Nonlinearity **17** (2004) 551–565 PII: S0951-7715(04)67016-2

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, Germany

E-mail: giannoulis@mathematik.uni-stuttgart.de and mielke@mathematik.uni-stuttgart.de

Received 1 August 2003, in final form 21 November 2003
Published 19 December 2003
Online at stacks.iop.org/Non/17/551 (DOI: 10.1088/0951-7715/17/2/011)

Recommended by E S Titi

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; i.e. we show that solutions that 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.

Mathematics Subject Classification: 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 100 years ago, but many problems remain open even today. Most prominent is the question of how to obtain irreversible thermodynamics as a macroscopic limit from microscopic models that are reversible (Hamiltonian). For a survey of the methods and results of the mathematical justification of nonequilibrium statistical mechanics, we refer to, e.g., [Spo91, Bol96].

In this paper, we consider another part of this field that is far from thermodynamic fluctuations. We are interested in reversible, macroscopic limits of atomistic models that 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 such as travelling 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 1960s 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 much larger spatial and temporal scales, such that the modulation equation is a macroscopic equation.

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

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

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-neighbour interaction, and W is an external potential that might arise through the 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, proposition 7.1] and [SW00], where $W \equiv 0$ and solutions of the form

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

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-wavelength limit.

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

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

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

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

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

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

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

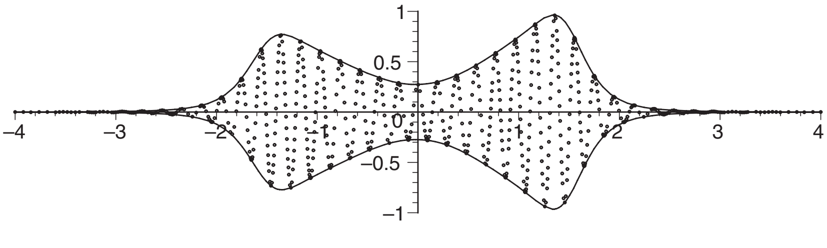

Figure 1. A modulated initial datum $(x_j(0))_{j\in\mathbb{Z}}$ $(\cdots\cdots)$ together with the envelopes $\pm 2|A(0, \cdot)|$.

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)

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

where ρ can be calculated explicitly. Here $\tau = \varepsilon^2 t$ is the macroscopic time variable 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 $t \mapsto (x_j(t))_{j\in\mathcal{J}}$ that 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

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

satisfies the estimate

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

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 [GM04]. In [Kal89, proposition 7.2], results are stated without proof, providing estimates such as (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 Galilean 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 that 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 that are close to the linear one. Only the extremely long timescales allow for the accumulation of the nonlinear effects which are inherent to NLSE.

On shorter timescales, 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 [Mie04], where Wigner measures are used to describe the energy transport in multidimensional lattices. For larger nonlinear microscopic oscillations we refer to [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 modelled by

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

where $x_j = x_j(t) \in \mathbb{R}$, $t \geqslant 0$, $j \in \mathbb{Z}$ and $\partial_j^{\pm} x := \pm(x_{j\pm1} - x_j)$ (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

$$V(d) = \frac{v_1}{2} d^2 + \tilde{V}(d), \qquad W(y) = \frac{w_1}{2} y^2 + \tilde{W}(y)$$

with

$$\tilde{V}(d) = \frac{v_2}{3} d^3 + \frac{v_3}{4} d^4 + \mathrm{O}(d^5) \qquad \text{and} \qquad \tilde{W}(y) = \frac{w_2}{3} y^3 + \frac{w_3}{4} y^4 + \mathrm{O}(y^5).$$

The linear part of (2.1) reads

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

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

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

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 $\omega, \omega', \omega''$ to denote $\omega(\vartheta), \omega'(\vartheta), \omega''(\vartheta)$, respectively. The associated basic mode, $\boldsymbol{E}(t, j) := \mathrm{e}^{\mathrm{i}(\omega t + \vartheta j)}$, 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:

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

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. $\boldsymbol{E}^n$). Hence, we insert the multiple scale ansatz

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

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 $n \in \mathbb{Z}$, $|n| \leqslant k$. It will be sufficient to consider only terms with $k \leqslant 3$, but it is instructive to keep the full generality.

For the left-hand side we obtain

$$\begin{aligned}\ddot{X}_j^{(A)} = \sum_{k,n} \varepsilon^k [&-(n\omega)^2 A_{k,n} - 2\varepsilon n\mathrm{i}\omega c \partial_\xi A_{k,n} + \varepsilon^2(c^2 \partial_\xi^2 A_{k,n} + 2n\mathrm{i}\omega \partial_\tau A_{k,n}) \\ &-2\varepsilon^3 c \partial_\xi \partial_\tau A_{k,n} + \varepsilon^4 \partial_\tau^2 A_{k,n}]\boldsymbol{E}^n,\end{aligned} \tag{2.4}$$

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

With

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

the linear part of the right-hand side reads

$$\begin{aligned} L_j X^{(A)} &= v_1(X_{j+1}^{(A)} - 2X_j^{(A)} + X_{j-1}^{(A)}) - w_1 X_j^{(A)} \\ &= \sum_{k,n} \varepsilon^k \{v_1[A_{k,n}(\tau, \xi + \varepsilon)\mathrm{e}^{\mathrm{i}n\vartheta} - 2A_{k,n} + A_{k,n}(\tau, \xi - \varepsilon)\mathrm{e}^{-\mathrm{i}n\vartheta}] - w_1 A_{k,n}\}\boldsymbol{E}^n. \end{aligned}$$

From the expansion

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

we obtain

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

with $r_{k,n}=(v_1/6)[\mathrm{e}^{\mathrm{i}n\vartheta}\partial_\xi^3A_{k,n}(\tau,\xi+\theta_{k,n}^{+}\varepsilon)-\mathrm{e}^{-\mathrm{i}n\vartheta}\partial_\xi^3A_{k,n}(\tau,\xi-\theta_{k,n}^{-}\varepsilon)]$. Here we used the fact 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 $\varepsilon^k\boldsymbol{E}^n$ of the linear part of the right-hand side of the microscopic equation (2.1) with $x=X^{(A)}$. It remains for us to obtain a similar expansion for the nonlinear part, $N_j(X^{(A)}):=\tilde V'(\partial_j^+X^{(A)})-\tilde V'(\partial_j^-X^{(A)})-\tilde W'(X_j^{(A)})$. We start by deriving an expansion only in terms of ε^k. At first we note

$$N_j(X^{(A)})=v_2[(\partial_j^+X^{(A)})^2-(\partial_j^-X^{(A)})^2]-w_2(X_j^{(A)})^2+v_3[(\partial_j^+X^{(A)})^3-(\partial_j^-X^{(A)})^3]$$
$$-w_3(X_j^{(A)})^3+\hat V'(\partial_j^+X^{(A)})-\hat V'(\partial_j^-X^{(A)})-\hat W'(X_j^{(A)})$$

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

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

respectively, where

$$a_1^{\pm}=\pm(\mathrm{e}^{\pm\mathrm{i}\vartheta}-1)A_{1,1}\boldsymbol{E}+\text{c.c.},\qquad b_1^{\pm}=\partial_\xi A_{1,0}+(\mathrm{e}^{\pm\mathrm{i}\vartheta}\partial_\xi A_{1,1}\boldsymbol{E}+\text{c.c.}),$$

$$a_2^{\pm}=\pm(\mathrm{e}^{\pm\mathrm{i}\vartheta}-1)A_{2,1}\boldsymbol{E}\pm(\mathrm{e}^{\pm2\mathrm{i}\vartheta}-1)A_{2,2}\boldsymbol{E}^2+\text{c.c.},$$

$$r_1^{\pm}=\pm\sum_{n=-1}^{1}\mathrm{e}^{\pm\mathrm{i}n\vartheta}\frac{1}{2}\partial_\xi^2A_{1,n}(\tau,\xi\pm\hat\theta_{1,n}^{\pm}\varepsilon)\boldsymbol{E}^n+\sum_{n=-2}^{2}\mathrm{e}^{\pm\mathrm{i}n\vartheta}\partial_\xi A_{2,n}(\tau,\xi\pm\tilde\theta_{1,n}^{\pm}\varepsilon)\boldsymbol{E}^n$$
$$\pm\sum_{k\geqslant3}\varepsilon^{k-3}\sum_{n=-k}^{k}[A_{k,n}(\tau,\xi\pm\varepsilon)\mathrm{e}^{\pm\mathrm{i}n\vartheta}-A_{k,n}]\boldsymbol{E}^n\qquad\text{with } \hat\theta_{1,n}^{\pm},\tilde\theta_{1,n}^{\pm}\in(0,1),$$

$$a_1=A_{1,0}+(A_{1,1}\boldsymbol{E}+\text{c.c.}),\qquad a_2=A_{2,0}+(A_{2,1}\boldsymbol{E}+A_{2,2}\boldsymbol{E}^2+\text{c.c.}),$$

$$r_1=\sum_{k\geqslant3}\varepsilon^{k-3}\sum_{n=-k}^{k}A_{k,n}\boldsymbol{E}^n.$$

Insertion into the nonlinearity gives

$$N_j(X^{(A)})=\varepsilon^2\{v_2[(a_1^+)^2-(a_1^-)^2]-w_2a_1^2\}+\varepsilon^3\{2v_2[a_1^+(b_1^++a_2^+)-a_1^-(b_1^-+a_2^-)]$$
$$+v_3[(a_1^+)^3-(a_1^-)^3]-2w_2a_1a_2-w_3a_1^3\}+\varepsilon^4(r_2^+-r_2^--r_2)$$
$$+\hat V'(\partial_j^+X^{(A)})-\hat V'(\partial_j^-X^{(A)})-\hat W'(X_j^{(A)}) \tag{2.7}$$

with

$$r_2^{\pm}=2v_2a_1^{\pm}r_1^{\pm}+3v_3(a_1^{\pm})^2(b_1^{\pm}+a_2^{\pm}+\varepsilon r_1^{\pm})+(v_2+3v_3\varepsilon a_1^{\pm})(b_1^{\pm}+a_2^{\pm}+\varepsilon r_1^{\pm})^2$$
$$+v_3\varepsilon^2(b_1^{\pm}+a_2^{\pm}+\varepsilon r_1^{\pm})^3,$$

$$r_2=2w_2a_1r_1+3w_3a_1^2(a_2+\varepsilon r_1)+(w_2+3w_3\varepsilon a_1)(a_2+\varepsilon r_1)^2+w_3\varepsilon^2(a_2+\varepsilon r_1)^3,$$

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

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

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

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 the fact that the nonlinearity generates only terms of the power $k \geqslant 2$. Thus, we obtain from (2.4), (2.6) and (2.7) for $k = 1$ and $n = 0, 1$

for $\varepsilon^1 \boldsymbol{E}^0$: $\quad 0 = -w_1 A_{1,0}$;
for $\varepsilon^1 \boldsymbol{E}^1$: $\quad -\omega^2 A_{1,1} = -\omega^2 A_{1,1}$.

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

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

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 $\varepsilon^2 \boldsymbol{E}^n$, $n = 0, 1, 2$,

for $\varepsilon^2 \boldsymbol{E}^0$: $\quad 0 = -w_1 A_{2,0} - 2w_2|A_{1,1}|^2$;
for $\varepsilon^2 \boldsymbol{E}^1$: $\quad -\omega^2 A_{2,1} - 2\mathrm{i}\omega c \partial_\xi A_{1,1} = -\omega^2 A_{2,1} + 2\mathrm{i}\omega\omega' \partial_\xi A_{1,1}$;
for $\varepsilon^2 \boldsymbol{E}^2$: $\quad -4\omega^2 A_{2,2} = -\omega(2\vartheta)^2 A_{2,2} - (v_2 s_1 c_1 + w_2)A_{1,1}^2$.

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

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

The equation for $\varepsilon^2 \boldsymbol{E}^1$ 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 $\varepsilon^2 \boldsymbol{E}^2$ 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 $\varepsilon^3 \boldsymbol{E}^n$ for $n = 0, 1, 2, 3$, we obtain

for $\varepsilon^3 \boldsymbol{E}^0$: $\quad 0 = -w_1 A_{3,0} + (2v_2 c_1 \bar{A}_{1,1}\partial_\xi A_{1,1} - 2w_2 \bar{A}_{1,1}A_{2,1} + \text{c.c.})$;

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

for $\varepsilon^3 \boldsymbol{E}^2$:
$$-4\omega^2 A_{3,2} - 4\mathrm{i}\omega c \partial_\xi A_{2,2} = -\omega(2\vartheta)^2 A_{3,2} + 2\mathrm{i}\omega(2\vartheta)\omega'(2\vartheta)\partial_\xi A_{2,2} + 2v_2 c_1(c_1 - 3)A_{1,1}\partial_\xi A_{1,1} - 2(v_2 s_1 c_1 + w_2)A_{1,1}A_{2,1};$$

for $\varepsilon^3 \boldsymbol{E}^3$:
$$-9\omega^2 A_{3,3} = -\omega(3\vartheta)^2 A_{3,3} + 2[v_2 s_1(c_1 + s_1^2) - w_2]A_{1,1}A_{2,2} + [v_3 c_1^2(3 - c_1) - w_3]A_{1,1}^3.$$

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

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

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

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

where $\omega(3\vartheta)^2 - 9\omega^2 = -8w_1 + v_1 c_1^2(c_1 - 6) = -8[w_1 + v_1(2 + \cos^3\vartheta - 3\cos\vartheta)] < 0$ 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 $\varepsilon^3 \boldsymbol{E}^0$ and $\varepsilon^3 \boldsymbol{E}^2$, 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*

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

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 that start in this form will maintain it on suitably long timescales.

Let us now consider the case where the nonlinearity in our oscillator chain model (2.1), $N_j(x) = \tilde{V}'(\partial_j^+ x) - \tilde{V}'(\partial_j^- x) - \tilde{W}'(x_j)$, 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 $\varepsilon^3 \boldsymbol{E}^0$ and $\varepsilon^3 \boldsymbol{E}^2$ we obtain (or can set) $A_{3,0} = A_{3,2} = 0$. The NLSE (2.12) reads

$$\mathrm{i}\partial_\tau A_{1,1} = \frac{1}{2}\omega''\partial_\xi^2 A_{1,1} + \rho|A_{1,1}|^2 A_{1,1} \qquad \text{with } \rho := -\frac{3(v_3 c_1^2 + w_3)}{2\omega}, \tag{2.14}$$

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

$$A_{3,3} = \Psi A_{1,1}^3 \qquad \text{with } \Psi := \frac{v_3 c_1^2(3 - c_1) - w_3}{\omega(3\vartheta)^2 - 9\omega^2}. \tag{2.15}$$

Up to this stage there have been 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 $k \geqslant 4$, $|n| \leqslant k$. Hence, setting deliberately also $A_{3,1} = A_{2,1} = 0$ and $A_{k,\cdot} = 0$ for $k \geqslant 4$, the general multiple scale ansatz, $X_j^{(A)}(t)$, in (2.3) obtains the special form

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

with $\tau = \varepsilon^2 t$, $\xi = \varepsilon(j + \omega' t)$ and leads to the 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)

$$\mathrm{i}\partial_\tau A = \frac{1}{2}\omega''\partial_\xi^2 A + \rho|A|^2 A \qquad \text{with } \rho = -\frac{3(v_3 c_1^2 + w_3)}{2\omega}, \quad c_1 = 2(1-\cos\vartheta)$$

then

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

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 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 $j \in \mathcal{J}$ as 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

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

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

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

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

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

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

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

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

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

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

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

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\tilde{x}(t)\|_Y^2 = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{j\in\mathbb{Z}}[\dot{x}_j^2 + v_1(x_{j+1}-x_j)^2 + w_1 x_j^2]$$
$$= 2\sum_{j\in\mathbb{Z}}\dot{x}_j[\ddot{x}_j - v_1(x_{j+1}-2x_j+x_{j-1}) + w_1 x_j] = 2\sum_{j\in\mathbb{Z}}\dot{x}_j\cdot 0$$

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

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* $V(d) = (v_1/2)d^2 + \mathrm{O}(d^4)$ *and* $W(y) = (w_1/2)y^2 + \mathrm{O}(y^4)$ *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 $\tilde{x}$ *of (3.2) with an initial condition* $\tilde{x}(0)$ *satisfying*

$$\|\tilde{x}(0) - \tilde{X}^A(0)\|_Y \leqslant d\varepsilon^{3/2} \tag{3.5}$$

fulfils the estimate

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

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

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

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

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

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 $k \geqslant 4$, $n = -k, \ldots, k$ and $A_{1,1} = A$, $A_{3,3} = \Psi A^3$. Hence, proceeding as 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 $\varepsilon^k\boldsymbol{E}^n$ of (2.8) with Z^A instead of $X^{(A)}$ the expansion

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

with

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

and

$$\begin{aligned}
a_1 &= A\boldsymbol{E} + \text{c.c.}, & r_1 &= \Psi A^3\boldsymbol{E}^3 + \text{c.c.},\\
a_1^\pm &= \pm(\mathrm{e}^{\pm\mathrm{i}\vartheta} - 1)A\boldsymbol{E} + \text{c.c.}, & b_1^\pm &= \mathrm{e}^{\pm\mathrm{i}\vartheta}\partial_\xi A\boldsymbol{E} + \text{c.c.},
\end{aligned}$$

$$r_1^\pm = \pm\mathrm{e}^{\pm\mathrm{i}\vartheta}\tfrac{1}{2}\partial_\xi^2 A(\tau, \xi \pm \hat{\theta}_{1,1}^\pm\varepsilon)\boldsymbol{E} + \Psi[\pm(\mathrm{e}^{\pm\mathrm{i}3\vartheta} - 1)A^3 + \varepsilon\mathrm{e}^{\pm\mathrm{i}3\vartheta}\partial_\xi A^3(\tau, \xi \pm \tilde{\theta}_{3,3}^\pm\varepsilon)]\boldsymbol{E}^3 + \text{c.c.},$$

where $\theta_{1,1}^\pm, \tilde{\theta}_{3,3}^\pm, \hat{\theta}_{1,1}^\pm \in (0, 1)$. In accordance with 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

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

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

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

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

$$\|\tilde{Y}^A(t)\|_Y \leqslant \varepsilon^{5/2} |\Psi| C_2 C_{\mathrm{Sob}} C_A^3 \qquad \text{for } \varepsilon \leqslant \varepsilon_0 \quad \text{and} \quad t \leqslant \frac{\tau_0}{\varepsilon^2} \tag{3.10}$$

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

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

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

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

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

From the mean value theorem, we obtain

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

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

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

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

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

$$|d_j^\pm|, |y_j| \leqslant 7C_{\mathrm{Sob}} C_A =: \tilde{C}$$

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

Now, we use the cubic form of the nonlinearity. Since $\tilde{V}''(d) = 3v_3 d^2 + \mathrm{O}(d^3)$ and $\tilde{W}''(y) = 3w_3 y^2 + \mathrm{O}(y^3)$, we can, if necessary, decrease ε_0 further, to obtain

$$|M_j| \leqslant \varepsilon^2 \frac{\hat{C}}{\sqrt{3}}(|R_{j+1}| + |R_j| + |R_{j-1}|) \qquad \text{with } \hat{C} := 4\sqrt{3}(2v_3 + w_3)\tilde{C}^2$$

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

$$\|(0, M)\|_Y = \|M\| \leqslant \varepsilon^2 \hat{C} \|\tilde{R}\|_Y \qquad \text{for } \varepsilon \leqslant \varepsilon_0, \quad \varepsilon^2 t \leqslant \tau_0, \quad \|\tilde{R}\|_Y \leqslant D. \tag{3.12}$$

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

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

From assumption (3.5) of theorem 3.2 and (3.10), it follows that

$$\|\tilde{R}(0)\|_Y \leqslant \varepsilon^{-3/2}(\|\tilde{x}(0) - \tilde{X}^A(0)\|_Y + \|\tilde{Y}^A(0)\|_Y) \leqslant 2d$$

for $\varepsilon \leqslant \varepsilon_0$ 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 $t \geqslant 0$, we obtain

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

By Gronwall's inequality, it follows that

$$\|\tilde{R}(t)\|_Y \leqslant (2d + \varepsilon^2 tC_\rho)\mathrm{e}^{\varepsilon^2 t\hat{C}} \qquad \text{for } t \leqslant \frac{\tau_0}{\varepsilon^2} \quad \text{with } \varepsilon \leqslant \varepsilon_0.$$

Of course, this estimate is only valid as long as $\|\tilde{R}(t)\|_Y \leqslant D$. Hence, we now choose $D = (2d + \tau_0 C_\rho)\mathrm{e}^{\tau_0\hat{C}}$. 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 $\varepsilon \leqslant \varepsilon_0$. This estimate together with (3.10) proves the desired result since $\|\tilde{x}(t) - \tilde{X}^A(t)\|_Y \leqslant \varepsilon^{3/2}\|\tilde{R}(t)\|_Y + \|\tilde{Y}^A(t)\|_Y$. □

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*

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

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

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

Integrating over x', the estimate $(a+b)^2 \leqslant 2(a^2+b^2)$ and the Cauchy–Schwarz inequality yield

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

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

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

The substitution $\xi = \varepsilon x$ yields

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

$$= \frac{1}{\varepsilon} \int_{\xi\in\mathbb{R}} \left(|\phi(\xi)|^2 + \varepsilon^2 \left| \frac{\mathrm{d}}{\mathrm{d}\xi}\phi(\xi) \right|^2 \right) \mathrm{d}\xi \leqslant \frac{1}{\varepsilon} \|\phi\|^2_{\mathrm{H}^1(\mathbb{R})}$$

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

4. The periodic case

The modulation theory can also be applied to finite chains if the number of atoms is sufficiently large. We impose periodicity conditions for the discrete system and obtain an NLSE 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

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

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

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

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:

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

with $\rho = -3(v_3 c_1^2 + w_3)/2\omega$, 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

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

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, $\tilde{x}(0) = (x(0), \dot{x}(0)) \in \mathbb{R}^m \times \mathbb{R}^m$, satisfies $\|\tilde{x}(0) - \tilde{X}^A(0)\|_{\mathbb{R}^m\times\mathbb{R}^m} \leqslant d\varepsilon^{3/2}$, then x satisfies $\|\tilde{x}(t) - \tilde{X}^A(t)\|_{\mathbb{R}^m\times\mathbb{R}^m} \leqslant C\varepsilon^{3/2}$ 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 where the solutions of $(4.1)_m$ are compared with X^{A_m}, where A_m solves $(4.3)_{\ell,\Theta_m}$ with

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

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 $v_1 = w_1 = 1$, $\tilde{V} \equiv 0$, $\tilde{W}(x_j) = \frac{1}{4}x_j^4$ and $\ell = 2\pi$, $\vartheta = \pi/2$. The associated NLSE $(4.3)_{2\pi,0}$ reads as

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

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

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 the NLSE, we reconstructed $|A(\tau,\cdot)|$ and $\mathrm{Re}A(\tau,\cdot)$ via the formulae

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

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

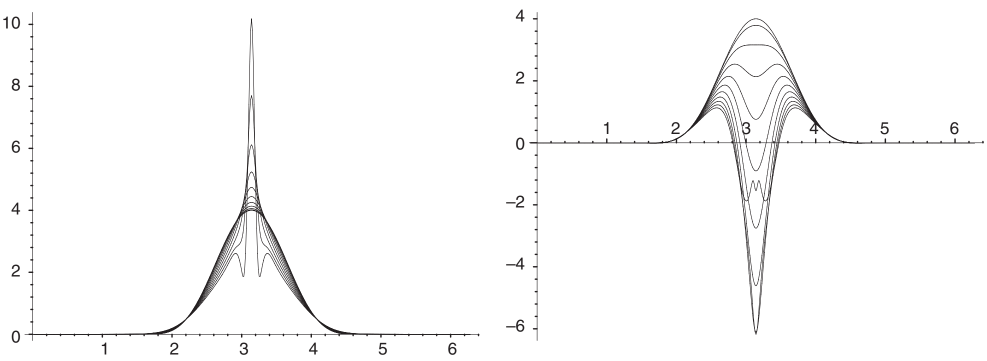

Figure 2. Solution of NLSE (4.5): $|A(\kappa\tau_0,\cdot)|$ (left) and $\mathrm{Re}A(\kappa\tau_0,\cdot)$ (right), $\kappa = 0.1, \ldots, 1$.

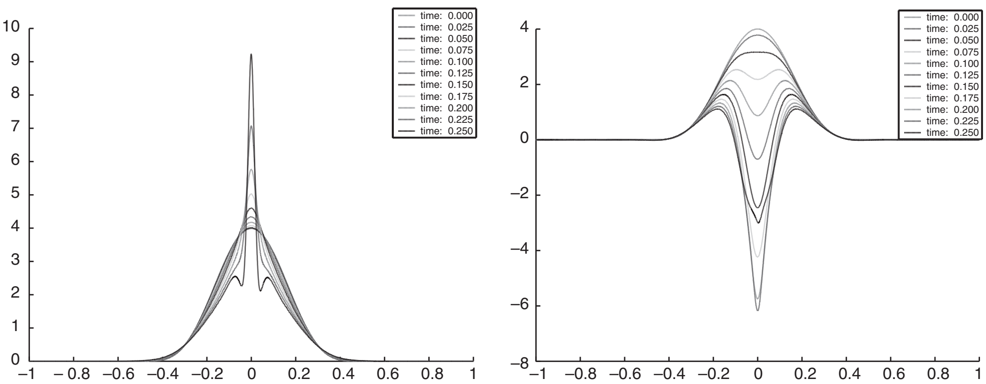

Figure 3. $|A|$ (left) and $\mathrm{Re}A$ (right) from the solution of $(4.1)_m$ with $m = 4000$.

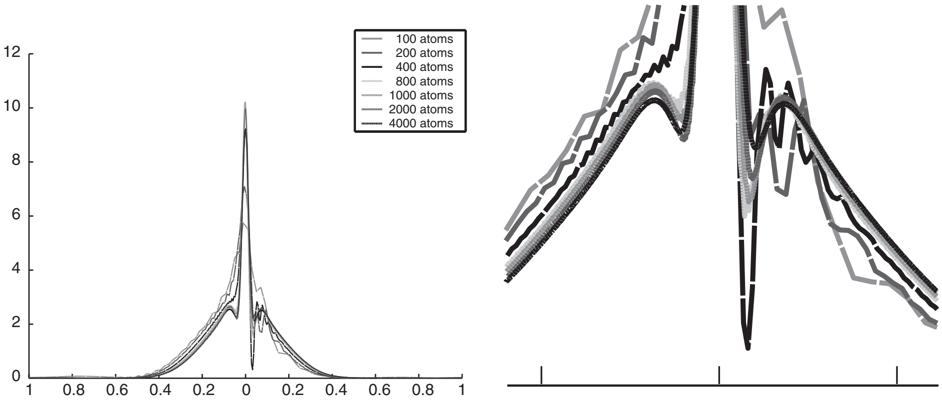

Figure 4. Comparison $|A(0.25, \cdot)|$ for $m = 100, \ldots, 4000$ (to the right: magnification).

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 asymmetry disappears for $m \to \infty$.

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 independent 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 with 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, Modelling and Simulation of Multiscale Problems' through grant Mi 459/3–2. We thank David Remy for doing the numerical computations.

References

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

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

[DK00] Dreyer W and Kunik M 2000 Cold, thermal and oscillator closure of the atomic chain *J. Phys. A: Math. Gen.* **33** 2097–129

Dreyer W and Kunik M 2000 *J. Phys. A: Math. Gen.* **33** 2458 (Corrigendum)

[DKV95] Deift P, Kriecherbauer T and Venakides S 1995 Forced lattice vibrations. I *Commun. Pure Appl. Math* **48** 1187–249

Deift P, Kriecherbauer T and Venakides S 1995 Forced lattice vibrations. II *Commun. Pure Appl. Math* **48** 1251–98

[FM02] Friesecke G and Matthies K 2002 Atomic-scale localization of high-energy solitary waves on lattices *Physica* D **171** 211–20

[FP99] Friesecke G and Pego R L 1999 Solitary waves on FPU lattices: I. Qualitative properties, renormalization and continuum limit *Nonlinearity* **12** 1601–27

[FP02] Friesecke G and Pego R L 2002 Solitary waves on FPU lattices: II. Linear implies nonlinear stability *Nonlinearity* **15** 1343–59

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

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

[GM04] Giannoulis J and Mielke A 2004 in preparation

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

[IK00] Iooss G and Kirchgässner K 2000 Travelling waves in a chain of coupled nonlinear oscillators *Commun. Math. Phys.* **211** 439–64

[Ioo00] Iooss G 2000 Travelling waves in the Fermi–Pasta–Ulam lattice *Nonlinearity* **13** 849–66

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

[Kal89] Kalyakin L A 1989 Long wave asymptotics. Integrable equations as asymptotic limits of non-linear systems *Russ. Math. Surv.* **44** 3–42

[KSM92] Kirrmann P, Schneider G and Mielke A 1992 The validity of modulation equations for extended systems with cubic nonlinearities *Proc. R. Soc. Edin.* A **122** 85–91

[MA94] MacKay R S and Aubry S 1994 Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators *Nonlinearity* **7** 1623–43

[Mie02] Mielke A 2002 The Ginzburg–Landau equation in its role as a modulation equation *Handbook of Dynamical Systems II* ed B Fiedler (Amsterdam: Elsevier) pp 759–834

[Mie04] Mielke A 2004 Macroscopic behavior of microscopic oscillations in harmonic lattices, in preparation

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

[Paz83] Pazy A 1983 *Semigroups of Linear Operators and Applications to Partial Differential Equations* (*Applied Mathematical Sciences* vol 44) (New York: Springer)

[Sch98] Schneider G 1998 Justification of modulation equations for hyperbolic systems via normal forms *NoDEA Nonlinear Differ. Eqns Appl.* **5** 69–82

[SW00] Schneider G and Wayne C E 2000 Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi–Pasta–Ulam model *Int. Conf. on Differential Equations* vol 1, ed B Fiedler *et al* (Singapore: World Scientific) pp 390–404

[Spo91] Spohn H 1991 *Large Scale Dynamics of Interacting Particles* (New York: Springer)

[Tem88] Temam R 1988 *Infinite-Dimensional Dynamical Systems in Mechanics and Physics* (New York: Springer)