Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Dispersive evolution of pulses in oscillator chains with general interaction potentials

Johannes Giannoulis, Alexander Mielke

submitted: 3rd August 2005

Weierstraß–Institut
für Angewandte Analysis
und Stochastik
Mohrenstraße 39
10117 Berlin, Germany
E-Mail: giannoul@wias-berlin.de
mielke@wias-berlin.de

No. 1050
Berlin 2005

2000 *Mathematics Subject Classification.* 34E13, 34C20, 37K60, 70F45, 70K70.

Key words and phrases. Nonlinear oscillator chain, multiscale ansatz, nonlinear Schrödinger equation, justification of modulation equations, normal form transformation, nonresonance conditions.

Research partially supported by the DFG Priority Program *Analysis, Modeling and Simulation of Multiscale Problems* under Mi 459/3.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We consider the dispersive evolution of a single pulse in a nonlinear oscillator chain embedded in a background field. We assume that each atom of the chain interacts pairwise with an arbitrary but finite number of neighbours. The pulse is modeled as a macroscopic modulation 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. Since the prototype of a discrete model is a lattice (modeling, e.g., a crystal), it is natural that starting with the seminal work of Fermi, Pasta and Ulam [FPU55], a lot of interest and work is attracted to the simplest, one-dimensional representant, viz. the monoatomic, infinite oscillator chain:

$$\ddot{x}_j = \sum_{m=1}^{M} [V_m'(x_{j+m} - x_j) - V_m'(x_j - x_{j-m})] - W'(x_j), \quad j \in \mathbb{Z}, \tag{1.1}$$

where $x_j(t) \in \mathbb{R}$ is the deviation of an atom from its rest position $j \in \mathbb{Z}$ at time $t \geq 0$, due to the interaction potentials V_m with its m-th neighbours and an on-site potential W, coupling the atoms to a background field.

Here, 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 [Mie02] 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. [Spo91, Bol96]), 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., [DKKZ96, DKV95]). Second, a big body of work is concentrated on the dynamics of special types of solutions like solitons, breathers or wave trains [FrW94, MaA94, Kon96, FrP99, Ioo00, IoK00, FrM02, FrP02, Jam03, FrP04a, FrP04b, IoJ05]. Third, one can be interested in the response of the oscillator chain to a simple initial disturbance [BCS01] or to Riemann initial data [DKKZ96, DKV95].

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 [Mie05] 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. 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 hyperboling scaling is used in [FiV99, DHM05, Her04], where for $W \equiv 0$ the modulations of large-amplitude travelling waves are considered, and the derived macroscopic limit is the so-called Whitham modulation equation. Supported by numerical investigations in [Her04], the validity of this equation is discussed in detail and for special cases it is rigorously justified in [DHM05]. A similar modulation ansatz has been used in [HLM94] 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 $A_j(t) \in \mathbb{C}$.

However, closest to our work is the justification of the Korteweg-de Vries equation as the long wave-length limit in [Kal89, FrP99, ScW00]. There, for $W \equiv 0$ small-amplitude solutions of the 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 $\partial_\tau U + \kappa_1 U \partial_\xi U + \kappa_2 \partial_\xi^3 U = 0$.

Like [GiM04] the present work is 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 "conjugate complex" and) A satisfies the nonlinear Schrödinger equation $\mathrm{i}\partial_\tau A = \gamma_1 \partial_\xi^2 A + \gamma_2 |A|^2 A$. Our aim is to generalize [GiM04] 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 neighbours. To be more specific, we consider potentials of the form

$$V_m(d) := \frac{\alpha_{m,1}}{2} d^2 + \frac{\alpha_{m,2}}{3} d^3 + \frac{\alpha_{m,3}}{4} d^4 + \mathcal{O}(d^5), \quad W(x) := \frac{\beta_1}{2} x^2 + \frac{\beta_2}{3} x^3 + \frac{\beta_3}{4} x^4 + \mathcal{O}(x^5) \quad (1.2)$$

for $m = 1, \ldots, M$. (In particular, [GiM04] relates to the case $M = 1$ and $\alpha_{1,2} = 0 = \beta_2$, which leads to a much simpler analysis.) We investigate solutions which are microscopically periodic in space and time. The linearized model is given by

$$\ddot{x}_j = L_j(x) := \sum_{m=1}^{M} \alpha_{m,1}(x_{j+m} - 2x_j + x_{j-m}) - \beta_1 x_j. \tag{1.3}$$

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

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

Throughout, we require that a stability condition holds:

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

and we take $\omega(\vartheta) > 0$. In the case of interactions only between nearest neighbours ($M = 1$, $V := V_1$, $\alpha_k := \alpha_{1,k}$) (SC) is equivalent to $\min\{\beta_1, 4\alpha_1 + \beta_1\} > 0$ (cf. (2.14)). 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.

Our aim is to understand the macroscopic evolution of solutions, which are modulations of the microscopic pattern, given by a modulation function $A : [0, \infty) \times \mathbb{R} \to \mathbb{C}$:

$$x_j(t) = (X^A_\varepsilon)_j(t) + \mathcal{O}(\varepsilon^2) \quad \text{with} \quad (X^A_\varepsilon)_j(t) := \varepsilon A(\varepsilon^2 t, \varepsilon(j - c_{\mathrm{gr}} t))\mathbf{E}(t, j) + \text{c.c.} \tag{1.4}$$

with $\varepsilon \leq \varepsilon_0$ for some $\varepsilon_0 > 0$. We let $\tau = \varepsilon^2 t$ and $\xi = \varepsilon(j - c_{\mathrm{gr}} t)$ for the macroscopic time and space variable, respectively. Since the solutions given through (1.4) 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 effects. In the hyperbolic scaling $\tau = \varepsilon t$ with $\xi = \varepsilon j$ one only sees hyperbolic transport effects, but no dispersion.

Inserting such an ansatz into (1.1), it turns out that this provides a useful approximation for solutions of (1.1) only if the group velocity c_{gr} equals $-\omega'$, and A satisfies the associated nonlinear Schrödinger equation (NLSE)

$$2\mathrm{i}\omega\partial_\tau A = \omega\omega''\partial_\xi^2 A + \rho|A|^2 A, \tag{1.5}$$

where ρ can be calculated explicitly (cf. (2.12)). A formal derivation of (1.5) is obtained by assuming that solutions in the form (1.4) exist (cf. Section 2).

The mathematical justification 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 (1.4) stay in this form over intervals $[0, \tau_0/\varepsilon^2]$ of positive macroscopic length $\tau_0 > 0$. More precisely, Theorem 4.1 states the following: Given a sufficiently smooth solution A of NLSE (1.5), $\tau_0 > 0$ and $d > 0$, there exist $\varepsilon_0 > 0$ and $C > 0$ such that for all $\varepsilon \leq \varepsilon_0$ any solution x of (1.1) with

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

satisfies the estimate

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

We prove this result in principle by the same approach we used in our previous paper [GiM04] on this subject. There, we considered the situation of only nearest-neighbour interactions and restricted the justification of the NLSE on the case of cubic leading terms of the nonlinearity in (1.1) (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 [KSM92], relying on a Gronwall type argument. Thus, in the general case of quadratic leading terms treated here, this Gronwall type argument can not be used directly. We circumvent this difficulty by a method which was developed in [Sch98] 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 (1.1)) 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 4.1, by using the Gronwall type argument mentioned above. Transforming this result back into the variable $\widetilde{x}$, we obtain Theorem 4.1.

The construction of the normal form transform 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 $\vartheta_0 \in (-\pi, \pi]$:

$$\exists\, C^{\mathrm{NR}}_{\vartheta_0} > 0 : \qquad \inf_{s,t=1,2;\ \theta\in(-\pi,\pi]} \left|\omega(\vartheta_0)+(-1)^s\omega(\theta)+(-1)^t\omega(\vartheta_0-\theta)\right| \geq C^{\mathrm{NR}}_{\vartheta_0} > 0. \qquad (\mathrm{NR3})_{\vartheta_0}$$

Our first result is Theorem 4.1 which is proved under a strengthened version of $(\mathrm{NR3})_{\vartheta_0}$, which we call *uniform nonresonance condition*

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

In the case of nearest-neighbour 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 $\min\{\beta_1, (16/3)\alpha_1+\beta_1\} > 0$ (cf. (2.17)), which is slightly sharper than the stability condition (SC) $\min\{\beta_1, 4\alpha_1+\beta_1\} > 0$. However, for $\alpha_1 \geq 0$, 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

$$X^{A,2}_\varepsilon := \varepsilon A\mathbf{E} + \varepsilon^2\left(\frac{\beta_2}{\delta_0}|A|^2+A_{2,1}\mathbf{E}+\frac{a}{\delta_2}A^2\mathbf{E}^2\right) + \text{c.c.}, \tag{1.6}$$

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 (1.5) for $\tau \in [0, \tau_0]$ and $A_{2,1} : [0, \tau_0] \times \mathbb{R} \to \mathbb{C}$ solves the equation

$$2\mathrm{i}\omega\partial_\tau A_{2,1} = \omega\omega''\partial_\xi^2 A_{2,1}+\rho(2|A|^2A_{2,1}+A^2\overline{A}_{2,1})-2\omega'\partial_\tau\partial_\xi A-\frac{\mathrm{i}}{6}(\omega^2)'''\partial_\xi^3 A+2e|A|^2\partial_\xi A \tag{1.7}$$

where again e can be given explicitly (cf. (2.13)). 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.

2 The formal derivation of the NLSE

The formal derivation of the NLSE as a modulation equation for the oscillator chain model (1.1) with $M = 1$ has been presented in full detail in [GiM04, 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 [Kon96].

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. (1.7)), we take the opportunity to present the formal derivation in a more general way, culminating in the equation system (2.11), 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. (2.1)) for arbitrary $p \in \mathbb{N}$.

Since we want to study the macroscopic evolution of modulated solutions of the form (1.4), it is naturally to insert such an ansatz into our microscopic model (1.1) 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 (1.1) will generate higher harmonic terms (with factors $\mathbf{E}^n$) having scaling parameters ε^k, $k \in \mathbb{N}$. Hence, we insert into (1.1) the multiple scale ansatz

$$X_\varepsilon^{A,p} := \sum_{k=1}^{p} \varepsilon^k \sum_{n=-k}^{k} A_{k,n}\mathbf{E}^n \tag{2.1}$$

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

$$(\ddot{X}_\varepsilon^{A,p})_j = \sum_{m=1}^{M}\{V_m'[(X_\varepsilon^{A,p})_{j+m} - (X_\varepsilon^{A,p})_j] - V_m'[(X_\varepsilon^{A,p})_j - (X_\varepsilon^{A,p})_{j-m}]\} - W'[(X_\varepsilon^{A,p})_j] \tag{2.2}$$

in terms of $\varepsilon^k\mathbf{E}^n$. Then, by equating the left- and right-hand side coefficients of each of these terms for $k = 1, \ldots, p,\ n = 0, \ldots, k$ 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 for the left hand side

$$\ddot{X}_\varepsilon^{A,p} = \sum_{k=1}^{p} \varepsilon^k \sum_{q=1}^{k}\sum_{n=-q}^{q}\sum_{\mu+2\nu=k-q} c_{n\mu\nu}\partial_\tau^\nu\partial_\xi^\mu A_{q,n}\mathbf{E}^n + \varepsilon^{p+1}r_\varepsilon^{D,p} \tag{2.3}$$

with

$$\varepsilon^{p+1}r_\varepsilon^{D,p} := \sum_{k=p+1}^{p+4} \varepsilon^k \sum_{q=1}^{p}\sum_{n=-q}^{q}\sum_{\mu+2\nu=k-q} c_{n\mu\nu}\partial_\tau^\nu\partial_\xi^\mu A_{q,n}\mathbf{E}^n, \tag{2.4}$$

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

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

We introduce for convenience the expression

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

By Taylor expansion we obtain

$$\partial_j^{\pm m} X_\varepsilon^{A,p} = \pm \sum_{k=1}^{2p} \varepsilon^k \sum_{q=\max\{1,k-p\}}^{\min\{k,p\}} \sum_{n=-q}^{q} f_{q(k-q)n}^{\pm m} \mathbf{E}^n \tag{2.5}$$

with

$$f_{q0n}^{\pm m} = (\mathrm{e}^{\pm \mathrm{i}mn\vartheta_0} - 1)A_{q,n}, \qquad f_{qrn}^{\pm m} = \mathrm{e}^{\pm \mathrm{i}mn\vartheta_0} \frac{(\pm m)^r}{r!} \partial_\xi^r A_{q,n} \quad \text{for } r = 1, \ldots, p-1,$$
$$f_{qpn}^{\pm m} = \mathrm{e}^{\pm \mathrm{i}mn\vartheta_0} \frac{(\pm m)^p}{p!} \partial_\xi^p A_{q,n}(\tau, \xi \pm \theta_{pqn}^{\pm \varepsilon m} \varepsilon m) \quad \text{with } \theta_{pqn}^{\pm \varepsilon m} \in (0,1).$$

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

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

we obtain for the linear part of the right hand side

$$\begin{aligned} L_j X_\varepsilon^{A,p} &= \sum_{m=1}^{M} \alpha_{m,1}(\partial_j^{+m} X_\varepsilon^{A,p} - \partial_j^{-m} X_\varepsilon^{A,p}) - \beta_1 (X_\varepsilon^{A,p})_j \\ &= \sum_{k=1}^{p} \varepsilon^k \sum_{q=1}^{k} \sum_{n=-q}^{q} \frac{(-\mathrm{i})^{k-q+2}}{(k-q)!} \left. \frac{\mathrm{d}^{k-q} \omega^2(\vartheta)}{\mathrm{d}\vartheta^{k-q}} \right|_{\vartheta = n\vartheta_0} \partial_\xi^{k-q} A_{q,n} \mathbf{E}^n + \varepsilon^{p+1} r_\varepsilon^{L,p} \end{aligned} \tag{2.6}$$

with

$$\varepsilon^{p+1} r_\varepsilon^{L,p} := \sum_{k=p+1}^{2p} \varepsilon^k \sum_{q=k-p}^{p} \sum_{n=-q}^{q} \sum_{m=1}^{M} \alpha_{m,1}(f_{q(k-q)n}^{+m} + f_{q(k-q)n}^{-m}) \mathbf{E}^n \tag{2.7}$$

Splitting $\partial_j^{\pm m} X_\varepsilon^{A,p} = \partial_j^{\pm m} X_\varepsilon^{A,\leq p} + \partial_j^{\pm m} X_\varepsilon^{A,>p}$ with

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

the nonlinear part of the right hand side of equation (2.2) reads

$$N_j(X_\varepsilon^{A,p}) = \sum_{s=2}^{p}\{\sum_{m=1}^{M}\alpha_{m,s}[(\partial_j^{+m}X_\varepsilon^{A,\le p})^s - (\partial_j^{-m}X_\varepsilon^{A,\le p})^s] - \beta_s(X_\varepsilon^{A,p})_j^s\}$$

$$+\sum_{s=2}^{p}\sum_{m=1}^{M}\alpha_{m,s}\sum_{\sigma=1}^{s}\binom{s}{\sigma}[(\partial_j^{+m}X_\varepsilon^{A,\le p})^{s-\sigma}(\partial_j^{+m}X_\varepsilon^{A,>p})^\sigma - (\partial_j^{-m}X_\varepsilon^{A,\le p})^{s-\sigma}(\partial_j^{-m}X_\varepsilon^{A,>p})^\sigma]$$

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

with

$$v_{m,p}(d) := V_m'(d) - \sum_{s=1}^{p}\alpha_{m,s}d^s, \quad w_p(x) := W'(x) - \sum_{s=1}^{p}\beta_s x^s \tag{2.8}$$

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 $(i)_s := (i_1,\ldots,i_s)$ with $i_t\in\{1,\ldots,p\}$, $|(i)_s| := \sum_{t=1}^s i_t$ 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

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

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

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

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

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

$$\sum_{s=2}^{p}\alpha_{m,s}(\partial_j^{\pm m}X_\varepsilon^{A,\le p})^s = \sum_{k=2}^{p}\varepsilon^k\sum_{n=-k}^{k}\sum_{s=2}^{k}\alpha_{m,s}\sum_{\substack{|(i,\nu)_s|=(k,n),\\ (|\nu|)_s\le(i)_s}}\sum_{\substack{(\max\{1,|\nu|\})_s\\ \le(q)_s\le(i)_s}}d_{(i-q,\nu)_s}^{\pm m}\partial_\xi^{(i-q)_s}A_{(q,\nu)_s}\mathbf{E}^n$$

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

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

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

This leads to

$$N_j(X_\varepsilon^{A,p}) = \sum_{k=2}^{p} \varepsilon^k \sum_{n=-k}^{k} \sum_{s=2}^{k} \sum_{\substack{|(i,\nu)_s|=(k,n),\\ (|\nu|)_s\le(i)_s}} \sum_{\substack{(\max\{1,|\nu|\})_s\\ \le(q)_s\le(i)_s}} (-1) e^{M,s}_{(i-q,\nu)_s} \partial_\xi^{(i-q)_s} A_{(q,\nu)_s} \mathbf{E}^n + \varepsilon^{p+1} r_\varepsilon^{N,p} \tag{2.9}$$

with

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

and

$$\begin{aligned} \varepsilon^{p+1} r_\varepsilon^{N,p} := & \sum_{k=p+1}^{p^2} \varepsilon^k \sum_{s=[(k-1)/p]+1}^{k} \sum_{|(i)_s|=k} [\sum_{m=1}^{M} \alpha_{m,s} (b^{+m}_{(i)_s} - b^{-m}_{(i)_s}) - \beta_s a_{(i)_s}] \\ & + \sum_{s=2}^{p} \sum_{m=1}^{M} \alpha_{m,s} \sum_{\sigma=1}^{s} \binom{s}{\sigma} [(\partial_j^{+m} X_\varepsilon^{A,\le p})^{s-\sigma} (\partial_j^{+m} X_\varepsilon^{A,>p})^{\sigma} \\ & \qquad - (\partial_j^{-m} X_\varepsilon^{A,\le p})^{s-\sigma} (\partial_j^{-m} X_\varepsilon^{A,>p})^{\sigma}] \\ & + \sum_{m=1}^{M} [v_{m,p}(\partial_j^{+m} X_\varepsilon^{A,p}) - v_{m,p}(\partial_j^{-m} X_\varepsilon^{A,p})] - w_p[(X_\varepsilon^{A,p})_j]. \end{aligned} \tag{2.10}$$

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 ommited since they are just the complex conjugates of the terms for $n = 1,\dots,k$), we obtain the equations that determine the functions $A_{k,n}$

$$\begin{aligned} \delta_n(\vartheta_0) A_{k,n} = & \sum_{q=\max\{1,n\}}^{k-1} \{ \sum_{\mu+2\nu=k-q} c_{n\mu\nu} \partial_\tau^\nu \partial_\xi^\mu A_{q,n} + \frac{(-\mathrm{i})^{k-q}}{(k-q)!} \left. \frac{\mathrm{d}^{k-q}\omega^2(\vartheta)}{\mathrm{d}\vartheta^{k-q}} \right|_{\vartheta=n\vartheta_0} \partial_\xi^{k-q} A_{q,n} \} \\ & + \sum_{s=2}^{k} \sum_{\substack{|(i,\nu)_s|=(k,n),\\ (|\nu|)_s\le(i)_s}} \sum_{\substack{(\max\{1,|\nu|\})_s\\ \le(q)_s\le(i)_s}} e^{M,s}_{(i-q,\nu)_s} \partial_\xi^{(i-q)_s} A_{(q,\nu)_s} \end{aligned} \tag{2.11}$$

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 (2.1) with $p = 3$ and $p = 4$ in which we are interested here. Note, that it holds $A = A_{1,1}$, $A_{k,-n} = \overline{A}_{k,n}$. 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 $A_{1,0} = 0$)

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

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 $\omega(\vartheta_0) \neq 0$ 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

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

with

$$\begin{aligned}
\rho &:= 2(\gamma_{2011}^2-\beta_2^2)/\delta_2 + 4\beta_2^2/\beta_1 - 3(\gamma_{3002}+\beta_3), \\
b &:= \gamma_{1111}a/\delta_2+3\gamma_{2101}-\gamma_{2102}, \\
c &:= 2(3\gamma_{2011}-\gamma_{2012}+\beta_2)a/\delta_2-3\gamma_{3002}+\gamma_{3003}+\beta_3.
\end{aligned} \tag{2.12}$$

The function $A_{3,1}$ remains undetermined. Note, that the equation for $k = 3$, $n = 1$ is the nonlinear Schrödinger equation (1.5) which determines the evolution of A. Thus, if we are interested only in the formal derivation of this equation we can insert in (1.1) the improved approximation (2.1) 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 (1.1) and calculate by the formalism (2.11) the functions $A_{4,n}$: By using the previous results, we obtain

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

with

$$
\begin{aligned}
d_1 &:= [\gamma_{1201}-\omega(\vartheta_0)\omega''(\vartheta_0)]\beta_2/\beta_1,\\
d_2 &:= 2\beta_2|a/\delta_2|^2 + 4\beta_2^3/\beta_1^2 + 6\beta_3\beta_2(1/\delta_2-2/\beta_1) + 6\beta_4,\\
e &:= 2[2(3\gamma_{2101}-\gamma_{2102})\gamma_{2011}+\gamma_{1111}(\gamma_{2011}^2-\beta_2^2)/\delta_2]/\delta_2 + 3\gamma_{3111},\\
f_1 &:= [3\gamma_{1201}-\gamma_{1202}-2\omega(\vartheta_0)\omega''(\vartheta_0)]a/\delta_2 + 2\gamma_{1111}b/\delta_2 + \gamma_{2211} - 2\gamma_{2210},\\
f_2 &:= -4\beta_2^2 a/\beta_1\delta_2 + 2(\beta_2-3\gamma_{2011}+\gamma_{2012})c/\delta_3 + 6(\beta_3-\gamma_{3021})a/\delta_2 - 6\beta_3\beta_2/\beta_1 + 4(\gamma_{4012}+\beta_4),\\
g &:= -3\gamma_{1130}c/\delta_3 + 2(3\gamma_{2103}-16\gamma_{2102}+18\gamma_{2101})a/\delta_2 + 4(3\gamma_{2011}-\gamma_{2012}+\beta_2)b/\delta_2\\
&\quad + 3(2\gamma_{3111}-\gamma_{3112}),\\
h &:= (\gamma_{2031}-2\gamma_{2030}+\beta_2)a^2/\delta_2^2 + 2(\gamma_{2031}-\gamma_{2012}+6\gamma_{2011}+\beta_2)c/\delta_3 + 3(2\gamma_{3021}-\gamma_{3022}+\beta_3)a/\delta_2\\
&\quad + \gamma_{4013}-2\gamma_{4012}+\beta_4.
\end{aligned}
\tag{2.13}
$$

The function $A_{4,1}$ remains undetermined. Since the equation for $k=4$, $n=1$ determines $A_{2,1}$, we know 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 2.1 *If the microscopic oscillator chain equation (1.1) has for all* $\varepsilon \in (0, \varepsilon_0)$ *solutions of the form*

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

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 (1.5). Analogously, if (1.1) has for all* $\varepsilon \in (0,\varepsilon_0)$ *solutions of the form*

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

where $\delta_n(\vartheta_0) := n^2\omega^2(\vartheta_0) - \omega^2(n\vartheta_0)$ *and* $a := 4\mathrm{i}\sum_{m=1}^M \alpha_{m,2}\sin(m\vartheta_0)[1-\cos(m\vartheta_0)] + \beta_2$, *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 (1.5) and equation (1.7), respectively.*

We call this result a formal derivation, since the existence of solutions satisfying such expansions is not clear at all. The purpose of the justification of the NLSE (and (1.7) in the second case) is to show that solutions which start in these forms will maintain them on suitably long time scales.

It should be noted that in order to determine the functions $A_{k,n}$ of the improved approximation $X_\varepsilon^{A,p}$ (cf. (2.1)) by the formalism (2.11), it has to hold $\delta_0(\vartheta_0) = -\omega^2(0) \neq 0$ and $\delta_n(\vartheta_0) = n^2\omega^2(\vartheta_0) - \omega^2(n\vartheta_0) \neq 0$. Under the stability condition (SC): $\omega^2(\vartheta) > 0$ for all $\vartheta \in (-\pi, \pi]$, this is satisfied if the nonresonance condition of second order

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

holds.

Proposition 2.2 *In the case of nearest-neighbour interactions ($M = 1$ with $\alpha_1 := \alpha_{1,1}$) it holds:*

$$(\mathrm{SC}) \Longleftrightarrow \min\{\beta_1, 4\alpha_1+\beta_1\} > 0, \tag{2.14}$$

$$\textit{For } n = 2,3,4\textit{:} \qquad \forall\, \vartheta \in (-\pi,\pi]:\ \delta_n(\vartheta) > 0 \Longleftrightarrow \min\{\beta_1, \frac{2^{n+2}}{n^2-1}\alpha_1 + \beta_1\} > 0, \tag{2.15}$$

$$(\mathrm{SC}) \text{ and } (\mathrm{NR3})_{\vartheta_0} \Longrightarrow (\mathrm{NR2})^4_{\vartheta_0}, \tag{2.16}$$

$$(\mathrm{SC}) \text{ and } (\mathrm{NR3})_{\mathrm{unif}} \Longleftrightarrow \min\{\beta_1, (16/3)\alpha_1+\beta_1\} > 0. \tag{2.17}$$

Remark 2.1 The stability condition (SC) restricts us by (2.14) to the harmonic coefficients $\beta_1 > 0$ and $\alpha_1 > -(1/4)\beta_1$. For $\alpha_1 > -(3/16)\beta_1$ we obtain by (2.15) and (2.17) that $(\mathrm{NR3})_{\mathrm{unif}}$ implies $(\mathrm{NR2})^3_{\vartheta_0}$. From (2.15) 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 $-(3/16)\beta_1 \geq \alpha_1 > -(15/64)\beta_1$, and $2\omega(\vartheta_0) \neq \omega(2\vartheta_0)$ and $4\omega(\vartheta_0) \neq \omega(4\vartheta_0)$ in the case $-(15/64)\beta_1 \geq \alpha_1 > -(1/4)\beta_1$. By (2.17), both conditions follow from $(\mathrm{NR3})_{\vartheta_0}$.

Proof : Equivalence (2.14) follows immediately from (DR) and (SC). For $n = 2,3,4$ it holds

$$\delta_n(\vartheta) = n^2\omega^2(\vartheta) - \omega^2(n\vartheta) = 2^n\alpha_1 g_n(\cos\vartheta) + (n^2-1)\beta_1 \tag{2.18}$$

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 (2.15).

By (SC), $(\mathrm{NR2})^4_{\vartheta_0}$ is equivalent to $\delta_n(\vartheta_0) > 0$ for $n = 2,3,4$, and thus, by (2.18), to

$$\frac{\alpha_1}{\beta_1} > \frac{-(n^2-1)}{2^n g_n(\cos\vartheta_0)} =: f_n(\cos\vartheta_0). \tag{2.19}$$

In Figure 1 we ploted $f_n(c)$ for $n = 2$ (black), $n = 3$ (dark grey) and $n = 4$ (light grey) over $c \in [-1, 0.45]$ (left) and $c \in [-1, 0]$ (right). Note, that as α_1/β_1 approaches $-1/4$ from above we have to take care that it remains above $f_n(\cos\vartheta_0)$ for our fixed ϑ_0. Note also, that f_3 approximates $-1/4$ from below, and by (SC) it holds $\alpha_1/\beta_1 > f_3(c)$ for all $c \in [-1, 1)$.

Ad (2.16): Setting $s = 2$, $t = 1$, $\theta = -\vartheta_0$ into $(\mathrm{NR3})_{\vartheta_0}$ we obtain $|2\omega(\vartheta_0) - \omega(2\vartheta_0)| \geq C^{\mathrm{NR}}_{\vartheta_0} > 0$. By (2.14) and (2.15) it holds $\delta_3(\vartheta) = 9\omega^2(\vartheta) - \omega^2(3\vartheta) > 0$ for all $\vartheta \in (-\pi,\pi]$. Finally, $4\omega(\vartheta_0) = \omega(4\vartheta_0)$ is equivalent to $\omega(\vartheta_0) + \omega(3\vartheta_0) - \omega(4\vartheta_0) = \omega(3\vartheta_0) - 3\omega(\vartheta_0)$. By $9\omega^2(\vartheta) - \omega^2(3\vartheta) > 0$ and $\omega(\vartheta) > 0$ for all $\vartheta \in (-\pi,\pi]$, this means that $\omega(\vartheta_0) + \omega(3\vartheta_0) - \omega(4\vartheta_0) < 0$. But $\omega(0) > 0$. Hence, by the continuity of ω, there exists a $\vartheta \in \mathbb{T}$ with $|\vartheta| \in (0, 3|\vartheta_0|)$, such that $\omega(\vartheta_0) + \omega(\vartheta) - \omega(\vartheta_0 - \vartheta) = 0$, which contradicts $(\mathrm{NR3})_{\vartheta_0}$.

Ad (2.17): By (SC) and $\omega(\vartheta) = \omega(-\vartheta) > 0$, $(\mathrm{NR3})_{\mathrm{unif}}$ can be reduced to

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

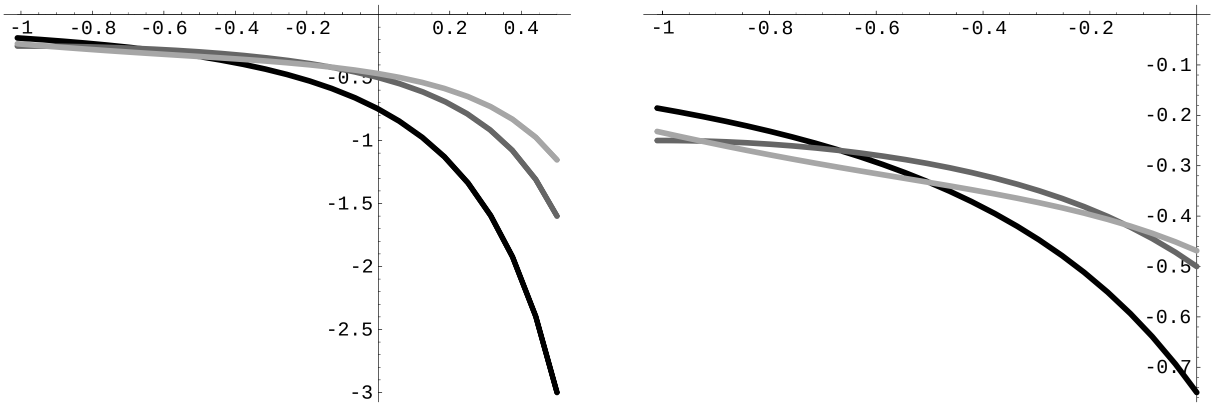

Figure 1: Plots over $c \in [-1, 0.45]$ (left) and $c \in [-1, 0]$ (right) of $f_n(c)$ defined by (2.19) for $n = 2$ (black), $n = 3$ (dark grey), $n = 4$ (light grey).

For $\alpha_1 = 0$ we take $C^{\rm NR}_{\rm unif} = \beta_1^{1/2} > 0$. 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, $(\text{NR3})_{\rm unif}$ 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 holds $\mu_- = \omega(0)$ and $\mu_+ = \omega(\pm\pi)$ but there exists no $\vartheta \in \mathbb{T}$ with $\vartheta = \vartheta \mp \pi = 0$. In this case, we show $\omega(\vartheta) + \omega(\vartheta-\theta) - \omega(\theta) \geq C^{\rm NR}_{\rm unif} > 0$ for all $\vartheta, \theta \in (-\pi, \pi]$, and (2.17) 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 $[(\beta+\gamma)(\beta_0+\gamma)]^{1/2} - \eta \geq \gamma$, 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 $\min_{\alpha\in[0,1]}\{(\alpha+\gamma)^{1/2}-\alpha^{1/2}\} = (1+\gamma)^{1/2}-1$, 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 $\mu_+-\mu_-\to\infty$ $(\alpha_1\to\infty)$. $\square$

3 The normal form transform

The oscillator chain model (1.1) can be rewritten as the first-order ordinary differential equation

$$\dot{\widetilde{x}} = \widetilde{L}\widetilde{x} + \widetilde{Q}(\widetilde{x},\widetilde{x}) + \widetilde{M}(\widetilde{x}) \tag{3.1}$$

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 (1.3), and

$$[Q(x,y)]_j := \sum_{m=1}^{M}\alpha_{m,2}[(x_{j+m}-x_j)(y_{j+m}-y_j)-(x_j-x_{j-m})(y_j-y_{j-m})] - \beta_2 x_j y_j, \tag{3.2}$$

$$[M(x)]_j := \sum_{m=1}^{M}[v_{m,2}(x_{j+m}-x_j) - v_{m,2}(x_j-x_{j-m})] - w_2(x_j), \tag{3.3}$$

with $v_{m,2}$ and w_2 defined in (2.8).

On the Banach space Y we use the energy norm

$$\|(x,y)\|_Y^2 := \|x\|_E^2 + \|y\|_{\ell^2}^2 \quad \text{with} \quad \|x\|_E^2 := \sum_{m=1}^{M}\alpha_{m,1}\sum_{j\in\mathbb{Z}}|x_{j+m}-x_j|^2 + \beta_1\|x\|_{\ell^2}^2 \tag{3.4}$$

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):

$$\mu_-^2\|x\|_{\ell^2}^2 \leq \|x\|_E^2 \leq \mu_+^2\|x\|_{\ell^2}^2 \tag{3.5}$$

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 (1.3) preserves this norm: The solutions $\widetilde{x}: t \mapsto \widetilde{x}(t) = \mathrm{e}^{t\widetilde{L}}\widetilde{x}(0)$ of (1.3) satisfy $\|\widetilde{x}(t)\|_Y = \|\widetilde{x}(0)\|_Y$ for all $t \in \mathbb{R}$ (cf. [GiM04, Proposition 3.1]).

We introduce the normal form transformation $F: Y \to Y$ with

$$\widetilde{y} = F(\widetilde{x}) := \widetilde{x} + B(\widetilde{x}, \widetilde{x}) \tag{3.6}$$

where the bilinear form $B: Y \times Y \to Y$ remains to be determined. Applying this transformation on (3.1) we obtain

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + \overline{Q}(\widetilde{x}, \widetilde{x}) + \overline{M}(\widetilde{x}) \tag{3.7}$$

with

$$\overline{Q}(\widetilde{x}, \widetilde{x}) := -\widetilde{L}B(\widetilde{x}, \widetilde{x}) + B(\widetilde{L}\widetilde{x}, \widetilde{x}) + B(\widetilde{x}, \widetilde{L}\widetilde{x}) + \widetilde{Q}(\widetilde{x}, \widetilde{x}), \tag{3.8}$$

$$\overline{M}(\widetilde{x}) := B(\widetilde{Q}(\widetilde{x}, \widetilde{x}) + \widetilde{M}(\widetilde{x}), \widetilde{x}) + B(\widetilde{x}, \widetilde{Q}(\widetilde{x}, \widetilde{x}) + \widetilde{M}(\widetilde{x})) + \widetilde{M}(\widetilde{x}). \tag{3.9}$$

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

$$B_2(\widetilde{x}, \widetilde{y}) := B_1(\widetilde{L}\widetilde{x}, \widetilde{y}) + B_1(\widetilde{x}, \widetilde{L}\widetilde{y}), \tag{3.10}$$

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 $B_1(\widetilde{x}, \widetilde{x}) = B_1(x, \dot{x}; x, \dot{x})$,

$$LB_1(x, \dot{x}; x, \dot{x}) - B_1(Lx, L\dot{x}; x, \dot{x}) - 2B_1(\dot{x}, Lx; \dot{x}, Lx) - B_1(x, \dot{x}; Lx, L\dot{x}) = Q(x, x). \tag{3.11}$$

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 (1.3) it holds $\widehat{L}\widehat{x} := \widehat{Lx}: \ \vartheta \mapsto -\omega^2(\vartheta)\widehat{x}(\vartheta)$. 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 (3.2) the Fourier transform

$$[\widehat{Q}(\widehat{x},\widehat{y})](\vartheta) := \widehat{Q(x,y)}(\vartheta) = \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\vartheta-\theta)q(\vartheta,\theta)\widehat{y}(\theta)\,\mathrm{d}\theta \tag{3.12}$$

with

$$q(\vartheta,\theta) := 2\mathrm{i}\sum_{m=1}^{M} \alpha_{m,2}\,[\sin(m\vartheta) - \sin(m(\vartheta-\theta)) - \sin(m\theta)] - \beta_2. \tag{3.13}$$

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 (3.11)

$$-\omega^2\widehat{B}_1(\widehat{x},\dot{\widehat{x}};\widehat{x},\dot{\widehat{x}}) + \widehat{B}_1(\omega^2\widehat{x},\omega^2\dot{\widehat{x}};\widehat{x},\dot{\widehat{x}}) - 2B_1(\dot{\widehat{x}},-\omega^2\widehat{x};\dot{\widehat{x}},-\omega^2\widehat{x}) + \widehat{B}_1(\widehat{x},\dot{\widehat{x}};\omega^2\widehat{x},\omega^2\dot{\widehat{x}}) = \widehat{Q}(\widehat{x},\widehat{x})$$

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

$$\left[\omega^2(\vartheta-\theta)+\omega^2(\theta)-\omega^2(\vartheta)\right]\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22}\end{pmatrix} - 2\begin{pmatrix}\omega^2(\vartheta-\theta)\omega^2(\theta)b_{22} & -\omega^2(\vartheta-\theta)b_{21} \\ -\omega^2(\theta)b_{12} & b_{11}\end{pmatrix} = \begin{pmatrix} q & 0 \\ 0 & 0\end{pmatrix}$$

holds for all $(\vartheta,\theta) \in (-\pi,\pi]^2$. This yields

$$b_{11} = \frac{\alpha q}{\alpha^2-\beta^2}, \qquad b_{22} = \frac{2q}{\alpha^2-\beta^2}, \qquad b_{12} = b_{21} = 0, \tag{3.14}$$

with $\alpha(\vartheta,\theta) := \omega^2(\vartheta-\theta) + \omega^2(\theta) - \omega^2(\vartheta)$ and $\beta(\vartheta,\theta) := 2\omega(\vartheta-\theta)\omega(\theta)$, 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

$$B_1(\widetilde{x},\widetilde{y}) = b_1(x_1,y_1) + b_2(x_2,y_2) \quad \text{for } \widetilde{x} = (x_1,x_2),\ \widetilde{y} = (y_1,y_2), \tag{3.15}$$

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

$$[\widehat{b}_i(\widehat{x}_i,\widehat{y}_i)](\vartheta) = \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}_i(\vartheta-\theta)b_{ii}(\vartheta,\theta)\widehat{y}_i(\theta)\,\mathrm{d}\theta \quad \text{for } \vartheta \in \mathbb{T} \tag{3.16}$$

with the b_{ii} determined by (3.14). From (3.10) we obtain $B_2 : Y\times Y \to \ell^2$:

$$B_2(\widetilde{x},\widetilde{y}) = b_1(x_2,y_1) + b_1(x_1,y_2) + b_2(Lx_1,y_2) + b_2(x_2,Ly_1) \quad \text{for } \widetilde{x} = (x_1,x_2),\ \widetilde{y} = (y_1,y_2). \tag{3.17}$$

This determines $B : Y\times Y \to Y$ with $B = (B_1,B_2)$.

Since

$$(\alpha^2-\beta^2)(\vartheta,\theta) = [\omega(\vartheta-\theta)-\omega(\theta)-\omega(\vartheta)]\,[\omega(\vartheta-\theta)-\omega(\theta)+\omega(\vartheta)] \\ \times [\omega(\vartheta-\theta)+\omega(\theta)-\omega(\vartheta)]\,[\omega(\vartheta-\theta)+\omega(\theta)+\omega(\vartheta)], \tag{3.18}$$

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:

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

(In the case of only nearest-neighbour interactions $M = 1$ $(\mathrm{NR3})_{\mathrm{unif}}$ is equivalent to $\min\{\beta_1, (16/3)\alpha_{1,1}+\beta_1\} > 0$, cf. (2.17) in Proposition 2.2.)

Since $B = (B_1, B_2)$ is given via (3.15), (3.17), 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 3.1 *For a bilinear form $b : \ell^2\times\ell^2 \to \ell^2$ with the Fourier transform*

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

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

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

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, it holds

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

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 (3.18), and the analyticity of ω and q given by (3.13), the $b_{ii}(\vartheta,\theta)$, $i=1,2$, defined by (3.14) 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 (3.15) and B_2 (3.17), and $\|\widetilde{y}\|_\infty := \max\{\|y_1\|_{\ell^\infty}, \|y_2\|_{\ell^\infty}\}$, we obtain

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

where we used $\|Lx\|_{\ell^2} \le C\|x\|_{\ell^2}$ and $\|Lx\|_{\ell^\infty} \le C\|x\|_{\ell^\infty}$. Thus, setting $c := \max\{c_1, c_2\}$, and using (3.5) and $\|x_1\|_{\ell^2} + \|x_2\|_{\ell^2} \le \mu\|\widetilde{x}\|_Y$ with $\mu := (\mu_- + 1)/\mu_-$, we obtain

$$\|B(\widetilde{x},\widetilde{y})\|_Y^2 \le \mu_+^2 \|B_1(\widetilde{x},\widetilde{y})\|_{\ell^2}^2 + \|B_2(\widetilde{x},\widetilde{y})\|_{\ell^2}^2 \le C_B^2 \|\widetilde{x}\|_Y^2 \|\widetilde{y}\|_\infty^2 \quad \text{for } \widetilde{x},\widetilde{y} \in Y \tag{3.19}$$

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

$$\|B(\widetilde{x},\widetilde{y})\|_Y \le \mu C_B \|\widetilde{x}\|_Y \|\widetilde{y}\|_Y \quad \text{for } \widetilde{x},\widetilde{y} \in Y. \tag{3.20}$$

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. (3.14)). By

$$[\widehat{b_i}(\widehat{x},\widehat{y})](\vartheta) = \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\vartheta-\theta) b_{ii}(\vartheta,\theta) \widehat{y}(\theta)\,\mathrm{d}\theta = \frac{1}{2\pi}\int_{\mathbb{T}} \widehat{x}(\tau) b_{ii}(\vartheta,\theta) \widehat{y}(\vartheta-\theta)\,\mathrm{d}\theta = [\widehat{b_i}(\widehat{y},\widehat{x})](\vartheta)$$

it follows $b_i(x,y) = b_i(y,x)$ for $x,y \in \ell^2$. By (3.15), (3.17), 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 $(\mathrm{NR3})_{\mathrm{unif}}$ 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 (3.15), (3.17), the system

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + \overline{M}(\widetilde{x}) \quad \text{with} \quad \overline{M}(\widetilde{x}) = 2B(\widetilde{Q}(\widetilde{x},\widetilde{x}) + \widetilde{M}(\widetilde{x}), \widetilde{x}) + \widetilde{M}(\widetilde{x}). \tag{3.21}$$

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 centre $F(0) = 0$, 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

$$\|\rho(\widetilde{y}_1) - \rho(\widetilde{y}_2)\|_Y \le C_\rho \|\widetilde{y}_1 - \widetilde{y}_2\|_Y \quad \text{for } \widetilde{y}_1, \widetilde{y}_2 \in Y \text{ with } \|\widetilde{y}_1\|_Y, \|\widetilde{y}_2\|_Y < \varepsilon_\rho. \tag{3.22}$$

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, (3.20) gives $F \in \mathrm{C}^1(Y,Y)$, since

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

Thus, for sufficiently small $\widetilde{x} \in Y$ the system (3.21) reads

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + N(\widetilde{y}) \quad \text{with} \quad N(\widetilde{y}) := \overline{M}(\rho(\widetilde{y})). \tag{3.23}$$

By the definition of $\overline{M}$ and $\rho(\widetilde{y}) = \widetilde{y} + \mathcal{O}(\widetilde{y}^2)$, the nonlinearity N of the transformed system (3.23) has only cubic and higher order terms, but no quadratic ones. This is the crucial motivation in applying the normal form transformation F on the system (3.1), since it enables us to apply the Gronwall type argument already used in [GiM04] for the justification of the NLSE associated to systems with cubic leading terms in their nonlinearity (cf. Section 4.2).

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 (2.2) for $k = 1, \ldots, p$ (and $n = 0, \ldots, k$). Hence, for the improved approximation $X_\varepsilon^{A,p}$ with the $A_{k,n}$ calculated in Section 2 the residual terms have by (2.3), (2.6) and (2.9) the form

$$\operatorname{res}(X_\varepsilon^{A,p}) := \ddot{X}_\varepsilon^{A,p} - LX_\varepsilon^{A,p} - N(X_\varepsilon^{A,p}) = \varepsilon^{p+1}(r_\varepsilon^{D,p} - r_\varepsilon^{L,p} - r_\varepsilon^{N,p}) \tag{4.1}$$

From (2.4) we obtain that there exists a $C_D > 0$, depending on $\omega : \vartheta \mapsto \omega(\vartheta)$, ϑ, and ε_0 such that for $\varepsilon \le \varepsilon_0$ it holds

$$|(r_\varepsilon^{D,p})_j(t)| \le C_D \sum_{k=p+1}^{p+4} \sum_{q=1}^{p} \sum_{n=-q}^{q} \sum_{\substack{\mu+2\nu=k-q,\\ \mu+\nu\le 2}} |\partial_\tau^\nu \partial_\xi^\mu A_{q,n}(\tau,\xi)| \tag{4.2}$$

with $\tau = \varepsilon^2 t$, $\xi = \varepsilon(j+\omega'(\vartheta)t)$. Analogously, from (2.7) and (2.10) we obtain that there exists a $C > 0$, depending on ω, ϑ, $V_m, W \in \mathrm{C}^{p+2}(\mathbb{R})$ $(m = 1, \ldots, M)$ and ε_0, such that for $\varepsilon \le \varepsilon_0$ and $\varepsilon^2 t \in [0, \tau_0]$ it holds

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

with $\theta_{pqn}^{\pm\varepsilon m} \in (0,1)$ (cf. (2.5)). 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 $r = 0, \ldots, p$, $l = 1, \ldots, p$, $n = 0, \ldots, l$ and all $\tau \in [0, \tau_0]$ (and C depends also on d). By Sobolev's imbedding 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 [GiM04]

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

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 (4.2) 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 (4.1) and $\operatorname{res}(\widetilde{X}_\varepsilon^{A,p}) := (0, \operatorname{res}(X_\varepsilon^{A,p}))$ we finally obtain

$$\|\mathrm{res}(\widetilde{X}_\varepsilon^{A,p})(t)\|_Y = \|\mathrm{res}(X_\varepsilon^{A,p})(t)\|_{\ell^2} \leq \tilde{C}_r \varepsilon^{p+1/2} \quad \text{for } \varepsilon \leq \varepsilon_0 \text{ and } \varepsilon^2 t \leq \tau_0. \tag{4.3}$$

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., [Paz83, Tem88]) 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 initialy $A(0,\cdot) \in \mathrm{H}^6(\mathbb{R})$ 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 $V_m, W, \omega, \theta, A, \tau_0$ (and $A_{2,1}$ if $p = 4$), such that

$$\|X_\varepsilon^{A,p}(t)\|_{\ell^\infty},\ \|\dot{X}_\varepsilon^{A,p}(t)\|_{\ell^\infty},\ \|X_\varepsilon^{A,p}(t)\|_{\ell^2},\ \|\dot{X}_\varepsilon^{A,p}(t)\|_{\ell^2} \leq \varepsilon^{1/2} C \text{ for } \varepsilon \leq \varepsilon_0 \leq 1 \text{ and } \varepsilon^2 t \leq \tau_0,$$

which leads by the definitions (3.4) and $\|(x,y)\|_\infty := \max\{\|x\|_{\ell^\infty}, \|y\|_{\ell^\infty}\}$ to

$$\|\widetilde{X}_\varepsilon^{A,p}(t)\|_\infty \leq \varepsilon C, \quad \|\widetilde{X}_\varepsilon^{A,p}(t)\|_Y \leq \varepsilon^{1/2} C_1 \quad \text{for } \varepsilon \leq \varepsilon_0 \leq 1 \text{ and } \varepsilon^2 t \leq \tau_0 \tag{4.4}$$

with $C_1 := C(\mu_+^2 + 1)^{1/2}$ (cf. (3.5)). Analogously, there exist $C_2, C_2', C_3 > 0$ such that

$$\|\widetilde{X}_\varepsilon^{A,p}(t) - \widetilde{X}_\varepsilon^{A}(t)\|_Y \leq C_2 \varepsilon^{3/2},\ \|\widetilde{X}_\varepsilon^{A,p}(t) - \widetilde{X}_\varepsilon^{A}(t)\|_\infty \leq C_2' \varepsilon^2,\ \|\widetilde{X}_\varepsilon^{A,p}(t) - \widetilde{X}_\varepsilon^{A,2}(t)\|_Y \leq C_3 \varepsilon^{5/2} \tag{4.5}$$

for $\varepsilon \leq \varepsilon_0 \leq 1$, $\varepsilon^2 t \leq \tau_0$.

4.2 Justification under uniform nonresonance $(\mathrm{NR3})_{\mathrm{unif}}$

We consider the transformed system (3.23) $\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + N(\widetilde{y})$ 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

$$\begin{aligned} \mathrm{res}(\widetilde{Y}_\varepsilon^{A,3}) &:= \dot{\widetilde{Y}}_\varepsilon^{A,3} - \widetilde{L}\widetilde{Y}_\varepsilon^{A,3} - N(\widetilde{Y}_\varepsilon^{A,3}) \\ &= \dot{\widetilde{X}}_\varepsilon^{A,3} + 2B(\dot{\widetilde{X}}_\varepsilon^{A,3}, \widetilde{X}_\varepsilon^{A,3}) - \widetilde{L}\widetilde{X}_\varepsilon^{A,3} - \widetilde{L}B(\widetilde{X}_\varepsilon^{A,3}, \widetilde{X}_\varepsilon^{A,3}) - \overline{M}(\widetilde{X}_\varepsilon^{A,3}) \\ &= \dot{\widetilde{X}}_\varepsilon^{A,3} + 2B(\mathrm{res}(\widetilde{X}_\varepsilon^{A,3}), \widetilde{X}_\varepsilon^{A,3}) - \widetilde{L}\widetilde{X}_\varepsilon^{A,3} - \widetilde{Q}(\widetilde{X}_\varepsilon^{A,3}, \widetilde{X}_\varepsilon^{A,3}) - \widetilde{M}(\widetilde{X}_\varepsilon^{A,3}) \\ &= \mathrm{res}(\widetilde{X}_\varepsilon^{A,3}) + 2B(\mathrm{res}(\widetilde{X}_\varepsilon^{A,3}), \widetilde{X}_\varepsilon^{A,3}), \end{aligned} \tag{4.6}$$

where we used (3.8) (with $\overline{Q} = 0$) and (3.9). From (4.3) and (4.4), it follows by (3.19)

$$\|\mathrm{res}(\widetilde{Y}_\varepsilon^{A,3})(t)\|_Y \leq C_r \varepsilon^{7/2} \quad \text{for } \varepsilon \leq \varepsilon_0,\ \varepsilon^2 t \leq \tau_0 \quad \text{with } C_r := (1 + 2C_B C \varepsilon_0)\widetilde{C}_r \tag{4.7}$$

Inserting the error $\varepsilon^{3/2}\widetilde{R} := \widetilde{y} - \widetilde{Y}_\varepsilon^{A,3}$ between a solution $\widetilde{y}$ of the transformed system (3.23) and the transformed approximation $\widetilde{Y}_\varepsilon^{A,3}$ into (3.23), we obtain by the definition of the residual term $\mathrm{res}(\widetilde{Y}_\varepsilon^{A,3})$ the differential equation for the error

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

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 te

$$\widetilde{R}(t) = \mathrm{e}^{t\widetilde{L}}\widetilde{R}(0) + \varepsilon^{-3/2}\int_0^t \mathrm{e}^{(t-s)\widetilde{L}}\left[N(\widetilde{Y}_\varepsilon^{A,3}(s)+\varepsilon^{3/2}\widetilde{R}(s)) - N(\widetilde{Y}_\varepsilon^{A,3}(s)) - \mathrm{res}(\widetilde{Y}_\varepsilon^{A,3}(s))\right] \mathrm{d}s. \tag{4.8}$$

Assuming $\|\widetilde{x}(0)-\widetilde{X}_\varepsilon^{A}(0)\|_Y \le d\varepsilon^{3/2}$, we obtain by (4.5): $\|\widetilde{x}(0)-\widetilde{X}_\varepsilon^{A,3}(0)\|_Y \le (d+C_2)\varepsilon^{3/2}$, and thus by (4.4): $\|\widetilde{x}(0)+\widetilde{X}_\varepsilon^{A,3}(0)\|_Y \le [(d+C_2)\varepsilon_0+2C_1]\varepsilon^{1/2}$ 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 (3.20),

$$\|\widetilde{R}(0)\|_Y = \varepsilon^{-3/2}\|\widetilde{y}(0)-\widetilde{Y}_\varepsilon^{A,3}(0)\|_Y \le \widetilde{d} \quad \text{for } \varepsilon \le \varepsilon_0 < 1 \tag{4.9}$$

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 it holds

$$\|N(\widetilde{Y}_\varepsilon^{A,3}(t)+\varepsilon^{3/2}\widetilde{R}(t)) - N(\widetilde{Y}_\varepsilon^{A,3}(t))\|_Y \le C_N\varepsilon^2\|\varepsilon^{3/2}\widetilde{R}(t)\|_Y \tag{4.10}$$
$$\text{for } \varepsilon \le \varepsilon_0,\ \varepsilon^2 t \le \tau_0,\ \|\widetilde{R}(t)\|_Y \le D.$$

Then, by $\|\mathrm{e}^{t\widetilde{L}}\|_{Y\to Y} = 1$, (4.7), (4.9) and (4.10), equation (4.8) 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 $\|\widetilde{R}(s)\|_Y \le D$ 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 (4.10), we have obtained

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

For $\|\widetilde{y}(t)-\widetilde{Y}_\varepsilon^{A,3}(t)\|_Y$, $\|\widetilde{Y}_\varepsilon^{A,3}(t)\|_Y < \varepsilon_\rho/2$ it holds $\|\widetilde{y}(t)\|_Y < \varepsilon_\rho$. 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 (3.22)

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

Hence, we finally obtain by (4.5)

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

Thus, except for the proof of (4.10) 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 (1.5) as a macroscopic limit for the oscillator chain model (1.1).

Theorem 4.1 *Assume that V_m, $W \in \mathrm{C}^5(\mathbb{R})$ in (1.1) have the form (1.2) 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 (1.5) with $A(0, \cdot) \in \mathrm{H}^6(\mathbb{R})$ and let X^A_ε be the formal approximation given in (1.4) with $c_{\mathrm{gr}} = -\omega'$. 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 (3.1) with an initial condition $\widetilde{x}(0)$ satisfying

$$\|\widetilde{x}(0)-\widetilde{X}^A_\varepsilon(0)\|_Y \le d\varepsilon^{3/2},$$

fulfills the estimate

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

(In the case of nearest-neighbour interactions $(\mathrm{NR2})^3_{\vartheta_0}$ is implied by (SC) and $(\mathrm{NR3})_{\mathrm{unif}}$, cf. Remark 2.1.)

Proof of (4.10): 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 (3.23) and (3.9)

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

From (3.19) $\|B(\widetilde{x}, \widetilde{y})\|_Y \le C_B\|\widetilde{x}\|_\infty\|\widetilde{y}\|_Y$ it follows

$$\begin{aligned} \|N(y+\Delta)-N(y)\|_Y &\le 2C_B\|\widetilde{x}_1\|_\infty \left(\|\widetilde{Q}(\widetilde{x}_1, \widetilde{x}_1)-\widetilde{Q}(\widetilde{x}_2, \widetilde{x}_2)\|_Y+\|\widetilde{M}(\widetilde{x}_1)-\widetilde{M}(\widetilde{x}_2)\|_Y\right) \\ &\quad + 2C_B \left(\|\widetilde{Q}(\widetilde{x}_2, \widetilde{x}_2)\|_\infty+\|\widetilde{M}(\widetilde{x}_2)\|_\infty\right) \|\widetilde{x}_1-\widetilde{x}_2\|_Y + \|\widetilde{M}(\widetilde{x}_1) - \widetilde{M}(\widetilde{x}_2)\|_Y. \end{aligned} \tag{4.11}$$

By the definition of M (3.3), we have

$$\begin{aligned} \|M(x_1)-M(x_2)\|^2_{\ell^2} &= \sum_{j\in\mathbb{Z}} \left([M(x_1)]_j-[M(x_2)]_j\right)^2 \\ &= \sum_{j\in\mathbb{Z}} \Bigg(\sum_{m=1}^{M}[v_{m,2}(x^{(1)}_{j+m}-x^{(1)}_j)-v_{m,2}(x^{(1)}_j-x^{(1)}_{j-m})-v_{m,2}(x^{(2)}_{j+m}-x^{(2)}_j)+v_{m,2}(x^{(2)}_j-x^{(2)}_{j-m})] \\ &\qquad -w_2(x^{(1)}_j)+w_2(x^{(2)}_j)\Bigg)^2 \end{aligned}$$

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. (2.8)), 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 $\|x_1\|_{\ell^\infty}, \|x_2\|_{\ell^\infty} \le \delta/\mu_-$ it holds

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

Hence, we obtain

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

and thus, by $\widetilde{M}(\widetilde{x}) = (0, M(x))$, the definitions of $\|\cdot\|_Y$ and $\|\cdot\|_\infty$, and (3.5)

$$\|\widetilde{M}(\widetilde{x}_1)-\widetilde{M}(\widetilde{x}_2)\|_Y \le C_M(\|\widetilde{x}_1\|^2_\infty+\|\widetilde{x}_2\|^2_\infty)\|\widetilde{x}_1-\widetilde{x}_2\|_Y \quad \text{for } \|\widetilde{x}_1\|_Y, \|\widetilde{x}_2\|_Y \le \delta \tag{4.12}$$

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

$$\|\widetilde{M}(\widetilde{x})\|_\infty \le \|\widetilde{M}(\widetilde{x})\|_Y \le C_M\|\widetilde{x}\|^2_\infty\|\widetilde{x}\|_Y \quad \text{for } \|\widetilde{x}\|_Y \le \delta. \tag{4.13}$$

Analogously, since

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

we obtain, by $\widetilde{Q}(\widetilde{x},\widetilde{x}) = (0, Q(x,x))$, the definition of Q (3.2), and (3.5)

$$\|\widetilde{Q}(\widetilde{x}_1,\widetilde{x}_1)-\widetilde{Q}(\widetilde{x}_2,\widetilde{x}_2)\|_Y \le C_Q\|\widetilde{x}_1+\widetilde{x}_2\|_\infty\|\widetilde{x}_1-\widetilde{x}_2\|_Y \tag{4.14}$$

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

$$\|\widetilde{Q}(\widetilde{x},\widetilde{x})\|_\infty = \|Q(x,x)\|_{\ell^\infty} \le (6\sum_{m=1}^{M}|\alpha_{m,2}|+|\beta_2|)\|x\|^2_{\ell^\infty} \le (6\sum_{m=1}^{M}|\alpha_{m,2}|+|\beta_2|)\|\widetilde{x}\|^2_\infty. \tag{4.15}$$

Inserting (4.12)–(4.15) in (4.11), we obtain

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

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 (3.22)

$$\|N(y+\Delta)-N(y)\|_Y \le C_nC_\rho(\|\widetilde{x}_1\|^2_\infty+\|\widetilde{x}_2\|^2_\infty)\|\Delta\|_Y \quad \text{for } \|y\|_Y+\|\Delta\|_Y < \varepsilon_\rho \tag{4.16}$$

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

For $y := \widetilde{Y}^{A,3}_\varepsilon(t) = \widetilde{X}^{A,3}_\varepsilon(t) + B(\widetilde{X}^{A,3}_\varepsilon(t), \widetilde{X}^{A,3}_\varepsilon(t))$, we obtain from (4.4) by (3.19)

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

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 (4.4), (3.22) 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 (4.16), we obtain (4.10) with $C_N := 5C_n C_\rho C^2$, which is independent of D. $\square$

4.3 Justification under the nonresonance condition $(\mathrm{NR3})_{\vartheta_0}$

In the previous section we proved Theorem 4.1 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 [Sch98]. By the representation (3.18) of $\gamma := \alpha^2 - \beta^2$, $(\mathrm{NR3})_{\vartheta_0}$ implies that there exist $c, \delta > 0$ such that

$$|\gamma(\vartheta,\theta)| > c \quad \text{for all } (\vartheta,\theta) \in \mathbb{T}\times\mathbb{T} \text{ with } |\vartheta - \vartheta_0| < 2\delta \tag{4.17}$$

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

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

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

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

Thus, it holds

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

where

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

Now, we define $B' = (B_1', B_2') : Y\times Y \to Y$ by (3.14)–(3.17) 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 (3.12), 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.,

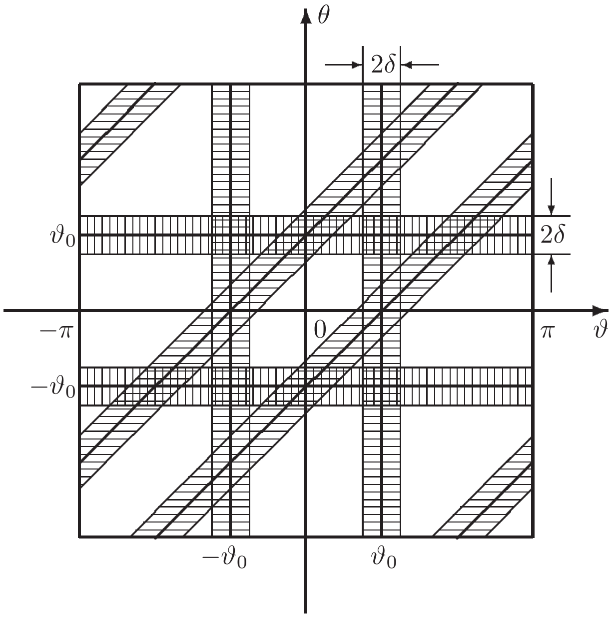

Figure 2: Sketch of $\Gamma(\vartheta_0,\delta)$ for $\vartheta_0 \in (-\pi,\pi]$, $\delta>0$

$q(\vartheta,\theta)=q(\vartheta,\vartheta-\theta)$, $b_{ii}(\vartheta,\theta)=b_{ii}(\vartheta,\vartheta-\theta)$), 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 $(\mathrm{NR3})_{\mathrm{unif}}$, it holds

$$2B'(\widetilde{L}\widetilde{x},\widetilde{x})-\widetilde{L}B'(\widetilde{x},\widetilde{x})+\widetilde{Q}'(\widetilde{x},\widetilde{x})=0 \tag{4.18}$$

(cf. (4.18)). Thus, applying to (3.1) the normal form transformation $F':Y\to Y$ with

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

we obtain

$$\dot{\widetilde{y}}=\widetilde{L}\widetilde{y}+\widetilde{S}(\widetilde{x},\widetilde{x})+\overline{M}'(\widetilde{x}) \tag{4.19}$$

with

$$\begin{aligned}\widetilde{S}(\widetilde{x},\widetilde{y})&:=\widetilde{Q}(\widetilde{x},\widetilde{y})-\widetilde{Q}'(\widetilde{x},\widetilde{y}),\\ \overline{M}'(\widetilde{x})&:=2B'(\widetilde{Q}(\widetilde{x},\widetilde{x})+\widetilde{M}(\widetilde{x}),\widetilde{x})+\widetilde{M}(\widetilde{x}).\end{aligned} \tag{4.20}$$

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 3.1

$$\|b'_i(x,y)\|_{\ell^2}\le c'_i\|x\|_{\ell^2}\|y\|_{\ell^\infty}\quad(i=1,2),\qquad \|S(x,y)\|_{\ell^2}\le c_s\|x\|_{\ell^2}\|y\|_{\ell^\infty}\quad\text{for } x,y\in\ell^2. \tag{4.21}$$

and thus

$$\|B'(\widetilde{x},\widetilde{y})\|_Y \leq C'_B\|\widetilde{x}\|_Y\|\widetilde{y}\|_\infty, \quad \|\widetilde{S}(\widetilde{x},\widetilde{y})\|_Y \leq C_S\|\widetilde{x}\|_Y\|\widetilde{y}\|_\infty \qquad \text{for } \widetilde{x},\widetilde{y} \in Y \tag{4.22}$$

with $C_S = c_s/\mu_-$ and C'_B defined by C_B with $c,\ c_i$ replaced by $c',\ c'_i$ (cf. (3.19)), which yields also

$$\|B'(\widetilde{x},\widetilde{y})\|_Y \leq \mu C'_B\|\widetilde{x}\|_Y\|\widetilde{y}\|_Y, \quad \|\widetilde{S}(\widetilde{x},\widetilde{y})\|_Y \leq \mu C_S\|\widetilde{x}\|_Y\|\widetilde{y}\|_Y \qquad \text{for } \widetilde{x},\widetilde{y} \in Y \tag{4.23}$$

(cf. (3.20)). 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

$$\|\rho'(\widetilde{y}_1) - \rho'(\widetilde{y}_2)\|_Y \leq C'_\rho\|\widetilde{y}_1 - \widetilde{y}_2\|_Y \quad \text{for } \widetilde{y}_1,\widetilde{y}_2 \in Y \text{ with } \|\widetilde{y}_1\|_Y, \|\widetilde{y}_2\|_Y < \varepsilon'_\rho. \tag{4.24}$$

(cf. (3.22)). Hence, for sufficiently small $\widetilde{x} \in Y$ the system (4.19) reads in analogy to (3.23)

$$\dot{\widetilde{y}} = \widetilde{L}\widetilde{y} + S'(\widetilde{y},\widetilde{y}) + N'(\widetilde{y}) \quad \text{with } S'(\widetilde{y},\widetilde{y}) := \widetilde{S}(\rho'(\widetilde{y}),\rho'(\widetilde{y})),\ N'(\widetilde{y}) := \overline{M}'(\rho'(\widetilde{y})). \tag{4.25}$$

However, unlike (3.23), the system (4.25) posseses also quadratic nonlinear terms $S'(\widetilde{y},\widetilde{y}) \neq 0$. Thus, it is clear that in order to prove a result similar to Theorem 4.1 we have to alternate some parameters in the method used in Section 4.2, which was taylored to the case of only cubic and higher order nonlinear terms. To this end, we start with an abstract form of our method, where some parameters are unspecified, and deduce the needed values of these parameters.

In Section 2 we carried out the formal derivation of the NLSE for the multiple scale ansatz $X_\varepsilon^{A,p}$ given in (2.1) 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 (2.1) in our original microscopic model (1.1) 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

$$\|\mathrm{res}(\widetilde{X}_\varepsilon^{A,p})(t)\|_Y \leq \tilde{C}_r\varepsilon^{p+1/2} \quad \text{for } \varepsilon \leq \varepsilon_0 \text{ and } \varepsilon^2 t \leq \tau_0, \tag{4.3}$$

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 (1.5) and $A_{2,1}$ of (1.7), which guarantees that there exist constants $C_A, \tau_0 > 0$ such that

$$\max_{k+2l\leq 7}\|\partial_\tau^l\partial_\xi^k A(\tau,\cdot)\|_{\mathrm{L}^2(\mathbb{R})} + \max_{k+2l\leq 6}\|\partial_\tau^l\partial_\xi^k A_{2,1}(\tau,\cdot)\|_{\mathrm{L}^2(\mathbb{R})} \leq C_A \quad \text{for } \tau \leq \tau_0 \tag{4.26}$$

(cf. Section 4.1).

Then, considering the transformed system (4.25) 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

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

where we used (4.18) and (4.20) with $\widetilde{x} = \widetilde{X}_\varepsilon^{A,p}$. From (4.3) and (4.4), it follows by $(4.22)_1$

$$\|\operatorname{res}(\widetilde{Y}_\varepsilon^{A,p})(t)\|_Y \le \widetilde{C}_r'\varepsilon^{p+1/2} \quad \text{for } \varepsilon \le \varepsilon_0,\ \varepsilon^2 t \le \tau_0 \quad \text{with } \widetilde{C}_r' := (1+2C_B'C\varepsilon_0)\widetilde{C}_r. \tag{4.27}$$

Inserting into (4.25) the error $\varepsilon^\alpha \widetilde{R}' := \widetilde{y} - \widetilde{Y}_\varepsilon^{A,p}$ between a solution $\widetilde{y}$ of (4.25) and the transformed approximation $\widetilde{Y}_\varepsilon^{A,p}$, we obtain the differential equation for the error

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

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

$$\begin{aligned}
\|\widetilde{R}'(t)\|_Y \le \|\widetilde{R}'(0)\|_Y + & \\
+ \varepsilon^{-\alpha}\int_0^t \Big(& \|S'(\widetilde{Y}_\varepsilon^{A,p}(s)+\varepsilon^\alpha\widetilde{R}'(s), \widetilde{Y}_\varepsilon^{A,p}(s)+\varepsilon^\alpha\widetilde{R}'(s)) - S'(\widetilde{Y}_\varepsilon^{A,p}(s), \widetilde{Y}_\varepsilon^{A,p}(s))\|_Y + \\
& + \|N'(\widetilde{Y}_\varepsilon^{A,p}(s)+\varepsilon^\alpha\widetilde{R}'(s)) - N'(\widetilde{Y}_\varepsilon^{A,p}(s))\|_Y + \|\operatorname{res}(\widetilde{Y}_\varepsilon^{A,p}(s))\|_Y \Big)\, \mathrm{d}s.
\end{aligned} \tag{4.28}$$

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

$$\|\widetilde{R}'(0)\|_Y \le d', \tag{4.29}$$

$$\|\operatorname{res}(\widetilde{Y}_\varepsilon^{A,p})(t)\|_Y \le C_r'\varepsilon^{\alpha+2}, \tag{4.30}$$

$$\|N'(\widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha\widetilde{R}'(t)) - N'(\widetilde{Y}_\varepsilon^{A,p}(t))\|_Y \le C_N'\varepsilon^{\alpha+2}\|\widetilde{R}'(t)\|_Y, \tag{4.31}$$

$$\|S'(\widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha\widetilde{R}'(t), \widetilde{Y}_\varepsilon^{A,p}(t)+\varepsilon^\alpha\widetilde{R}(t)') - S'(\widetilde{Y}_\varepsilon^{A,p}(t), \widetilde{Y}_\varepsilon^{A,p}(t))\|_Y \le C_S'\varepsilon^{\alpha+2}\|\widetilde{R}'(t)\|_Y \tag{4.32}$$

hold for $\varepsilon \le \varepsilon_0$, $\varepsilon^2 t \le \tau_0$, $\|\widetilde{R}'(t)\|_Y \le D'$. Inserting these estimates in (4.28), 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 $\|\widetilde{y}(t) - \widetilde{Y}_\varepsilon^{A,p}(t)\|_Y$, $\|\widetilde{Y}_\varepsilon^{A,p}(t)\|_Y < \varepsilon'_\rho/2$ inequality (4.24) yields

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

and for $\alpha \in (2, 5/2]$ we obtain by $(4.5)_3$

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

Let us now verify the estimates (4.29)–(4.32), and thereby deduce the required values of α and p: Since N' consists of the cubic and higher order nonlinear terms, the proof of (4.31) follows along the same lines as that of (4.10) in Section 4.2: By (4.20), $(4.22)_1$, (4.24), (4.25), 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

$$\|\widetilde{Y}_\varepsilon^{A,p}(t)\|_Y + \|\varepsilon^\alpha \widetilde{R}'(t)\|_Y \leq \varepsilon'_\rho \tag{4.33}$$

the estimate

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

with $C'_n > 0$ depending only on C'_B, V_m, W and ε'_ρ (cf. (4.16)). Moreover, for $\alpha > 1$ we can show, that for given $D' > 0$ there exists an $\varepsilon_0 > 0$ such that (4.33) 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 (4.31).

In order to prove (4.32), 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 (4.24) it holds under condition (4.33) $\|x'_1 - \widetilde{x}'_2\|_Y \leq \varepsilon^\alpha C'_\rho \|\widetilde{R}'(t)\|_Y$. This yields by $(4.23)_2$

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

and by $(4.5)_2$ and $(4.22)_2$

$$\|\widetilde{S}(\widetilde{X}_\varepsilon^{A,4}(t) - \widetilde{X}_\varepsilon^{A}(t), x'_1 - \widetilde{x}'_2)\|_Y \leq \varepsilon^{2+\alpha} C_S C'_2 C'_\rho \|\widetilde{R}'(t)\|_Y$$

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

$$\|\widetilde{S}(\widetilde{X}_\varepsilon^{A}(t), \widetilde{z})\|_Y \leq C_S C_P \varepsilon^2 \|\widetilde{z}\|_Y \quad \text{for } \widetilde{z} \in Y \text{ and } \varepsilon \leq \varepsilon_0,\ \varepsilon^2 t \leq \tau_0. \tag{4.34}$$

Then, for given $D' > 0$ there exists a $\varepsilon_0 > 0$ satisfying (4.33) and

$$\begin{aligned}
&\|S'(\widetilde{Y}_\varepsilon^{A,p}(t) + \varepsilon^\alpha \widetilde{R}'(t), \widetilde{Y}_\varepsilon^{A,p}(t) + \varepsilon^\alpha \widetilde{R}'(t)) - S'(\widetilde{Y}_\varepsilon^{A,p}(t), \widetilde{Y}_\varepsilon^{A,p}(t))\|_Y \\
&\leq \varepsilon^{\alpha+2} C_S C'_\rho (\varepsilon_0^{\alpha-2} \mu C'_\rho D' + 2C'_2 + 2C_P) \|\widetilde{R}'(t)\|_Y \quad \text{for } \varepsilon \leq \varepsilon_0,\ \varepsilon^2 t \leq \tau_0,\ \|\widetilde{R}'(t)\|_Y \leq D'.
\end{aligned}$$

Thus, in order to obtain (4.32) 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 (4.30) for the residual $\mathrm{res}(\widetilde{Y}_\varepsilon^{A,p})$ follows from (4.27) 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 (4.29) 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 $(4.23)_1$, $(4.4)_2$, 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 $(4.5)_3$ 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 (4.34), we have proven the following result based on the nonresonance condition $(\mathrm{NR3})_{\vartheta_0}$

Theorem 4.2 *Assume that V_m, $W \in \mathrm{C}^6(\mathbb{R})$ in (1.1) have the form (1.2) 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 (1.5) with $A(0, \cdot) \in \mathrm{H}^7(\mathbb{R})$ and of (1.7) with $A_{2,1}(0, \cdot) \in \mathrm{H}^6(\mathbb{R})$, respectively, and let $X_\varepsilon^{A,2}$ be the formal approximation (1.6). 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 (3.1) with an initial condition $\widetilde{x}(0)$ satisfying

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

fulfills the estimate

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

(In the case of nearest-neighbour interactions $(\mathrm{NR2})_{\vartheta_0}^4$ is implied by (SC) and $(\mathrm{NR3})_{\vartheta_0}$, cf. (2.16) in Proposition 2.2.)

Proof of (4.34): 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 $(4.21)_2$. 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 (4.17). 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 (4.21)

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

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

$$\|\widehat{\mathcal{P}}\widehat{X}_\varepsilon^A(t)\|_{\mathrm{L}^2(\mathbb{T})} \leq C_p\varepsilon^2 \quad \text{for } \varepsilon \leq \varepsilon_0,\ \varepsilon^2 t \leq \tau_0. \tag{4.35}$$

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

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

Hence, by (4.26) we obtain (4.35), and thus (4.34), with $C_p := \varepsilon_0^{1/2} 2C_A/(1{-}\cos\delta)$. □

References

[BCS01] 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(1), 131–148, 149–171, 2001.

[Bol96] C. Boldrighini. Macroscopic limits of microscopic systems. *Rend. Mat. Appl. (7)*, 16(1), 1–107, 1996.

[DKKZ96] P. Deift, S. Kamvissis, T. Kriecherbauer, and X. Zhou. The Toda rarefaction problem. *Comm. Pure Appl. Math.*, 49(1), 35–83, 1996.

[DHM05] W. Dreyer, M. Herrmann, and A. Mielke. Micro–macro transitions for the atomic chain via Whitham's modulation equation *Nonlinearity*, submitted, May 2005

[DrK00] W. Dreyer and M. Kunik. Cold, thermal and oscillator closure of the atomic chain. *J. Phys. A*, 33(10), 2097–2129, 2000. (Corrigendum: J. Phys. A 33 (2000) 2458).

[DKV95] P. Deift, T. Kriecherbauer, and S. Venakides. Forced lattice vibrations. I, II. *Comm. Pure Appl. Math.*, 48, 1187–1249, 1251–1298, 1995.

[FPU55] 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.

[FiV99] A.-M. Filip and S. Venakides. Existence and modulation of traveling waves in particle chains. *Comm. Pure Appl. Math.*, 52(6), 693–735, 1999.

[FrM02] G. Friesecke and K. Matthies. Atomic-scale localization of high-energy solitary waves on lattices. *Physica D*, 171(4), 211–220, 2002.

[FrP99] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices: I. Qualitative properties, renormalization and continuum limit. *Nonlinearity*, 12, 1601–1627, 1999.

[FrP02] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices: II. Linear implies nonlinear stability. *Nonlinearity*, 15, 1343–1359, 2002.

[FrP04a] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices: III. Howland-type Floquet theory. *Nonlinearity*, 17, 207–227, 2004.

[FrP04b] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices: IV. Proof of stability at low energy. *Nonlinearity*, 17, 229–251, 2004.

[FrW94] G. Friesecke and J. A. D. Wattis. Existence theorem for solitary waves on lattices. *Comm. Math. Phys.*, 161(2), 391–418, 1994.

[GiM04] J. Giannoulis and A. Mielke. The nonlinear Schrödinger equation as a macroscopic limit for an oscillator chain with cubic nonlinearities. *Nonlinearity*, 17, 551–565, 2004.

[HLM94] M. H. Hays, C. D. Levermore, and P. D. Miller. Macroscopic lattice dynamics. *Physica D*, 79(1), 1–15, 1994.

[Her04] M. Herrmann. *Ein Mikro-Makro-Übergang für die nichtlineare atomare Kette mit Temperatur.* Dissertation, Humboldt-Universität zu Berlin, 2004

[IoJ05] G. Iooss and G. James. Localized waves in nonlinear oscillator chains. *Chaos*, 15(1), 015113, 15 pp., 2005.

[IoK00] G. Iooss and K. Kirchgässner. Travelling waves in a chain of coupled nonlinear oscillators. *Comm. Math. Phys.*, 211(2), 439–464, 2000.

[Ioo00] G. Iooss. Travelling waves in the Fermi-Pasta-Ulam lattice. *Nonlinearity*, 13(3), 849–866, 2000.

[Jam03] G. James. Centre manifold reduction for quasilinear discrete systems. *J. Nonlinear Sci.*, 13(1), 27–63, 2003

[Kal89] L. A. Kalyakin. Long wave asymptotics. Integrable equations as asymptotic limits of non-linear systems. *Russian Math. Surveys*, 44(1), 3–42, 1989.

[KSM92] 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, 85–91, 1992.

[Kon96] V. V. Konotop. Small-amplitude envelope solitons in nonlinear lattices. *Phys. Rev. E*, 53, 2843–2858, 1996.

[MaA94] R. S. MacKay and S. Aubry. Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators. *Nonlinearity*, 7(6), 1623–1643, 1994.

[Mie05] A. Mielke. Macroscopic behavior of microscopic oscillations in harmonic lattices via Wigner-Husimi transforms. *Arch. Ration. Mech. Anal.*, submitted, May 2005.

[Mie02] 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.

[Paz83] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[Sch98] G. Schneider. Justification of modulation equations for hyperbolic systems via normal forms. *NoDEA Nonlinear Differential Equations Appl.*, 5(1), 69–82, 1998.

[ScW00] 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.

[Spo91] H. Spohn. *Large scale dynamics of interacting particles*. Springer-Verlag, New York, 1991.

[Tem88] R. Temam. *Infinite-dimensional dynamical systems in mechanics and physics*. Springer-Verlag, New York, 1988.