The nonlinear Schrödinger equation as a macroscopic limit for an oscillator chain with cubic nonlinearities

Johannes Giannoulis and Alexander Mielke

Institut für Analysis, Dynamik und Modellierung,

Universität Stuttgart, Pfaffenwaldring 57, D–70569 Stuttgart

{giannoulis,mielke}@mathematik.uni-stuttgart.de

4 August 2003; revised: 20 November 2003

Abstract

We consider the nonlinear model of an infinite oscillator chain embedded in a background field. We start from an appropriate modulation ansatz of the space-time periodic solutions to the linearized (microscopic) model and derive formally the associated (macroscopic) modulation equation, which turns out to be the nonlinear Schrödinger equation. Then we justify this necessary condition rigorously for the case of nonlinearities with cubic leading terms, that is, we show that solutions which have the form of the assumed ansatz for $t = 0$ preserve this form over time-intervals with a positive macroscopic length. Finally, we transfer this result to the analogous case of a finite, but large periodic chain and illustrate it by a numerical example.

Key words and phrases: nonlinear infinite oscillator chain, macroscopic limit, nonlinear Schrödinger equation, justification of modulation equations

Mathematics Subject Classification 2000: 37K60; 70F45, 70K70, 34E13, 35Q55.

1 Introduction

One of the most challenging problems in multiscale analysis is that of finding continuum models for discrete, atomistic models. In statistical physics these questions were already addressed one hundred years ago, but many problems remain open until today. Most prominently is the question of how to obtain irreversible thermodynamics as a macroscopic limit from microscopic models which are reversible (Hamiltonian). For a survey on the methods and results of the mathematical justification of nonequilibrium statistical mechanics we refer e.g. to [Spo91, Bol96].

In this paper we consider another part of this field which is far from thermodynamic fluctuations. We are interested in reversible, macroscopic limits of atomistic models which are obtained by choosing well-prepared initial conditions. One chooses the initial data in a specified class of functions and hopes to obtain an evolution within this function class. The associated evolution equation will be called the macroscopic limit problem.

This point of view is quite different from the traditional mathematical approach to large discrete systems, where specific solution classes are investigated like traveling fronts, pulses, wave trains or breathers [FW94, FP99, Ioo00, IK00, Jam03, FP02]. Instead, our approach is very close to the theory of modulation equations which evolved in the late 1960's for problems in fluid mechanics (see [Mie02] for a recent 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 a much larger space and time scale, such that the modulation equation is a macroscopic equation.

To be more specific we consider a one-dimensional discrete system of the form

$$\ddot{x}_j = V'(x_{j+1}-x_j) - V'(x_j-x_{j-1}) - W'(x_j), \quad j \in \mathcal{J}, \tag{1.1}$$

where the index set $\mathcal{J}$ is either $\mathbb{Z}$ or the finite cyclic group $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$, $m \in \mathbb{N}$. These are the equations of motion for the deviations x_j from the rest position j of (a chain of) atoms with equal mass 1. V is the potential for the nearest-neighbor interaction and W is an external potential which might arise through embedding of the atomic chain in a background field. Special solution classes for (1.1) with $\mathcal{J} = \mathbb{Z}$ are investigated in [MA94, IK00, Jam03, FM02]. Closest to our work is the justification of the KdV limit in [Kal89] (Prop. 7.1) and [SW00], where $W \equiv 0$ and solutions of the form

$$x_j(t) = \varepsilon^2 U(\varepsilon^3 t, \varepsilon(x-ct)) + \mathcal{O}(\varepsilon^4) \tag{1.2}$$

are studied, where U satisfies the macroscopic limit equation

$$\partial_\tau U + \kappa_1 U \partial_\xi U + \kappa_2 \partial_\xi^3 U = 0, \tag{1.3}$$

which is the Korteweg-de Vries equation. These solutions appear to be constant on a microscopic level, i.e., on bounded sets for t and j when $\varepsilon \ll 1$. Thus, this case is called the long wave-length limit.

We investigate solutions which are microscopically periodic in space and time. Assuming $V(d) = \frac{v_1}{2} d^2 + \mathcal{O}(d^3)$ and $W(y) = \frac{w_1}{2} y^2 + \mathcal{O}(y^3)$ we find the linearized system

$$\ddot{x}_j = v_1(x_{j+1}-2x_j+x_{j-1}) - w_1 x_j, \quad j \in \mathcal{J},$$

where we always assume $\min\{w_1, w_1+4v_1\} > 0$ in order to obtain stability. The linear system has the solutions

$$x_j(t) = \mathrm{e}^{i(\widetilde{\omega} t + \vartheta j)} \text{ with } \widetilde{\omega}^2 = \omega(\vartheta)^2 := 2v_1(1-\cos\vartheta) + w_1.$$

Fixing ϑ and hence $\widetilde{\omega} = \omega(\vartheta)$, we study modulated solutions of the type

$$x_j(t) = X_j^A(t) + \mathcal{O}(\varepsilon^2) \quad \text{with} \quad X_j^A(t) := \varepsilon A(\varepsilon^2 t, \varepsilon(j-ct)) \mathrm{e}^{i(\widetilde{\omega} t + \vartheta j)} + \text{c.c.}. \tag{1.4}$$

In Figure 1 such a sequence $(x_j(0))_{j\in\mathbb{Z}}$ is displayed together with the envelopes $\pm 2|A(0,\cdot)|$.

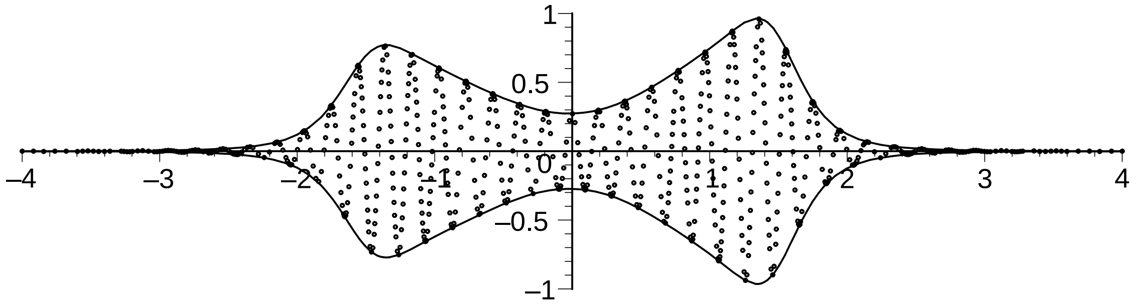


Figure 1: A modulated initial datum $(x_j(0))_{j\in\mathbb{Z}}$ (dots) together with the envelopes $\pm 2|A(0,\cdot)|$.

In Section 2 we show that this provides a useful approximation for solutions of (1.1) only if the group velocity c equals $-\omega'(\vartheta)$ and A satisfies the associated nonlinear Schrödinger equation (NLSE)

$$\mathrm{i}\partial_\tau A = \frac{1}{2}\omega''(\vartheta)\partial_\xi^2 A + \rho|A|^2 A, \tag{1.5}$$

where ρ can be calculated explicitly. Here $\tau = \varepsilon^2 t$ is the macroscopic time and $\xi = \varepsilon(j-ct)$ is the macroscopic space variable. This derivation of (1.5) is formal, since we assumed that solutions in the form (1.4) exist.

In Section 3 we justify the ansatz (1.4) by showing that solutions $t \mapsto (x_j(t))_{j\in\mathcal{J}}$ which start at $t = 0$ in the form (1.4) stay in this form over intervals $[0, \tau_0/\varepsilon^2]$, which have a positive macroscopic length.

Theorem 3.2 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 any solution x of (1.1) with

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

satisfies the estimate

$$\|(x(t), \dot{x}(t)) - (X^A(t), \dot{X}^A(t))\|_{\ell^2\times\ell^2} \le C\varepsilon^{3/2} \text{ for } t \in [0, \tau_0/\varepsilon^2]. \tag{1.6}$$

An essential, technical assumption of our theory is that the nonlinearity in (1.1) starts with cubic terms, i.e. $V'''(0) = W'''(0) = 0$. For such systems a relatively easy proof for the justification of NLSE was developed in [KSM92]. We believe that the same is true without this assumption, however, the proof will be much more difficult (see, e.g., [Sch98]) and is postponed to future work [GM03]. In [Kal89] (Prop. 7.2) results are stated without proof providing estimates like (1.6) under much stronger conditions, namely that the nonlinearity has to be analytic and that the solution A of (1.5) has to be analytic and rapidly decaying.

Moreover, the case without the stabilizing background potential W is also more difficult, since Galileian invariance may interact with our modulated patterns. In that case more complicated modulation equations are to be expected.

In Section 4 we provide an analogous result for the case of a finite, but large periodic chain. Moreover, we present numerical results which compare the macroscopic limit equation NLSE (1.5), posed on $(0, 2\pi)$ with periodicity conditions, with (1.1) for $\mathcal{J} = \mathbb{Z}_m$, where $m = 100, \ldots, 4000$ corresponds approximately to $\varepsilon = 0.06, \ldots, 0.0016$.

Note that our solutions given through (1.4) as well as those in (1.2) are small and thus lead to dynamics which are close to the linear one. Only the extremely long time scales allow for the accumulation of the nonlinear effects which are inherent to NLSE.

On shorter time scales, namely for $\tau = \varepsilon t$ with $\xi = \varepsilon j$, one only sees hyperbolic transport effects, but no dispersion. For the linear case we refer to [Mie03], where Wigner measures are used to describe the energy transport in multi-dimensional lattices. For larger nonlinear microscopic oscillations see [HLM94, DKV95, FV99, DK00, DH02], where the Whitham equation is derived to describe the associated modulations.

2 Formal derivation of the NLSE

In this section we formally derive the NLSE as a macroscopic limit, also called modulation equation. We give the calculations in full detail, since so far such an analysis is not yet standard. Moreover, we want to present the results in such a way, that they can be used for the rigorous analysis in the next section. Here we treat the general case, where quadratic nonlinearities are also allowed. The oscillator chain is modeled by

$$\ddot{x}_j = V'\left(\partial_j^+ x\right) - V'\left(\partial_j^- x\right) - W'(x_j), \quad j \in \mathbb{Z}, \tag{2.1}$$

where $x_j = x_j(t) \in \mathbb{R}$, $t \geq 0$, $j \in \mathbb{Z}$, and $\partial_j^\pm x := \pm(x_{j\pm 1} - x_j)$ (implying $\partial_j^+ x = \partial_{j+1}^- x$, $\partial_j^- x = \partial_{j-1}^+ x$). The potentials $V, W \in \mathrm{C}^5(\mathbb{R})$ are of the form

$$V(d) = \tfrac{v_1}{2} d^2 + \widetilde{V}(d), \quad W(y) = \tfrac{w_1}{2} y^2 + \widetilde{W}(y),$$
$$\text{with } \widetilde{V}(d) = \tfrac{v_2}{3} d^3 + \tfrac{v_3}{4} d^4 + \mathcal{O}(d^5) \text{ and } \widetilde{W}(y) = \tfrac{w_2}{3} y^3 + \tfrac{w_3}{4} y^4 + \mathcal{O}(y^5).$$

The linear part of (2.1) reads

$$\ddot{x}_j = L_j x := v_1\left(\partial_j^+ x - \partial_j^- x\right) - w_1 x_j = v_1\left(x_{j+1} - 2x_j + x_{j-1}\right) - w_1 x_j, \quad j \in \mathbb{Z}, \tag{2.2}$$

and has the basic solutions $x_j(t) = \mathrm{e}^{\mathrm{i}(\widetilde{\omega} t + \vartheta j)}$, if the dispersion relation

$$\widetilde{\omega}^2 = \omega(\vartheta)^2 := -\left[v_1\left(\mathrm{e}^{\mathrm{i}\vartheta} - 2 + \mathrm{e}^{-\mathrm{i}\vartheta}\right) - w_1\right] = 2v_1(1 - \cos\vartheta) + w_1$$

is satisfied. We always assume $\min\{w_1, w_1 + 4v_1\} > 0$, such that $\omega(\vartheta)^2 > 0$ for all ϑ. Subsequently we fix a value $\vartheta \in (-\pi, \pi]$, and write shortly ω, ω', ω'' to denote $\omega(\vartheta)$, $\omega'(\vartheta)$, $\omega''(\vartheta)$, respectively. The associated basic mode $\mathbf{E}(t,j) := \mathrm{e}^{\mathrm{i}(\omega t + \vartheta j)}$ is considered to be the microscopic pattern.

Our aim is to understand the macroscopic evolution of modulations of the microscopic pattern, which are given by a modulation function A:

$$x_j(t) = X_j^A(t) + \mathcal{O}(\varepsilon^2) \quad \text{with} \quad X_j^A(t) = \varepsilon A(\varepsilon^2 t, \varepsilon(j - ct))\mathbf{E}(t,j) + \text{c.c.},$$

where $\tau = \varepsilon^2 t$ and $\xi = \varepsilon(j - ct)$ play the role of a macroscopic time and space variable, respectively. Inserting such an ansatz into the nonlinear problem (2.1) will generate higher harmonic terms (i.e., $\mathbf{E}^n$). Hence, we insert the multiple scale ansatz

$$X_j^{(A)}(t) := \sum_{k \in \mathbb{N}} \varepsilon^k \sum_{n=-k}^{k} A_{k,n}(\tau, \xi)\mathbf{E}(t,j)^n \qquad (j \in \mathbb{Z},\ t \geq 0) \tag{2.3}$$

in (2.1), where $A_{k,n}(\tau,\xi)\in\mathbb{C}$ and $A_{k,-n}=\overline{A_{k,n}}$ (implying $A_{k,0}\in\mathbb{R}$ for all $k\in\mathbb{N}$). In the following we will use the abbreviation $\sum_{k,n}$ for the summation over $k\in\mathbb{N}$ and $n\in\mathbb{Z}$, $|n|\le k$. It will be sufficient to consider only terms with $k\le 3$, but it is instructive to keep the full generality.

For the left hand side we obtain

$$\ddot{X}_j^{(A)}=\sum_{k,n}\varepsilon^k\big[-(n\omega)^2A_{k,n}-2\varepsilon n\mathrm{i}\omega c\partial_\xi A_{k,n}+\varepsilon^2\left(c^2\partial_\xi^2A_{k,n}+2n\mathrm{i}\omega\partial_\tau A_{k,n}\right)\\ -2\varepsilon^3c\partial_\xi\partial_\tau A_{k,n}+\varepsilon^4\partial_\tau^2A_{k,n}\big]\mathbf{E}^n, \tag{2.4}$$

where the arguments (τ,ξ) of $A_{k,n}$ are omitted.

With

$$\partial_j^\pm X^{(A)}(t)=\pm\sum_{k,n}\varepsilon^k\left[A_{k,n}(\tau,\xi\pm\varepsilon)\mathrm{e}^{\pm\mathrm{i}n\vartheta}-A_{k,n}\right]\mathbf{E}(t,j)^n \tag{2.5}$$

the linear part of the right hand side reads

$$\begin{aligned}L_jX^{(A)}&=v_1\left(X_{j+1}^{(A)}-2X_j^{(A)}+X_{j-1}^{(A)}\right)-w_1X_j^{(A)}\\&=\sum_{k,n}\varepsilon^k\left\{v_1\left[A_{k,n}(\tau,\xi+\varepsilon)\mathrm{e}^{\mathrm{i}n\vartheta}-2A_{k,n}+A_{k,n}(\tau,\xi-\varepsilon)\mathrm{e}^{-\mathrm{i}n\vartheta}\right]-w_1A_{k,n}\right\}\mathbf{E}^n.\end{aligned}$$

From the expansion

$$A_{k,n}(\tau,\xi\pm\varepsilon)=A_{k,n}\pm\varepsilon\partial_\xi A_{k,n}+\varepsilon^2\frac{1}{2}\partial_\xi^2A_{k,n}\pm\varepsilon^3\frac{1}{6}\partial_\xi^3A_{k,n}\left(\tau,\xi\pm\theta_{k,n}^\pm\varepsilon\right),\quad\theta_{k,n}^\pm\in(0,1),$$

we obtain

$$L_jX^{(A)}=\sum_{k,n}\varepsilon^k\Big\{-\omega(n\vartheta)^2A_{k,n}+\varepsilon[2\mathrm{i}\omega(n\vartheta)\omega'(n\vartheta)]\partial_\xi A_{k,n}\\ +\varepsilon^2[\omega'(n\vartheta)^2+\omega(n\vartheta)\omega''(n\vartheta)]\partial_\xi^2A_{k,n}+\varepsilon^3r_{k,n}\Big\}\mathbf{E}^n \tag{2.6}$$

with $r_{k,n}=\frac{v_1}{6}\left[\mathrm{e}^{\mathrm{i}n\vartheta}\partial_\xi^3A_{k,n}\left(\tau,\xi+\theta_{k,n}^+\varepsilon\right)-\mathrm{e}^{-\mathrm{i}n\vartheta}\partial_\xi^3A_{k,n}\left(\tau,\xi-\theta_{k,n}^-\varepsilon\right)\right]$. Here we used that $\omega^2=2v_1(1-\cos\vartheta)+w_1$ implies $\omega\omega'=v_1\sin\vartheta$ and $(\omega')^2+\omega\omega''=v_1\cos\vartheta$.

With (2.6) we have obtained an expansion in terms of $\varepsilon^k\mathbf{E}^n$ of the linear part of the right hand side of the microscopic equation (2.1) with $x=X^{(A)}$. It remains to obtain a similar expansion for the nonlinear part $N_j(X^{(A)}):=\widetilde{V}'(\partial_j^+X^{(A)})-\widetilde{V}'(\partial_j^-X^{(A)})-\widetilde{W}'(X_j^{(A)})$. We start by deriving an expansion only in terms of ε^k. At first we note

$$\begin{aligned}N_j(X^{(A)})=&\,v_2\left[(\partial_j^+X^{(A)})^2-(\partial_j^-X^{(A)})^2\right]-w_2(X_j^{(A)})^2+v_3\left[(\partial_j^+X^{(A)})^3-(\partial_j^-X^{(A)})^3\right]\\&-w_3(X_j^{(A)})^3+\widehat{V}'(\partial_j^+X^{(A)})-\widehat{V}'(\partial_j^-X^{(A)})-\widehat{W}'(X_j^{(A)})\end{aligned}$$

with $\widehat{V}'(d)=\widetilde{V}'(d)-v_2d^2-v_3d^3=\mathcal{O}(d^4)$ and $\widehat{W}'(y)=\widetilde{W}'(y)-w_2y^2-w_3y^3=\mathcal{O}(y^4)$. With (2.5) and (2.3) we obtain

$$\partial_j^\pm X^{(A)}=\varepsilon\left(a_1^\pm+\varepsilon b_1^\pm\right)+\varepsilon^2a_2^\pm+\varepsilon^3r_1^\pm\quad\text{and}\quad X_j^{(A)}=\varepsilon a_1+\varepsilon^2a_2+\varepsilon^3r_1,$$

respectively, where

$$
\begin{aligned}
a_1^{\pm} &= \pm\left(\mathrm{e}^{\pm\mathrm{i}\vartheta}-1\right) A_{1,1}\mathbf{E} + \text{c.c.}, \qquad b_1^{\pm} = \partial_\xi A_{1,0} + \left(\mathrm{e}^{\pm\mathrm{i}\vartheta}\partial_\xi A_{1,1}\mathbf{E} + \text{c.c.}\right),\\
a_2^{\pm} &= \pm\left(\mathrm{e}^{\pm\mathrm{i}\vartheta}-1\right) A_{2,1}\mathbf{E} \pm \left(\mathrm{e}^{\pm2\mathrm{i}\vartheta}-1\right) A_{2,2}\mathbf{E}^2 + \text{c.c.},\\
r_1^{\pm} &= \pm\sum_{n=-1}^{1} \mathrm{e}^{\pm\mathrm{i}n\vartheta}\frac{1}{2}\partial_\xi^2 A_{1,n}\left(\tau,\xi\pm\widehat{\theta}_{1,n}^{\pm}\varepsilon\right)\mathbf{E}^n + \sum_{n=-2}^{2}\mathrm{e}^{\pm\mathrm{i}n\vartheta}\partial_\xi A_{2,n}\left(\tau,\xi\pm\widetilde{\theta}_{1,n}^{\pm}\varepsilon\right)\mathbf{E}^n\\
&\qquad \pm\sum_{k\geq3}\varepsilon^{k-3}\sum_{n=-k}^{k}\left[A_{k,n}(\tau,\xi\pm\varepsilon)\mathrm{e}^{\pm\mathrm{i}n\vartheta}-A_{k,n}\right]\mathbf{E}^n \quad \text{with} \quad \widehat{\theta}_{1,n}^{\pm},\widetilde{\theta}_{1,n}^{\pm}\in(0,1),\\
a_1 &= A_{1,0} + \left(A_{1,1}\mathbf{E} + \text{c.c.}\right), \qquad a_2 = A_{2,0} + \left(A_{2,1}\mathbf{E} + A_{2,2}\mathbf{E}^2 + \text{c.c.}\right),\\
r_1 &= \sum_{k\geq3}\varepsilon^{k-3}\sum_{n=-k}^{k} A_{k,n}\mathbf{E}^n.
\end{aligned}
$$

Insertion into the nonlinearity gives

$$
\begin{aligned}
N_j(X^{(A)}) = {}& \varepsilon^2\left\{v_2\left[(a_1^+)^2-(a_1^-)^2\right]-w_2 a_1^2\right\}\\
&+\varepsilon^3\left\{2v_2\left[a_1^+(b_1^++a_2^+)-a_1^-(b_1^-+a_2^-)\right]+v_3\left[(a_1^+)^3-(a_1^-)^3\right]-2w_2a_1a_2-w_3a_1^3\right\}\\
&+\varepsilon^4\left(r_2^+-r_2^--r_2\right)+\widehat{V}'(\partial_j^+X^{(A)})-\widehat{V}'(\partial_j^-X^{(A)})-\widehat{W}'(X_j^{(A)})
\end{aligned}
\tag{2.7}
$$

with

$$
\begin{aligned}
r_2^{\pm} &= 2v_2a_1^{\pm}r_1^{\pm}+3v_3(a_1^{\pm})^2(b_1^{\pm}+a_2^{\pm}+\varepsilon r_1^{\pm})+(v_2+3v_3\varepsilon a_1^{\pm})(b_1^{\pm}+a_2^{\pm}+\varepsilon r_1^{\pm})^2+v_3\varepsilon^2(b_1^{\pm}+a_2^{\pm}+\varepsilon r_1^{\pm})^3,\\
r_2 &= 2w_2a_1r_1+3w_3a_1^2\left(a_2+\varepsilon r_1\right)+\left(w_2+3w_3\varepsilon a_1\right)\left(a_2+\varepsilon r_1\right)^2+w_3\varepsilon^2\left(a_2+\varepsilon r_1\right)^3,
\end{aligned}
$$

where the last three terms in (2.7) are of order $\mathcal{O}(\varepsilon^4)$, since we have $\widehat{V}'(d) = \mathcal{O}(d^4)$, $\widehat{W}'(y) = \mathcal{O}(y^4)$ and $X_j^{(A)}, \partial_j^{\pm}X^{(A)} = \mathcal{O}(\varepsilon)$.

The general procedure for deriving modulation equations consists in equating the left hand side and right hand side coefficients of each term $\varepsilon^k\mathbf{E}^n$ in equation (2.1) with $x = X^{(A)}$

$$
\ddot{X}_j^{(A)} = L_jX^{(A)} + N_j(X^{(A)})
\tag{2.8}
$$

separately. Thereby, we can omit the equations for $n < 0$ since they are the complex conjugates of the equations for $n > 0$. We start with $k = 1$, where we use that the nonlinearity generates only terms of the power $k \geq 2$. Thus, we obtain from (2.4), (2.6), and (2.7) for $k = 1$ and $n = 0, 1$

$$
\begin{aligned}
&\text{for } \varepsilon^1\mathbf{E}^0: && 0 = -w_1A_{1,0};\\
&\text{for } \varepsilon^1\mathbf{E}^1: && -\omega^2A_{1,1} = -\omega^2A_{1,1}.
\end{aligned}
$$

From $w_1 > 0$ we conclude $A_{1,0} = 0$, and $A_{1,1}$ remains free at this stage. Using $A_{1,0} = 0$, calculation of the terms appearing in (2.7) yields

$$
\begin{aligned}
N_j(X^{(A)}) = &-\varepsilon^2\left[w_2|A_{1,1}|^2\mathbf{E}^0+\left(v_2s_1c_1+w_2\right)A_{1,1}^2\mathbf{E}^2+\text{c.c.}\right]\\
&+\varepsilon^3\Big\{\left[2v_2c_1\bar{A}_{1,1}\partial_\xi A_{1,1}-2w_2\bar{A}_{1,1}A_{2,1}\right]\mathbf{E}^0
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \left[2\left(v_2 s_1 c_1 - w_2\right)\bar{A}_{1,1}A_{2,2} - 2w_2 A_{1,1}A_{2,0} - 3\left(v_3 c_1^2 + w_3\right)|A_{1,1}|^2 A_{1,1}\right]\mathbf{E} \\
&\quad + \left[2v_2 c_1(c_1-3)A_{1,1}\partial_\xi A_{1,1} - 2\left(v_2 s_1 c_1 + w_2\right)A_{1,1}A_{2,1}\right]\mathbf{E}^2 \\
&\quad + \left[2\left[v_2 s_1(c_1+s_1^2) - w_2\right]A_{1,1}A_{2,2} + \left[v_3 c_1^2(3-c_1) - w_3\right]A_{1,1}^3\right]\mathbf{E}^3 + \text{c.c.}\big\} \\
&+ \varepsilon^4\left(r_2^+ - r_2^- - r_2\right) + \widehat{V}'(\partial_j^+ X^{(A)}) - \widehat{V}'(\partial_j^- X^{(A)}) - \widehat{W}'(X_j^{(A)})
\end{aligned}
\tag{2.9}
$$

with $s_1 := 2\mathrm{i}\sin\vartheta$, $c_1 := 2(1-\cos\vartheta)$. For $k=2$ we obtain from (2.4), (2.6), and (2.9), by comparing the terms associated with $\varepsilon^2\mathbf{E}^n$, $n=0,1,2$,

$$
\begin{aligned}
&\text{for } \varepsilon^2\mathbf{E}^0: && 0 = -w_1 A_{2,0} - 2w_2|A_{1,1}|^2; \\
&\text{for } \varepsilon^2\mathbf{E}^1: && -\omega^2 A_{2,1} - 2\mathrm{i}\omega c\partial_\xi A_{1,1} = -\omega^2 A_{2,1} + 2\mathrm{i}\omega\omega'\partial_\xi A_{1,1}; \\
&\text{for } \varepsilon^2\mathbf{E}^2: && -4\omega^2 A_{2,2} = -\omega(2\vartheta)^2 A_{2,2} - \left(v_2 s_1 c_1 + w_2\right)A_{1,1}^2.
\end{aligned}
$$

With $w_1 > 0$ the equation for $\varepsilon^2\mathbf{E}^0$ gives

$$
A_{2,0} = -\frac{2w_2}{w_1}|A_{1,1}|^2. \tag{2.10}
$$

The equation for $\varepsilon^2\mathbf{E}^1$ yields $c = -\omega'$. To proceed further in the general case we have to assume the nonresonance condition $4\omega^2 \neq \omega(2\vartheta)^2$, such that the equation for $\varepsilon^2\mathbf{E}^2$ implies

$$
A_{2,2} = \frac{v_2 s_1 c_1 + w_2}{4\omega^2 - \omega(2\vartheta)^2}A_{1,1}^2. \tag{2.11}
$$

However, in the case of cubic nonlinearities (where $v_2 = w_2 = 0$) we do not need this nonresonance condition, since we may simply set $A_{2,2} = 0$. The function $A_{2,1}$ remains free at this stage. In the same manner, by equating the left hand side and right hand side coefficients of the terms $\varepsilon^3\mathbf{E}^n$ for $n = 0,1,2,3$ we obtain

$$
\begin{aligned}
&\text{for } \varepsilon^3\mathbf{E}^0: \quad 0 = -w_1 A_{3,0} + \left(2v_2 c_1 \bar{A}_{1,1}\partial_\xi A_{1,1} - 2w_2\bar{A}_{1,1}A_{2,1} + \text{c.c.}\right); \\
&\text{for } \varepsilon^3\mathbf{E}^1: \quad -\omega^2 A_{3,1} - 2\mathrm{i}\omega c\partial_\xi A_{2,1} + c^2\partial_\xi^2 A_{1,1} + 2\mathrm{i}\omega\partial_\tau A_{1,1} \\
&\qquad\qquad = -\omega^2 A_{3,1} + 2\mathrm{i}\omega\omega'\partial_\xi A_{2,1} + [(\omega')^2 + \omega\omega'']\partial_\xi^2 A_{1,1} + 2(v_2 s_1 c_1 - w_2)\bar{A}_{1,1}A_{2,2} \\
&\qquad\qquad\quad -2w_2 A_{1,1}A_{2,0} - 3(v_3 c_1^2 + w_3)|A_{1,1}|^2 A_{1,1}; \\
&\text{for } \varepsilon^3\mathbf{E}^2: \quad -4\omega^2 A_{3,2} - 4\mathrm{i}\omega c\partial_\xi A_{2,2} = -\omega(2\vartheta)^2 A_{3,2} + 2\mathrm{i}\omega(2\vartheta)\omega'(2\vartheta)\partial_\xi A_{2,2} \\
&\qquad\qquad\qquad\qquad + 2v_2 c_1(c_1-3)A_{1,1}\partial_\xi A_{1,1} - 2(v_2 s_1 c_1 + w_2)A_{1,1}A_{2,1}; \\
&\text{for } \varepsilon^3\mathbf{E}^3: \quad -9\omega^2 A_{3,3} = -\omega(3\vartheta)^2 A_{3,3} + 2[v_2 s_1(c_1+s_1^2) - w_2]A_{1,1}A_{2,2} + [v_3 c_1^2(3-c_1) - w_3]A_{1,1}^3.
\end{aligned}
$$

From the equation for $\varepsilon^3\mathbf{E}^1$ we obtain with $c = -\omega'$, (2.10), and (2.11) the nonlinear Schrödinger equation

$$
\mathrm{i}\partial_\tau A_{1,1} = \frac{1}{2}\omega''\partial_\xi^2 A_{1,1} + \left[\frac{(v_2 s_1 c_1)^2 - w_2^2}{\omega\left(v_1 c_1^2 + 3w_1\right)} + \frac{2w_2^2}{\omega w_1} - \frac{3\left(v_3 c_1^2 + w_3\right)}{2\omega}\right]|A_{1,1}|^2 A_{1,1}. \tag{2.12}
$$

From the equation for $\varepsilon^3\mathbf{E}^3$ we obtain with (2.11)

$$
A_{3,3} = \frac{1}{\omega(3\vartheta)^2 - 9\omega^2}\left\{2\left[v_2 s_1(c_1+s_1^2) - w_2\right]\frac{v_2 s_1 c_1 + w_2}{v_1 c_1^2 + 3w_1} + \left[v_3 c_1^2(3-c_1) - w_3\right]\right\}A_{1,1}^3, \tag{2.13}
$$

where

$$\omega(3\vartheta)^2-9\omega^2 = -8w_1+v_1c_1^2(c_1-6) = -8\left[w_1+v_1\left(2+\cos^3\vartheta-3\cos\vartheta\right)\right] < 0$$

for all $\vartheta \in (-\pi,\pi]$, since $f(\vartheta) = 2+\cos^3\vartheta-3\cos\vartheta$ has (global) minimum 0 and maximum 4 and we have assumed $\min\{w_1, w_1+4v_1\} > 0$. The amplitude functions $A_{3,0}$ and $A_{3,2}$ can be calculated from $A_{1,1}$ by the equations for $\varepsilon^3\mathbf{E}^0$ and $\varepsilon^3\mathbf{E}^2$, respectively, if additionally $A_{2,1}$ is specified. However, here this will not be needed. The function $A_{3,1}$ remains free.

Thus, we have established the following result.

Theorem 2.1 *If the microscopic oscillator chain equation (2.1) has for all $\varepsilon \in (0,\varepsilon_0)$ solutions of the form*

$$x_j(t) = X_j^A(t) + \mathcal{O}(\varepsilon^2) \quad \textit{with} \quad X_j^A(t) = \varepsilon A(\varepsilon^2 t, \varepsilon(j-ct))\mathbf{E}(t,j) + \text{c.c.},$$

where $A : [0,\tau_0]\times\mathbb{R}\to\mathbb{C}$ is a smooth function, then A has to satisfy the NLSE (2.12).

We call this result a formal derivation, since the existence of solutions satisfying this ansatz is not clear at all. The purpose of the next section is to show that solutions which start in this form will maintain it on suitably long time scales.

Let us now consider the case where the nonlinearity in our oscillator chain model (2.1) $N_j(x) = \widetilde{V}'(\partial_j^+ x) - \widetilde{V}'(\partial_j^- x) - \widetilde{W}'(x_j)$ has no quadratic terms, i.e., the case $v_2 = w_2 = 0$. In this case we obtain from (2.10) and (2.11) (or can set, if $4\omega^2 = \omega(2\vartheta)^2$) $A_{2,0} = A_{2,2} = 0$, and from the equations for $\varepsilon^3\mathbf{E}^0$ and $\varepsilon^3\mathbf{E}^2$ we obtain (or can set) $A_{3,0} = A_{3,2} = 0$. The nonlinear Schrödinger equation (2.12) reads

$$\mathrm{i}\partial_\tau A_{1,1} = \frac{1}{2}\omega''\partial_\xi^2 A_{1,1} + \rho\,|A_{1,1}|^2 A_{1,1} \quad \text{with} \quad \rho := -\frac{3\,(v_3c_1^2+w_3)}{2\omega}, \tag{2.14}$$

where $c_1 = 2(1-\cos\vartheta)$, and from (2.13) we obtain

$$A_{3,3} = \Psi A_{1,1}^3 \quad \text{with} \quad \Psi := \frac{v_3c_1^2(3-c_1)-w_3}{\omega(3\vartheta)^2-9\omega^2}. \tag{2.15}$$

Up to this stage there are no conditions posed on $A_{3,1}$ and $A_{2,1}$, or on $A_{k,n}$ for $(k,n) \in \mathbb{N}\times\mathbb{Z}$ with $k \ge 4$, $|n| \le k$. Hence, setting deliberately also $A_{3,1} = A_{2,1} = 0$ and $A_{k,\cdot} = 0$ for $k \ge 4$, the general multiple scale ansatz $X_j^{(A)}(t)$ in (2.3) obtains the special form

$$Z_j^A(t) := \varepsilon A(\tau,\xi)\mathbf{E}(t,j) + \varepsilon^3\Psi A^3(\tau,\xi)\mathbf{E}(t,j)^3 + \text{c.c.} \qquad (j\in\mathbb{Z},\ t\ge 0) \tag{2.16}$$

with $\tau = \varepsilon^2 t$, $\xi = \varepsilon(j+\omega' t)$, and leads to the *nonlinear Schrödinger equation (NLSE)* (2.14) with $A_{1,1} = A$, which is the sought-after *macroscopic* or *modulation equation.*

3 Justification of the NLSE

We will prove rigorously that if A is a solution of the NLSE (2.14)

$$\mathrm{i}\partial_\tau A = \frac{1}{2}\omega''\partial_\xi^2 A + \rho\,|A|^2 A \quad \text{with} \quad \rho = -\frac{3\,(v_3c_1^2+w_3)}{2\omega}, \quad c_1 = 2(1-\cos\vartheta),$$

then

$$X_j^A(t) = \varepsilon A(\tau,\xi)\mathbf{E}(t,j) + \text{c.c.} \qquad (j\in\mathbb{Z},\, t\geq 0) \tag{3.1}$$

with $\tau = \varepsilon^2 t$, $\xi = \varepsilon(j{+}\omega' t)$, is a reasonable approximation to solutions of the oscillator chain model (2.1) with cubic leading terms of the nonlinearities (i.e. $V'''(0) = W'''(0) = 0$). Up to now no systematic theory for the justification of modulation equations for discrete systems has been developed. For this reason we give all the estimates in full detail. In particular, it is not enough to estimate errors at each point $j \in \mathcal{J}$ like in the formal derivation of the previous section. We rather need estimates in suitable Banach spaces. To this end, we transform (2.1) into the first-order ordinary differential equation

$$\dot{\widetilde{x}} = \mathcal{L}\widetilde{x} + \mathcal{N}(\widetilde{x}) \qquad \text{with} \quad \widetilde{x} := (x,\dot{x}) \tag{3.2}$$

in the Banach space $Y := \ell^2{\times}\ell^2$, with $\mathcal{L}$ and $\mathcal{N}$ given by

$$[\mathcal{L}\widetilde{x}]_j := (\dot{x}_j, L_j x) \quad \text{with} \qquad L_j x = v_1\left(\partial_j^+ x{-}\partial_j^- x\right) - w_1 x_j, \tag{3.3}$$

$$[\mathcal{N}(\widetilde{x})]_j := (0, N_j(x)) \quad \text{with} \quad N_j(x) = \widetilde{V}'\left(\partial_j^+ x\right) - \widetilde{V}'\left(\partial_j^- x\right) - \widetilde{W}'\left(x_j\right). \tag{3.4}$$

In addition to the standard norm on the Banach space Y we use the energy norm $\|\cdot\|_Y$ with $\|(x,y)\|_Y^2 := \|x\|_E^2 + \|y\|^2$, where $\|\cdot\|$ denotes the standard ℓ^2-Norm, i.e., $\|y\|^2 := \|y\|_{\ell^2}^2 = \sum_{j\in\mathbb{Z}} |y_j|^2$, and $\|\cdot\|_E$ denotes the energy norm

$$\|x\|_E^2 := \sum_{j\in\mathbb{Z}}\left(v_1\left|\partial_j^+ x\right|^2 + w_1|x_j|^2\right) = v_1\sum_{j\in\mathbb{Z}}\left|\partial_j^+ x\right|^2 + w_1\|x\|^2.$$

For $\min\{w_1, w_1{+}4v_1\} > 0$ the norms $\|\cdot\|$ and $\|\cdot\|_E$ are equivalent with

$$\min\{w_1, w_1{+}4v_1\}\|x\|^2 \leq \|x\|_E^2 \leq \max\{w_1, w_1{+}4v_1\}\|x\|^2.$$

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

$$H(x,\dot{x}) = \frac{1}{2}\|\dot{x}\|^2 + \sum_{j\in\mathbb{Z}}\Big[V(\partial_j^+ x) + W(x_j)\Big]$$

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 following result states the well-known fact that the flow of the linearized system (2.2) preserves this norm.

Proposition 3.1 *The solutions $\widetilde{x} : t \mapsto \widetilde{x}(t) = \mathrm{e}^{t\mathcal{L}}\widetilde{x}(0)$ of (2.2) satisfy $\|\widetilde{x}(t)\|_Y = \|\widetilde{x}(0)\|_Y$ for all $t\in\mathbb{R}$.*

Proof: Since $x_j \in \mathbb{R}$ for all $j\in\mathbb{Z}$, we have by definition

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|\widetilde{x}(t)\|_Y^2 &= \frac{\mathrm{d}}{\mathrm{d}t}\sum_{j\in\mathbb{Z}}\left[\dot{x}_j^2 + v_1\left(x_{j+1}{-}x_j\right)^2 + w_1 x_j^2\right] \\ &= 2\sum_{j\in\mathbb{Z}}\dot{x}_j\left[\ddot{x}_j - v_1\left(x_{j+1}{-}2x_j{+}x_{j-1}\right) + w_1 x_j\right] = 2\sum_{j\in\mathbb{Z}}\dot{x}_j\cdot 0,\end{aligned}$$

since the linear system $\dot{\widetilde{x}} = \mathcal{L}\widetilde{x}$ reads $\ddot{x}_j - v_1\,(x_{j+1}-2x_j+x_{j-1}) + w_1 x_j = 0$, $j \in \mathbb{Z}$. $\square$

The following theorem constitutes our justification of the validity of the NLSE (2.14) as a macroscopic limit for the oscillator chain model (2.1) with cubic nonlinearities (i.e., with $v_2 = w_2 = 0$ in the potentials V, W).

Theorem 3.2 *Assume that $V,\ W \in \mathrm{C}^5(\mathbb{R})$ in (2.1) satisfy $V(d) = \frac{v_1}{2}d^2 + \mathcal{O}(d^4)$ and $W(y) = \frac{w_1}{2}y^2 + \mathcal{O}(y^4)$ with $\min\{w_1, w_1{+}4v_1\} > 0$. Let $A : [0,\tau_0]\times\mathbb{R} \to \mathbb{C}$ be a solution of the NLSE (2.14) with $A(0,\cdot) \in \mathrm{H}^5(\mathbb{R})$ and let X^A be the formal approximation (3.1). 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 (3.2) with an initial condition $\widetilde{x}(0)$ satisfying

$$\left\|\widetilde{x}(0) - \widetilde{X}^A(0)\right\|_Y \le d\varepsilon^{3/2}, \tag{3.5}$$

fulfills the estimate

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

Proof: Using the standard theory of semilinear wave equations [Tem88, Paz83], there exists $C_A > 0$ such that the solution A of NLSE satisfies

$$\|\partial_\xi^k \partial_\tau^l A(\tau,\cdot)\|_{\mathrm{L}^2(\mathbb{R})} \le C_A \quad \text{ for } \tau \in [0,\tau_0] \text{ and } k,l \in \mathbb{N}_0 \text{ with } k{+}2l \le 5. \tag{3.6}$$

Inserting the approximation (2.16) $Z^A = X^A{+}Y^A$ with $X^A = \varepsilon A\mathbf{E} + \text{c.c.}$, $Y^A := \varepsilon^3\Psi A^3\mathbf{E}^3 + \text{c.c.}$ into (3.2), we obtain the residual term

$$\widetilde{\rho}^A := (0,\rho^A) := \dot{\widetilde{Z}}^A - \mathcal{L}\widetilde{Z}^A - \mathcal{N}(\widetilde{Z}^A) \quad \text{with} \quad \rho_j^A = \ddot{Z}_j^A - L_j Z^A - N_j(Z^A). \tag{3.7}$$

By (2.3) and (2.16), Z^A equals $X^{(A)}$ with $A_{1,0} = A_{2,0} = A_{2,1} = A_{2,2} = A_{3,0} = A_{3,1} = A_{3,2} = A_{k,n} = 0$ for $k \ge 4$, $n = -k,\dots,k$, and $A_{1,1} = A$, $A_{3,3} = \Psi A^3$. Hence, proceeding like in the previous section, and using (2.4) with $c = -\omega'$, (2.6) and (2.9), we obtain by formal comparison of the coefficients of the terms $\varepsilon^k\mathbf{E}^n$ of (2.8), with Z^A instead of $X^{(A)}$, the expansion

$$\begin{aligned}
\rho_j^A = \ddot{Z}_j^A - L_j Z^A - N_j(Z^A) &= \varepsilon^4\left\{\left[\varepsilon\partial_\tau^2 A{+}2\omega'\partial_\xi\partial_\tau A{-}r_{1,1}\right]\mathbf{E} + \text{c.c.}\right\}\\
&+ \varepsilon^4\Psi\left\{\left[\varepsilon^3\partial_\tau^2 A^3{+}2\varepsilon^2\omega'\partial_\xi\partial_\tau A^3{+}\varepsilon\left((\omega')^2\partial_\xi^2 A^3{+}6\mathrm{i}\omega\partial_\tau A^3\right){+}6\mathrm{i}\omega\omega'\partial_\xi A^3 - \widetilde{r}_{3,3}\right]\mathbf{E}^3 + \text{c.c.}\right\}\\
&+ \varepsilon^4(r_2^-{-}r_2^+{+}r_2) - \widehat{V}'(\partial_j^+ Z^A) + \widehat{V}'(\partial_j^- Z^A) + \widehat{W}'(Z_j^A)
\end{aligned} \tag{3.8}$$

with

$$\begin{aligned}
r_{1,1} &= \frac{v_1}{6}\left[\mathrm{e}^{\mathrm{i}\vartheta}\partial_\xi^3 A\left(\tau,\xi{+}\theta_{1,1}^+\varepsilon\right) - \mathrm{e}^{-\mathrm{i}\vartheta}\partial_\xi^3 A\left(\tau,\xi{-}\theta_{1,1}^-\varepsilon\right)\right],\\
\widetilde{r}_{3,3} &= v_1\left[\mathrm{e}^{\mathrm{i}3\vartheta}\partial_\xi A^3\left(\tau,\xi{+}\widetilde{\theta}_{3,3}^+\varepsilon\right) - \mathrm{e}^{-\mathrm{i}3\vartheta}\partial_\xi A^3\left(\tau,\xi{-}\widetilde{\theta}_{3,3}^-\varepsilon\right)\right],\\
r_2^\pm &= v_3\left[3(a_1^\pm)^2(b_1^\pm{+}\varepsilon r_1^\pm){+}3\varepsilon a_1^\pm(b_1^\pm{+}\varepsilon r_1^\pm)^2{+}\varepsilon^2(b_1^\pm{+}\varepsilon r_1^\pm)^3\right],\\
r_2 &= w_3\left(3\varepsilon a_1^2 r_1{+}3\varepsilon^3 a_1 r_1^2{+}\varepsilon^5 r_1^3\right)
\end{aligned}$$

and

$$a_1 = A\mathbf{E}+\text{c.c.},\;\; a_1^\pm = \pm\left(\mathrm{e}^{\pm\mathrm{i}\vartheta}-1\right)A\mathbf{E}+\text{c.c.},\;\; b_1^\pm = \mathrm{e}^{\pm\mathrm{i}\vartheta}\partial_\xi A\mathbf{E}+\text{c.c.},\;\; r_1 = \Psi A^3\mathbf{E}^3+\text{c.c.},$$
$$r_1^\pm = \pm\mathrm{e}^{\pm\mathrm{i}\vartheta}\frac{1}{2}\partial_\xi^2 A(\tau,\xi\pm\widehat{\theta}_{1,1}^\pm\varepsilon)\mathbf{E}+\Psi\left[\pm\left(\mathrm{e}^{\pm\mathrm{i}3\vartheta}-1\right)A^3+\varepsilon\mathrm{e}^{\pm\mathrm{i}3\vartheta}\partial_\xi A^3(\tau,\xi\pm\widetilde{\theta}_{3,3}^\pm\varepsilon)\right]\mathbf{E}^3+\text{c.c.},$$

where $\theta_{1,1}^\pm, \widetilde{\theta}_{3,3}^\pm, \widehat{\theta}_{1,1}^\pm \in (0,1)$. In accordance to the formal derivation of the previous section, in (3.8) there appear no terms of order ε^k, $k=1,2,3$, since we assumed $v_2 = w_2 = 0$, $c = -\omega'$, and that A solves the NLSE (2.14) with Ψ given by (2.15).

From (3.8) we obtain the estimate

$$\left|\rho_j^A(t)\right| \le \varepsilon^4 C_1\left(1+\max_{m+2n\le 4}\|\partial_\xi^m\partial_\tau^n A(\tau,\cdot)\|_\infty\right)\max_{k+2l\le 4}\sup_{|s|\le 1}\left|\partial_\xi^k\partial_\tau^l A\left(\tau,\varepsilon(j+\omega' t+s)\right)\right|$$

for all $\varepsilon \le \varepsilon_0$, $\tau = \varepsilon^2 t \le \tau_0$, and $j \in \mathbb{Z}$. The coefficient $C_1 > 0$ depends only on ε_0, C_A, V and W. (Recall that $\widehat{V}'(d) = \mathcal{O}(d^4)$, $\widehat{W}'(y) = \mathcal{O}(y^4)$.) Applying the subsequent Proposition 3.3 to $\phi = \partial_\xi^k\partial_\tau^l A(\tau,\cdot)$, and using the Sobolev embedding $\|u\|_\infty \le C_{\text{Sob}}\|u\|_{\mathrm{H}^1(\mathbb{R})}$ for $u \in \mathrm{H}^1(\mathbb{R})$, as well as (3.6), we obtain for $\widetilde{\rho}^A = (0,\rho^A)$ the estimate

$$\left\|\widetilde{\rho}^A(t)\right\|_Y \le \varepsilon^{7/2}C_1\left(1+C_{\text{Sob}}C_A\right)3\sqrt{8}C_A =: \varepsilon^{7/2}C_\rho \qquad \text{for } \varepsilon \le \varepsilon_0 \text{ and } t \le \tau_0/\varepsilon^2. \quad (3.9)$$

From $\widetilde{Y}_j^A = (Y_j^A, \dot{Y}_j^A) = \varepsilon^3\Psi(A^3\mathbf{E}^3+\text{c.c.}, 3A^2(\varepsilon^2\partial_\tau A+\varepsilon\omega'\partial_\xi A+\mathrm{i}\omega A)\mathbf{E}^3+\text{c.c.})$ we similarly obtain

$$\|\widetilde{Y}^A(t)\|_Y \le \varepsilon^{5/2}|\Psi|C_2\, C_{\text{Sob}}\, C_A^3 \qquad \text{for } \varepsilon \le \varepsilon_0 \text{ and } t \le \tau_0/\varepsilon^2 \quad (3.10)$$

with $C_2 > 0$ depending only on ε_0, V, W.

Above we have estimated the residual term $\widetilde{\rho}^A$. Now we have to show that this implies that the error between the approximation $\widetilde{Z}^A$ and the true solutions $\widetilde{x}$ remains small on the interval $[0,\tau_0]$. For the error $\widetilde{x}-\widetilde{Z}^A$ we use the ansatz $\widetilde{R} = (R,\dot{R}) = \varepsilon^{-3/2}(\widetilde{x}-\widetilde{Z}^A)$. Hence, it is our aim to show that $\widetilde{R}$ remains bounded independent of $\varepsilon \in (0,\varepsilon_0)$. From (3.2) and (3.7) we obtain

$$\dot{\widetilde{R}} = \mathcal{L}\widetilde{R}+(0,M)-\varepsilon^{-3/2}\widetilde{\rho}^A \quad (3.11)$$

with $(0,M) := \varepsilon^{-3/2}\left[\mathcal{N}(\varepsilon^{3/2}\widetilde{R}+\widetilde{Z}^A)-\mathcal{N}(\widetilde{Z}^A)\right]$. By definition (3.4) we have

$$\begin{aligned}\varepsilon^{3/2}M_j = {}& \widetilde{V}'\left(\varepsilon^{3/2}\partial_j^+R+\partial_j^+Z^A\right)-\widetilde{V}'\left(\partial_j^+Z^A\right)-\widetilde{V}'\left(\varepsilon^{3/2}\partial_j^-R+\partial_j^-Z^A\right)+\widetilde{V}'\left(\partial_j^-Z^A\right)\\ &-\widetilde{W}'\left(\varepsilon^{3/2}R_j+Z_j^A\right)+\widetilde{W}'\left(Z_j^A\right).\end{aligned}$$

From the mean value theorem we obtain

$$M_j = \widetilde{V}''\left(\varepsilon\, d_j^+\right)\partial_j^+R-\widetilde{V}''\left(\varepsilon\, d_j^-\right)\partial_j^-R-\widetilde{W}''\left(\varepsilon\, y_j\right)R_j$$

with $d_j^\pm := \vartheta_j^\pm\varepsilon^{1/2}\partial_j^\pm R+\frac{1}{\varepsilon}\partial_j^\pm Z^A$, $y_j := \vartheta_j\varepsilon^{1/2}R_j+\frac{1}{\varepsilon}Z_j^A$, where $\vartheta_j^\pm, \vartheta_j \in (0,1)$.

From $Z^A = \varepsilon A\mathbf{E}+\varepsilon^3\Psi A^3\mathbf{E}^3+\text{c.c.}$, (3.6), and Sobolev's imbedding theorem we obtain

$$\begin{aligned}\left|d_j^\pm\right|,\ |y_j| &\le \varepsilon^{1/2}\left(|R_{j+1}|+|R_j|+|R_{j-1}|\right)+6\|A(\tau,\cdot)\|_\infty+\varepsilon^2 6|\Psi|\|A(\tau,\cdot)\|_\infty^3\\ &\le \varepsilon^{1/2}3\|\widetilde{R}\|_Y+6C_{\text{Sob}}C_A+\varepsilon^2 6|\Psi|(C_{\text{Sob}}C_A)^3 \quad \text{for all } j\in\mathbb{Z},\ \varepsilon\le\varepsilon_0,\ \varepsilon^2 t\le\tau_0.\end{aligned}$$

Thus, for given $D>0$ there exists a sufficiently small $\varepsilon_0>0$, such that the estimate

$$\left|d_j^{\pm}\right|,\ |y_j| \le 7C_{\mathrm{Sob}}C_A =: \widetilde{C}$$

holds for all $j\in\mathbb{Z}$, $\varepsilon\le\varepsilon_0$, $\varepsilon^2 t\le\tau_0$, and $\|\widetilde{R}\|_Y\le D$.

Now, we use the cubic form of the nonlinearity. Since $\widetilde{V}''(d)=3v_3d^2+\mathcal{O}(d^3)$ and $\widetilde{W}''(y)=3w_3y^2+\mathcal{O}(y^3)$, we can, if necessary, decrease ε_0 further, to obtain

$$|M_j|\le\varepsilon^2\frac{\widehat{C}}{\sqrt{3}}\left(|R_{j+1}|+|R_j|+|R_{j-1}|\right)\quad\text{with}\quad \widehat{C}:=4\sqrt{3}(2v_3+w_3)\widetilde{C}^2$$

for $\varepsilon\le\varepsilon_0$, $\varepsilon^2t\le\tau_0$, $\|\widetilde{R}\|_Y\le D$ and, thus,

$$\|(0,M)\|_Y=\|M\|\le\varepsilon^2\widehat{C}\|\widetilde{R}\|_Y\quad\text{for } \varepsilon\le\varepsilon_0,\ \varepsilon^2t\le\tau_0,\ \|\widetilde{R}\|_Y\le D. \tag{3.12}$$

The semigroup associated to the linear problem $\dot{\widetilde{R}}=\mathcal{L}\widetilde{R}$ is given by $G(t)=\mathrm{e}^{t\mathcal{L}}$. By the variation of constants formula, (3.11) can be transformed into

$$\widetilde{R}(t)=G(t)\widetilde{R}(0)+\int_0^t G(t-s)\left[(0,M(s))-\varepsilon^{-3/2}\widetilde{\rho}^A(s)\right]\mathrm{d}s.$$

From Assumption (3.5) of the theorem and (3.10) it follows

$$\|\widetilde{R}(0)\|_Y\le\varepsilon^{-3/2}\left(\left\|\widetilde{x}(0)-\widetilde{X}^A(0)\right\|_Y+\left\|\widetilde{Y}^A(0)\right\|_Y\right)\le 2d$$

for $\varepsilon\le\varepsilon_0$ and sufficiently small ε_0. Using this estimate and (3.9), (3.12) as well as Proposition 3.1, which gives $\|G(t)\|_{Y\to Y}=1$ for all $t\ge0$, we obtain

$$\|\widetilde{R}(t)\|_Y\le 2d+\varepsilon^2\left(\int_0^t\widehat{C}\|\widetilde{R}(s)\|_Y\mathrm{d}s+tC_\rho\right)\quad\text{for } \varepsilon\le\varepsilon_0,\ \varepsilon^2t\le\tau_0,\ \|\widetilde{R}\|\le D.$$

By Gronwall's inequality, it follows

$$\|\widetilde{R}(t)\|_Y\le\left(2d+\varepsilon^2tC_\rho\right)\mathrm{e}^{\varepsilon^2t\widehat{C}}\quad\text{for } t\le\tau_0/\varepsilon^2 \text{ with } \varepsilon\le\varepsilon_0.$$

Of course, this estimate is only valid as long as $\|\widetilde{R}(t)\|_Y\le D$. Hence, we now choose $D=(2d+\tau_0C_\rho)\mathrm{e}^{\tau_0\widehat{C}}$. Then, decreasing $\varepsilon_0>0$ sufficiently in the manner leading to (3.12), it follows that the Gronwall estimate holds for all $t\in[0,\tau_0/\varepsilon^2]$ with $\varepsilon\le\varepsilon_0$. This estimate together with (3.10) proves the desired result, since $\|\widetilde{x}(t)-\widetilde{X}^A(t)\|_Y\le\varepsilon^{3/2}\|\widetilde{R}(t)\|_Y+\|\widetilde{Y}^A(t)\|_Y$. □

In the above proof we used the following result, which is a sharpened version of the first estimate in [SW00, Lemma 3.9].

Proposition 3.3 *For $\phi\in\mathrm{H}^1(\mathbb{R})$, $\varepsilon\in(0,1)$, and $c\in\mathbb{R}$ we have the estimate*

$$\sum_{j\in\mathbb{Z}}\sup_{|s|\le1}\left|\phi\left(\varepsilon(j+c+s)\right)\right|^2\le\frac{8}{\varepsilon}\|\phi\|^2_{\mathrm{H}^1(\mathbb{R})}.$$

Proof: Let $\phi \in \mathrm{H}^1(\mathbb{R})$, $j \in \mathbb{Z}$, and $x, x' \in (j{+}c{-}1, j{+}c{+}1)$. From the fundamental theorem of calculus we obtain

$$|\phi(x)| \leq |\phi(x')| + \int_{j+c-1}^{j+c+1} |\phi'(\xi)| \, \mathrm{d}\xi.$$

Integration over x', the estimate $(a{+}b)^2 \leq 2(a^2{+}b^2)$, and Cauchy–Schwarz inequality yield

$$\begin{aligned} |\phi(x)| &\leq \textstyle\int_{j+c-1}^{j+c+1} \left(|\phi(\xi)|{+}|\phi'(\xi)|\right) \mathrm{d}\xi \leq \sqrt{2} \int_{j+c-1}^{j+c+1} \left(|\phi(\xi)|^2{+}|\phi'(\xi)|^2\right)^{1/2} \mathrm{d}\xi \\ &\leq 2 \left(\textstyle\int_{j+c-1}^{j+c+1} \left(|\phi(\xi)|^2{+}|\phi'(\xi)|^2\right) \mathrm{d}\xi \right)^{1/2}, \end{aligned}$$

and, hence, $\sup_{|s|\leq 1} \left|\phi\left(\varepsilon(j{+}c{+}s)\right)\right|^2 \leq 4 \left\|\phi(\varepsilon \cdot)\right\|^2_{H^1((j+c-1, j+c+1))}$. Summing over $j \in \mathbb{Z}$, we obtain

$$\sum_{j\in\mathbb{Z}} \sup_{|s|\leq 1} \left|\phi\left(\varepsilon(j{+}c{+}s)\right)\right|^2 \leq 8 \|\phi(\varepsilon \cdot)\|^2_{\mathrm{H}^1(\mathbb{R})}.$$

The substitution $\xi = \varepsilon x$ yields

$$\|\phi(\varepsilon\, \cdot)\|^2_{\mathrm{H}^1(\mathbb{R})} = \int_{x\in\mathbb{R}} \left(|\phi(\varepsilon x)|^2 + \left|\tfrac{\mathrm{d}}{\mathrm{d}x}\phi(\varepsilon x)\right|^2\right) \mathrm{d}x = \tfrac{1}{\varepsilon} \int_{\xi\in\mathbb{R}} \left(|\phi(\xi)|^2 + \varepsilon^2 \left|\tfrac{\mathrm{d}}{\mathrm{d}\xi}\phi(\xi)\right|^2\right) \mathrm{d}\xi \leq \tfrac{1}{\varepsilon} \|\phi\|^2_{\mathrm{H}^1(\mathbb{R})}$$

for $\varepsilon \in (0, 1)$, which is the desired estimate. $\square$

4 The periodic case

The modulation theory can be also applied to finite chains if the number of atoms is sufficiently large. We impose periodicity conditions for the discrete system and obtain a nonlinear Schrödinger equation with generalized periodicity conditions. We follow here the analogous approach of [MSZ00], where modulations in the Swift-Hohenberg equation were described via a Ginzburg-Landau equation.

We denote by $(4.1)_m$ the oscillator chain

$$\ddot{x}_j = V'(\partial_j^+ x) - V'(\partial_j^- x) - W'(x_j), \quad j \in \mathbb{Z}_m, \tag{4.1}$$

where $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ is the cyclic group with m elements. We use exactly the same ansatz X^A as in (3.1), namely

$$X_j^A(t) = \varepsilon A(\tau, \xi)\mathbf{E}(t, j) + \text{c.c.} \quad \text{with } \tau = \varepsilon^2 t,\ \xi = \varepsilon(j{-}ct). \tag{4.2}$$

However, to obtain periodicity in j we need $A(\tau, \xi{+}\varepsilon m)\mathrm{e}^{\mathrm{i}\vartheta m} = A(\tau, \xi)$. Thus, we pose NLSE on the interval $(0, \ell)$ with a generalized boundary condition:

$$\left.\begin{array}{l} \mathrm{i}\partial_\tau A = \frac{1}{2}\omega''\partial_\xi^2 A + \rho|A|^2 A, \\ A(\tau, \xi{+}\ell)\mathrm{e}^{\mathrm{i}\Theta} = A(\tau, \xi) \end{array}\right\} \text{ for } \xi \in \mathbb{R} \text{ and } \tau \in [0, \tau_0] \tag{4.3}$$

with $\rho = -\frac{3\left(v_3 c_1^2 + w_3\right)}{2\omega}$, where $c_1 = 2(1{-}\cos\vartheta)$. To indicate the parameters we write $(4.3)_{\ell,\Theta}$ for the NLSE with generalized periodicity $\mathrm{e}^{\mathrm{i}\Theta}$ on the interval $(0, \ell)$. Of course, it suffices to solve NLSE on $(0,\ell)$ with the boundary conditions $A(\tau,\ell)\mathrm{e}^{\mathrm{i}\Theta} = A(\tau,0)$ and $\partial_\xi A(\tau,\ell)\mathrm{e}^{\mathrm{i}\Theta} = \partial_\xi A(\tau,0)$.

To make $(4.1)_m$ and $(4.3)_{\ell,\Theta}$ compatible via the ansatz (4.2) we need to have

$$\ell = \varepsilon m \text{ and } \Theta = \vartheta m \bmod 2\pi. \tag{4.4}$$

For given ϑ and Θ the relation (4.4) has infinitely many solutions (m,ε) if and only if ϑ and Θ are rational multiples of 2π and there exists $m_0 \in \mathbb{N}$ with $\Theta = \vartheta m_0 \bmod 2\pi$.

Theorem 4.1 *Assume $\ell > 0$ and $\Theta,\vartheta \in 2\pi\mathbb{Q}\cap\mathbb{S}^1$, such that (4.4) has a solution $(m,\varepsilon) \in \mathbb{N}\times(0,\infty)$. Moreover, let $A \in \mathrm{C}([0,\tau_0],\mathrm{H}^5_{\mathrm{loc}}(\mathbb{R},\mathbb{C}))$ solve NLSE $(4.3)_{\ell,\Theta}$.*

Then, for each $d > 0$ there exist $\varepsilon_0 > 0$ and $C > 0$, such that the following holds: If (m,ε) solves (4.4) with $\varepsilon \in (0,\varepsilon_0)$ and if x is a solution of $(4.1)_m$ whose initial datum $\widetilde{x}(0) = (x(0),\dot{x}(0)) \in \mathbb{R}^m\times\mathbb{R}^m$ satisfies $\|\widetilde{x}(0) - \widetilde{X}^A(0)\|_{\mathbb{R}^m\times\mathbb{R}^m} \le d\varepsilon^{3/2}$, then x satisfies $\|\widetilde{x}(t) - \widetilde{X}^A(t)\|_{\mathbb{R}^m\times\mathbb{R}^m} \le C\varepsilon^{3/2}$ for $t \in [0,\tau_0/\varepsilon^2]$.

The proof of this result is identical to the one on the infinite chain. We just have to replace sums over $\mathbb{Z}$ by sums over $\mathbb{Z}_m$ and integrals over $\mathbb{R}$ by integrals over $\mathbb{R}/\ell\mathbb{Z}$.

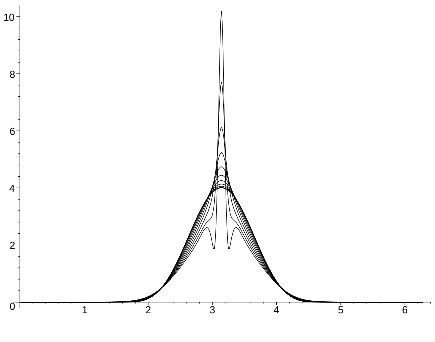
By using classical perturbation analysis for NLSE we may generalize the result to the case that the solutions of $(4.1)_m$ are compared with X^{A_m}, where A_m solves $(4.3)_{\ell,\Theta_m}$ with

$$|\Theta_m - \Theta| \le d\varepsilon = \frac{d\ell}{m}, \quad \|A_m(0)\|_{\mathrm{H}^5(\mathbb{R})} \le C \quad \text{and} \quad \|A_m(0) - A(0)\|_{\mathrm{H}^5(\mathbb{R})} \le d\varepsilon = \frac{d\ell}{m}.$$

Finally, we illustrate the result by a numerical example. This example shows that reasonable approximation properties can be expected for sufficiently small ε. To make the numerics as simple as possible we have chosen $v_1 = w_1 = 1$, $\widetilde{V} \equiv 0$, $\widetilde{W}(x_j) = \frac{1}{4}x_j^4$ and $\ell = 2\pi$, $\vartheta = \frac{\pi}{2}$. The associated NLSE $(4.3)_{2\pi,0}$ reads

$$\begin{aligned} &\mathrm{i}\partial_\tau A = -\tfrac{1}{6\sqrt{3}}\partial_\xi^2 A - \tfrac{3}{2\sqrt{3}}|A|^2 A, \\ &A(\tau,\xi+2\pi) = A(\tau,\xi), \quad A(0,\xi) = \begin{cases} [1+\cos(\xi-\pi)]^2 & \text{for } |\xi-\pi| \le \frac{\pi}{2}, \\ 0 & \text{else.} \end{cases} \end{aligned} \tag{4.5}$$

We solved this problem numerically for $\tau \in [0,\tau_0]$ with $\tau_0 = 0.25$, see Figure 2.



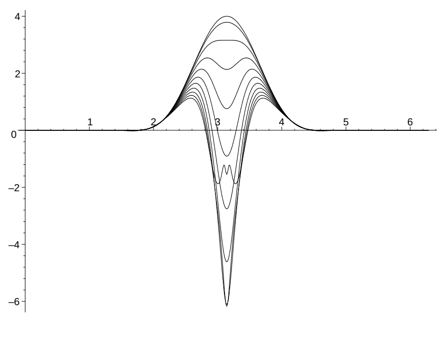


Figure 2: Solution of NLSE (4.5): $|A(\kappa\tau_0,\cdot)|$ (left) and $\mathrm{Re}\,A(\kappa\tau_0,\cdot)$ (right), $\kappa = 0.1,\ldots,1$.

Since $\Theta = 0$, any $m \in 4\mathbb{N}$ and $\varepsilon = 2\pi/m$ satisfy (4.4). We solved $(4.1)_m$ for several m from 100 to 4000 with the initial condition obtained from (4.2) and $A(0,\cdot)$ from (4.5). To compare the discrete solutions with NLSE we reconstructed $|A(\tau,\cdot)|$ and $\operatorname{Re} A(\tau,\cdot)$ via the formulae

$$\begin{aligned}
|A(\varepsilon^2 t, \varepsilon(j+\omega' t))| &= \frac{1}{2\varepsilon}\left[x_j(t)^2 + \frac{1}{\omega^2}\dot{x}_j(t)^2\right]^{1/2}, \\
\operatorname{Re} A(\varepsilon^2 t, \varepsilon(j+\omega' t)) &= \frac{1}{2\varepsilon}\left[x_j(t)\cos(\omega t+\vartheta j) - \dot{x}_j(t)\frac{\sin(\omega t+\vartheta j)}{\omega}\right].
\end{aligned}$$

These functions are plotted in Figure 3 for different $\tau \in [0, \tau_0]$ and $m = 4000$.

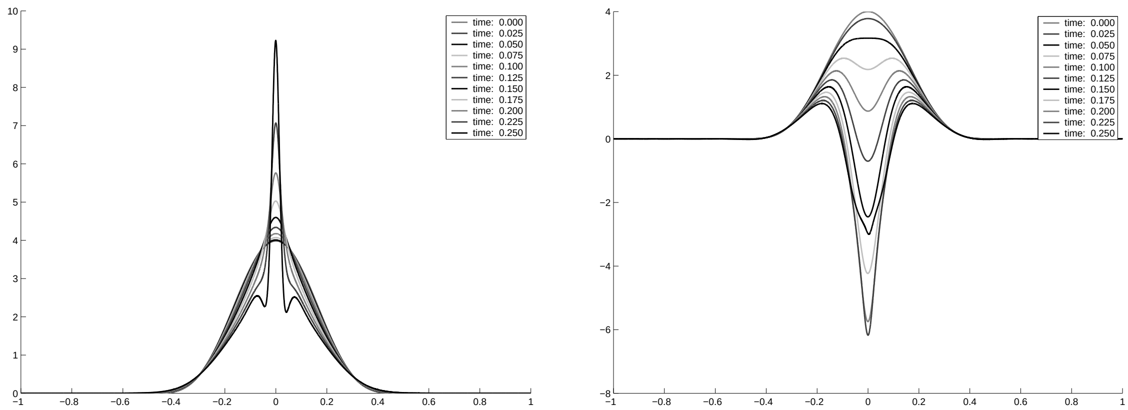


Figure 3: $|A|$ (left) and $\operatorname{Re} A$ (right) from the solution of $(4.1)_m$ with $m = 4000$.

Finally, in Figure 4 we compare the solutions at the final macroscopic time $\tau = \tau_0$ for different values of m. Note that the initial pulse $A(0,\cdot)$ has a symmetric shape. However, in the discrete system $(4.1)_m$ the pulse travels with microscopic speed $c = -\omega' = -1/\sqrt{3}$ to the left. This certainly breaks the symmetry. Figure 4 shows clearly that the symmetry is broken and that the unsymmetry disappears for $m \to \infty$.

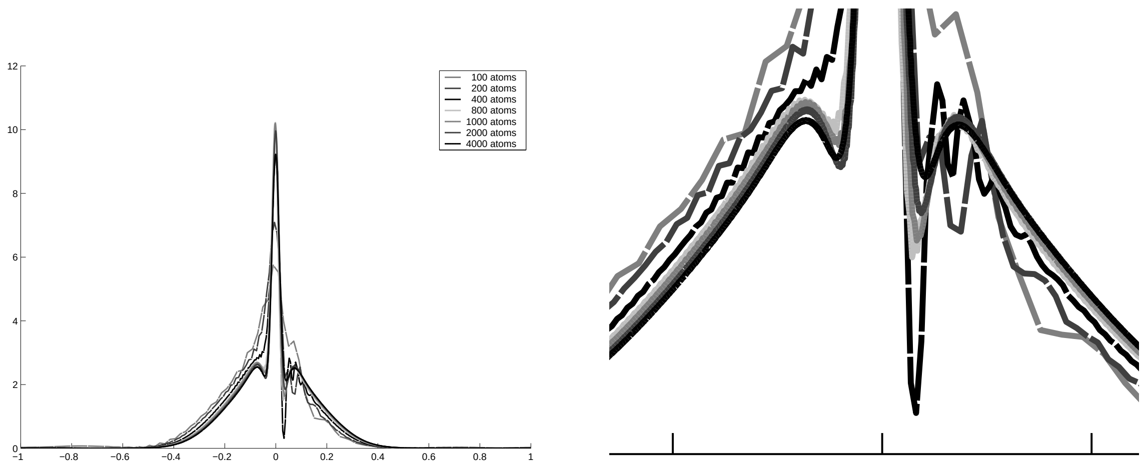


Figure 4: Comparison $|A(0.25,\cdot)|$ for $m = 100, \ldots, 4000$ (to the right: magnification).

It should be noted that the numerical effort for the calculation of $A(\tau_0,\cdot)$ from the discrete system grows like m^3: On the one hand the size of the system is proportional to m. Moreover, the time steps can be chosen independently of m, since the right hand sides are uniformly bounded (in fact, because of $\varepsilon = \ell/m$ the solutions are smaller and smaller). On the other hand the macroscopic time is $\tau = \varepsilon^2 t$. To reach τ_0 we need to integrate the microscopic time from 0 to $m^2\tau_0/\ell^2$ for each of the m atoms. (In comparison to this, the numerical effort for the calculation of $A(\tau_0, \cdot)$ from the NLSE (4.5) is obviously independent of m.) During this time the pulse travels $m\tau_0|c|/\ell^2$ times around $\mathbb{Z}_m$. For $m = 4000$, $\tau_0 = 0.25$, $c = -1/\sqrt{3}$ and $\ell = 2\pi$ this means that the pulse travels around $\mathbb{Z}_{4000}$ more than 14 times!

Acknowledgments: This research was supported by the DFG Priority Programme 1095 "*Analysis, Modeling and Simulation of Multiscale Problems*" through grant *Mi 459/3–2*. We thank David Remy for doing the numerical computations.

References

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

[DH02] W. Dreyer and M. Herrmann. Towards rigorous micro–macro transitions: the microscopic oscillator motion. AMS MSP Preprint 51, March 2002.

[DK00] 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.

[FM02] G. Friesecke and K. Matthies. Atomic-scale localization of high-energy solitary waves on lattices. *Physica D*, 171(4), 211–220, 2002.

[FP99] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices: I. Qualitative properties, renormalization and continuum limit. *Nonlinearity*, 12, 1601–1627, 1999.

[FP02] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices: II. Linear implies nonlinear stability. *Nonlinearity*, 15, 1343–1359, 2002.

[FV99] A.-M. Filip and S. Venakides. Existence and modulation of traveling waves in particle chains. *Comm. Pure Appl. Math.*, 52(6), 693–735, 1999.

[FW94] G. Friesecke and J. A. D. Wattis. Existence theorem for solitary waves on lattices. *Comm. Math. Phys.*, 161(2), 391–418, 1994.

[GM03] J. Giannoulis and A. Mielke. In preparation, 2003.

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

[IK00] 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.

[MA94] 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.

[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.

[Mie03] A. MIELKE. Macroscopic behavior of microscopic oscillations in harmonic lattices. In preparation, 2003.

[MSZ00] A. MIELKE, G. SCHNEIDER, and A. ZIEGRA. Comparison of inertial manifolds and application to modulated systems. *Math. Nachr.*, 214, 53–69, 2000.

[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.

[SW00] 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.