke, *The nonlinear Schrödinger equation as a macroscopic limit for an oscillator chain with cubic nonlinearities*, Nonlinearity, **17** (2004), 551–565.

[20] M. H. Hays, C. D. Levermore, and P. D. Miller, *Macroscopic lattice dynamics*, Physica D, **79** (1994), no. 1, 1–15.

[21] M. Herrmann, Ein Mikro-Makro-Übergang für die nichtlineare atomare Kette mit Temperatur, Dissertation, Humboldt-Universität zu Berlin, 2004

[22] G. Iooss and G. James, *Localized waves in nonlinear oscillator chains*, Chaos, **15** (2005), no. 1, 015113.

[23] G. Iooss and K. Kirchgässner, *Travelling waves in a chain of coupled nonlinear oscillators*, Comm. Math. Phys., **211** (2000), no. 2, 439–464.

[24] G. Iooss, *Travelling waves in the Fermi-Pasta-Ulam lattice*, Nonlinearity, **13** (2000) no. 3, 849–866.

[25] G. James, *Centre manifold reduction for quasilinear discrete systems*, J. Nonlinear Sci., **13** (2003), no. 1, 27–63.

[26] L. A. Kalyakin, *Long wave asymptotics. Integrable equations as asymptotic limits of nonlinear systems*, Russian Math. Surveys, **44** (1989), no. 1, 3–42.

[27] P. Kirrmann, G. Schneider and A. Mielke, *The validity of modulation equations for extended systems with cubic nonlinearities*, Proc. Roy. Soc. Edinburgh Sect. A, **122** (1992), 85–91.

[28] V. V. Konotop, *Small-amplitude envelope solitons in nonlinear lattices*, Phys. Rev. E, **53** (1996), 2843–2858.

[29] R. S. MacKay and S. Aubry, *Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators*, Nonlinearity, **7** (1994), no. 6, 1623–1643.

[30] A. Mielke, *Macroscopic behavior of microscopic oscillations in harmonic lattices via Wigner-Husimi transforms*, Arch. Ration. Mech. Anal., to appear.

[31] A. Mielke, *The Ginzburg–Landau equation in its role as a modulation equation*, In B. Fiedler, editor, Handbook of Dynamical Systems II, pages 759–834. Elsevier Science B.V., 2002.

[32] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, volume 44 of Applied Mathematical Sciences, Springer-Verlag, New York, 1983.

[33] A. Ponno and D. Bambusi, *Korteweg-de Vries equation and energy sharing in Fermi-Pasta-Ulam*, Chaos, **15** (2005), no. 1, 015107.

[34] M. Remoissenet, *Low-amplitude breather and envelope solitons in quasi-one-dimensional physical models*, Physical Review B, **33** (1986), no. 4, 2386–2392.

[35] G. Schneider, *Justification of modulation equations for hyperbolic systems via normal forms*, NoDEA Nonlinear Differential Equations Appl., **5** (1998), no. 1, 69–82.

[36] G. Schneider and C. E. Wayne. Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model. In B. Fiedler, K. Gröger, and J. Sprekels, editors, *International Conference on Differential Equations*, Volume 1, pages 390–404. World Scientific, 2000.

[37] H. Spohn, Large scale dynamics of interacting particles. Springer-Verlag, New York, 1991.

[38] R. Temam, *Infinite-dimensional dynamical systems in mechanics and physics*, Springer-Verlag, New York, 1988.

Received July 2005; revised November 2005.

E-mail address: giannoulis@wias-berlin.de
E-mail address: mielke@wias-berlin.de