Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Continuum descriptions for the dynamics in discrete lattices: derivation and justification

Johannes Giannoulis[1], Michael Herrmann[2], Alexander Mielke[1,2]

submitted: 26th April 2006

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

[2] Institut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin-Adlershof
Germany
E-Mail: michaelherrmann@math.hu-berlin.de

No. 1126
Berlin 2006

2000 *Mathematics Subject Classification.* 34C20, 34E13, 37K60, 37K05, 70F45, 70K70.

Key words and phrases. Discrete lattice systems, modulation equations, Hamiltonian structure, multiscale ansatz, dispersive wave propagation.

Research supported by the DFG Priority Program SPP 1095 *Analysis, Modeling and Simulation of Multiscale Problems* under Mi 459/3.

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

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

Abstract

The passage from microscopic systems to macroscopic ones is studied by starting from spatially discrete lattice systems and deriving several continuum limits. The lattice system is an infinite-dimensional Hamiltonian system displaying a variety of different dynamical behavior. Depending on the initial conditions one sees quite different behavior like macroscopic elastic deformations associated with acoustic waves or like propagation of optical pulses. We show how on a formal level different macroscopic systems can be derived such as the Korteweg-de Vries equation, the nonlinear Schrödinger equation, Whitham's modulation equation, the three-wave interaction model, or the energy transport equation using the Wigner measure. We also address the question how the microscopic Hamiltonian and the Lagrangian structures transfer to similar structures on the macroscopic level. Finally we discuss rigorous analytical convergence results of the microscopic system to the macroscopic one by either weak-convergence methods or by quantitative error bounds.

1 Introduction

A major topic in the area of multiscale problems is the derivation of macroscopic, continuum models from microscopic, discrete ones. The prototype of a discrete many-particle system is a periodic lattice for modeling a crystal. Starting from the seminal work of Fermi, Pasta, and Ulam ([FPU55]), a lot of interest and work has been attracted to the study of the statical and dynamical behavior of ordered discrete systems. In the dynamical situation one is interested in macroscopic limits that are obtained by choosing well-prepared initial conditions: We choose the initial data in a specified class of functions and want to obtain an evolution equation within this function class, which we call the macroscopic limit problem. This approach is motivated by the theory of modulation equations, which evolved in the late 1960's for problems in fluid mechanics (see e.g. [Mie02] for a survey on this subject). If the linearized model has a space-time periodic solution, one asks how initial modulations of this pattern evolve in time. The modulations occur on much larger spatial and temporal scales; thus the modulation equation is a macroscopic equation.

In mathematically rigorous terms this can be described by studying the following coarse graining diagram:

$$
\begin{array}{cccc}
 & \textbf{microscopic} & \xrightarrow{\mathcal{S}_\varepsilon} & \textbf{macroscopic} \\
\text{initial data } t=0 & \mathbf{z}_\varepsilon^0 & \xrightarrow{\quad\varepsilon \to 0\quad} & \mathbf{A}^0 \\
\boxed{\textbf{time evolution}} & \Big\downarrow\, t>0 & & \tau>0\, \Big\downarrow \\
 & \mathbf{z}_\varepsilon(\tau/\varepsilon^\sigma) & \xrightarrow{\quad\varepsilon \to 0\quad} & \mathbf{A}(\tau) \\
 & \text{discrete, atomistic} & \boxed{\textbf{coarse graining}} & \text{continuum}
\end{array}
$$

Here $\mathbf{z}_\varepsilon : [0, \tau_*/\varepsilon^\sigma] \to Z_\varepsilon$ denotes the solution of the microscopic model depending on the microscopic time t and $\mathbf{A} : [0, \tau_*] \to Z_0$ is the solution of the macroscopic model. In the best case the diagram commutes, i.e., if the coarse graining $\mathcal{S}_\varepsilon \mathbf{z}_\varepsilon(\tau/\varepsilon^\sigma) \to \mathbf{A}(\tau)$ holds at time $\tau = 0$, then it also holds for all $\tau \in [0, \tau_*]$. Examples of such results will be Theorems 5.1, 5.2, 7.1 and 7.2.

Before establishing these results, we survey methods to derive macroscopic models on the formal level by using a suitable multiscale ansatz and expanding the coefficients of equal powers of the small parameter and of the harmonics of the microscopic fluctuation to 0. The emphasis is to survey the theory and to explain the main techniques and results on simple models like the FPU chain or the Klein-Gordon chain, see Section 2.3.

Naturally, our survey can only cover a small part of the rich subject of dynamics in discrete systems. We will totally omit any of the works on static solutions for lattices, see e.g. [FJ00, BG02b, BG02a, Sch05a, Ble05, MBL06]. Moreover, there is a huge body of work concerning the understanding of special solution classes like traveling or standing pulses with or without periodic modulations, see [FW94, MA94, Kon96, FP99, Ioo00, IK00, FM02, FP02, FM03, Jam03, FP04a, FP04b, IJ05, DHM06]. The response of oscillator chains to a simple initial disturbance or to Riemann initial data is studied in [DKV95, DKKZ96, DK00, BCS01, DHR06], where in particular completely integrable systems like the Toda lattice are of interest. Finally in the framework of non-equilibrium statistical mechanics (cf. for a survey e.g. [Spo91, Bol96]) one is interested in highly disordered systems, where only statistical averages satisfy nice macroscopic equations.

2 The discrete models

In the first subsection we write down the class of systems that can be treated with the methods surveyed below. This includes general polyatomic lattices in any space dimension. The interactions can be general and can occur between several atoms, not just pair potentials, and can have arbitrary range. In the second subsection we treat the linearizations, which simplify a lot and can be treated in particular by Fourier transform methods. There, the central structure are the different dispersion relations, which will be used heavily in the subsequent analysis. Finally we present two simple model problems that represent most of the interesting features. These models will be addressed in most of the following results to illustrate the general results.

2.1 General lattices systems

We model a perfectly period crystal based on a d-dimensional Bravais lattice Γ embedded into $\mathbb{R}^d$. This lattice is homeomorphic to the additive group $\mathbb{Z}^d$ but might have a different metric structure. Each lattice point $\gamma \in \Gamma$ denotes a unit cell in the actual crystal and, hence, the vectors $x_\gamma \in \mathbb{R}^m$ and $\dot{x}_\gamma$ are collections of all the relevant positions and velocities, respectively, of the atoms inside this unit cell. By $(\mathbf{x}, \dot{\mathbf{x}}) \in \ell_2(\Gamma)^m \times \ell_2(\Gamma)^m$ we denote the state of the system, where $\mathbf{x} = (x_\gamma)_{\gamma\in\Gamma}$ and $\dot{\mathbf{x}} = (\dot{x}_\gamma)_{\gamma\in\Gamma}$. By $M \in \mathbb{R}^{m\times m}$ we denote the mass matrix for each cell, which is assumed to be symmetric and positive definite. The total kinetic energy in the crystal is

$$\mathcal{K}(\dot{\mathbf{x}}) = \tfrac{1}{2}\langle\!\langle M\dot{\mathbf{x}}, \dot{\mathbf{x}}\rangle\!\rangle \overset{\text{def}}{=} \textstyle\sum_{\gamma\in\Gamma} \tfrac{1}{2}\langle M\dot{x}_\gamma, \dot{x}_\gamma\rangle.$$

The potential energy $\mathcal{V}(\mathbf{x})$ is obtained by adding up all contributions acting on one cell via a single potential $V_{\text{cell}} : \ell_2(\Gamma)^m \to \mathbb{R}$ given the forces of the state $\mathbf{x}$ on the cell at $\gamma = 0$:

$$\mathcal{V}(\mathbf{x}) = \textstyle\sum_{\alpha\in\Gamma} V_{\text{cell}}(T_\alpha \mathbf{x}).$$

Here T_α is the translation operator with $T_\alpha\mathbf{x} = (x_{\alpha+\gamma})_{\gamma\in\Gamma}$. In the case of finite-range interaction the potential V_{cell} only depends on finitely many components, e.g., $V_{\text{cell}}(\mathbf{x}) = V_0(x_0) + \sum_{0<|\gamma|\leq R} V_\gamma(x_\gamma - x_0)$ for pair interactions.

The Newtonian equations for this lattice model are given as

$$M\ddot{x}_\gamma = -\mathrm{D}_{x_\gamma}\mathcal{V}(\mathbf{x}) = -\textstyle\sum_{\alpha\in\Gamma} \nabla_{x_{\gamma-\alpha}} V_{\text{cell}}(T_\alpha\mathbf{x}) \quad \text{for } \gamma \in \Gamma. \tag{1}$$

Of course this system is invariant under the translations T_α, $\alpha \in \Gamma$, and has the total energy $\mathcal{E}(\mathbf{x}, \dot{\mathbf{x}}) = \mathcal{K}(\dot{\mathbf{x}}) + \mathcal{V}(\mathbf{x})$ as first integral. Moreover, it is a canonical Hamiltonian system with momenta $\mathbf{p} = M\dot{\mathbf{x}}$, Hamiltonian function $\mathcal{H}$, and symplectic form $\boldsymbol{\omega}_{\text{can}}$:

$$\begin{aligned} &\mathcal{H}(\mathbf{x}, \mathbf{p}) = \tfrac{1}{2}\langle\!\langle M^{-1}\mathbf{p}, \mathbf{p}\rangle\!\rangle + \mathcal{V}(\mathbf{x}) \text{ and} \\ &\boldsymbol{\omega}_{\text{can}}\big((\mathbf{v}_1, \mathbf{q}_1), (\mathbf{v}_2, \mathbf{q}_2)\big) = \langle\!\langle \mathbf{v}_1, \mathbf{q}_2\rangle\!\rangle - \langle\!\langle \mathbf{v}_2, \mathbf{q}_1\rangle\!\rangle. \end{aligned} \tag{2}$$

Clearly, the Newtonian equations (1) are equivalent to the Hamiltonian equations $\dot{\mathbf{x}} = \partial_{\mathbf{p}}\mathcal{H}(\mathbf{x}, \mathbf{p})$, $\dot{\mathbf{p}} = -\partial_{\mathbf{x}}\mathcal{H}(\mathbf{x}, \mathbf{p})$. Moreover, they can be obtained as the Euler-Lagrange equation for the Lagrangian

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \mathcal{K}(\dot{\mathbf{x}}) - \mathcal{V}(\mathbf{x}). \tag{3}$$

2.2 Linear systems and dispersion relations

Linearization leads us to linearized systems, where the potential V is a quadratic form. The linear equation takes the form

$$M\ddot{x}_\gamma = -\sum_{\beta\in\Gamma} A_\beta x_{\gamma+\beta} = \sum_{\alpha\in\Gamma} A_{\gamma-\alpha}x_\alpha \quad \text{for } \gamma\in\Gamma, \tag{4}$$

where the interaction matrices satisfy the symmetry condition $A_\beta = A_{-\beta}^\top$ and a decay condition like $\|A_\beta\| \le C\mathrm{e}^{-b|\beta|}$. The quadratic potential energy then reads $\mathcal{V}(\mathbf{x}) = \frac{1}{2}\sum_{\alpha,\gamma\in\Gamma}\langle A_{\gamma-\alpha}x_\alpha, x_\gamma\rangle$.

An essential feature of such harmonic lattices is the presence of many traveling wave solutions in the form of plain waves:

$$x_\gamma(t) = \mathrm{e}^{\mathrm{i}(\theta\cdot\gamma+\omega t)}\Phi \quad \text{where } \theta\in\mathbb{R}^d_* \text{ and } (\mathbb{A}(\theta)-\omega^2 M)\Phi = 0. \tag{5}$$

The wave vectors θ are taken from the torus $\mathcal{T}_\Gamma$, which is obtained by factoring $\mathbb{R}^d_* = \mathrm{Lin}(\mathbb{R}^d)$ with respect to the dual lattice. The symbol matrix $\mathbb{A}(\theta)$ reads

$$\mathbb{A}(\theta) = \sum_{\beta\in\Gamma}\mathrm{e}^{\mathrm{i}\theta\cdot\beta}A_\beta \in \mathbb{C}^{m\times m} \quad \text{for } \theta\in\mathcal{T}_\Gamma.$$

Hence, $\mathbb{A}(\theta)$ is Hermitian, and we always impose the basic assumption of stability in the form $\mathbb{A}(\theta)\ge 0$ for all $\theta\in\mathcal{T}_\Gamma$.

Plane-wave solutions as in (5) exist if ω and θ satisfy the *dispersion relation*

$$0 = \mathrm{Disp}(\omega,\theta) \overset{\text{def}}{=} \det\Big(\omega^2 M - \mathbb{A}(\theta)\Big).$$

Under our stability condition, there are always m non-negative eigenvalue curves

$$\omega = \Omega_k(\theta), \quad k=1,\dots,m,$$

which we order such that $0\le\Omega_1\le\Omega_2\le\cdots\le\Omega_m$. The index k is called the band index. Two velocities will be important below, the *phase velocity* c_{ph} and the *group velocity* c_{gr}:

$$c_{\mathrm{ph}} = c_{\mathrm{ph},k}(\theta) \overset{\text{def}}{=} \tfrac{\Omega_k(\theta)}{|\theta|^2}\,\theta \quad \text{and} \quad c_{\mathrm{gr}} = c_{\mathrm{gr},k}(\theta) \overset{\text{def}}{=} \nabla\Omega_k(\theta).$$

The dynamics of the linear system is completely determined by M and the symbol matrix $\mathbb{A}:\mathcal{T}_\Gamma\to\mathbb{C}^{m\times m}_{\ge 0}$. This is easily seen by transforming (5) into wave vector space. For this define $\mathbf{X}(\theta) = \mathbb{F}\mathbf{x} : \mathcal{T}_\Gamma\to\mathbb{C}^m$ via $\mathbb{F}\mathbf{x} \overset{\text{def}}{=} \sum_\gamma \mathrm{e}^{-\mathrm{i}\theta\cdot\gamma}x_\gamma$, then $\mathbf{X}(t) = \mathbb{F}\mathbf{x}(t):\mathcal{T}_\Gamma\to\mathbb{C}^m$ satisfies the equation

$$M\partial_t^2\mathbf{X}(t,\theta) = -\mathbb{A}(\theta)\mathbf{X}(t,\theta) \tag{6}$$

if and only if x satisfies (5). However, the latter equation is an ODE for each fixed $\theta\in\mathcal{T}_\Gamma$.

For studying the qualitative behavior of the solutions in the subsequent sections, this is not sufficient, and we need to understand the back-transform for large times t. Then, the smoothness properties of the dispersion relations will be important, see Sections 5.3 and 6.2.

2.3 The chain models of FPU and KG

To illustrate our abstract theory we will frequently refer to the simple scalar and one-dimensional case, viz., $\Gamma = \mathbb{Z} \subset \mathbb{R}$ and $x_j \in \mathbb{R}$. The models have the general form

$$\ddot{x}_j = -V_0'(x_j) + \sum_{k=1}^K \big(V_k'(x_{j+k}-x_j) - V_k'(x_j-x_{j-k})\big), \quad j \in \mathbb{Z}. \tag{7}$$

Here V_0 is called the on-site potential that couples the atoms to a background field. The interaction is assumed to be pairwise and involves K neighboring atoms.

The *Fermi-Pasta-Ulam chain* (FPU) is obtained by omitting the on-site potential and choosing $K = 1$:

$$\ddot{x}_j = V_1'(x_{j+1}-x_j) - V_1'(x_j-x_{j-1}), \quad j \in \mathbb{Z}. \tag{8}$$

The importance of this model is its Galilean invariance, i.e., for all $\xi, c \in \mathbb{R}$ the transformation $(\mathbf{x}, \dot{\mathbf{x}}) \mapsto (x_j+\xi + ct, \dot{x}_j+c)_{j\in\mathbb{Z}}$ leaves (8) invariant.

Another simple class is obtained by assuming again $K = 1$ with linear nearest-neighbor interaction and a nonlinear background potential. In analogy to the Klein-Gordon equation this model is called *Klein-Gordon chain* (KG):

$$\ddot{x}_j = x_{j+1} - 2x_j + x_{j-1} - V_0'(x_j), \quad j \in \mathbb{Z}. \tag{9}$$

In these two models the dispersion relation has the structure

$$0 = \mathrm{Disp}(\omega, \theta) = \omega^2 - a - 2b(1-\cos\theta) \text{ with } a = V_0''(0) \text{ and } b = V_1''(0),$$

where $a, b \geq 0$ is equivalent to our stability condition. The solution reads

$$\omega = \Omega(\theta) = \big(a + 2b(1-\cos\theta)\big)^{1/2},$$

which is smooth for $a > 0$. For $a = 0$ we find $\Omega(\theta) = \sqrt{b}\,2\,|\sin(\theta/2)|$, which is not differentiable at $\theta = 0$, but the two limits $\pm\sqrt{b}$ of Ω' at $\theta = 0$ are the macroscopic wave speeds.

3 Formal derivation of continuum models

3.1 General multiscale approach

We discuss here the derivation of macroscopic models that appear for solutions having a relatively small amplitude, but we refer to [DHR06] and Section 3.7 for results on large amplitude solutions.

The basic ansatz relies on modulations of basic plane waves $\mathrm{e}^{\mathrm{i}(\omega t+\theta\cdot\gamma)}\Phi$ on large spatial scales and suitably chosen slow time scales. We choose $\varepsilon > 0$ to be the small parameter that relates the microscopic and the macroscopic temporal and spatial scales, i.e., we set

$$\tau = \varepsilon^s t \quad \text{and} \quad y = \varepsilon\gamma \in \mathbb{R}^d \quad \text{for } \gamma \in \Gamma \subset \mathbb{R}^d.$$

Of course, there are cases where different scalings in different spatial directions are useful, but for simplicity we restrict ourselves to this case.

We now choose a finite set of wave vectors $\theta_1, \ldots, \theta_N \in \mathcal{T}_\Gamma$ and associated band indices $k_1, \ldots, k_N \in \{1, \ldots, m\}$ and consider the associated plane waves

$$x_\gamma(t) = \mathbf{E}_n(t,\gamma)\mathbf{\Phi}_n, \quad \text{where } \mathbf{E}_n(t,\gamma) \stackrel{\text{def}}{=} \mathrm{e}^{\mathrm{i}(\omega_n t + \theta_n\cdot\gamma)}$$
$$\text{with } \omega_n = \Omega_{k_n}(\theta_n) \text{ and } \mathbf{\Phi}_n = \Phi_{k_n}(\theta_n).$$

This may include the case $\theta = 0$ and $\omega = 0$, which relates to the macroscopic limit of solutions without microstructure.

The two-scale method now starts from the ansatz

$$\begin{aligned}
&(x_\gamma(t), \dot{x}_\gamma(t)) = R_\varepsilon(\mathbf{A})_\gamma(t), \text{ where } \mathbf{A} = (A_1, \ldots, A_N) \text{ and} \\
&R_\varepsilon(\mathbf{A})_\gamma(t) = \textstyle\sum_{n=1}^N \varepsilon^{\sigma_n} A_n(\varepsilon^s t, \varepsilon\gamma)\mathbf{E}_n(t,\gamma)\mathbf{\Phi}_n \\
&\qquad + \textstyle\sum_{n,k=1}^N \varepsilon^{\sigma_n+\sigma_k}\mathbf{\Psi}_{n,k}(\varepsilon^s t, \varepsilon\gamma)\mathbf{E}_n\mathbf{E}_k \\
&\qquad + \textstyle\sum_{n,k,l=1}^N \varepsilon^{\sigma_n+\sigma_k+\sigma_l}\mathbf{\Psi}_{n,k,l}(\varepsilon^s t, \varepsilon\gamma)\mathbf{E}_n\mathbf{E}_k\mathbf{E}_l + \text{h.o.t.}
\end{aligned} \tag{10}$$

Here the powers $s, \sigma_1, \ldots, \sigma_N \in \mathbb{R}$ have to be chosen appropriately. We refer to the variety of different models that can be obtained in this way. To obtain real-valued solutions one chooses $A_n = \overline{A}_{N-n}$ and similarly for the higher order terms. In cases with $\theta_n \neq 0$ the functions A_n are the modulating amplitudes of the basic periodic plane wave.

The aim is to derive suitable equations for $A_1, \ldots, A_N$, which make this ansatz (10) consistent with the discrete model (1). The obtained equations are partial differential equations combined with some algebraic relations. These equations are called the macroscopic equations, because they are posed in terms of the macroscopic variables $\tau = \varepsilon^s t$ and $y = \varepsilon\gamma$. Inserting the ansatz (10) into the nonlinear system (1) we have to expand both sides in terms of the products $\varepsilon^{\widetilde{q}}\Pi_{n=1}^N \mathbf{E}_n^{q_n}$ with $\widetilde{q} = \sum_{n=1}^N \sigma_n q_n$. Here we have to expand difference quotients $x_{\gamma+\alpha} - x_\gamma$ in terms of spatial derivative of A_n. Moreover, the resonances between the plane waves are important to allow for nontrivial nonlinear interaction. They are characterized by vectors $q \in \mathbb{N}^N$ such that $\Pi_{n=1}^N \mathbf{E}_n^{q_n} \equiv 1$, see Section 7.2 for a general theory.

We arrive at a hierarchy of equations that can be parametrized by the multi-index $\mathbf{q} = (q_1, ..., q_N) \in \mathbb{N}^N$. These equations decompose into two groups. If the term $\mathbf{E}_\mathbf{q} = \Pi_{n=1}^N \mathbf{E}_n^{q_n}$ is nonresonant, i.e., different from all the terms $\mathrm{e}^{\mathrm{i}(\omega t + \theta\cdot\gamma)}$ that satisfy the dispersion relation, then the equation for $\Psi_\mathbf{q}(\tau, y)$ is uniquely solvable. The resonant groups associate with the terms $\mathbf{E}_\mathbf{q} = \Pi_{n=1}^N \mathbf{E}_n^{q_n}$ that equal one of the terms $\mathrm{e}^{\mathrm{i}(\Omega_j(\theta)t + \theta\cdot\gamma)}$, which without loss of generality is already in our list, let us say $\mathbf{E}_m$. Naturally, the coefficient $\mathbf{\Psi}_\mathbf{q}$ cannot be determined uniquely, because the plane wave $\mathbf{E}_m\mathbf{\Phi}_m$ solves the linear problem. Thus, by Fredholm's alternative we obtain a solvability condition for the terms on the left-hand side that contains only lower order terms that are already determined. This gives either a PDE or an algebraic equation on the previously chosen functions. Moreover, the general solution contains a new scalar function $B_{\mathbf{q}}$, namely $\mathbf{\Psi}_{\mathbf{q}} = B_{\mathbf{q}}\mathbf{\Phi}_m + \mathbf{\Psi}^0_{\mathbf{q}}$.

We refer to [Mie02, GM04, GM06] for a more detailed description of this procedure. In fact, without doing any explicit calculation on the specific discrete lattice system (1) it is possible to describe the form of the macroscopic equations as follows:

$$\begin{array}{ll} \text{If } \omega_k \neq 0: & \partial_\tau A_k = \sum_{\mathbf{q}\in M_k(s)} c_{\mathbf{q}} \Pi_{n=1}^N A_n^{q_n}, \\ \text{If } \omega_k = 0: & \partial_\tau^2 A_k = \sum_{\mathbf{q}\in M_k(2s)} \widetilde{c}_{\mathbf{q}} \Pi_{n=1}^N A_n^{q_n}, \end{array} \tag{11}$$

where $M_k(s) \stackrel{\text{def}}{=} \{\, \mathbf{q} \mid \sigma_k{+}s{=}\sum_1^N \sigma_n q_n,\ 0{=}\sum_1^N \omega_n q_n,\ \sum_1^N q_n\theta_n = 0 \text{ on } \mathcal{T}_\Gamma \,\}$. For more details see [Gia06, GMS06] and Section 7.2.

The following Sections 3.2 to 3.7 treat a list of examples, which highlight the generality of the approach.

3.2 The quasilinear wave equation

A simple but important macroscopic model for FPU chains results by the following multiscale ansatz with hyperbolic scaling:

$$x_j(t) = \varepsilon^{-1} X(\varepsilon t, \varepsilon j), \quad \tau = \varepsilon t, \quad y = \varepsilon j. \tag{12}$$

Note that here x_j denotes the spatial position of atom j rather than its displacement. We insert the ansatz (12) into (8) and eliminate the relative displacements by the Taylor expansion $x_{j\pm1} - x_j \approx \pm\varepsilon\partial_y X(\varepsilon t, \varepsilon j)$. Using $\partial_\tau = \varepsilon\partial_t$ we can identify the macroscopic modulation equations as the nonlinear wave equation

$$\partial_{\tau\tau} X - \partial_y V_1'(\partial_y X) = 0. \tag{13}$$

Via $r = \partial_y X$ and $v = \partial_\tau X$ it transforms into the quasilinear first-order system

$$\partial_\tau r - \partial_y v = 0, \quad \partial_\tau v - \partial_y V_1'(r) = 0. \tag{14}$$

These equations describe the macroscopic evolution of non-oscillatory solutions of FPU. However, due to the nonlinearity V_1' smooth solutions of (14) can form shocks in finite times, and in this case the quasilinear wave equation is not longer an appropriate macroscopic model for FPU. This problem is addressed in [DHR06].

3.3 The Korteweg-de Vries equation

Another example for macroscopic modulation equations, see [SW00, FP99], relies on the KdV-ansatz

$$x_j(t) = \varepsilon\, U\big(\varepsilon^3 t,\, \varepsilon(j{+}ct)\big) \tag{15}$$

with scaling $\tau = \varepsilon^3 t$, $y = \varepsilon(j{+}ct)$. We insert the ansatz into (8) and use Taylor expansion up to order $O(\varepsilon^6)$. Comparing the leading order terms we find that c is given by $c^2 = V_1''(0)$. Since the next order terms all cancel, the modulation equation is determined by the terms corresponding to ε^5, and finally we obtain

$$2\,c\,\partial_{\tau y}U - \tfrac{1}{12}\,c^2\,\partial_y U\,\partial_{yy}U - V_1'''(0)\,\partial_{yyyy}U = 0, \tag{16}$$

which is a KdV equation for $\partial_y U$.

3.4 The nonlinear Schrödinger equation

We consider the scalar, d-dimensional lattice (1) (i.e., $d \in \mathbb{N}$ and $m = 1$)

$$\ddot{x}_\gamma = \textstyle\sum_{0<|\beta|\leq R}[V_\beta'(x_{\gamma+\beta}{-}x_\gamma){-}V_\beta'(x_\gamma{-}x_{\gamma-\beta})] - V_0'(x_\gamma), \quad \gamma \in \Gamma, \tag{17}$$

and are interested in the macroscopic deformations of a modulated plane wave solution of the linearized system

$$x_\gamma(t) = \varepsilon A(\tau,y)\mathbf{E}(t,\gamma) + \text{c.c.} + O(\varepsilon^2) \quad \text{with} \quad \mathbf{E}(t,\gamma) = \mathrm{e}^{\mathrm{i}(\omega t + \theta\gamma)} \tag{18}$$

(c.c.: conjugate complex) for a fixed wave vector $\theta \in \mathcal{T}_\Gamma$ with frequency ω satisfying the dispersion relation $\omega^2 = \Omega^2(\theta) > 0$.

Since the system is dispersive and nonlinear and the amplitude A is weakly scaled by $0 < \varepsilon \ll 1$, we need a slow macroscopic time scale $\tau = \varepsilon^2 t$ comparing to the macroscopic space scale $y = \varepsilon(\gamma{-}c_{\text{gr}}t)$, in order to see the evolution of A as time passes. This is the so called *dispersive scaling.* The choice of y also reflects that we are moving with the pulse at its microscopical group velocity $c_{\text{gr}} = \nabla_\theta\Omega(\theta)$. By this scaling it turns out that the evolution of A is given by the nonlinear Schrödinger equation

$$\mathrm{i}\partial_\tau A = \mathrm{Div}_y(\frac{1}{2}\mathrm{D}_\theta^2\Omega(\theta)\nabla_y A) + \rho|A|^2 A. \tag{nlS}$$

For the justification of this equation we refer to Section 7.1.

3.5 Three-wave interaction

For the lattice (17) we are now interested in a macroscopic description for the evolution of the amplitudes A_n, $n = 1,2,3$, of three nonlinearly interacting modulated plane waves with different wave numbers θ_n and frequencies ω_n, where $\omega_n^2 = \Omega^2(\theta_n)$. Thus, ansatz (10) takes the special form

$$x_\gamma(t) = \varepsilon \textstyle\sum_{n=1}^3 A_n(\tau,y)\mathbf{E}_n(t,\gamma) + \text{c.c.} + O(\varepsilon^2) \text{ with } \mathbf{E}_n(t,\gamma) = \mathrm{e}^{\mathrm{i}(\omega_n t + \theta_n\cdot\gamma)}$$

but now using the hyperbolic scaling $\tau = \varepsilon t$, $y = \varepsilon\gamma$ again. It turns out that, if the wave vectors θ_n and frequencies ω_n are in resonance, viz.,

$$\theta_1 + \theta_2 + \theta_3 = 0 \mod \mathcal{T}_\Gamma \quad \text{and} \quad \omega_1 + \omega_2 + \omega_3 = 0, \tag{19}$$

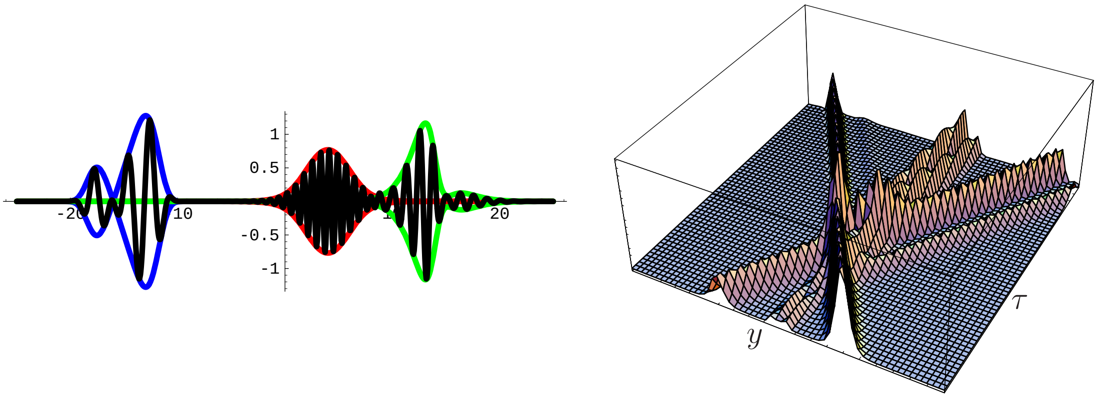

Figure 3.1: Left: typical initial condition. Right: energy distribution in space-time.

the amplitudes A_n, $n = 1, 2, 3$, satisfy the so called three-wave interaction equations

$$\begin{cases} \omega_1 \partial_\tau A_1 = \omega_1 \nabla_\theta \Omega(\theta_1) \cdot \nabla_y A_1 + c \overline{A}_2 \overline{A}_3, \\ \omega_2 \partial_\tau A_2 = \omega_2 \nabla_\theta \Omega(\theta_2) \cdot \nabla_y A_2 + c \overline{A}_1 \overline{A}_3, \\ \omega_3 \partial_\tau A_3 = \omega_3 \nabla_\theta \Omega(\theta_3) \cdot \nabla_y A_3 + c \overline{A}_1 \overline{A}_2, \end{cases} \tag{20}$$

with $c = 2 \sum_{0<|\beta|\leq R} V_\beta'''(0) \sum_{n=1}^{3} \sin(\theta_n \cdot \beta) + \mathrm{i} V_0'''(0)$. Each equation consists of a transport part via the group velocity and a nonlinear coupling to the two other modes. Figure 3.1 illustrates the behavior. Without the resonance condition (19) being fulfilled, nonlinear terms would not arise and the pulses would just pass through each other. For the justification of this equation we refer to 7.2.

3.6 Coupled systems

While the two examples above apply to a system with or without background potential V_0, we are now looking at systems with Galilean invariance, where the canonical example is FPU from (8). The aim is to understand the coupling between macroscopic deformations and microscopically oscillating pulses. Since in general the macroscopic wave speeds and the microscopic group velocity are different, we use the hyperbolic time scale $\tau = \varepsilon t$. Ansatz (10) reduces to

$$x_\gamma(t) = \varepsilon^\alpha X(\tau, y) + \varepsilon^\beta A(\tau, y) \mathbf{E} + \varepsilon^\beta \overline{A}(\tau, y) \overline{\mathbf{E}} + \text{h.o.t.}$$

with $\mathbf{E} = \mathrm{e}^{\mathrm{i}(\omega t + \theta \cdot \gamma)}$ and $\omega = \Omega(\theta)$. Here α and β might be different and depend on the nonlinearities as well as the scaling of the initial data. We treat the case of the FPU chain with $V_1(r) = \frac{a}{2} r^2 + \frac{b}{3} r^3 + \frac{c}{4} r^4$ with $a > 0$ and $b, c \in \mathbb{R}$.

As a first example we consider the case $\alpha = 0$, $\beta = 1$ and find the system

$$\partial_\tau^2 X = c_{\mathrm{m}}^2 \partial_\xi^2 X, \quad \mathrm{i} \partial_\tau A = \mathrm{i} c_{\mathrm{gr}} \partial_y A - \rho_0 (\partial_y X)\, A$$

with $c_m := \Omega'(0) = \sqrt{a}$, $c_{\rm gr} := \Omega'(\theta)$ and $\rho_0 := 2b\Omega(\theta)/a$. Since the contributions X and A scale differently, the coupling of X and A takes place only in one equation. We have the two conserved quantities

$$\mathbb{H}(A) = \int_{\mathbb{R}} \omega^2 |A|^2 \,\mathrm{d}y \quad \text{and} \quad \mathbb{E}(X, X_\tau) = \int_{\mathbb{R}} \tfrac{1}{2} X_\tau^2 + \tfrac{c_{\rm m}^2}{2} X_y^2 \,\mathrm{d}y.$$

The second example has $\alpha = 0$ and $\beta = 1/2$, which leads to the system

$$X_{\tau\tau} = \left(c_{\rm m}^2 X_y + \rho_1 |A|^2\right)_y, \quad 2\mathrm{i}\omega A_\tau = \mathrm{i}\omega c_{\rm gr} A_y - \left(\rho_1 X_y + 2\rho_2 |A|^2\right) A,$$

where $\rho_1 := 2b(1-\cos\theta)$ and $\rho_2 := 3c(1-\cos\theta)^2$. This system is a Lagrangian and Hamiltonian system in the sense to be discussed in Section 4. The Lagrangian reads

$$\begin{aligned}&\mathbb{L}(X, A, X_\tau, A_\tau)\\ &= \int_{\mathbb{R}} \omega \operatorname{Im}\left(\overline{A}(2A_\tau - c_{\rm gr} A_y)\right) + \tfrac{1}{2} X_\tau^2 - \tfrac{c_{\rm m}^2}{2} X_y^2 - |A|^2 \left(\rho_1 X_y + \rho_2 |A|^2\right) \mathrm{d}y.\end{aligned}$$

There are two first integrals

$$\begin{aligned}\mathbb{H}(A) &= \int_{\mathbb{R}} \omega^2 |A|^2 \,\mathrm{d}y,\\ \mathbb{E}(X, A, X_\tau) &= \int_{\mathbb{R}} \omega c_{\rm gr} \operatorname{Im}(\overline{A} A_y) + \tfrac{1}{2} X_\tau^2 + \tfrac{c_m^2}{2} X_y^2 + |A|^2 \left(\rho_1 X_y + \rho_2 |A|^2\right) \mathrm{d}y.\end{aligned}$$

The symplectic structure of the associated Hamiltonian system for (X, A, X_τ) is non-canonical and can easily be deduced as in Section 4.2.

3.7 Whitham's modulation equation

In [Whi74] Whitham studies certain nonlinear PDEs and relying on the hyperbolic scaling he develops a theory that is capable to describe the macroscopic evolution of large microscopic oscillations. Here we apply Whitham's approach to three different chain models. We start with KG, cf. (7), and make the following multiscale ansatz

$$x_j(t) = \mathfrak{X}\big(\varepsilon t,\, \varepsilon j,\, \varepsilon^{-1}\,\Theta(\varepsilon t,\, \varepsilon j)\big), \tag{21}$$

where $\mathfrak{X}$ is assumed to be 2π-periodic with respect to the phase variable $\phi = \varepsilon^{-1}\Theta$. In this ansatz both the wave number θ and the frequency ω depend on the macroscopic coordinates $(\tau,\, y)$ and are defined by the *modulated phase* Θ via $\theta = \partial_y \Theta$ and $\omega = \partial_\tau \Theta$. It can be shown that to leading order the function $\mathfrak{X}$ must satisfy the following nonlinear advance-delay-differential equation

$$\omega^2\, \partial_\phi^2 \mathfrak{X} = \nabla_{-\theta} \nabla_{+\theta} \mathfrak{X} - V_0'(\mathfrak{X}) \tag{22}$$

with $(\nabla_{\pm\theta}\mathfrak{X})(\phi) = \pm\mathfrak{X}(\phi \pm \theta) \mp \mathfrak{X}(\phi)$. As usual we refer to solutions of this equation as *traveling waves*. The existence problem for solutions of (22) with small amplitudes is investigated in [IK00]. For convex potentials V_0 we can provide existence of solutions by adapting an idea from [FV99], compare with the similar problem for FPU in [DHR06]. According to (22), the action L of a traveling wave is given by

$$L(\theta,\, \omega) = \tfrac{1}{2\pi} \int_0^{2\pi} \tfrac{\omega^2}{2} \left(\partial_\phi \mathfrak{X}\right)^2 - \left(\nabla_{+\theta} \mathfrak{X}\right)^2 - V_0(\mathfrak{X})\, \mathrm{d}\phi. \tag{23}$$

To identify the macroscopic modulation equations it is convenient to use the Lagrangian formalism, see [Whi74] and Section 4.1, because direct expansions in powers of ε turn out to be quite complicate. With some simple averaging the total action of the chain can be expressed by a functional $\mathbb{L}$, which depends on $(\Theta,\, \partial_\tau\Theta)$ only, and we can derive the modulation equations by the principle of least action. It comes out that the modulation equations are equivalent to the following nonlinear system of conservations laws

$$\partial_\tau\theta - \partial_y\omega = 0, \quad \partial_\tau S + \partial_y g = 0, \tag{24}$$

where $S = \partial_\omega L$ and $g = \partial_\theta L$. In particular, the system (24) is closed by the equation of state (23) and the Gibbs equation $\mathrm{d}L = S\mathrm{d}\omega + g\mathrm{d}\theta$.

The modulation theory for FPU, see [DHM06, DHR06] and the references therein, is more complicate than in the KG case due to the nonlinearity of V_1, and the Galilean invariance of (8). In particular, we must combine (12) and (21) as follows

$$x_j(t) = \varepsilon^{-1}\, X(\varepsilon t,\, \varepsilon j) + \mathfrak{X}\big(\varepsilon t,\, \varepsilon j,\, \varepsilon^{-1}\,\Theta(\varepsilon t,\, \varepsilon j)\big), \tag{25}$$

where as before the profile function $\mathfrak{X}$ is assumed to be 2π-periodic with respect to $\phi = \varepsilon^{-1}\Theta$. This ansatz gives rise to four important macroscopic fields, namely the wave number $\theta = \partial_y\Theta$, the frequency $\omega = \partial_\tau\Theta$, the specific length $r = \partial_y X$, and the macroscopic velocity $v = \partial_\tau X$. To leading order, the profile function X must satisfy the traveling wave equation

$$\omega^2\, \partial_\phi^2\, \mathfrak{X} = \nabla_{-\theta} V_1'(\nabla_{+\theta}\, \mathfrak{X}). \tag{26}$$

For convex potentials V_1 the existence of solutions can be proved by an convex optimization problem, see [DHR06], and rigorous results without convexity assumptions can be found in [FW94, PP00, Ioo00]. The derivation of the modulation equations for (25) again relies on Lagrangian reduction, see for instance [Her04, DHM06], and leads to the following nonlinear system

$$\partial_\tau r - \partial_y v = 0, \quad \partial_\tau v + \partial_y p = 0, \quad \partial_\tau\theta - \partial_y\omega = 0, \quad \partial_\tau S + \partial_y g = 0. \tag{27}$$

These equations can be interpreted as the macroscopic conservation laws of mass, momentum, wave number and entropy. As for KG, the constitutive relations for (27) result from a careful investigation of the thermodynamic properties of traveling waves. More precisely, it can be shown, at least formally, that (26) provides an equation of state $U = U(r,\, \theta,\, S)$ as well as the universal Gibbs equation $\mathrm{d}E = \omega\mathrm{d}S - p\mathrm{d}r - g\mathrm{d}\theta + v\mathrm{d}v$, where U and $E = \frac{1}{2}v^2 + U$ denote the internal and total energy, respectively.

The third example is the discrete nonlinear Schrödinger equation

$$-\mathrm{i}\,\dot{a}_j + a_{j+1} - 2\,a_j + a_{j-1} + \varrho\,|a_j|^2\, a_j = 0, \quad j \in \mathbb{Z}, \tag{28}$$

with complex valued a_j and real parameter ϱ. This equation has exact solutions (traveling waves) of the form $a_j = B\,\mathrm{e}^{\mathrm{i}(\theta j + \omega t)}$ with real amplitude B if ω obeys the nonlinear dispersion relation $\omega + \cos\theta - 2 + \varrho B^2 = 0$. The modulation theory for several variants of (28) was studied in [HLM94] and bases on the multiscale ansatz $a_j(t) = B(\varepsilon t,\, \varepsilon j)\, \mathrm{e}^{\mathrm{i}\,\Theta(\varepsilon t, \varepsilon j)/\varepsilon}$, where as before we set $\theta = \partial_y \Theta$ and $\omega = \partial_\tau \Theta$. One obtains the macroscopic balance laws

$$\partial_\tau\big(A^2\big) - \partial_y\big(2\,A^2\,\sin\theta\big) = 0, \quad \partial_\tau\theta + \partial_y\big(\varrho\,A^2 + \cos\theta\big) = 0, \tag{29}$$

where the second evolution equation is equivalent to $\partial_\tau\theta - \partial_y\omega = 0$. We mention that (29) can also be derived by means of Hamiltonian or Lagrangian reduction discussed in the next section, see [HLM94] for the details.

4 Hamiltonian and Lagrangian structures

The derivation of macroscopic equations for discrete models (or continuous models with microstructure) can be seen as a kind of reduction of the infinite dimensional system to a simpler subclass. If we choose well ordered initial conditions, we hope that the solution will stay in this order and evolve according to a slow evolution with macroscopic effects only. We may interprete this as a kind of (approximate) invariant manifold, and the macroscopic equation describes the evolution on this manifold, the functions $A_1, ..., A_N$ defining kind of coordinates. For such a reduction procedure it is a natural question how the original Hamiltonian and Lagrangian structures, as described in Section 2.1, "reduce" to the macroscopic equation. Here we just survey the main ideas and some examples and refer to [GHM06] for the full details.

Before addressing this question we first address the exact reduction of a Hamiltonian and Lagrangian systems to exactly invariant manifolds (cf. e.g. [Mie91]). First consider the Lagrangian setting for $\mathcal{L}$ defined on $\mathrm{T}Q$. Assume that we have an invariant manifold $\mathbb{M} \subset \mathrm{T}Q$ given in the form

$$\mathbb{M} = \{\, (q, \dot q) = S(p, \dot p) \in \mathrm{T}Q \mid (p, \dot p) \in \mathrm{T}P \,\}.$$

Then, we may define the reduced Lagrangian $L^{\mathrm{red}} = \mathcal{L} \circ S : \mathrm{T}P \to \mathbb{R}$. An easy calculation proves that any solution p of the reduced Lagrangian system

$$0 = -\tfrac{\mathrm{d}}{\mathrm{d}t}\big(\partial_{\dot p} L^{\mathrm{red}}(p, \dot p)\big) + \partial_p L^{\mathrm{red}}(p, \dot p)$$

leads to a solution $(q, \dot q) = S(p, \dot p)$ of the original Lagrangian system. Vice versa, any solution of the latter system that also lies in $\mathbb{M}$ solves the reduced Lagrangian system.

In the Hamiltonian case the tangent bundle structure of $Z = \mathrm{T}Q$ is generalized to a general symplectic structure $\boldsymbol{\omega}$ on the state space Z. Together with the Hamiltonian $\mathcal{H}$ the Hamiltonian system reads

$$\boldsymbol{\Omega}(z)\dot z = \mathrm{D}\mathcal{H}(z) \quad \text{or} \quad \dot z = \mathbf{J}(z)\mathrm{D}\mathcal{H}(z),$$

where $\boldsymbol{\omega}_z(v_1, v_2) = \langle \boldsymbol{\Omega}(z)v_1, v_2\rangle$ and $\mathbf{J}(z) = \boldsymbol{\Omega}(z)^{-1} : \mathrm{T}_z^*Z \to \mathrm{T}_zZ$. For a symplectic, flow-invariant submanifold $\mathbb{M} = \{\, z = R(y) \in Z \mid y \in Y\,\}$ we define the reduced symplectic structure $\boldsymbol{\Omega}^{\mathrm{red}}$ and the reduced Hamiltonian H^{red} via

$$\boldsymbol{\Omega}^{\mathrm{red}}(y) = \mathrm{D}R(y)^*\boldsymbol{\Omega}(R(y))\mathrm{D}R(y) \quad \text{and} \quad H^{\mathrm{red}}(y) = \mathcal{H}(R(y)).$$

Using the flow-invariance of $\mathbb{M}$ it is easy to see that any solution of the reduced Hamiltonian system $\boldsymbol{\Omega}^{\mathrm{red}}(y)\dot{y} = \mathrm{D}H^{\mathrm{red}}(y)$ solves the original system and vice versa if starting on $\mathbb{M}$.

Our applications will of course use the ansatz R_ε from (10) for the reduction, which can be seen as an approximation of an invariant manifold.

4.1 Lagrangian reduction

The multiscale ansatz (10) discussed above was chosen such that it is formally consistent and in many cases it is possible to justify the ansatz by a rigorous error analysis, as surveyed in Sections 6.1 and 7. Hence, we consider the multiscale ansatz as a parametrization of an (approximate) invariant manifold. Inserting the ansatz (10) into the Lagrangian $\mathcal{L}$ defined in (3) we obtain a reduced Lagrangian in the form

$$L^{\mathrm{red}}(\varepsilon, \mathbf{A}, \partial_\tau\mathbf{A}) = \varepsilon^\rho\mathbb{L}(\mathbf{A}, \partial_\tau\mathbf{A}) + O(\varepsilon^{\rho+1}), \text{ where } \mathbf{A} = (A_1, \ldots, A_N).$$

Here L^{red} is still an infinite sum over $\gamma \in \Gamma$. However, when expanding in powers of ε, the multiscale ansatz leads to a limit that is an integral over the macroscopic space variable $y \in \mathbb{R}^d$. The infinite sum can be considered as a Riemann sum for the spatial integral.

Since L^{red} is independent of τ, the solutions of the reduced Euler-Lagrange equation conserve the associated energy $\mathbb{E}$ obtain as

$$\mathbb{E}(\mathbf{A}, \mathbf{A}_\tau) = \langle\!\langle \partial_\tau\mathbf{A}, \partial_{\mathbf{A}_\tau}\mathbb{L}(\mathbf{A}, \mathbf{A}_\tau)\rangle\!\rangle - \mathbb{L}(\mathbf{A}, \mathbf{A}_\tau).$$

It is proved in [GHM06] that the Lagrangian equation for $\mathbf{A}$ associated with the lowest order term $\mathbb{L}$ of the reduced Lagrangian $L^{\mathrm{red}}(\varepsilon, \cdot)$ really provides exactly the macroscopic equation (11) derived in Section 3.1.

Here we illustrate this result using a simple example based on the Klein-Gordon chain (9) with the potential $V_0(x) = \frac{a}{2}x^2 + \frac{b}{4}x^4$. We consider a single modulated pulse in the form

$$x_j(t) = \varepsilon^{1/2}A(\varepsilon t, \varepsilon j)\mathbf{E} + \varepsilon^{1/2}\overline{A}(\varepsilon t, \varepsilon j)\overline{\mathbf{E}} \quad \text{with } \mathbf{E} = \mathrm{e}^{\mathrm{i}(\omega t + \theta j)}, \tag{30}$$

where $\omega = \Omega(\theta)$. Inserting this ansatz into $\mathcal{L}$ and using $\vartheta = \mathrm{e}^{\mathrm{i}\theta} - 1$ we find

$$\begin{aligned} L^{\mathrm{red}}(\varepsilon, A, A_\tau) = {} & \textstyle\sum_{\mathbb{Z}} \Big(\frac{\varepsilon}{2}\omega^2|A\mathbf{E} - \overline{A}\,\overline{\mathbf{E}}|^2 + \varepsilon^2\mathrm{i}\omega\big(A\mathbf{E} - \overline{A}\,\overline{\mathbf{E}}\big)\big(A_\tau\mathbf{E} + \overline{A}_\tau\overline{\mathbf{E}}\big) \\ & - \tfrac{\varepsilon}{2}|A\vartheta\mathbf{E} + \overline{A}\,\overline{\vartheta}\,\overline{\mathbf{E}}|^2 - \varepsilon^2\big(A\vartheta\mathbf{E} + \overline{A}\,\overline{\vartheta}\,\overline{\mathbf{E}}\big)\big(A_y\mathbf{E} + \overline{A}_y\overline{\mathbf{E}}\big) \\ & - \tfrac{\varepsilon a}{2}|A\mathbf{E} + \overline{A}\,\overline{\mathbf{E}}|^2 - \tfrac{\varepsilon^2 b}{4}|A\mathbf{E} + \overline{A}\,\overline{\mathbf{E}}|^4 + O(\varepsilon^3)\Big) \\ = {} & \varepsilon\mathbb{L}(A, A_\tau) + O(\varepsilon^2) \quad \text{with} \\ \mathbb{L}(A, A_\tau) = {} & \textstyle\int_{\mathbb{R}} \mathrm{i}\omega\big(A\overline{A}_\tau - \overline{A}A_\tau\big) - \big(\vartheta A\overline{A}_y + \overline{\vartheta}\,\overline{A}A_y\big) - \frac{3b}{2}|A|^4\,\mathrm{d}y. \end{aligned}$$

The important observation for this calculation is that the lowest order terms cancel, which can be seen as a manifestation of equipartition of kinetic and potential energy in the plane waves. Moreover, the terms involving $\mathbf{E}^k$ with $k \neq 0$ also drop out by periodicity. This averaging is a formal procedure here, but we will see in the next subsection that in a two-scale setting with an extra phase variable it can be made exact.

Using $\vartheta - \overline{\vartheta} = 2\mathrm{i}\sin\theta = 2\mathrm{i}\omega(\theta)\omega'(\theta)$ the Euler-Lagrange equation reads

$$0 = -\partial_\tau\big(\partial_{\overline{A}_\tau}\mathbb{L}\big) - \partial_y\big(\partial_{\overline{A}_y}\mathbb{L}\big) + \partial_{\overline{A}}\mathbb{L} = -2\mathrm{i}\omega A_\tau + 2\mathrm{i}\omega\omega'\, A_y - 3b|A|^2 A. \tag{31}$$

Of course, this is exactly the desired macroscopic modulation equation, which can be obtained as in Section 3.1. Moreover, because of invariance in τ, there is a first integral, namely the associated energy

$$\mathbb{E}(A, \partial_\tau A) = \textstyle\int_{\mathbb{R}} \mathrm{i}\omega\omega' \left(\overline{A}A_y - A\overline{A}_y\right) + \frac{3b}{2}|A|^4 \,\mathrm{d}y.$$

4.2 Hamiltonian reduction

In the Hamiltonian setting we might also try the derive the reduced Hamiltonian by inserting the multiscale ansatz (10) into the Hamiltonian $\mathcal{H}$ defined in (2). We obtain

$$\widetilde{H}(\varepsilon, \mathbf{A}, \partial_\tau\mathbf{A}) = \varepsilon^\varrho \mathbb{H}(\mathbf{A}, \partial_\tau\mathbf{A}) + O(\varepsilon^{\varrho+1}).$$

In the example of the previous subsection we immediately find $\varrho = 0 < \rho = 1$ and $\mathbb{H}(\mathbf{A}, \partial_\tau\mathbf{A}) = \int_{\mathbb{R}} 2\omega^2|A|^2\,\mathrm{d}y$. Moreover, the symplectic form can be reduced and we obtain

$$\mathbf{\Omega}_\varepsilon^{\mathrm{red}} = \mathbf{\Omega}_0 + O(\varepsilon) \quad \text{with } \mathbf{\Omega}_0 = 2\mathrm{i}\omega.$$

It is easy to see that the function $\mathbb{H}$ is also a first integral of the macroscopic system (31). However, it is not the desired energy $\mathbb{E}$, and the flow associated with the Hamiltonian system $\mathbf{\Omega}_0\partial_\tau A = \mathrm{D}\mathbb{H}(A)$ is the phase translation $A(0,\cdot) \mapsto \mathrm{e}^{-2\mathrm{i}\omega t}A(t,\cdot)$. The discrepancy is easily understood, because in $\mathcal{H}$ the leading terms of the kinetic and potential theory are added while they cancel in $\mathcal{L}$. Note that $\mathbb{H}$ is associated with the phase symmetry of (31) that is not present in the original discrete system. It is introduced into the problem via the multiscale ansatz and it manifests itself only in the limit.

Thus, to treat the Hamiltonian limit correctly it is suitable to embed the discrete Hamiltonian system into a continuous one that has the corresponding symmetries. In this systems we can compensate for drifts in the phases via the phase velocity and for drifts with the group velocities by going into suitably moving frames. On the level of Hamiltonians this leads to a subtraction of the corresponding first integrals. The terms balance in exactly the right way such that the same cancellations occur as in the Lagrangian setting. This is the content of the following classical result in the theory of Hamiltonian systems with symmetry.

Proposition 4.1. *Let $(Z, \mathcal{H}, \mathbf{\Omega})$ be a Hamiltonian system, which is equi-variant with respect to the one-parameter symmetry group $(T_\alpha)_{\alpha\in\mathbb{R}}$ with associated first integral $\mathcal{I}$. Then $z : [0,T] \to Z$ solves $\mathbf{\Omega}\dot{z} = \mathrm{D}\mathcal{H}(z)$ if and only if $\widetilde{z} : t \mapsto T_{ct}z(t)$ solves $\mathbf{\Omega}\dot{\widetilde{z}} = \mathrm{D}\widetilde{\mathcal{H}}_c(\widetilde{z})$, where $\widetilde{\mathcal{H}}_{c,\omega} = \mathcal{H} - c\mathcal{I}$.*

We illustrate the idea in the pulse propagation problem treated in the previous subsection. The continuous Hamiltonian system is defined on the cylinder space-phase $\Xi = \mathbb{R} \times \mathbb{S}^1$ and has the configuration space $\mathrm{L}^2(\Xi)$. For functions $u \in \mathrm{L}^2(\Xi)$ we consider the system

$$\begin{aligned} &\partial_t^2 u = \Delta_{(1,0)}u - au + bu^3 \quad \text{with } a > 0 \text{ and} \\ &\Delta_{(\varepsilon,\delta)}u(\eta,\phi) = u(\eta{+}\varepsilon, \phi{+}\delta) - 2u(\eta,\phi) + u(\eta{-}\varepsilon, \phi{-}\delta). \end{aligned} \tag{32}$$

Introducing $p = \partial_\tau u$ this is a canonical Hamiltonian system with

$$H^{\mathrm{cont}}(u,p) = \int_\Xi \tfrac{1}{2}p^2 + \tfrac{1}{2}\big(\nabla_{(1,0)}u\big)^2 + \tfrac{a}{2}u^2 + \tfrac{b}{4}u^4 \,\mathrm{d}\eta\,\mathrm{d}\phi. \tag{33}$$

Here the important fact is that this system contains the KG chain exactly, because the system decouples completely into an uncountable family of KG chains just displaced by $(\eta,\phi) \in [0,1) \times \mathbb{S}^1$. Moreover, (32) is invariant under translations in the spatial direction η as well as in the phase direction ϕ. This leads to the two first integrals

$$I^{\mathrm{sp}}(u,p) = \int_\Xi p\,\partial_\eta u \,\mathrm{d}\eta\,\mathrm{d}\phi \quad \text{and} \quad I^{\mathrm{ph}}(u,p) = \int_\Xi p\,\partial_\phi u \,\mathrm{d}\eta\,\mathrm{d}\phi. \tag{34}$$

The flows associated with the canonical symplectic structure and with one of these first integral leads to the transport along the corresponding direction with constant speed one.

Using the symmetry of H^{cond} we can go into a frame moving with the phase speed $c_{\mathrm{ph}} = \omega/\theta$. According to Proposition 4.1 the corresponding Hamiltonian is $H^{\mathrm{ph}}(u,p) = H^{\mathrm{cont}}(u,p) - \omega I^{\mathrm{ph}}(u,p)$. Into this Hamiltonian we insert the suitably adjusted multiscale ansatz (30), namely

$$\begin{aligned} u(t,\eta,\phi) &= \varepsilon^{1/2}A(\varepsilon t, \varepsilon\eta)\mathbf{E}_{\mathrm{ph}} + \varepsilon^{1/2}\overline{A}(\varepsilon t, \varepsilon\eta)\overline{\mathbf{E}}_{\mathrm{ph}}, \\ p(t,\eta,\phi) &= \varepsilon^{1/2}\mathrm{i}\omega\big(A(\varepsilon t, \varepsilon\eta)\mathbf{E}_{\mathrm{ph}} - \overline{A}(\varepsilon t, \varepsilon\eta)\overline{\mathbf{E}}_{\mathrm{ph}}\big) \\ &\quad + \varepsilon^{3/2}\big(\partial_\tau A(\varepsilon t, \varepsilon\eta)\mathbf{E}_{\mathrm{ph}} + \partial_\tau \overline{A}(\varepsilon t, \varepsilon\eta)\overline{\mathbf{E}}_{\mathrm{ph}}\big), \end{aligned}$$

where $\mathbf{E}_{\mathrm{ph}} = \mathrm{e}^{\mathrm{i}(\phi+\theta\varepsilon)}$ does no longer depend on time. Through the subtraction of the properly chosen multiple of the corresponding first integral we exactly obtain the cancellation of the leading terms. Moreover, integration over $\phi \in \mathbb{S}^1$ makes all terms $\mathbf{E}_{\mathrm{ph}}^k$ with $k \neq 0$ exactly 0. Hence, the resulting reduced Hamiltonian has the expansion

$$H_\varepsilon^{\mathrm{red}}(A, \partial_\tau A) = \varepsilon\mathbb{E}(A) + O(\varepsilon^2)$$

with $\mathbb{E}$ from above. A simple calculation shows that $\mathbf{\Omega}_0\partial_\tau A = \mathrm{D}\mathbb{E}(A)$ is exactly the macroscopic equation (31).

4.3 Derivation of KdV from the FPU chain

Here we apply both the Lagrangian and Hamiltonian reduction from above to the FPU chain with KdV-multiscale ansatz, see (15). For simplicity we restrict to the infinite chain with $V_1(0) = V_1'(0) = 0$, and we always assume that all arising integrals do exist.

Following the idea in [BP06] we embed the discrete system into a continuous one. For this example we choose the continuous configuration space Q to be $\mathrm{L}^2(\mathbb{R})$ and identify each discrete configuration $(x_j)_{j\in\mathbb{Z}}$ with an piecewise linear function $w = w(\eta) \in \mathrm{L}^2(\mathbb{R})$ defined by $x_j = w(j)$. Since (15), i.e. $w(t,\eta) = \varepsilon U(\varepsilon^3 t, \varepsilon(\eta{+}ct))$, describes slow macroscopic modulations without fast oscillations, there is no need for introding phase variables. The Lagrangian $\mathcal{L}$ of the continuous system is given by $\mathcal{L}(w,\, \dot w) = \mathcal{K}(\dot w) - \mathcal{V}(w)$, with

$$\mathcal{V}(w) = \textstyle\int_{\mathbb{R}} V_1(\nabla_+ w)\,\mathrm{d}\eta, \quad \mathcal{K}(\dot w) = \textstyle\int_{\mathbb{R}} \dot w^2\,\mathrm{d}\eta \tag{35}$$

with $(\nabla_+ w)(\eta) = w(\eta{+}1) - w(\eta)$. The continuous system is invariant under the group of translations, and this gives rise to a further conserved quantity $\mathcal{I}$. Exploiting Noether's theorem we find the first integral $\mathcal{I}(w,\, \dot w) = \int_{\mathbb{R}} \dot w\, \partial_\eta w\,\mathrm{d}\eta$, which has no counterpart in the discrete microscopic FPU chain.

Inserting the ansatz (15) into the energies and using $\int_{\mathbb{R}} \partial_y U \partial_{yy} U \mathrm{d}y = 0$, $\int_{\mathbb{R}} \partial_y U \partial_{yyy} U \mathrm{d}y = -\int_{\mathbb{R}} (\partial_{yy} U)^2 \mathrm{d}y$, and $c^2 = V_1''(0)$ we find

$$\begin{aligned}
\mathcal{K}(\dot w) &= \varepsilon^3\, \tfrac{1}{2}\, \mathbb{H}(U) + \varepsilon^5\, \mathbb{I}(U,\, \partial_\tau U) + O(\varepsilon^7),\\
\mathcal{V}(w) &= \varepsilon^3\, \tfrac{1}{2}\, \mathbb{H}(U) + \varepsilon^5\, \mathbb{E}(U) + O(\varepsilon^7),\\
\mathcal{I}(w,\, \dot w) &= \varepsilon^3\, c^{-1}\, \mathbb{H}(U) + \varepsilon^5\, c^{-1}\, \mathbb{I}(U,\, \partial_\tau U) + O(\varepsilon^7),
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{H}(U) &= c^2 \textstyle\int_{\mathbb{R}} (\partial_y U)^2\,\mathrm{d}y, \quad \mathbb{I}(U,\, \partial_\tau U) = c \int_{\mathbb{R}} \partial_\tau U\, \partial_y U \mathrm{d}y,\\
\mathbb{E}(U) &= -\tfrac{1}{24}\, c^2 \textstyle\int_{\mathbb{R}} (\partial_{yy} U)^2 \mathrm{d}y + \tfrac{1}{6}\, V_1'''(0) \int_{\mathbb{R}} (\partial_y U)^3 \mathrm{d}y.
\end{aligned}$$

Consequently, with $\mathbb{L} = \mathbb{I} - \mathbb{E}$ we find

$$\begin{aligned}
\mathcal{L}(w,\, \dot w) &= \varepsilon^5\, \mathbb{L}(U,\, \partial_\tau U) + O(\varepsilon^7),\\
\mathcal{H}(w,\, \dot w) &= \varepsilon^3\, \mathbb{H}(U) + \varepsilon^5\, \mathbb{I}(U,\, \partial_\tau U) + \varepsilon^5\, \mathbb{E}(U) + O(\varepsilon^7),\\
\mathcal{H}(w,\, \dot w) - c\,\mathcal{I}(w,\, \dot w) &= \varepsilon^5\, \mathbb{E}(U) + O(\varepsilon^7),
\end{aligned}$$

and it follows that the reduced Lagrangian equation equals (16).

In the next step we reduce the Hamiltonian structure. For the microscopic continuous system the canonical momentum is given by $p = \dot w$ with Hamiltonian $\mathcal{H}(w,p) = \mathcal{K}(p) + \mathcal{V}(w)$. For (w,p) the multiscale ansatz (15) means

$$(w,p) = R_\varepsilon(U)(\eta) = \big(\varepsilon U(\varepsilon\eta),\, \varepsilon^4\, \partial_\tau U(\varepsilon\eta) + \varepsilon^2\, c\, \partial_y U(\varepsilon\eta)\big),$$

where the last term is due to the frame moving with speed c. Reduction of the canonical symplectic form $\mathbf{\Omega}$ with $\langle \mathbf{\Omega}\,(w,\, p),\, (\tilde w,\, \tilde p)\rangle = \int_{\mathbb{R}} w\tilde p {-} \tilde w p\,\mathrm{d}\eta$ leads to

$$\begin{aligned}
&\langle \mathbf{\Omega}\, R_\varepsilon(U), R_\varepsilon(\tilde U)\rangle = \varepsilon^2 \langle \mathbf{\Omega}^{\mathrm{red}} U, \tilde U\rangle + O(\varepsilon^4) \text{ with}\\
&\langle \mathbf{\Omega}^{\mathrm{red}} U, \tilde U\rangle = c \textstyle\int_{\mathbb{R}} \big(U \partial_y \tilde U {-} \tilde U \partial_y U\big)\,\mathrm{d}\eta = -2\, c \int_{\mathbb{R}} \partial_y U\, \tilde U\,\mathrm{d}\eta.
\end{aligned}$$

From this we conclude $\mathbf{\Omega}^{\rm red} = -2\,c\,\partial_y$. Note that $\mathbf{\Omega}^{\rm red}$ is defined on $\mathrm{L}^2(\mathbb{R})$, whereas $\mathbf{\Omega}$ lives on $\mathrm{L}^2(\mathbb{R})\times\mathrm{L}^2(\mathbb{R})$. This dimension reduction is natural, because the multiscale ansatz (15) yields a coupling of w and p in leading order. Finally it follows immediately that the reduced Hamiltonian equation $\mathbf{\Omega}^{\rm red}U_\tau = \mathrm{D}\mathbb{E}(U)$ is again equivalent to (16).

4.4 Derivation of nlS from the KG chain

We consider the KG chain (9) with $V_0(x) = \frac{a}{2}x^2 + \frac{b}{4}x^4$. The sum of the kinetic and potential energy gives the Hamiltonian

$$H(\mathbf{x},\dot{\mathbf{x}}) = \textstyle\sum_{j\in\mathbb{Z}} \left(\frac12\dot{x}_j^2+\frac12(x_{j+1}-x_j)^2+\frac{a}{2}x_j^2+\frac{b}{4}x_j^4\right).$$

Since we are interested in modulated pulses, we proceed as in Section 4.2 and embed the discrete chain on $\mathbb{Z}$ into the cylinder $\Xi = \mathbb{R}\times\mathbb{S}^1$ leading to the continuous Hamiltonian system (32) with Hamiltonian $H^{\rm cont}$ in (33).

Again we have the two symmetries of spatial translations $T^{\rm sp}$ and phase translations $T^{\rm ph}$ leading to the two first integrals $I^{\rm sp}$ and $I^{\rm ph}$ given in (34). However, we proceed differently, because we are interested in a dispersive ansatz $u(t,\eta) = \varepsilon A(\varepsilon^2 t, \varepsilon(\eta+ct))\mathbf{E} + \text{c.c.} + \text{h.o.t.}$, where $c = c_{\rm gr}$, cf. (18). Thus, we apply Proposition 4.1 using the symmetry transformation

$$(\widetilde{u},\widetilde{p}) = T^{\rm sp}_{ct}T^{\rm ph}_{(\omega-c\theta)t}(u,p), \qquad \widetilde{\mathcal{H}} = \mathcal{H} - cI^{\rm sp} - (\omega - c\theta)I^{\rm ph}.$$

The associated canonical Hamiltonian system $\mathbf{\Omega}^{\rm can}(\widetilde{u},\widetilde{p}) = \mathrm{D}\widetilde{\mathcal{H}}(\widetilde{u},\widetilde{p})$ on $\mathrm{L}(\Xi)^2$ is still fully equivalent to a family of uncoupled KG chains.

Inserting the scaling exposes the macroscopic behavior. For this define

$$(u(\eta,\phi), p(\eta,\phi)) = (\varepsilon U(\varepsilon\eta, \phi-\theta\eta), \varepsilon P(\varepsilon\eta,\phi-\theta\eta)),$$

which keeps the canonical structure, if we move a factor the ε, which arises from the transformation rule $\mathrm{d}y = \varepsilon\,\mathrm{d}\eta$, into a the time parametrization $\tau = \varepsilon^2 t$. We obtain the new Hamiltonian

$$\begin{aligned}\mathcal{H}_\varepsilon(U,P) = \textstyle\int_\Xi \frac{1}{2\varepsilon^2}\Big(&\big[P-\omega U_\phi - \varepsilon c U_y\big]^2 + \big(\nabla_{(\varepsilon,\theta)}U\big)^2 \\ &+aU^2 - \big[\omega P U_\phi + \varepsilon c P U_y\big]^2\Big) + \tfrac{b}{4}U^4\,\mathrm{d}y\,\mathrm{d}\phi,\end{aligned}$$

where $\nabla_{(\varepsilon,\theta)}U(y,\phi) = U(y+\varepsilon,\phi+\theta) - U(y,\phi)$. Now we see that the suitably transformed version of the modulational ansatz (18), viz.,

$$(U(y,\phi), P(y,\phi)) = R_\varepsilon(A)(y,\phi) = (\mathrm{Re}\,A(y)\mathrm{e}^{\mathrm{i}\phi}, \omega\,\mathrm{Re}\,A(y)\mathrm{e}^{\mathrm{i}\phi}) + O(\varepsilon),$$

leads to the expansion

$$\mathcal{H}_\varepsilon(R_\varepsilon(A)) = \mathbb{H}_{\rm nlS}(A) + O(\varepsilon) \text{ with } \mathbb{H}_{\rm nlS}(A) = \textstyle\int_\mathbb{R} \omega\omega''|A_y|^2 + \frac{3b}{8}|A|^4\,\mathrm{d}y$$

and the reduced symplectic structure $\mathbf{\Omega}^{\rm red} = 2\mathrm{i}\omega$. Thus, we recover the one-dimensional version of nlS given in Section 3.4.

5 Weak convergence methods

For static problems there is a rich literature concerning the Γ-convergence of potential energy functionals of discrete models to continuum models (cf. [FJ00, FT02, BG02a, BG02b, MBL06]). Here we want to summarize some first results for dynamic problems that rely on weak convergence.

5.1 An abstract weak convergence result

In [Mie06a] it was shown that linear elastodynamics can be derived from a general linear lattice model as described in Section 2. However, this result used exact periodicity and linearity in an essential way. The abstract approach presented here will be discussed in [Mie06b] in full details. Its main advantage lies in the flexibility, which allows for applications in nonlinear and macroscopically heterogeneous settings.

We consider a family of Hamiltonian systems parametrized by $\varepsilon \in [0, 1]$,

$$\mathbf{\Omega}_\varepsilon(z)\dot{z} = \mathrm{D}\mathcal{H}_\varepsilon(z), \tag{36}$$

and we are interested in the limit behavior for $\varepsilon \to 0$. Again, ε measures the ratio between the microscopic and the macroscopic spatial scales, viz., $y = \varepsilon\gamma$.

We consider the situation that all $\mathcal{H}_\varepsilon$ are defined on one reflexive Banach space Z, but may take the value $+\infty$ outside the subspace Z_ε. It is a question of general interest to characterize the further conditions on the convergence of $\mathcal{H}_\varepsilon$ to $\mathcal{H}_0$ and of $\mathbf{\Omega}_\varepsilon$ to $\mathbf{\Omega}_0$ such that suitable limits z of solutions z_ε of (36) are solutions of the limit problem (36) for $\varepsilon = 0$. A first guess would be that $\mathcal{H}_0$ is the Γ-limit of $\mathcal{H}_\varepsilon$, i.e.

$$\begin{array}{lll} \text{(G1)} & z_\varepsilon \rightharpoonup z \quad \Longrightarrow \quad \mathcal{H}_0(z) \leq \liminf\limits_{\varepsilon\to 0} \mathcal{H}_\varepsilon(z_\varepsilon), \\ \text{(G2)} & \forall\, z \in Z\ \exists (\widetilde{z}_\varepsilon)_{\varepsilon\in(0,1)} :\ \widetilde{z}_\varepsilon \rightharpoonup z \text{ and } \mathcal{H}_0(z) = \lim\limits_{\varepsilon\to 0} \mathcal{H}_\varepsilon(\widetilde{z}_\varepsilon). \end{array}$$

However, we will see below that it cannot be expected in general.

We assume that the subspaces $Z_\varepsilon \subset Z$ are closed and that $\mathcal{H}_\varepsilon \in \mathrm{C}^1(Z_\varepsilon, \mathbb{R})$ for $\varepsilon \in [0, 1]$. Moreover, there exist mappings $G_\varepsilon \in \mathrm{Lin}(Z_0, Z_\varepsilon)$ such that we have

$$Z_\varepsilon \ni z_\varepsilon \rightharpoonup z \in Z_0 \quad \Longrightarrow \quad G_\varepsilon^* \mathrm{D}\mathcal{H}_\varepsilon(z_\varepsilon) \rightharpoonup \mathrm{D}\mathcal{H}_0(z) \text{ in } Z_0^*. \tag{37}$$

Finally we assume that the symplectic operators $\mathbf{\Omega}_\varepsilon$ are independent of $z \in Z$ and that there exists a larger Banach space W such that Z embeds continuously and densely into W such that $\mathbf{\Omega}_\varepsilon : W \to Z^*$ has an inverse operator for all $\varepsilon \in [0, 1]$ with the norm bounded independently of ε. For the convergence we ask the condition

$$Z_\varepsilon \ni z_\varepsilon \rightharpoonup z \in Z_0 \quad \Longrightarrow \quad G_\varepsilon^* \mathbf{\Omega}_\varepsilon z_\varepsilon \rightharpoonup \mathbf{\Omega}_0 z \text{ in } Z^*. \tag{38}$$

Now we use the fact that solutions z_ε of (36) also solve the weak equation

$$\textstyle\int_0^T \langle \mathrm{D}\mathcal{H}_\varepsilon(z_\varepsilon(t)), \varphi_\varepsilon(t)\rangle + \langle \mathbf{\Omega}_\varepsilon z_\varepsilon(t), \dot{\varphi}_\varepsilon(t)\rangle \,\mathrm{d}t - \langle \mathbf{\Omega}_\varepsilon z_\varepsilon, \varphi_\varepsilon\rangle\big|_0^T = 0 \tag{39}$$

for all $\varphi_\varepsilon \in \mathrm{C}^1([0,T],Z_\varepsilon)$. Choosing $\varphi_\varepsilon(t) = G_\varepsilon\varphi(t)$ for some $\varphi \in \mathrm{C}^1([0,T],Z_0)$ and using suitable a priori bounds on z_ε in $\mathrm{C}^0([0,T],Z) \cap \mathrm{C}^1([0,T],W)$ it is possible to extract a weakly convergent subsequence with $z_\varepsilon(t) \rightharpoonup z(t)$ for some $z \in \mathrm{C}^0([0,T],Z_\mathrm{w}) \cap \mathrm{L}^\infty([0,T],W)$. By the assumptions (37) and (38) we pass to the limit in (39) and obtain

$$\int_0^T \langle \mathrm{D}\mathcal{H}_0(z),\varphi\rangle + \langle \mathbf{\Omega}_0 z,\dot\varphi\rangle\,\mathrm{d}t - \langle \mathbf{\Omega}_0 z,\varphi\rangle\big|_0^T = 0.$$

Under suitable assumptions it then follows that z solves (36) for $\varepsilon = 0$.

5.2 Elastodynamics

The program described in the previous subsection can be applied to polyatomic Klein–Gordon chains, which we also allow to have large-scale variations in the stiffness and masses. The KG chains under consideration are assumed to have a periodicity of N on the microscopic level, and all quantities may change also on the macroscopic scale $y = \varepsilon j$. For $k \in \mathbb{Z}_N = \{\, j \bmod N \mid j \in \mathbb{Z}\,\}$ we have given functions $m_k, a_k, b_k, c_k \in \mathrm{L}^\infty(\mathbb{R})$, which are all bounded from below by a positive constant. The KG chain is then given by the canonical Hamiltonian system on $\ell^2 \times \ell^2$

$$\begin{aligned}\mathcal{H}_\varepsilon^{\mathrm{discr}}(\mathbf{x},\mathbf{p}) = \sum_{j\in\mathbb{Z}} \Big(&\tfrac{p_j^2}{2m_{[j]}(\varepsilon j)} + \tfrac{a_{[j]}(\varepsilon j)}{2}(x_{j+1}-x_j)^2 \\ &+ \tfrac{\varepsilon^2 b_{[j]}(\varepsilon j)}{2}x_j^2 + \tfrac{\varepsilon^2 c_{[j]}(\varepsilon j)}{4}x_j^4\Big),\end{aligned} \tag{40}$$

where $[j] = j \bmod N$. To derive a suitable continuum model we embed $\ell^2 \times \ell^2$ into $Z = Z_0 = \mathrm{H}^1(\mathbb{R}) \times \mathrm{L}^2(\mathbb{R})$ via

$$\begin{aligned}&Z_\varepsilon = \{\, (u,v) \in Z \mid u|_{[\varepsilon j,\varepsilon j+\varepsilon]} \text{ affine, } v|_{(\varepsilon j-\varepsilon/2,\,\varepsilon j+\varepsilon/2)} \text{ constant}\,\}\text{and} \\ &(u,v) = E_\varepsilon(\mathbf{x},\mathbf{p}) \text{ with } (u(\varepsilon j), v(\varepsilon j)) = (x_j,p_j) \text{ for all } j \in \mathbb{Z}.\end{aligned} \tag{41}$$

The associated Hamiltonian $\mathcal{H}_\varepsilon$ coincides with $\mathcal{H}_\varepsilon^{\mathrm{discr}}$ up to a factor ε, which relates to the time rescaling, namely $\mathcal{H}_\varepsilon(u,v) =$

$$\int_\mathbb{R} \tfrac{v(y)^2}{2M(y,y/\varepsilon)} + \tfrac{A(y,y/\varepsilon)}{2}u'(y)^2\,\mathrm{d}y + \sum_{j\in\mathbb{Z}} \varepsilon\big(\tfrac{B(\varepsilon j,j)}{2}u(\varepsilon j)^2 + \tfrac{C(\varepsilon j,j)}{4}u(\varepsilon j)^4\big),$$

where $M(y,z) = m_{[k]}(y)$ for $z \in (k-1/2,\, k+1/2)$, $A(y,z) = a_{[k]}(y)$ for $z \in (k,\, k+1)$ for $k \in \mathbb{Z}$, with similar formulas for B and C.

The important step in the analysis is the construction of the operator $G_\varepsilon\colon Z_0 \to Z_\varepsilon$. We define $(u_\varepsilon, v_\varepsilon) = G_\varepsilon(u,v)$ via $v_\varepsilon(y) = \frac{M(y,y/\varepsilon)}{M^*(y)}v(y)$ and

$$\int_\mathbb{R} A(y,y/\varepsilon)u_\varepsilon'(y)\widetilde{u}'(y) + u_\varepsilon(y)\widetilde{u}(y)\,\mathrm{d}y = \int_\mathbb{R} A^*(y)u'\widetilde{u}' + u\widetilde{u}\,\mathrm{d}y$$

for all $\widetilde{u}$ with $(\widetilde{u},0) \in Z_\varepsilon$, see (41). Here A^* is the averaged stiffness and M^* the averaged masses

$$A^*(y) = \big(\tfrac{1}{N}\int_0^N A(y,z)^{-1}\,\mathrm{d}z\big)^{-1} \quad \text{and} \quad M^*(y) = \tfrac{1}{N}\int_0^N M(y,z)\,\mathrm{d}z.$$

It is then possible to prove the abstract conditions 37 and 38, which leads to the following results, cf. [Mie06b].

Theorem 5.1. *Let $E_\varepsilon : \ell^2 \times \ell^2 \to Z = \mathrm{H}^1(\mathbb{R})$ be the embedding in* (41). *Let $(x^\varepsilon, p^\varepsilon) : [0, T/\varepsilon] \to \ell^2 \times \ell^2$ be solutions of the canonical Hamiltonian system associated with $\mathcal{H}_\varepsilon^{\mathrm{discr}}$ in* (40)*. If for $\tau = 0$ we have*

$$\begin{pmatrix} I & 0 \\ 0 & M(\cdot, \cdot/\varepsilon) \end{pmatrix} E_\varepsilon \begin{pmatrix} \mathbf{x}^\varepsilon(\tau/\varepsilon) \\ \varepsilon \mathbf{p}^\varepsilon(\tau/\varepsilon) \end{pmatrix} \rightharpoonup \begin{pmatrix} u(\tau) \\ M^*(\cdot)v(\tau) \end{pmatrix} \quad \text{in } Z,$$

then this convergence holds for all $\tau \in [0, T]$, where $(u, v) : [0, T] \to Z$ is a solution of the macroscopic wave equation arising from the canonical Hamiltonian system with

$$\mathcal{H}_0(u, v) = \int_{\mathbb{R}} \tfrac{1}{2M^*(y)} v^2 + \tfrac{A^*(y)}{2}(u')^2 + \tfrac{B^*(y)}{2} u^2 + \tfrac{C^*(y)}{4} u^4 \,\mathrm{d}y,$$

where $B^*(y) = \frac{1}{N} \int_0^N B(y, z)\,\mathrm{d}z$ and $C^*(y) = \frac{1}{N} \int_0^N c(y, z)\,\mathrm{d}z$.

It should be noted that $\mathcal{H}_0$ is not the Γ-limit of $\mathcal{H}_\varepsilon$ when using canonical variables. However, if we use the Lagrangian coordinates $(u^\varepsilon, \dot{u}^\varepsilon) = (u^\varepsilon, M(\cdot, \cdot/\varepsilon)^{-1} p^\varepsilon)$, then it is the Γ-limit.

5.3 Energy transport via Wigner-Husimi measures

Waves in dispersive media travel with a speed that depends on their wave length. We now discuss this for the general linear model introduced in Section 2.2. Wave propagation is driven by the group velocity $c_{\mathrm{gr}} = \nabla \Omega_j(\theta)$, which depends on the wave vector $\theta \in \mathcal{T}_\Gamma$ and the band number $j \in \{1, \ldots, m\}$. Thus, at each macroscopic point $y \in \mathbb{R}^d$ we need to know how much energy is located in which band and in which wave-vector regime.

The relevant mathematical tool is the Wigner measure or the Husimi measure, which was used in [Gér91, LP93, MMP94, GMMP97, TP04] to study transport of oscillations (relating to energy, density, or other physical quantities). The case of discrete lattices is analyzed in detail in [Mac04, Mie06a]. For this we rewrite (6) into diagonal and rescaled form

$$\frac{\partial}{\partial \tau} U^\varepsilon(\tau, \theta) = \mathbb{B}(\varepsilon, \theta) U^\varepsilon(\tau, \theta) \text{ with } \mathbb{B}(\varepsilon, \theta) = \frac{\mathrm{i}}{\varepsilon} \mathrm{diag}(\Omega_1(\theta), ..., \Omega_m(\theta)). \tag{42}$$

The Wigner transform $W^\varepsilon[u^\varepsilon]$ of $u^\varepsilon = \mathbb{F}^{-1} U^\varepsilon$ is now defined as a matrix-valued distribution on $\mathbb{R} \times \mathcal{T}_\Gamma$. For the diagonal entries it is possible to pass to the limit $\varepsilon \to 0$ and one finds the Wigner measure $\mu_j^{\mathrm{W}} = \lim_{\varepsilon \to 0} (W^\varepsilon[u^3])_{jj}$. More precisely, we have the following result, see [Mie06a].

Theorem 5.2. *Let $u^\varepsilon : [0, T] \to \mathrm{L}^2(\mathcal{T}_\Gamma, \mathbb{C}^m)$ be a family of solutions for* (42) *with $\|u^\varepsilon(0)\|_{\mathrm{L}^2} \leq C$. Let $j \in \{1, ..., m\}$ and $S_j \subset \mathcal{T}_\Gamma$ be given such that $\Omega_j \in \mathrm{C}^1(\mathcal{T}_\Gamma \backslash S_j)$. If for $\tau = 0$ we have*

$$\lim_{\varepsilon \to 0} (W^\varepsilon[u^\varepsilon](\tau))_{jj} = \mu_j^{\mathrm{W}}(\tau) \text{ in } \mathcal{D}(\mathbb{R}^d \times \mathcal{T}_\Gamma) \text{ and } \mu_j^{\mathrm{W}}(0, \mathbb{R}^d \times S_j) = 0,$$

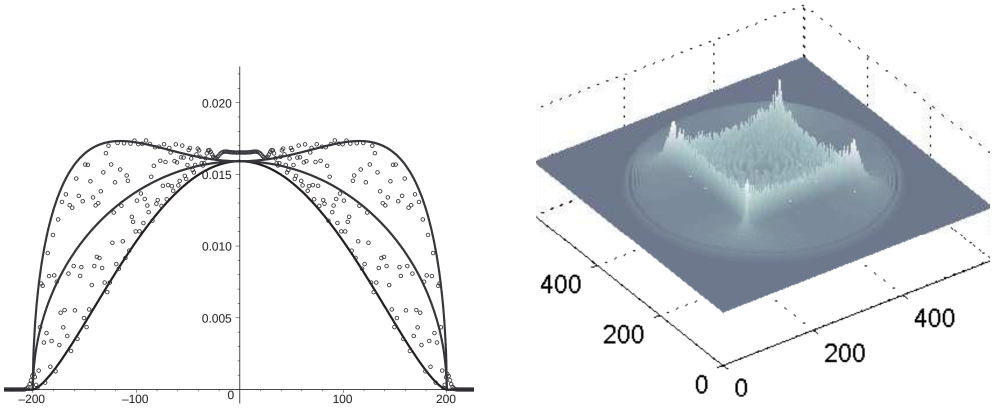

Figure 5.1: Right: energy distribution at $t = 200$ for the linear chain $\ddot{x}_j = x_{j+1} - 2x_j + x_{j-1}$ with initial data $x_j(0) = \delta_j$ and $\dot{x}_j(0) = 0$. Right: Energy distribution for the square lattice $\mathbb{Z}^2$ with simple nearest-neighbor interaction at time $t = 120$.

then this convergence holds for all $\tau \in [0, T]$*, where* $\mu_j^{\mathrm{W}} : [0, T] \to \mathcal{M}(\mathbb{R}^d \times \mathcal{T}_\Gamma)$ *is a solution of the energy-transport equation*

$$\partial_\tau \mu_j^{\mathrm{W}} = \nabla\Omega_j(\theta) \cdot \partial_y \mu_j^{\mathrm{W}} \quad \textit{on } [0, T] \times \mathbb{R}^d \times \mathcal{T}_\Gamma.$$

Using this result it is possible to obtain the energy distribution by integration over θ, namely

$$e(\tau, y)\mathrm{d}y = \sum_{j=1}^{m} \int_{\theta \in \mathcal{T}_\Gamma} \mu_j^{\mathrm{W}}(0,\, y - \nabla\Omega_j(\theta)\tau, \mathrm{d}\theta).$$

The above theorem is restricted to the case that no mass concentrates on the singular set S_j, where the dispersion relation is not smooth and, hence, the group velocity is not defined. However, using the Husimi measure as developed in [Mie06a] it is possible to treat this case also in some cases.

6 Quantitative estimates via Gronwall estimates

Another technique for the justification of continuum models uses quantitative estimates to control the error between the macroscopic equation and the microscopic equation. We present the abstract idea in Section 6.1 and apply in Section 7. This method can also be used to prove dispersive stability results as discussed Section 6.2.

We work totally in the original microscopic lattice model

$$\dot{z} = Lz + \mathcal{N}(z), \tag{43}$$

where Z is the Banach space for the state $z(t)$, and $L : Z \to Z$ is the linear part, which is assumed to generate a bounded semigroup $(\mathrm{e}^{Lt})_{t \geq 0}$, i.e.

$$\exists\, C_L > 0\ \forall\, t \geq 0\ \forall\, z \in Z : \quad \|\mathrm{e}^{Lt} z\| \leq C_L \|z\|. \tag{44}$$

We also rely on our standard assumption that the solution $z = 0$ is energetically stable, as the Hamiltonian energy is conserved. Here it means

$$\exists\, C_E > 0\ \forall \text{ sln. } z \text{ of (43) } \forall\, t \geq 0: \quad \|z(t)\| \leq C_E \|z(0)\|. \tag{45}$$

The nonlinearity $\mathcal{N}: \ Z \to Z$ is assumed to be locally Lipschitz. However, the essential features have to be addressed by using additional Banach spaces Y and W such $Y \subset Z \subset W$ with continuous embeddings and

$$\text{(i) } \forall\, z \in Z\text{: } \|z\|_W \leq \|z\| \qquad \text{(ii) } \forall\, \widetilde{z} \in Y\text{: } \|\widetilde{z}\| \leq \|\widetilde{z}\|_Y. \tag{46}$$

In applications to lattices we have in mind

$$Y = \ell_1(\Gamma, \mathbb{R}^k)^2, \quad Z = \ell_2(\Gamma, \mathbb{R}^k)^2, \quad W = \ell_\infty(\Gamma, \mathbb{R}^k)^2. \tag{47}$$

In Section 6.1 the importance is that $\mathcal{N}$ satisfies

$$\begin{gathered} \exists\, C > 0\ \exists\, \nu > 0\ \forall\, z_1, z_2 \in Z \text{ with } \|z_1\|_W, \|z_2\|_W \leq 1: \\ \|\mathcal{N}(z_1) - \mathcal{N}(z_2)\| \leq C_{\mathcal{N}} \big(\|z_1\|_W + \|z_2\|_W\big)^\nu \|z_1 - z_2\|. \end{gathered} \tag{48}$$

In Section 6.2 the importance of Y is the *dispersive decay estimate*

$$\exists\, \kappa \in (0,1)\ \exists\, C_W > 0\ \forall\, z \in Y\ \forall\, t > 0: \quad \|\mathrm{e}^{Lt} z\|_W \leq \frac{C_W}{(1+t)^\kappa} \|z\|_Y. \tag{49}$$

of the linear semigroup. For the nonlinearity we then use

$$\exists\, C_{\mathcal{N}} > 0\ \exists\, \alpha, \nu > 0\ \forall\, z \in Z: \quad \|\mathcal{N}(z)\|_Y \leq C_{\mathcal{N}} \|z\|_W^\nu \|z\|^\alpha. \tag{50}$$

With Y, Z and W as in (47) a standard nonlinearity $\mathcal{N}((x_\gamma)_{\gamma\in\Gamma}) = (n(x_\gamma))_{\gamma\in\Gamma}$ with $|n(\xi_1) - n(\xi_2)| \leq C(|\xi_1| + |\xi_2|)^\beta |\xi_1 - \xi_2|$ will satisfy (48) with $\nu = \beta$ and (50) with $\nu = \beta - 1$ and $\alpha = 2$.

6.1 Error control for approximate solutions

The basic idea is to construct an approximate solution z_{app}, which in fact will be given in the form $z_{\text{app}} = R_\varepsilon(\mathbf{A})$, and to derive an estimate for the associated error. For any $z \in \mathrm{C}^1([0,T], Z)$ we define the *residual* via

$$\mathrm{Res}(z)(t) = \dot{z}(t) - Lz(t) - \mathcal{N}(z(t)). \tag{51}$$

The following result shows that the smallness of the residual together with the stability condition (44) implies that the error between z_{app} and an exact solution is small.

Theorem 6.1. *Assume that the conditions* (44), (46i) *and* (48) *hold. Moreover, let* C_R, C_A, τ_*, σ, α, $\varrho > 0$ *be given as well as a family* $(z^\varepsilon_{\mathrm{app}})_{\varepsilon\in(0,1)}$ *of approximate solutions* $z^\varepsilon_{\mathrm{app}} \in \mathrm{C}^1([0,\tau_*/\varepsilon^\sigma], Z)$ *satisfying*

$$\|z^\varepsilon_{\mathrm{app}}(t)\|_W \le C_A\varepsilon^\alpha \quad \text{and} \quad \|\mathrm{Res}(z^\varepsilon_{\mathrm{app}})(t)\| \le C_R\varepsilon^\varrho \tag{52}$$

for all $t \in [0,\tau_*/\varepsilon^\sigma]$. *Moreover, assume*

$$\varrho > \alpha + \sigma \quad \text{and} \quad \nu\alpha \ge \sigma. \tag{53}$$

Then, for each $d > 0$ *there exist* $\varepsilon_0 \in (0,1)$ *and* $D > 0$ *such that for all* $\varepsilon \in (0,\varepsilon_0]$ *any exact solution* z *of* (43) *with* $\|z(0)-z^\varepsilon_{\mathrm{app}}(0)\| \le d\varepsilon^{\varrho-\sigma}$ *satisfies*

$$\|z(t) - z^\varepsilon_{\mathrm{app}}(t)\| \le D\varepsilon^{\varrho-\sigma} \text{ for } t \in [0,\tau_*/\varepsilon^\sigma]. \tag{54}$$

In (53) the case $\nu\alpha > \sigma$ is not really interesting, as in this regime the nonlinearity is not really active. In the first inequality ϱ may be as big as we like, what improves the order of approximation in (54) but does not allow us to extend the length of the time interval, i.e, to make σ bigger, because it is restricted by the second inequality.

Proof: For the construction of ε and D we define $C_1 = C_L(d + C_R\tau_*)$ and $C_2 = C_LC_{\mathcal{N}}(3C_A)^\nu$ and let $D = 2C_1\mathrm{e}^{C_2\tau_*}$ and $\varepsilon_0 = \min\{1,(C_A/D)^\delta\}$, where $\delta = 1/(\varrho - \alpha - \sigma) > 0$.

We write the exact solution z of (43) in the form $z(t) = z^\varepsilon_{\mathrm{app}}(t) + \varepsilon^\beta R(t)$ with $\beta = \varrho - \sigma$. Clearly $\|R(0)\| \le d$ and we have to show $\|R(t)\| \le D$ for all $t \in [0,\tau_*/\varepsilon^\sigma]$. Inserting this ansatz into (43) and applying the variation-of-constants formula we find

$$R(t) = \mathrm{e}^{Lt}R(0) + \int_0^t \tfrac{\mathrm{e}^{L(t-s)}}{\varepsilon^\alpha}\big(\mathcal{N}(z^\varepsilon_{\mathrm{app}}(s) + \varepsilon^\beta R(s)) - \mathcal{N}(z^\varepsilon_{\mathrm{app}}(s)) - \mathrm{Res}(z^\varepsilon_{\mathrm{app}})(s)\big)\,\mathrm{d}s.$$

Defining $r(t) = \|R(t)\|$ and using the available estimates give

$$r(t) \le C_Ld + \int_0^t C_L\big(C_{\mathcal{N}}[C_A\varepsilon^\alpha + C_A\varepsilon^\alpha + \varepsilon^\beta D]^\nu r(s) + C_R\varepsilon^{\varrho-\beta}\big)\,\mathrm{d}s,$$

where we assumed $r(s) \le D$ on $[0,t_D]$ and $t \le t_D$. Note that $d < D$ and r is continuous, which implies $t_D > 0$. We will show that $t_D = \tau_*/\varepsilon^\sigma$.

Assuming $\varepsilon \in (0,\varepsilon_0]$ we arrive at $r(t) \le C_Ld + C_LC_R\varepsilon^\sigma t + C_2\varepsilon^{\alpha\nu}\int_0^t r(s)\,\mathrm{d}s$. Because of $\varepsilon^\sigma t \le \tau_*$ we find $r(t) \le C_1 + \varepsilon^\sigma C_2\int_0^t r(s)\,\mathrm{d}s$ and Gronwall's lemma gives $r(t) \le C_1\mathrm{e}^{C_2\varepsilon^\sigma t} \le C_1\mathrm{e}^{C_2\tau_*} = D/2$ for all $t \in [0,t_D]$. However, this shows that $r(t)$ cannot reach D. As a consequence we may choose $t_D = \tau_*/\varepsilon^\sigma$ and we are done. ∎

6.2 Dispersive stability

Here we present conditions which guarantee that the dispersive decay estimate (49) for the linear semigroup can be transfered to the full nonlinear problem. We follow ideas from [Sch96, MSU01] and refer to [Pat06] for more satisfactory results.

Theorem 6.2. *Assume that* (45), (49), *and* (50) *hold with* $\nu\kappa > 1$. *Then, there exist* $C, \eta > 0$ *such that all solutions* z *of* (43) *with* $\|z(0)\|_Y \leq \eta$ *satisfy*

$$\|z(t)\|_W \leq \frac{C}{(1+t)^\kappa}\|z(0)\|_Y \quad \textit{for all} \quad t > 0. \tag{55}$$

Proof: We follow the ideas in [MSU01] Lemma 3 and adapt it to the more general case. We rely on $0 < \kappa < 1 < \nu\kappa$, which yield the estimate

$$\int_0^t \frac{\mathrm{d}s}{(1+s)^{\kappa\nu}(1+t-s)^\kappa} \leq \frac{c_{\nu,\kappa}}{(1+t)^\kappa} \text{ with } c_{\nu,\kappa} = \big(\frac{2^\kappa}{\kappa\nu-1} + \frac{2^{\kappa\nu}}{1-\kappa}\big). \tag{56}$$

This is easily obtained by estimating $\int_0^{t/2}$ and $\int_{t/2}^t$ separately. Using the variation-of-constants formula together with the available estimates we find

$$\|z(t)\|_W \leq \frac{C_W}{(1+t)^\kappa}\|z(0)\|_Y + \int_0^t \frac{C_W}{(1+t-s)^\kappa} C_{\mathcal{N}}\|z(s)\|_W^\nu \|z(s)\|^\alpha \,\mathrm{d}s.$$

With $r(t) = \max\{\, (1+s)^\kappa\|z(s)\|_W \mid s \in [0,t]\,\}$ and $\delta = \|z(0)\|_Y$ we obtain

$$(1+t)^\kappa\|z(t)\| \leq C_W\delta + \int_0^t \frac{C_W C_{\mathcal{N}} C_E^\alpha r(t)^\nu \delta^\alpha}{(1+t-s)^\kappa(1+s)^{\kappa\nu}}\,\mathrm{d}s.$$

Employing (55) and using that r is nondecreasing we find

$$r(t) \leq C_W\delta + C_*\delta^\alpha r(t)^\nu \text{ for all } t \geq 0, \text{ where } C_* = c_{\nu,\kappa}C_W C_{\mathcal{N}} C_E^\alpha.$$

We now choose η such that $C_*\eta^\alpha(3C_W\eta)^\nu \leq C_W\eta$ and claim that $r(t)$ remains less than $3C_W\delta$ if $\|z(0)\|_Y = \delta \leq \eta$, i.e., the desired assertion holds with $C = 3C_W$. Let $t_W = \sup\{\, t \geq 0 \mid \forall\, s \in [0,t] :\ r(s) \leq 3C_W\delta\,\}$, then for $t \in [0, t_W]$ and $0 < \delta \leq \eta$ we have

$$r(t) = C_W\delta + C_*\delta^\alpha(3C_W\delta)^\nu \leq 2C_W\delta < 3C_W\delta.$$

Since r is also continuous, we conclude $t_W = \infty$. ∎

The typical application of the above result involves the spaces $Y = \ell_1$ and $Z = \ell_2$ and $W = \ell_\infty$. Hence, for a nonlinearity with $\mathcal{N}(\mathbf{x}) = (n(x_j))_{j\in\Gamma}$ and $|n(x_j)| \leq C_n|x_j|^\beta$ we have (50) with $\alpha = 2$ and $\nu = \beta{-}2$. Moreover, the theory in [Pat06] provides explicity values of κ, which can be determined directly for the properties of the dispersion relations $\omega = \Omega_m(\theta)$ discussed in Section 2.2. For this note that e^{Lt} can be written as a discrete convolution

$$\mathrm{e}^{Lt}(\mathbf{x}, \dot{\mathbf{x}}) = \big(\textstyle\sum_{\alpha\in\Gamma} G_{\gamma-\alpha}(t)(x_\alpha, \dot{x}_\alpha)\big)_{\gamma\in\Gamma},$$

where the Green's functions $G_\gamma(t) \in \mathbb{R}^{2m\times 2m}$ satisfy $G_0(0) = \mathrm{id}$ and $G_\gamma(0) = 0$ for $\gamma \neq 0$. Each component of each $G_\gamma(t)$ can be calculated via oscillatory integrals of the type

$$\int_{\theta\in\mathcal{T}_\Gamma} \mathrm{e}^{\mathrm{i}(\Omega_k(\theta)t+\theta\cdot\gamma)} g(\theta)\,\mathrm{d}\theta$$

with given smooth functions g. Uniform decay properties in $\gamma \in \Gamma$ for such integrals strongly depend on the non-degeneracy of $\mathrm{D}^2\Omega_k(\theta)$. Integrating over balls in $\mathcal{T}_\Gamma$, where $\det \mathrm{D}^2\Omega_k(\theta)$ is bounded away from 0, we easily obtain a decay like $t^{-d/2}$. However, due to periodicity, degeneracies must occur, and the uniform decay is always worse.

For instance, the one-dimensional FPU and the KG chains from Section 2.3 lead to $\kappa = 1/3$, because $\Omega : \mathbb{S}^1 \to \mathbb{R}$ has turning points and the third derivatives is nonzero in these points. As a consequence the above method leads to the following very preliminary dispersive decay result.

Proposition 6.3. *Consider the KG chain* (9) *with* V_0 *of the form* $V_0'(x) = ax + O(|x|^\beta)$ *for* $|x| \to 0$ *with* $a > 0$ *and* $\beta > 5$. *Then, there exists* $\delta > 0$ *and* $C > 0$ *such that for each initial condition* $(\mathbf{x}(0), \dot{\mathbf{x}}(0))$ *we have*

$$\|(\mathbf{x}(0), \dot{\mathbf{x}}(0))\|_{\ell_1} \leq \delta \implies \|(\mathbf{x}(t), \dot{\mathbf{x}}(t))\|_{\ell_\infty} \leq \tfrac{C\|(\mathbf{x}(0),\dot{\mathbf{x}}(0))\|_{\ell_1}}{(1+t)^{1/3}} \quad \textit{for all } t \geq 0.$$

This result is still very weak in terms of the restriction on β, and we refer to [Pat06] for improved results . See also [Zua05, IZ05] for related dispersive decay results in discrete approximations of PDEs.

7 Justification of modulation equations

In this section we provide rigorous justification results for two examples. In contrast to Section 5 we will use the quantitative estimates provided in Section 6.1. The ideas are based on the justification theory developed for general modulation equations, see [KSM92, Sch94, Sch98] and the surveys [MSU01, Mie02]. In particular, we mention the papers [Sch95, Sch05b], which contain examples, where the modulation equations, derived formally as in Section 3, *fail to predict* the dynamics of the microscopic system correctly. Thus, the justification results are needed to validate the formally obtained macroscopic equations.

To explain the main ideas and still stay sufficiently simple we consider for both subsequent examples the d-dimensional, scalar model (17). The main observation about the multiscale ansatz $x_\gamma^{A,\varepsilon} = \varepsilon^\sigma A(\varepsilon\gamma)\mathbf{E} + \text{c.c.}$ is that it satisfies the estimates

$$\|(x_\gamma^{A,\varepsilon})_{\gamma\in\Gamma}\|_{\ell_2} \leq C_s\varepsilon^{\sigma-d/2}\|A\|_{\mathrm{H}^s} \text{ and } \|(x_\gamma^{A,\varepsilon})_{\gamma\in\Gamma}\|_{\ell_\infty} \leq C_s\varepsilon^\sigma\|A\|_{\mathrm{H}^s},$$

for any $s > d/2$. Thus, our solutions $z = (\mathbf{x}, \dot{\mathbf{x}})$ will be small only in $W = \ell_\infty(\Gamma)^2$ but may be large in $Z = \ell_2(\Gamma)^2$. However, for using the abstract approach provided in Theorem 6.1 we need to make the residual of the approximate solution $z_{\text{app}} = R_\varepsilon(\mathbf{A})$ small in Z. This means that the order of approximation of the formal ansatz R_ε in (10) has to be taken sufficiently high depending on the dimension d.

7.1 Nonlinear Schrödinger equation

We want to justify the nonlinear Schrödinger equation

$$\mathrm{i}\partial_\tau A = \mathrm{div}_y(\frac{1}{2}\mathrm{D}_\theta^2\Omega(\theta)\nabla_y A) + \rho|A|^2 A \qquad \text{(nlS)}$$

as a macroscopic modulation equation for the microscopic lattice system (17), for the formal derivation see Section 3.4. We use the dispersive scaling $\tau = \varepsilon^2 t$ and $y = \varepsilon(t - c_{\rm gr} t)$ for the basic periodic pattern $\mathbf{E} = {\rm e}^{{\rm i}(\omega t + \theta\cdot\gamma)}$, where $\omega = \Omega(\theta)$ and $c_{\rm gr} = \Omega'(\theta)$. To derive an evolution equation for the macroscopic modulation amplitude $A : [0,\infty)\times\mathbb{R}^d \to \mathbb{C}$ we have to use the improved ansatz

$$x_\gamma(t) = R_\varepsilon^K(A)_\gamma(t) := \sum_{k=1}^K \varepsilon^k \sum_{n=-k}^k A_{k,n}(\tau, y)\mathbf{E}^n,$$

where all the coefficient functions $A_{k,n}$ can be calculated formally if the non-resonance condition of order K holds, namely

$$n^2\Omega(\theta)^2 \neq \Omega(n\theta)^2 \text{ for } n = 0, 2, 3, ..., K. \tag{57}$$

Of course, we have $A = A_{1,1}$, where A satisfies (nlS). The other coefficient functions satisfy $A_{k,-n} = \overline{A}_{k,n}$ and are either algebraic expressions of functions $\partial_\tau^r \partial_y^s A_{p,n}^q$ with $r+2|s|+pq = k$, $p \leq k-1$ or (for $n = 1$, where the non-resonance condition fails) they satisfy some linear inhomogeneous Schrödinger-type equations.

Since all coefficients of the terms $\varepsilon^k\mathbf{E}^n$ with $k = 1, ..., K$ are equated to 0, the residual of the ansatz $z_{\rm app} = (R_\varepsilon^K(A), \frac{\rm d}{{\rm d}t}R_\varepsilon^K(A)) : [0, \tau_*/\varepsilon^2] \to Z = \ell(\Gamma)^2$ satisfies

$$\|{\rm Res}(z_{\rm app})(t)\|_{\ell_\infty} \leq C\varepsilon^{K+1}\|A\|_{{\rm H}^s} \text{ and } \|{\rm Res}(z_{\rm app})(t)\|_{\ell_2} \leq C\varepsilon^{K+1-d/2}\|A\|_{{\rm H}^s}$$

for any suitable $s > K+2+d/2$. Thus, we have all the ingredients to apply Theor em 6.1. However, we note that the dispersive time scale $\tau = \varepsilon^2 t$ needs $\sigma = 2$, while the amplitude $\|z_{\rm app}(t)\|_{\ell_\infty} \sim \varepsilon^\alpha$ with $\alpha = 1$. Now condition (53) only holds for $\nu \geq 2$. Thus, the nonlinearity $\mathcal{N}$ needs to be cubic (cf. (48)). The following result realizes this condition by assuming $V_\beta'''(0) = 0$, see [GM04] for the case $d = 1$.

Theorem 7.1. *Let $K \in \mathbb{N}$ with $K > 2+d/2$ and assume that the scalar d-dimensional lattice model (17) has potentials $V_\beta \in {\rm C}^{K+2}(\mathbb{R})$ with $V_\beta(0) = V_\beta'(0) = V_\beta'''(0) = 0$. Choose a wave vector $\theta \in \mathcal{T}_\Gamma$ satisfying the non-resonance conditions (57). Let $A \in {\rm C}([0,\tau_*], {\rm H}^{K+3}(\mathbb{R}^d, \mathbb{C})) \cap {\rm C}^1([0,\tau_*], {\rm H}^{K+1}(\mathbb{R}^d, \mathbb{C}))$ be an arbitrary solution of (nlS). Then, for each $d > 0$ there exist $\varepsilon_0 \in (0,1)$ and $D > 0$ such that for all $\varepsilon \in (0, \varepsilon_0]$ any exact solution $\mathbf{x}$ of (17) with*

$$\|(\mathbf{x}(0), \dot{\mathbf{x}}(0)) - (R_\varepsilon^{K-2}(A)(0), \dot{R}_\varepsilon^{K-2}(A)(0))\|_{\ell_2} \leq d\varepsilon^{K-1-d/2}$$

satisfies, for all $t \in [0, \tau_*/\varepsilon^2]$,

$$\|(\mathbf{x}(t), \dot{\mathbf{x}}(t)) - (R_\varepsilon^{K-2}(A)(t), \dot{R}_\varepsilon^{K-2}(A)(t))\|_{\ell_2} \leq D\varepsilon^{K-1-d/2}.$$

The condition $V_\beta'''(0) = 0$ allows us to apply the simple abstract result of Section 6.1. However, this condition is not necessary. In the case of nonlinearities that also have a quadratic part it is still possible to derive a similar result if we impose more restrictive non-resonance conditions. To treat that case one uses ideas from the theory of normal forms to transform the system via a near identity transform into a system that has the same linear part but no quadratic part in the nonlinearity. We refer to [Sch98, GM06] for positive results and mention also [Sch05b] for an example, where the result fails due to fact that the more restrictive non-resonance condition is violated.

7.2 Interaction of several modulated pulses

We report on results in [Gia06] and consider the scalar d-dimensional model (17) for which we want to show how the three-wave interaction equations (20) can be justified in terms of explicit error estimates. Given are three wave vectors $\theta_n \in \mathcal{T}_\Gamma$ and associated frequencies ω_n with $\omega_n^2 = \Omega^2(\theta_n)$, which are in resonance, namely

$$\theta_1 + \theta_2 + \theta_3 = 0 \quad \text{in } \mathcal{T}_\Gamma, \qquad \omega_1 + \omega_2 + \omega_3 = 0. \tag{58}$$

Following [Gia06, GMS06] we use the following type of non-resonance condition for other combinations of these wave vectors. We set $\theta_{-n} := -\theta_n$ and $\omega_{-n} := -\omega_n$ and say that the mode system $\{(\theta_n, \omega_n) : n = 1, 2, 3\}$ is *closed of order* K, if for all $k \in \{1, ..., K\}$ and all $n_1, ..., n_k \in \widetilde{N} = \{-3, -2, -1, 1, 2, 3\}$ the following holds:

$$\big(\textstyle\sum_1^k \omega_{n_l}\big)^2 = \Omega\big(\textstyle\sum_1^k \theta_{n_l}\big)^2 \iff \begin{cases} \exists\, n_* \in \widetilde{N}: & \theta_{n_*} = \sum_1^k \theta_{n_l} \\ \quad\text{and} & \omega_{n_*} = \sum_1^k \omega_{n_l}. \end{cases} \tag{59}$$

Here we use the hyperbolic scaling $\tau = \varepsilon t$ and $y = \varepsilon\gamma$ and, as explained at the beginning of Section 7, we need the improved multiscale ansatz

$$\mathbf{x}(t) = R_\varepsilon^K(\mathbf{A})(t) = \textstyle\sum_{k=1}^K \varepsilon^k \sum_{n_1,\dots,n_k \in \widetilde{N}} B_{n_1,\dots,n_k}(\tau, y)\mathbf{E}_{n_1} \dots \mathbf{E}_{n_k} \tag{60}$$

with $\mathbf{A} = (A_1, A_2, A_3)$, $\mathbf{E}_n = \mathrm{e}^{\mathrm{i}(\omega_n t + \theta_n \cdot \gamma)}$, $B_n = A_n$ and $\overline{B}_{n_1,\dots,n_k} = B_{-n_1,\dots,-n_k}$. Thus, to leading order we have three wave packets, which we expect to travel with their group velocities and to have interactions with the other wave packets.

As explained in Section 3.1 it is possible to determine the coefficient functions $B_{n_1,\dots,n_k}$ in such a way that the approximate solution $z_{\mathrm{app}} = (R_\varepsilon^K(\mathbf{A})(t), \dot{R}_\varepsilon^K(\mathbf{A})(t))$ and the residual $\mathrm{Res}(z_{\mathrm{app}})$ satisfy

$$\|z_{\mathrm{app}}(t)\|_{\ell_\infty} \le C\varepsilon^\alpha \text{ with } \alpha = 1 \quad\text{and}\quad \|\mathrm{Res}(z_{\mathrm{app}})(t)\|_{\ell_2} \le C\varepsilon^{K+1-d/2}$$

if the triple $\mathbf{A} = (A_1, A_2, A_3) : [0, \tau_*] \to \mathrm{L}^2(\mathbb{R}^d, \mathbb{C})^3$ is a sufficiently smooth solution of the three-wave interaction equation (20). Since $\tau = \varepsilon^\sigma t$ with $\sigma = 1$, we may apply Theorem 6.1 with $\nu = 1$, which means that nonlinearities with quadratic parts are allowed.

The precise statement from [Gia06] reads as follows.

Theorem 7.2. *Let $K \in \mathbb{N}$ with $K > 1 + d/2$ and assume that the d-dimensional, scalar lattice model* (17) *has potentials $V_\beta \in \mathrm{C}^{K+2}(\mathbb{R})$ with $V_\beta(0) = V_\beta'(0) = 0$ for $|\beta| < R$. Assume that the mode system $\{(\theta_n, \omega_n) : n = 1, 2, 3\}$ satisfies the resonance condition* (58) *and is closed of order K (cf.* (59)*). Let $\mathbf{A} \in \mathrm{C}([0, \tau_*], \mathrm{H}^{K+2}(\mathbb{R}^d; \mathbb{C})) \cap \mathrm{C}^{K+1}([0, \tau_*], \mathrm{H}^1(\mathbb{R}^d; \mathbb{C}))$ be an arbitrary solution of* (20)*. Then, for each $d > 0$ there exist $\varepsilon_0 \in (0, 1)$ and $D > 0$ such that for all $\varepsilon \in (0, \varepsilon_0]$ any exact solution $\mathbf{x}$ of* (17) *with*

$$\|(\mathbf{x}(0), \dot{\mathbf{x}}(0)) - (R_\varepsilon^{K-1}(\mathbf{A})(0), \dot{R}_\varepsilon^{K-1}(\mathbf{A})(0))\|_{\ell_2} \le d\varepsilon^{K-d/2}$$

satisfies, for all $t \in [0, \tau_*/\varepsilon]$,

$$\|(\mathbf{x}(t), \dot{\mathbf{x}}(t)) - (R_\varepsilon^{K-1}(\mathbf{A})(t), \dot{R}_\varepsilon^{K-1}(\mathbf{A})(t))\|_{\ell_2} \le D\varepsilon^{K-d/2}.$$

The whole theory can be generalized in several aspects. First we may consider mode systems with N different wave vectors, where $N \geq 4$. Then, we obtain a system of N equations for $A_1, ..., A_N$, where only those quadratic terms $\overline{A}_{n_2}\overline{A}_{n_3}$ occur in the equation for $\partial_\tau A_{n_1}$ if the three modes $(\theta_{n_l}, \omega_{n_l})_{l=1,2,3}$ satisfy the resonance condition (58). Other triple interactions do not matter on this time scale either because the frequencies or the wave vectors do not resonate. Quadruple or higher interactions are too small in amplitude to influence the macroscopic behavior (cf. [Gia06]).

Second it is possible to do the very same analysis for systems rather than for a scalar problem. Of course, then we have to pay attention to the different frequency bands. We also refer to [GMS06], where multipulse interactions are treated for nonlinear Schrödinger equations with periodic potentials, see [CMS04, Spa06].

Similar phenomena arise in such different subjects as phonon collisions (cf. [Spo05]) and in surface water waves (cf. [SW03]).

Acknowledgments. This work was substantially supported by the DFG Priority Program SPP 1095 *Analysis, Modeling and Simulation of Multiscale Problems* under Mi 459/3. The authors are grateful to W. Dreyer, G. Friesecke, T. Kriecherbauer, C. Patz, G. Schneider, S. Teufel, F. Theil and H. Uecker for fruitful discussions.

References

[BCS01] A. M. Balk, A. V. Cherkaev, and L. I. Slepyan. Dynamics of chains with non-monotone stress-strain relations. I. Model and numerical experiments. II. Nonlinear waves and waves of phase transition. *J. Mech. Phys. Solids*, 49(1), 131–148, 149–171, 2001.

[BG02a] A. Braides and M. S. Gelli. Continuum limits of discrete systems without convexity hypotheses. *Math. Mech. Solids*, 7, 41–66, 2002.

[BG02b] A. Braides and M. S. Gelli. Limits of discrete systems with long-range interactions. *J. Convex Anal.*, 9, 363–399, 2002. Special issue on optimization (Montpellier, 2000).

[Ble05] T. Blesgen. Two-phase structures as singular limit of a one-dimensional discrete model. *Cubo*, 7(2), 69–79, 2005.

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

[BP06] D. Bambusi and A. Ponno. On metastability in FPU. *Commun. Math. Physics*, 264, 539–561, 2006.

[CMS04] R. Carles, P. A. Markowich, and C. Sparber. Semiclassical asymptotics for weakly nonlinear Bloch waves. *J. Statist. Phys.*, 117(1-2), 343–375, 2004.

[DHM06] W. Dreyer, M. Herrmann, and A. Mielke. Micro–macro transition for the atomic chain via Whitham's modulation equation. *Nonlinearity*, 19, 471–500, 2006.

[DHR06] W. Dreyer, M. Herrmann, and J. Rademacher. Pulses, traveling waves and modulational theory in oscillator chains. In A. Mielke, editor, *Analysis, Modeling and Simulation of Multiscale Problems*. Springer, 2006. To appear.

[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).

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

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

[FJ00] G. Friesecke and R. D. James. A scheme for the passage from atomic to continuum theory for thin films, nanotubes and nanorods. *J. Mech. Phys. Solids*, 48(6-7), 1519–1540, 2000.

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

[FM03] G. Friesecke and K. Matthies. Geometric solitary waves in a 2D mass-spring lattice. *Discrete Contin. Dyn. Syst. Ser. B*, 3, 105–114, 2003.

[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.

[FP04a] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices: III. Howland-type Floquet theory. *Nonlinearity*, 17, 207–227, 2004.

[FP04b] G. Friesecke and R. L. Pego. Solitary waves on FPU lattices: IV. Proof of stability at low energy. *Nonlinearity*, 17, 229–251, 2004.

[FPU55] E. Fermi, J. Pasta, and S. Ulam. Studies of nonlinear problems. Technical report, Los Alamos Scientific Laboratory of the University of California, May 1955. Report LA-1940.

[FT02] G. Friesecke and F. Theil. Validity and failure of the Cauchy–Born hypothesis in a two–dimensional mass–spring lattice. *J. Nonlin. Science*, 12, 445–478, 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.

[Gér91] P. Gérard. Microlocal defect measures. *Comm. Partial Differential Equations*, 16(11), 1761–1794, 1991.

[GHM06] J. Giannoulis, M. Herrmann, and A. Mielke. Redcution of Hamiltonians and Lagrangians in the passage from discrete to continuum models. *In preparation*, 2006.

[Gia06] J. Giannoulis. On the interaction of modulated pulses in discrete lattices. *In preparation*, 2006.

[GM04] J. Giannoulis and A. Mielke. The nonlinear Schrödinger equation as a macroscopic limit for an oscillator chain with cubic nonlinearities. *Nonlinearity*, 17, 551–565, 2004.

[GM06] J. Giannoulis and A. Mielke. Dispersive evolution of pulses in oscillator chains with general interaction potentials. *Discr. Cont. Dynam. Systems Ser. B*, 6, 493–523, 2006.

[GMMP97] P. Gérard, P. A. Markowich, N. J. Mauser, and F. Poupaud. Homogenization limits and Wigner transforms. *Comm. Pure Appl. Math.*, 50, 323–379, 1997. Erratum: Comm. Pure Appl. Math. 53 (2000) 280–281.

[GMS06] J. Giannoulis, A. Mielke, and C. Sparber. Interaction of modulated pulses in the nonlinear Schrödinger equation with periodic potential. *In preparation*, 2006.

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

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

[IJ05] G. Iooss and G. James. Localized waves in nonlinear oscillator chains. *Chaos*, 15, ?? 15 pp., 2005.

[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.

[IZ05] L. Ignat and E. Zuazua. Dispersive properties of a viscous numerical scheme for the Schrödinger equation. *C. R. Math. Acad. Sci. Paris*, 340, 529–534, 2005.

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

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

[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.

[LP93] P.-L. Lions and T. Paul. Sur les mesures de Wigner. *Rev. Mat. Iberoamericana*, 9, 553–618, 1993.

[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.

[Mac04] F. Macià. Wigner measures in the discrete setting: high-frequency analysis of sampling and reconstruction operators. *SIAM J. Math. Anal.*, 36(2), 347–383, 2004.

[MBL06] L. Mugnai, T. Blesgen, and S. Luckhaus. On gamma-limits of discrete free energy functionals. In A. Mielke, editor, *Analysis, Modeling and Simulation of Multiscale Problems*. Springer Verlag, 2006.

[Mie91] A. Mielke. *Hamiltonian and Lagrangian flows on center manifolds*. Springer-Verlag, Berlin, 1991. With applications to elliptic variational problems.

[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.

[Mie06a] A. Mielke. Macroscopic behavior of microscopic oscillations in harmonic lattices using Wigner-Husimi measures. *Arch. Rat. Mech. Analysis*, 2006. In print (WIAS Preprint 1027).

[Mie06b] A. Mielke. Weak convergence methods for the micro-macro transition from from disrcete to continuous systems. *In preparation*, 2006.

[MMP94] P. A. Markowich, N. J. Mauser, and F. Poupaud. A Wigner-function approach to (semi)classical limits: electrons in a periodic potential. *J. Math. Phys.*, 35, 1066–1094, 1994.

[MSU01] A. Mielke, G. Schneider, and H. Uecker. Stability and diffusive dynamics on extended domains. In B. Fiedler, editor, *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems*, pages 563–583. Springer–Verlag, 2001.

[Pat06] C. Patz. *Dispersive behavior in discrete lattices.* PhD thesis, In preparation, 2006.

[PP00] A. A. Pankov and K. Pflüger. Traveling waves in lattice dynamical systems. *Math. Meth. Appl. Sci.*, 23, 1223–1235, 2000.

[Sch94] G. Schneider. Error estimates for the Ginzburg-Landau approximation. *Z. angew. Math. Phys.*, 45(3), 433–457, 1994.

[Sch95] G. Schneider. Validity and limitation of the Newell-Whitehead equation. *Math. Nachr.*, 176, 249–263, 1995.

[Sch96] G. Schneider. Diffusive stability of spatial periodic solutions of the Swift-Hohenberg equation. *Comm. Math. Phys.*, 178(3), 679–702, 1996.

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

[Sch05a] A. Schlömerkemper. Mathematical derivation of the continuum limit of the magnetic force between two parts of a rigid crystalline material. *Arch. Ration. Mech. Anal.*, 176, 227–269, 2005.

[Sch05b] G. Schneider. Justification and failure of the nonlinear Schrödinger equation in case of non-trivial quadratic resonances. *J. Diff. Eqns.*, 216, 354–386, 2005.

[Spa06] C. Sparber. Effective mass theorems for nonlinear schrödinger equations. *SIAM J. Appl. Math.*, 66, 820–842, 2006.

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

[Spo05] H. Spohn. The phonon Boltzmann equation, properties and link to weakly anharmonic lattice dynamics. *J. Stat. Physics*, 2005. To appear (arXiv:math-ph/0505025).

[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.

[SW03] G. Schneider and C. E. Wayne. Estimates for the three-wave interaction of surface water waves. *European J. Appl. Math.*, 14, 547–570, 2003.

[TP04] S. Teufel and G. Panati. Propagation of Wigner functions for the Schrödinger equation with a perturbed periodic potential. In *Multiscale methods in quantum mechanics*, Trends Math., pages 207–220. Birkhäuser Boston, 2004.

[Whi74] G. B. Whitham. *Linear and nonlinear waves.* Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics.

[Zua05] E. Zuazua. Propagation, observation, control and numerical approximation of waves approximated by finite difference method. *SIAM Review*, 47, 197–243, 2005.