Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Lagrangian and Hamiltonian two-scale reduction

Johannes Giannoulis[1], Michael Herrmann[2] and Alexander Mielke[2,3]

submitted: 20th February 2008

[1] Technische Universität München
Zentrum Mathematik
Boltzmannstraße 3
85747 Garching bei München
Germany
E-Mail: giannoulis@ma.tum.de

[2] Humboldt-Universität zu Berlin
Institut für Mathematik
Unter den Linden 6
10099 Berlin
Germany
E-Mail: michaelherrmann@math.hu-berlin.de

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

No. 1302
Berlin 2008

Key words and phrases. Reduced dynamics, multi-scale problems, Hamiltonian structures, modulation equations.

PACS. 02.30.Mv, 45.20.Jj, 52.35.-g.

This work was partially supported by the Deutsche Forschungsgemeinschaft (DFG) within the Priority Program *Analysis, Modeling and Simulation of Multiscale Problems* (SPP 1095) under Mi 459/3-3 and within MATHEON under D14.

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

Studying high-dimensional Hamiltonian systems with microstructure, it is an important and challenging problem to identify reduced macroscopic models that describe some effective dynamics on large spatial and temporal scales. This paper concerns the question how reasonable macroscopic Lagrangian and Hamiltonian structures can by derived from the microscopic system.

In the first part we develop a general approach to this problem by considering non-canonical Hamiltonian structures on the tangent bundle. This approach can be applied to all Hamiltonian lattices (or Hamiltonian PDEs) and involves three building blocks: (i) the embedding of the microscopic system, (ii) an invertible two-scale transformation that encodes the underlying scaling of space and time, (iii) an elementary model reduction that is based on a Principle of Consistent Expansions.

In the second part we exemplify the reduction approach and derive various reduced PDE models for the atomic chain. The reduced equations are either related to long wave-length motion or describe the macroscopic modulation of an oscillatory microstructure.

1 Introduction

A major topic in the area of multi-scale problems is the derivation of reduced or effective macroscopic models for a given microscopic system. A prototype for this problem is the passage from discrete lattice systems to continuum models which describe the *effective dynamics* on much larger spatial and temporal scales. In this case, the microscopic dynamics is governed by a high dimensional system of ODEs, whereas the macroscopic models are related to the PDEs of continuum mechanics or thermodynamics.

In the dynamical setting this problem can be stated as follows: Choosing well-ordered microscopic initial data in a specified class of functions, one hopes that the solution will stay close to this class of functions. We can interpret the class of functions as an approximate invariant manifold and aim to derive reduced equations that govern the evolution on this manifold. Moreover, if the original dynamics is related to underlying Lagrangian or Hamiltonian structures, the question arises how these structures behave under the reduction procedure. This approach is closely related to the theory of modulation equations, see [Mie02, GHM06, SU07] for surveys, which describes how an oscillatory microstructure is modulated on the macroscopic space–time.

In mathematically rigorous terms the transition from a microscopic to a macroscopic scale can be described by a *coarse graining diagram*, which involves the *scaling parameter* ε, see Figure 1. The curve $t \mapsto z_\varepsilon(t) \in M_\varepsilon$ denotes the solution of the

$$
\begin{array}{cccc}
 & \textbf{microscopic} & \xrightarrow{\;\mathcal{T}_\varepsilon\;} & \textbf{macroscopic} \\
\text{initial data } t = 0 & z_\varepsilon^0 & \xrightarrow{\;\varepsilon \to 0\;} & Z^0 \\
\boxed{\textbf{time evolution}} & \Big\downarrow t > 0 & & \tau > 0 \Big\downarrow \\
 & z_\varepsilon(\tau/\varepsilon^\beta) & \xrightarrow{\;\varepsilon \to 0\;} & Z(\tau) \\
 & \text{discrete, atomistic} & \boxed{\textbf{coarse graining}} & \text{continuum}
\end{array}
$$

Figure 1: The coarse graining diagram

microscopic model, i.e., it depends on the microscopic time t, and takes values in the microscopic state space M_ε. On the other hand, the macroscopic trajectory $\tau \mapsto Z(\tau) \in N$ is parametrized by the macroscopic time τ, and describes the evolution of the macroscopic state $Z(\tau) \in N$. The two scales in this problem are linked by a suitable *two-scale ansatz*, which consists of the time scaling $\tau = \varepsilon^\beta t$, as well as a scaling transformation $\mathcal{T}_\varepsilon : M_\varepsilon \to N$, which in particular encodes the spatial scaling. In the best case the diagram commutes, i.e., if the coarse graining $z_\varepsilon(t) \rightsquigarrow Z\big(\varepsilon^\beta t\big)$ holds at time $\tau = 0$, then it holds true for a finite time interval $\tau \in [0,\, \tau_{\text{fin}}]$. Any reasonable micro-macro transition must provide an effective macroscopic evolution equation for the macroscopic configuration $Z(\tau) \in N$. We can not expect the macroscopic equation to provide exact solutions to the microscopic system, but we can hope that it gives rise to *approximate solutions* that satisfy the microscopic law of motion up to higher orders in ε.

In the standard approach of model reduction one inserts a reasonable two-scale ansatz into the microscopic law of motion and derives a macroscopic evolution equation by means of formal expansions with respect to the scaling parameter ε.

However, this standard approach ignores the underlying Lagrangian and Hamiltonian structures and therefore the following questions arise naturally: (*i*) Are there macroscopic Lagrangian and Hamiltonian structures that correspond to the reduced macroscopic equation? (*ii*) If yes, how can one derive them and what is their relation to the microscopic structures?

The main issue of this paper is to develop a general framework for micro-macro transitions that relies on a *two-scale reduction* of microscopic Lagrangian and Hamiltonian structures. To this end we split our approach into three steps, namely *embedding*, *exact two-scale transformation*, and *reduction*, which can be studied independently. Our point of view is strongly motivated by the investigation of microscopic lattice systems, where the micro-macro transition replaces a high dimensional system of ODEs by a small number of macroscopic PDEs. Nevertheless, our approach to Hamiltonian two-scale reduction can also be applied to microscopic PDEs, see §1.1 below. Note, that for us a Hamiltonian structure consists of a Hamiltonian (function) and a (non-canonical) symplectic form. Alternatively one could study the reduction of Poisson structures.

The Hamiltonian two-scale reduction for lattices always involves the scaling of space *and* time variables. There exists a lot of literature concerning solely the coupling of slow and fast time scales in Hamiltonian systems with finite dimension, or fixed spatial scales. The arising mathematical problems can be tackled by means of averaging and adiabatic invariants, see for instance [Jar93, TR99, NV05]. Moreover, a lot of work has been done to derive efficient schemes for the numerical integration of such systems, compare [CJLL06, HLW02], and references therein.

A second class of micro-macro transitions is related to the passage between different spatial scales. For instance, in the static case it is a challenging problem to derive elastic energies from atomistic lattice models, and to study the macroscopic convergence of microscopic ground states and energies, see [FJ00, BG02b, BG02a, FT02, BG06, The06, Sch06, BLM06]. Another kind of spatial reduction arises, when the microscopic model combines both large and short space scales. Close to our point of view, [GKMS95] considers the Euler equations for an incompressible fluid under gravity, and studies the limit of vanishing height. It can be shown that the underlying Poisson-structure converges to a limit that corresponds to the shallow water equation. Moreover, using similar methods the equations for shells and plates can be derived from the three-dimensional models of nonlinear elasticity, see [GKM96].

1.1 Motivating examples

Let us first discuss two simple examples related to microscopic PDEs which highlight the essential features that arise in the general setting. Below we will see that microscopic lattices can be treated similarly if viewed as embedded into systems with continuous space variable.

The first example concerns the passage from the Boussinesq equation to the Korteweg–de Vries (KdV) equation. Here, the microscopic dynamics is governed by

$$x_{tt} = x_{\eta\eta} - x_{\eta\eta\eta\eta} + x_\eta \, x_{\eta\eta}, \tag{1}$$

where the unknown function x depends on the *microscopic time* t and the *microscopic space variable* $\eta \in \mathbb{R}$. Notice that x_t and x_η abbreviate $\partial_t x$ and $\partial_\eta x$, respectively. For simplicity we ignore all boundary effects, so that the microscopic configuration space is $Q = L^2(\mathbb{R};\, \mathrm{d}\eta)$. One particular macroscopic model for (1) is related to the two-scale ansatz

$$x(t,\, \eta) = \varepsilon X\big(\varepsilon^3 t,\, \varepsilon(\eta + t)\big) \tag{2}$$

where $\tau = \varepsilon^3 t$ and $y = \varepsilon(\eta + t)$ denote the *macroscopic* time and space, respectively. The function X is the macroscopic configuration and for fixed τ it takes values in $P = L^2(\mathbb{R};\, \mathrm{d}y)$. The *scaling* parameter $\varepsilon > 0$ is assumed to be small and bridges the two appearing scales.

The standard approach for model reduction works as follows: We plug the two-scale ansatz (2) into the microscopic law of motion (1), use formal expansions with respect to ε and equate the terms of leading order. For the example at hand one easily derives

$$2X_{\tau y} = -X_{yyyy} + X_y X_{yy}, \tag{3}$$

which is a KdV equation for X_y. As already mentioned, this standard approach works very well but in general it is not clear at all whether the derived macroscopic equation has its own Lagrangian and Hamiltonian structures.

We proceed with the Lagrangian and Hamiltonian two-scale reduction for the Boussinesq example in order to illustrate the difficulties that may arise in the general setting as well as the proposed solutions. On the one hand, the microscopic Lagrangian $\mathcal{L}$ for (1) is given by $\mathcal{L} = \mathcal{K} - \mathcal{V}$ with kinetic energy $\mathcal{K}$ and potential energy $\mathcal{V}$ given by

$$\mathcal{K}(x_t) = \int_{\mathbb{R}} \tfrac{1}{2} {x_t}^2 \,\mathrm{d}\eta, \quad \mathcal{V}(x) = \int_{\mathbb{R}} \tfrac{1}{2} {x_\eta}^2 + \tfrac{1}{2} {x_{\eta\eta}}^2 + \tfrac{1}{6} {x_\eta}^3 \,\mathrm{d}\eta. \tag{4}$$

Identifying the momenta $\pi = \partial_{x_t}\mathcal{L}$ with the velocities x_t we find that the microscopic Hamiltonian $\mathcal{H}$ equals the energy $\mathcal{E} = \mathcal{K} + \mathcal{V}$. In particular, the microscopic law of motion (1) equals the Euler–Lagrange equations to $\mathcal{L}$, and is moreover equivalent to the canonical equations to $\mathcal{H}$, which correspond to the symplectic form

$$\Sigma = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{5}$$

with 1 being the identity map $Q \to Q$. On the other hand, the KdV equation (3) is the Euler-Lagrange equation to $\mathbb{L}^{\mathrm{red}}(X,\, X_\tau) = \mathbb{K}^{\mathrm{red}}(X,\, X_\tau) - \mathbb{V}^{\mathrm{red}}(X)$ with

$$\mathbb{K}^{\mathrm{red}}(X,\, X_\tau) = \int_{\mathbb{R}} X_\tau X_y \,\mathrm{d}y, \quad \mathbb{V}^{\mathrm{red}}(X) = \int_{\mathbb{R}} \tfrac{1}{2}(X_{yy})^2 + \tfrac{1}{6}(X_y)^3 \,\mathrm{d}y.$$

Since $\mathbb{L}^{\mathrm{red}}$ depends linearly on the macroscopic velocities X_τ, the reduced macroscopic Hamiltonian structure is non-canonical. In fact, the Hamiltonian $\mathbb{H}^{\mathrm{red}}$ equals the potential energy $\mathbb{V}^{\mathrm{red}}$ and the symplectic structure

$$\sigma^{\mathrm{red}}(\dot{X},\, \acute{X}) = \int_{\mathbb{R}} \dot{X} \acute{X}_y \,\mathrm{d}y \tag{6}$$

is a skew-symmetric 2-form on P. Consequently, the macroscopic law of motion is given by $\sigma^{\mathrm{red}}(X_\tau,\, \cdot) = \mathrm{d}\mathbb{H}^{\mathrm{red}}$, which is a dynamical system on P, and not on TP or T^*P.

In order to describe how the microscopic Hamiltonian structure reduces to the macroscopic one, we regard the two-scale ansatz (2) as a time dependent transformation $\mathbb{T}_{\mathrm{con}}(t,\, \varepsilon) : X \in P \to Q$ with parameter ε. Its canonical lift $\mathbb{T}_{\mathrm{vel}}(t,\, \varepsilon) : TP \to TQ$ to the corresponding tangent bundles reads

$$\mathbb{T}_{\mathrm{vel}}(t,\, \varepsilon) : (X,\, X_\tau) \rightsquigarrow (x,\, x_t), \quad (x,\, x_t)(\eta) = \big(\varepsilon X,\, \varepsilon^4 X_\tau + \varepsilon^2 X_y\big)(\varepsilon\eta + \varepsilon t). \tag{7}$$

This transformation comprises the crucial ingredients of our approach: For fixed $\varepsilon > 0$ this transformation is *exact*, this means invertible, but describes explicitly how the macroscopic structures depend on ε. Therefore, one can read-off the *effective* structures from the leading order terms in ε.

Applying the inverse transformation of (7) to the energies from (4) we find

$$\mathbb{K}(\varepsilon,\, X,\, X_\tau) = \varepsilon^3 \int_{\mathbb{R}} \tfrac{1}{2}\big(\varepsilon^2 X_\tau + X_y\big)^2 \,\mathrm{d}y, \quad \mathbb{V}(\varepsilon,\, X) = \varepsilon^3 \tfrac{1}{2} \int_{\mathbb{R}} {X_y}^2 + \varepsilon^2 {X_{yy}}^2 + \tfrac{1}{3}\varepsilon^2 {X_\tau}^3 \,\mathrm{d}y.$$

Both transformed energies are of order ε^3. However, the transformed Lagrangian $\mathbb{L}$ is of order ε^5, since the terms of order ε^3 vanish due to *cancelation* via $\mathbb{L} = \mathbb{K} - \mathbb{V}$. Thus, we find

$$\mathbb{L}(\varepsilon,\, X,\, X_\tau) = \varepsilon^5 \mathbb{L}^{\mathrm{red}}(X,\, X_\tau) + \mathcal{O}\big(\varepsilon^7\big).$$

The transformation of the Hamiltonian structure is not so simple, since the transformation (7) involves a moving frame. The macroscopic Hamiltonian $\mathbb{H}$, i.e. the Legendre transform of $\mathbb{L}$, is given by

$$\mathbb{H}(\varepsilon,\, X,\, X_\tau) = \mathbb{E}(\varepsilon,\, X,\, X_\tau) - \mathbb{I}(\varepsilon,\, X,\, X_\tau) = \varepsilon^5 \mathbb{H}^{\mathrm{red}}(X,\, X_\tau) + \mathcal{O}\big(\varepsilon^7\big).$$

Here, $\mathbb{E} = \mathbb{K} + \mathbb{V}$ is the transformed energy and $\mathbb{I}$ is the transform of $\mathcal{I}$, where $\mathcal{I}$ is the *conserved quantity* related to the moving frame by Noether's Theorem:

$$\mathcal{I}(x,\, x_t) = \int_{\mathbb{R}} x_t\, x_\eta \,\mathrm{d}\eta, \quad \mathbb{I}(\varepsilon,\, X,\, X_\tau) = \varepsilon^3 \int_{\mathbb{R}} \big(\varepsilon^2 X_\tau + X_y\big) X_y \,\mathrm{d}y.$$

We conclude that the transformation (7) provides both the Lagrangian and the Hamiltonian for (3) to leading order ε^5. Moreover, it can be shown that the symplectic form (5), considered as a 2-form on the *tangent bundle* TQ but *not* on the cotangent bundle T^*Q, transforms into

$$\Sigma = \varepsilon^5 \begin{pmatrix} -2\partial_y & 0 \\ 0 & 0 \end{pmatrix} + \varepsilon^7 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

which equals (6) to leading order ε^5. Finally, the KdV equation is invariant under shifts in the y-direction, and this symmetry gives rise to the conserved quantity

$$\mathbb{I}^{\mathrm{red}}(X) = \int_{\mathbb{R}} X_y^2 \,\mathrm{d}y,$$

which turns out to be the lowest order expansions of $\mathbb{K}(\varepsilon, \cdot)$ and $\mathbb{V}(\varepsilon, \cdot)$, namely

$$\varepsilon^3 \,\mathbb{I}^{\mathrm{red}}(X) = 2\,\mathbb{K}(\varepsilon,\, X,\, X_\tau) + \mathcal{O}\big(\varepsilon^5\big) = 2\,\mathbb{V}(\varepsilon,\, X,\, X_\tau) + \mathcal{O}\big(\varepsilon^5\big) = \mathbb{I}(\varepsilon,\, X,\, X_\tau) + \mathcal{O}\big(\varepsilon^5\big).$$

We conclude that the terms which vanish due to cancelation correspond to a macroscopic integral of motion.

As a second motivating example we study the macroscopic evolution of a modulated pulse in the Klein-Gordon (KG) equation

$$x_{tt} = x_{\eta\eta} - \Phi_0'(x) \tag{8}$$

with $x = x(t,\, \eta)$, $\eta \in \mathbb{R}$, and nonlinear on-site potential Φ_0. A modulated pulse is an (approximate) solution which satisfies the ansatz

$$x(\eta, t) = \varepsilon A\big(\varepsilon^2 t,\, \varepsilon\eta - \varepsilon c t\big)\mathtt{e}^{\mathtt{i}(\omega t + \theta\eta)} + \text{c.c.} + \mathcal{O}\big(\varepsilon^2\big). \tag{9}$$

Here c.c. denotes the complex conjugate, the frequency ω and the wave number θ are fixed parameters, and c is the moving-frame speed. The plane waves $\mathtt{e}^{\mathtt{i}(\omega t+\theta\eta)}$ in (9) describe an *oscillatory microstructure* whose amplitude A is modulated on the macroscopic scale $\tau = \varepsilon^2 t$ and $y = \varepsilon(\eta - ct)$.

A first necessary condition for (9) to yield approximate solutions is that ω and θ satisfy the dispersion relation $\omega^2 = \theta^2 + \Phi_0''(0)$ and $c = -\omega'$ is the associated *group velocity*. Moreover, the complex-valued amplitude A must satisfy the nonlinear Schrödinger (nlS) equation

$$2\omega\mathtt{i}A_\tau = \varrho_1 A_{yy} - \varrho_2 \left|A\right|^2 A,$$

where the constants ϱ_1 and ϱ_2 can be computed explicitly. The validity of this macroscopic model has been proven rigorously in [KSM92] on the level of the equation of motion.

As in the Boussinesq example, both the microscopic and macroscopic models have Lagrangian and Hamiltonian structures and so we are interested in the question how these are related to each other. The new feature in this example is the presence of microscopic oscillations and the key idea is to introduce an additional one-dimensional, periodic *phase variable* $\phi \in T^1 \simeq [0,\, 2\pi]$. This new degree of freedom enables us to find a suitable two-scale transformation such that *all* (transformed) oscillations are confined in the phase direction ϕ. This suggests the two-scale ansatz

$$x(t, \eta, \phi) = \varepsilon X(\varepsilon^2 t, \varepsilon(\eta - ct), \phi + \omega t + \theta\eta), \tag{10}$$

which is similar to (9) but gives rise to an invertible two-scale transformation.

The introduction of ϕ can be viewed as an *embedding* of the microscopic system, such that (8) becomes

$$x_{tt}(t,\, \eta,\, \phi) = x_{\eta\eta}(t,\, \eta,\, \phi) - \Phi_0'(x(t,\, \eta,\, \phi)).$$

This embedding does not affect the microscopic dynamics, since ϕ appears just as a parameter. The embedded system has Lagrangian $\mathcal{L} = \mathcal{K} - \mathcal{V}$ and Hamiltonian $\mathcal{H} = \mathcal{E} = \mathcal{K} + \mathcal{V}$ with

$$\mathcal{K}(x_t) = \int\limits_{\mathbb{R}\times T^1} \tfrac{1}{2}x_t^2 \,\mathrm{d}\eta\mathrm{d}\phi, \quad \mathcal{V}(x) = \int\limits_{\mathbb{R}\times T^1} \tfrac{1}{2}x_\eta^2 + \Phi_0(x)\,\mathrm{d}\eta\mathrm{d}\phi \tag{11}$$

and corresponds to the symplectic form (5). Moreover, we find two continuous symmetry groups related to shifts with respect to η and ϕ, which by Noether's theorem correspond to the conserved quantities (integrals of motion)

$$\mathcal{I}_{\text{space}}(x,\, x_t) = \int\limits_{\mathbb{R}\times T^1} x_t\, x_\eta \,\mathrm{d}\eta\mathrm{d}\phi, \qquad \mathcal{I}_{\text{phase}}(x,\, x_t) = \int\limits_{\mathbb{R}\times T^1} x_t\, x_\phi \,\mathrm{d}\eta\mathrm{d}\phi. \tag{12}$$

The second integral of motion arises only due to the embedding but plays a prominent role in the two-scale reduction, since it is needed for the derivation of the macroscopic Hamiltonian. In fact, the moving frame in (10) involves drifts in space and phase direction and the associated integral of motion reads

$$\mathcal{I} = -c\,\mathcal{I}_{\text{space}} + \omega\,\mathcal{I}_{\text{phase}}.$$

Like for the Boussinesq example, we can use the transformation (10) and our general approach described below in order to derive the Lagrangian and Hamiltonian structures for the nlS equation directly from their microscopic counterparts. It comes out, that the leading order terms determine the microstructure, the next-leading order terms give the moving frame speed, and finally, the next-next leading order terms provide the macroscopic law of motion. This will be explained in detail within §3.4.

1.2 General approach to Lagrangian and Hamiltonian two-scale reduction

The concepts arising in the above examples can be generalized to the following abstract framework for a Lagrangian and Hamiltonian two-scale reduction.

The first step concerns the **embedding** of the microscopic system. We have seen above that the treatment of models with microstructure requires the introduction of new phase variables ϕ. Moreover, for discrete models like chains we replace the particle index $j \in \mathbb{Z}$ by a continuous variable $\eta \in \mathbb{R}$. In all cases this embedding does not change the microscopic dynamics, but it gives rise to new *continuous symmetry groups* and hence to additional integrals of motion which contribute to the macroscopic Hamiltonian.

In what follows we always consider the Lagrangian $\mathcal{L}$ of the embedded system which is defined on the tangent bundle TQ of the microscopic configuration space Q. Then there exists an equivalent Hamiltonian structure on the *cotangent bundle* T^*Q corresponding to the canonical symplectic form. However, for the reduction step explained below it is essential to consider a Hamiltonian $\mathcal{H}$ as well as a symplectic form σ both of which are defined on the *tangent bundle* TQ. To this end we pull back the canonical structure from T^*Q to TQ via the fiber derivative of $\mathcal{L}$. This will be discussed in detail in §2.1.

The most important step in any two-scale reduction is the **transformation** of the embedded system. For this purpose we introduce two-scale transformations by composing elementary building blocks such as *(weak) symmetry transformations*, *moving-frame transformations*, and *scaling transformations*. The first two building blocks are well understood in classical mechanics, whereas our concept of scaling transformations seems to be new, since it involves the scaling of space *and* time. The starting point for any scaling transformation is a map $\mathcal{S}_{\mathrm{con}} : Q \to P$ bridging the microscopic and the macroscopic configuration spaces Q and P. The definition of such a map involves only the scalings of the space coordinates, but its lift $\mathcal{S}_{\mathrm{vel}} : TQ \to TP$ to the tangent bundles takes into account also the time scaling.

Two-scale transformations are in the heart of any two-scale reduction, because they provide a macroscopic Lagrangian $\mathbb{L}$, a macroscopic Hamiltonian $\mathbb{H}$, and a symplectic form σ (all defined on TP), which depend explicitly on the scaling parameter ε.

The **reduction** step starts with the formal expansions of the transformed Lagrangian and Hamiltonian structures with respect to the scaling parameter ε, i.e.,

$$\mathbb{L}(\varepsilon) = \varepsilon^{\kappa} \left(\mathbb{L}_0 + \varepsilon \mathbb{L}_1 + \varepsilon^2 \mathbb{L}_2 + \ldots \right),$$

and

$$\mathbb{H}(\varepsilon) = \varepsilon^{\kappa}\left(\mathbb{H}_0 + \varepsilon\mathbb{H}_1 + \varepsilon^2\mathbb{H}_2 + ...\right), \quad \sigma(\varepsilon) = \varepsilon^{\kappa}\left(\sigma_0 + \varepsilon\sigma_1 + \varepsilon^2\sigma_2 + ...\right).$$

A key feature of our approach is the *Principle of Consistent Expansions* which will be proved in §2.1 and guarantees that $(\mathbb{H}_i, \sigma_i)$ is the Hamiltonian structure corresponding to the Lagrangian $\mathbb{L}_i$. For this principle to hold it is crucial to consider the Hamiltonian structure on the *tangent* (and not on the *cotangent*) bundle.

For some examples the leading order Lagrangian $\mathbb{L}_0$ is non-degenerate. Then the effective macroscopic model is completely determined already by the leading order terms. However, whenever the two-scale ansatz involves an oscillatory microstructure the leading order terms turn out to be degenerate in the following sense: The leading order Lagrangian $\mathbb{L}_0$ is quasi-stationary, i.e., it does not depend on X_τ, and this implies $\mathbb{H}_0 = -\mathbb{L}_0$ and $\sigma_0 = 0$. Moreover, there exists a sub-manifold P_0 of P such that the gradient of $\mathbb{L}_0$ vanishes on P_0. In this case we restrict $\mathbb{L} - \mathbb{L}_0$ and $\mathbb{H} - \mathbb{H}_0$ and σ to TP_0, and derive the effective macroscopic model by expanding the restricted structures.

The reduction procedure concerns the *convergence of Lagrangian and Hamiltonian structures* as $\varepsilon \to 0$, but this does *not* necessarily imply the *convergence of solutions*. Therefore each reduced model must be *justified*. In the general setting the justification problem turns out to be very subtle and is not addressed in this paper. However, for all examples presented here we discuss the corresponding justification problem after having derived the reduced Lagrangian and Hamiltonian structures. We also refer to the surveys [Mie02, GHM06, SU07] and to [Mie08] for an abstract theory using Γ-convergence for Hamiltonian systems.

The abstract framework for the two-scale reduction will be developed in detail within §2, where we prove the transformation rules for Lagrangian and Hamiltonian structures and discuss the reduction procedure in the various cases. Finally, in §3 we apply this method to several micro-macro transitions for the atomic chain.

1.3 Two-scale reductions for the atomic chain

The nonlinear atomic chain consists of identical particles with unit mass. The atoms are coupled to a background field by the *on-site potential* Φ_0 and nearest neighbors interact via the *pair potential* Φ_1. The microscopic dynamics is governed by Newton's equations

$$\ddot{x}_j(t) = \Phi_1'\big(x_{j+1}(t) - x_j(t)\big) - \Phi_1'\big(x_j(t) - x_{j-1}(t)\big) - \Phi_0'(x_j(t)), \tag{13}$$

where $j \in \mathbb{Z}$ is the discrete particle index and $x_j(t) \in \mathbb{R}$ denotes the displacement of the j-th particle at time t. For $\Phi_0 \equiv 0$ and an-harmonic Φ_1 we obtain the Fermi–Pasta–Ulam (FPU) chain, while Klein-Gordon (KG) chains correspond to harmonic Φ_1 but have an-harmonic Φ_0.

A general micro-macro transition for the atomic chain is related to the two-scale ansatz

$$x_j(t) = \varepsilon^\alpha X\big(\varepsilon^\beta t,\, \varepsilon(j - c\,t)\big)$$

with macroscopic time $\tau = \varepsilon^\beta t$, macroscopic particle index $y = \varepsilon(j - ct)$ and macroscopic configuration X. Notice that y is assumed to be a continuous variable and can be interpreted as the *coordinate of a macroscopic material point*.

In the example part §3 we study the following micro-macro transitions and discuss how the Lagrangian and Hamiltonian structures that correspond to the effective macroscopic equations can be derived directly from the Lagrangian and Hamiltonian structure of the atomic chain. To this end we embed the atomic chain (13) into a microscopic system with continuous particle index $\eta \in \mathbb{R}$, see §3.1.

Quasi-linear wave equation In §3.2 we consider the FPU chain and rely on the two-scale ansatz

$$x_j(t) = \varepsilon^{-1} X(\varepsilon t,\, \varepsilon j), \tag{14}$$

which has no moving frame and corresponds to the *hyperbolic scaling* $\tau = \varepsilon t$ and $y = \varepsilon j$. In this case the macroscopic evolution satisfies the nonlinear wave equation

$$\partial_{\tau\tau} X - \partial_y \Big(\Phi_1'(\partial_y X)\Big) = 0.$$

KdV equation The second example, see §3.3, concerns the passage from FPU chains to a KdV equation by means of a two-scale ansatz similar to (2).

Modulated pulses and the nlS equation In analogy to the second motivating example, in §3.4 we study the macroscopic evolution of a modulated pulse in the KG *chain*. Similar to above, the two-scale ansatz reads

$$x_j(t) = \varepsilon A\big(\varepsilon^2 t,\, \varepsilon j - \varepsilon c t\big)\mathtt{e}^{\mathtt{i}(\omega t + \theta j)} + \text{c.c.} \tag{15}$$

and the macroscopic dynamics is described by an nlS equation. The only difference as compared to the case of the *continuous* KG equation (8) concerns the dispersion relation leading to different coefficients in the macroscopic equation.

Three-wave-interaction The fourth example, see §3.5, is the most involved one and describes how three modulated pulses interact if they are in resonance. This gives rise to the following ansatz

$$x_j(t) = \varepsilon \sum_{n=1}^{3} A_n(\varepsilon t,\, \varepsilon j)\mathtt{e}^{\mathtt{i}(\omega_n t + \theta_n j)} + \text{c.c.} \tag{16}$$

with three phases $\phi_n = \omega_n t + \theta_n j$ and three amplitudes A_n. All pairs $p_n = (\theta_n, \omega_n)$ satisfy the dispersion relation of the KG chain and are coupled via the *resonance condition* $p_1 + p_2 + p_3 = 0$ in $T^1 \times \mathbb{R}$, where $T^k = \mathbb{R}^k/_{(2\pi\mathbb{Z})^k}$ is the k-dimensional torus. This resonance condition shows that we have only two independent phases. Moreover, the amplitudes are coupled on the hyperbolic scaling $\tau = \varepsilon t$, $y = \varepsilon j$ via the *three-wave-interaction equations*

$$\mathtt{i}\begin{pmatrix} 2\omega_1 & 0 & 0 \\ 0 & 2\omega_2 & 0 \\ 0 & 0 & 2\omega_3 \end{pmatrix} \partial_\tau \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \mathtt{i}\begin{pmatrix} 2\omega_1\omega_1' & 0 & 0 \\ 0 & 2\omega_2\omega_2' & 0 \\ 0 & 0 & 2\omega_3\omega_3' \end{pmatrix} \partial_y \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} - \Phi_0'''(0) \begin{pmatrix} \overline{A}_2\overline{A}_3 \\ \overline{A}_1\overline{A}_3 \\ \overline{A}_1\overline{A}_2 \end{pmatrix}. \tag{17}$$

Finally, in §3.6 we present further examples for micro-macro transitions in the atomic chain. Although they fit into the general framework they are postponed to a forthcoming paper as they display additional complications.

2 Lagrangian and Hamiltonian two-scale reduction

In this section we describe the general framework for the two-scale reduction of Lagrangian and Hamiltonian structures and present our abstract results concerning two-scale transformations and the problem of model reduction. Since we are mainly interested in Hamiltonian PDEs and lattices we assume that the microscopic configuration space Q is a function space. Moreover, for simplicity we suppose Q to be a Hilbert space (usually some L^2–space) with inner product $\langle\cdot, \cdot\rangle$.

As a prototypical example for a microscopic Lagrangian we consider a *normal system*, where the Lagrangian $\mathcal{L}$ is the difference of quadratic kinetic energy $\mathcal{K}$ and potential energy $\mathcal{V}$. More precisely, a normal Lagrangian $\mathcal{L}$ satisfies

$$\mathcal{L}(x, x_t) = \mathcal{K}(x_t) - \mathcal{V}(x), \qquad \mathcal{K}(x_t) = \tfrac{1}{2}\langle x_t, M\, x_t\rangle \tag{18}$$

with symmetric mass matrix $M : Q \to Q$. However, our approach is not restricted to normal systems but can be applied to all microscopic Lagrangian structures. We start with some general remarks regarding Hamiltonian structures.

2.1 Hamiltonian structures for given Lagrangian

In classical mechanics we have (at least) two possibilities to introduce a Hamiltonian structure for a given Lagrangian $\mathcal{L} : TQ \to \mathbb{R}$, where a Hamiltonian structure consists of both a Hamiltonian (function) and a symplectic form.

The standard approach is related to the *canonical Hamiltonian structure* on the *cotangent bundle* T^*Q. For its definition we consider the Legendre transform $\mathcal{H} : TQ \to \mathbb{R}$ of $\mathcal{L}$ which is defined by $\mathcal{H}(x, x_t) = \langle\pi(x, x_t), x_t\rangle - \mathcal{L}(x, x_t)$. Here,

$\pi(x,\, x_t) = \mathfrak{F}\mathcal{L}|_x(x_t)$ is the canonical momentum associated to x_t and is determined by the *fiber derivative* of the Lagrangian $\mathcal{L}$. This fiber derivative is given by

$$\mathfrak{F}\mathcal{L} : TQ \to T^*Q, \quad (x,\, x_t) \mapsto (x,\, \partial_{x_t}\mathcal{L}(x,\, x_t)) = (x,\, \pi(x,\, x_t)).$$

In the next step we replace the velocity x_t by π, assuming this is possible, and rewriting $\mathcal{H} : TQ \to \mathbb{R}$ in terms of x and π we obtain the *canonical Hamiltonian*

$$\overline{\mathcal{H}} : T^*Q \to \mathbb{R}, \quad \overline{\mathcal{H}}(x,\, \pi) = \mathcal{H}\big(\mathfrak{F}\mathcal{L}^{-1}(x,\, \pi)\big).$$

The Lagrangian equation to $\mathcal{L}$, i.e. $\frac{\mathrm{d}}{\mathrm{d}t}\pi(x,\, x_t) = \partial_x\mathcal{L}(x,\, x_t)$, is equivalent to the canonical equations $x_t = \partial_\pi\overline{\mathcal{H}}$, $\pi_t = -\partial_x\overline{\mathcal{H}}$, which can be written as

$$\overline{\sigma}_{\mathrm{can}}|_{\overline{z}}(\overline{z}_t,\, \cdot) = \mathrm{d}\overline{\mathcal{H}}|_{\overline{z}}(\cdot).$$

Here, $\overline{z} = (x,\, \pi) \in T^*Q$, and $\overline{\sigma}_{\mathrm{can}}$ denotes the *canonical symplectic form* on T^*Q given by

$$\overline{\sigma}_{\mathrm{can}}|_{\overline{z}}\big(\dot{\overline{z}},\, \acute{\overline{z}}\big) = \langle \acute{\pi},\, \dot{x}\rangle - \langle \dot{\pi},\, \acute{x}\rangle,$$

with $\dot{\overline{z}} = (\dot{x},\, \dot{\pi})$ and $\acute{\overline{z}} = (\acute{x},\, \acute{\pi})$ being two independent tangent vectors from $T|_{\overline{z}}T^*Q$.

The second Hamiltonian structure lives on the *tangent bundle* TQ and consists of the Hamiltonian $\mathcal{H} : TQ \to \mathbb{R}$ and a non-canonical symplectic form $\sigma \in \Lambda^2(TQ)$ defined as the pull-back of $\overline{\sigma}_{\mathrm{can}}$ via $\mathfrak{F}\mathcal{L}$, i.e. $\sigma = (\mathfrak{F}\mathcal{L})^*\overline{\sigma}_{\mathrm{can}}$. This means

$$\sigma|_z(\dot{z},\, \acute{z}) = \langle D\pi|_z(\acute{z}),\, \dot{x}\rangle - \langle D\pi|_z(\dot{z}),\, \acute{x}\rangle, \tag{19}$$

where $z = (x,\, x_t)$, $\dot{z}, \acute{z} \in T|_zTQ$, and $D\pi|_z$ is the linearization of π in z. Assuming that $\mathfrak{F}\mathcal{L}$ is differentiable, it can be shown, see [AM78] for a proof, that the Lagrangian equation for $\mathcal{L}$ is equivalent to the Hamiltonian system

$$\sigma|_z(z_t,\, \cdot) = \mathrm{d}\mathcal{H}|_z(\cdot). \tag{20}$$

Remark 1. The symplectic form σ can be identified with a family

$$\Sigma : TQ \to \mathrm{Lin}(Q{\times}Q,\, Q{\times}Q)$$

of skew-symmetric and operator-valued matrices such that $\sigma|_z(\dot{z},\, \acute{z}) = \langle \Sigma|_z\dot{z},\, \acute{z}\rangle_{Q\times Q}$ for all states $z = (x,\, x_t) \in Q \times Q$ and arbitrary tangent vectors $\dot{z}, \acute{z} \in Q \times Q$. The components Σ_{ij}, $i, j = 1, 2$, of Σ are linear operators $Q \to Q$ and satisfy $\Sigma_{ij}^T = -\Sigma_{ji}$. Consequently, the Hamiltonian system (20) is equivalent to

$$\Sigma|_{(x,\, x_t)}\, \frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x \\ x_t \end{pmatrix} = \begin{pmatrix} \partial_x\mathcal{H}(x,\, x_t) \\ \partial_{x_t}\mathcal{H}(x,\, x_t) \end{pmatrix}.$$

Example 2. On each Hilbert space Q we can define the *metric Lagrangian* $\mathcal{L}_{\mathrm{met}}$ by $\mathcal{L}_{\mathrm{met}}(x,\, x_t) = \frac{1}{2}\langle x_t,\, x_t\rangle$. This implies $\mathcal{H}_{\mathrm{met}} = \mathcal{L}_{\mathrm{met}}$ and

$$\sigma_{\mathrm{met}}|_{(x,\, x_t)}\Big((\dot{x},\, \dot{x}_t),\, (\acute{x},\, \acute{x}_t)\Big) = \langle \acute{x}_t,\, \dot{x}\rangle - \langle \dot{x}_t,\, \acute{x}\rangle, \quad \Sigma_{\mathrm{met}}|_{(x,\, x_t)} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

where 1 denotes the identity map $Q \to Q$. In what follows we refer to $\sigma_{\rm met}$ and $\Sigma_{\rm met}$ as the *metric symplectic form* on TQ. Moreover, for a normal Lagrangian with (18) we find $\mathcal{H}(x,\,x_t) = \mathcal{K}(x_t) + \mathcal{V}(x)$ as well as

$$\Sigma = \begin{pmatrix} 0 & -M \\ M & 0 \end{pmatrix} = M\Sigma_{\rm met},$$

where we used $M = M^T$.

The tangent-bundle approach to Hamiltonian structures is more general than the canonical one via the cotangent bundle, because it works even if the map $x_t \mapsto \pi$ is not invertible, but has the disadvantage that the symplectic form σ depends explicitly on the Lagrangian $\mathcal{L}$. Consequently, the Hamiltonian equations on TQ do not arise in canonical form. For the examples from §1.1 and §1.3 we find $\pi(x,\,x_t) = x_t$ so that the Hamiltonian structures on TQ and T^*Q seem to be equal. However, both structures *transform differently* under scaling transformations, see Principle 3 and §2.2.

For a first motivation why we prefer the tangent-bundle and avoid the cotangent-bundle structures, let us study *trivial scalings*: Given a Lagrangian $\mathcal{L}$ on TQ, we consider the scaled Lagrangian $\mathcal{L}_\varepsilon = \varepsilon\mathcal{L}$, where $\varepsilon > 0$ is some artificial small constant. The scaling of $\mathcal{H}$ is given by $\mathcal{H}_\varepsilon = \langle \partial_{x_t}\mathcal{L}_\varepsilon,\, x_t\rangle - \mathcal{L}_\varepsilon = \varepsilon\mathcal{H}$, and similarly we find $\sigma_\varepsilon = \varepsilon\sigma$. On the other hand, the standard (canonical) approach applied to $\mathcal{L}_\varepsilon$ yields $\overline{\mathcal{H}}_\varepsilon(x,\,\pi) = \varepsilon\overline{\mathcal{H}}(x,\,\varepsilon^{-1}\pi)$, and the canonical equations

$$x_t = +\,\partial_\pi\overline{\mathcal{H}}_\varepsilon(x,\,\pi) = +\,\partial_\pi\overline{\mathcal{H}}\big(x,\,\varepsilon^{-1}\pi\big), \quad \pi_t = -\,\partial_x\overline{\mathcal{H}}_\varepsilon(x,\,\pi) = -\,\varepsilon\partial_x\overline{\mathcal{H}}\big(x,\,\varepsilon^{-1}\pi\big) \tag{21}$$

again correspond to $\overline{\sigma}_{\rm can}$, which does not depend on ε. Of course, as long as ε is fixed, both formulations are completely equivalent, since we can replace π by $\varepsilon\pi$ in (21). However, if we try to identify leading order dynamics by expansions in powers of ε we obtain very different results. In fact, $\mathcal{L}_\varepsilon$, $\mathcal{H}_\varepsilon$ and σ_ε scale in the same way and, hence, ε drops out in both the corresponding Lagrangian and Hamiltonian equations on TQ. On the other hand, for a normal Lagrangian, as given in (18), we find $\overline{\mathcal{H}}_\varepsilon = \frac{1}{2\varepsilon}\langle\pi,\,M^{-1}\pi\rangle + \varepsilon\mathcal{V}(x)$ and the formal expansion of $\overline{\mathcal{H}}_\varepsilon$ gives $\overline{H}^{\rm red}(\pi) = \frac{1}{2\varepsilon}\langle\pi,\,M^{-1}\pi\rangle$ as "leading order" Hamiltonian on T^*Q. In particular, the corresponding canonical equations $x_t = \varepsilon^{-1}M^{-1}\pi$ and $\pi_t = 0$ do not recover the original dynamics.

More generally, the key difference between tangent and cotangent Hamiltonian structures is related to the following **Principle of Consistent Expansions**:

Principle 3. *Suppose that the Lagrangian $\mathcal{L}$ obeys a (formal) expansion in powers of a parameter ε, i.e.,*

$$\mathcal{L}(\varepsilon) = \varepsilon^\kappa\left(\mathcal{L}_0 + \varepsilon\mathcal{L}_1 + \varepsilon^2\mathcal{L}_2 + \ldots\right). \tag{22}$$

Then the Hamiltonian structure on TQ obeys a corresponding expansion

$$\mathcal{H}(\varepsilon) = \varepsilon^\kappa \left(\mathcal{H}_0 + \varepsilon\mathcal{H}_1 + \varepsilon^2\mathcal{H}_2 + ...\right), \quad \sigma(\varepsilon) = \varepsilon^\kappa \left(\sigma_0 + \varepsilon\sigma_1 + \varepsilon^2\sigma_2 + ...\right)$$

and all expansions are consistent. *This means, for each order ε^i we have*

$$\mathcal{H}_i = \langle \partial_{x_t}\mathcal{L}_i,\, x_t\rangle - \mathcal{L}_i, \qquad \sigma_i = (\mathfrak{F}\mathcal{L}_i)^*\,\overline{\sigma}_{\mathrm{can}}.$$

Proof. Since the fiber-derivative operation acts linearly on the Lagrangian we find

$$\mathfrak{F}\mathcal{L} = \varepsilon^\kappa \left(\mathfrak{F}\mathcal{L}_0 + \varepsilon\mathfrak{F}\mathcal{L}_1 + \varepsilon^2\mathfrak{F}\mathcal{L}_2 + ...\right)$$

and this implies both the existence and consistency of the expansion of the Hamiltonian structure. □

The validity of Principle 3 is a remarkable property of the Hamiltonian structure on TQ and has no analogue on T^*Q. In fact, (22) implies a consistent expansion for the canonical momentum π, i.e. $\pi(\varepsilon) = \varepsilon^\kappa\left(\pi_0 + \varepsilon\pi_1 + \varepsilon^2\pi_2 + ...\right)$ with $\pi_i = \partial_{x_t}\mathcal{L}_i$, but replacing x_t by π we normally end up with a non-consistent expansion for the Hamiltonian $\overline{\mathcal{H}}$ on T^*Q.

In the context of this paper we do not apply Principle 3 to the microscopic Lagrangian or Hamiltonian structures, since usually these do not depend on scaling parameters. However, the two-scale transformations introduced in §2.2 strongly depend on ε and so do the *transformed* Lagrangian and Hamiltonian structures. Thus, for the purpose of model reduction the tangent framework turns out to be very convenient as it provides the consistency of the Lagrangian and Hamiltonian structures for *all* powers of ε.

2.2 Exact two-scale transformations

As mentioned in the introduction, any micro-macro transition relies on an exact two-scale transformation which obviously changes the Lagrangian and Hamiltonian structures. All of the two-scale transformations considered in this paper are superpositions of elementary building blocks, namely

1. (weak) symmetry transformations,
2. moving-frame transformations,
3. scalings of space and time coordinates.

In this section we aim to describe how each of these building blocks transforms the Lagrangian and Hamiltonian structures on TQ. The concepts of symmetry and moving-frame transformations are well established in the theory of Hamiltonian systems, but since they are usually studied on the cotangent bundle we start with the reformulation of standard results.

2.2.1 Linear transformations

Let $\mathcal{T}_{\mathrm{con}} : Q \to \tilde{Q}$ be a linear isomorphism between Q and another Hilbert space $\tilde{Q}$ with inverse $\tilde{\mathcal{T}}_{\mathrm{con}} : \tilde{Q} \to Q$. The canonical lifts of $\mathcal{T}_{\mathrm{con}}$ and $\tilde{\mathcal{T}}_{\mathrm{con}}$ to the tangent bundles are denoted by $\mathcal{T}_{\mathrm{vel}} : TQ \to T\tilde{Q}$ and $\tilde{\mathcal{T}}_{\mathrm{vel}} : T\tilde{Q} \to TQ$, respectively, and satisfy $\mathcal{T}_{\mathrm{vel}}(x,\, x_t) = (\mathcal{T}_{\mathrm{con}}x,\, \mathcal{T}_{\mathrm{con}}x_t)$ as well as $\tilde{\mathcal{T}}_{\mathrm{vel}} = (\mathcal{T}_{\mathrm{vel}})^{-1}$.

Remark 4. In what follows we use the inverse transformation $\tilde{\mathcal{T}}_{\mathrm{vel}}$ in order to pull back *forms* from Q (the pull-back with respect to $\tilde{\mathcal{T}}_{\mathrm{vel}}$ is the push-forward with respect to $\mathcal{T}_{\mathrm{vel}}$). In particular, we pull back functions $\mathcal{F}$ (0-forms) and symplectic forms σ (2-forms). The images under this operation are denoted by $\tilde{\mathcal{F}} = (\tilde{\mathcal{T}}_{\mathrm{vel}})^*\mathcal{F}$ and $\tilde{\sigma} = (\tilde{\mathcal{T}}_{\mathrm{vel}})^*\sigma$, and satisfy

$$\tilde{\mathcal{F}}(\tilde{z}) = \mathcal{F}(\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}), \qquad \tilde{\sigma}|_{\tilde{z}}(\dot{\tilde{z}},\, \acute{\tilde{z}}) = \sigma|_{\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}}(\tilde{\mathcal{T}}_{\mathrm{vel}}\dot{\tilde{z}},\, \tilde{\mathcal{T}}_{\mathrm{vel}}\acute{\tilde{z}}),$$

where $\tilde{z} \in T\tilde{Q}$ and $\dot{\tilde{z}},\, \acute{\tilde{z}} \in T|_{\tilde{z}}T\tilde{Q}$.

Theorem 5. *Let $\tilde{\mathcal{L}} = \mathcal{L} \circ \tilde{\mathcal{T}}_{\mathrm{vel}}$ be the transformed Lagrangian and $(\tilde{\mathcal{H}},\, \tilde{\sigma})$ the associated Hamiltonian structure on $T\tilde{Q}$. Then, $\tilde{\mathcal{H}}$ and $\tilde{\sigma}$ equal the transformed Hamiltonian and symplectic form, respectively.*

Proof. Let $z = (x,\, x_t)$ be given, and $\tilde{z} = \mathcal{T}_{\mathrm{vel}}z = (\tilde{x},\, \tilde{x}_t)$. The definition of $\tilde{\mathcal{L}}$ implies

$$\tilde{\pi}(\tilde{z}) = \partial_{\tilde{x}_t}\tilde{\mathcal{L}}(z) = (\tilde{\mathcal{T}}_{\mathrm{con}})'\pi(\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}), \quad \text{i.e.} \quad \langle \tilde{\pi}(\tilde{z}),\, \cdot \rangle_{\tilde{Q}} = \langle \pi(\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}),\, \tilde{\mathcal{T}}_{\mathrm{con}}\cdot \rangle_Q, \tag{23}$$

where $(\tilde{\mathcal{T}}_{\mathrm{con}})'$ is the adjoint operator to $\tilde{\mathcal{T}}_{\mathrm{con}}$. From this identity we derive

$$\tilde{\mathcal{H}}(\tilde{z}) = \langle \tilde{\pi}(\tilde{z}),\, \tilde{x}_t \rangle_{\tilde{Q}} - \tilde{\mathcal{L}}(\tilde{z}) = \langle \pi(\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}),\, \tilde{\mathcal{T}}_{\mathrm{con}}\tilde{x}_t \rangle_Q - \mathcal{L}(\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}) = \mathcal{H}(\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}),$$

as well as

$$\langle D\tilde{\pi}|_{\tilde{z}}(\dot{\tilde{z}}),\, \cdot \rangle_{\tilde{Q}} = \langle D\pi|_{\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}}(\tilde{\mathcal{T}}_{\mathrm{vel}}\dot{\tilde{z}}),\, \tilde{\mathcal{T}}_{\mathrm{con}}\cdot \rangle_Q \tag{24}$$

for all $\dot{\tilde{z}} \in T|_{\tilde{z}}T\tilde{Q}$. Finally, combining (24) with (19) for $z = \tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}$ we find

$$\begin{aligned}
\sigma|_{\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}}(\tilde{\mathcal{T}}_{\mathrm{vel}}\dot{\tilde{z}},\, \tilde{\mathcal{T}}_{\mathrm{vel}}\acute{\tilde{z}}) &= \langle D\pi|_{\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}}(\tilde{\mathcal{T}}_{\mathrm{vel}}\acute{\tilde{z}}),\, \tilde{\mathcal{T}}_{\mathrm{con}}\dot{\tilde{x}} \rangle_Q - \langle D\pi|_{\tilde{\mathcal{T}}_{\mathrm{vel}}\tilde{z}}(\tilde{\mathcal{T}}_{\mathrm{vel}}\dot{\tilde{z}}),\, \tilde{\mathcal{T}}_{\mathrm{con}}\acute{\tilde{x}} \rangle_Q \\
&= \langle D\tilde{\pi}|_{\tilde{z}}(\acute{\tilde{z}}),\, \dot{\tilde{x}} \rangle_{\tilde{Q}} - \langle D\tilde{\pi}|_{\tilde{z}}(\dot{\tilde{z}}),\, \acute{\tilde{x}} \rangle_{\tilde{Q}} = \tilde{\sigma}|_{\tilde{z}}(\dot{\tilde{z}},\, \acute{\tilde{z}}),
\end{aligned}$$

and the proof is complete. □

Corollary 6. *The following equivalences are satisfied:*

1. *A curve $t \mapsto x(t) \in Q$ satisfies the Lagrangian equation to $\mathcal{L}$ if and only if the transformed curve $t \mapsto \tilde{x}(t) = \mathcal{T}_{\mathrm{con}}x(t) \in \tilde{Q}$ satisfies the Lagrangian equation to $\tilde{\mathcal{L}}$.*

2. *A curve $t \mapsto z(t) \in TQ$ satisfies the Hamiltonian equation to $(\mathcal{H},\, \sigma)$ if and only if the transformed curve $t \mapsto \tilde{z}(t) = \mathcal{T}_{\mathrm{vel}}z(t) \in T\tilde{Q}$ satisfies the Hamiltonian equation to $(\tilde{\mathcal{H}},\, \tilde{\sigma})$.*

2.2.2 Weak symmetry transformations

We introduce the notion of a *weak symmetry transformation* which describes a certain class of linear and invertible operators from Q into Q. Although both the Lagrangian and Hamiltonian structures are not invariant they behave nicely under such transformations. In particular, each weak symmetry transformation changes neither the fiber derivative of $\mathcal{L}$ nor the symplectic form σ.

Definition 7. A *weak symmetry transformation* (with respect to the Lagrangian $\mathcal{L}$) is a linear isomorphism $\mathcal{T}_{\rm con} : Q \to Q$ with the following properties:

1. $\mathcal{T}_{\rm con}$ is unitary, this means $\langle \mathcal{T}_{\rm con} x,\, \tilde{x}\rangle = \langle x,\, \tilde{\mathcal{T}}_{\rm con}\tilde{x}\rangle$ for all $x,\, \tilde{x} \in Q$.
2. The canonical momentum $\pi = \partial_{x_t}\mathcal{L}$ commutes with $\mathcal{T}_{\rm con}$ in the sense that
$$\pi(\mathcal{T}_{\rm vel} z) = \mathcal{T}_{\rm con}\pi(z) \tag{25}$$
holds for all $z = (x,\, x_t) \in TQ$.

Moreover, $\mathcal{T}_{\rm con}$ is called a *symmetry transformation* if it respects the Lagrangian, i.e., $\mathcal{L} = \tilde{\mathcal{L}}$ in the sense of Theorem 5.

Remark 8. (*i*) Unitarity implies $\tilde{\mathcal{T}}_{\rm con} = \mathcal{T}_{\rm con}'$ and $\langle x_1,\, x_2\rangle = \langle \mathcal{T}_{\rm con} x_1,\, \mathcal{T}_{\rm con} x_2\rangle$ for all $x_1, x_2 \in Q$. (*ii*) Condition (25) is equivalent to $\pi = \tilde{\pi}$, see (23), and this implies $\mathfrak{F}\mathcal{L} = \mathfrak{F}\tilde{\mathcal{L}}$ and $\sigma = \tilde{\sigma}$. (*iii*) Each symmetry transformation satisfies $\mathcal{H} = \tilde{\mathcal{H}}$ and $\sigma = \tilde{\sigma}$. (*iv*) $\mathcal{L} = \tilde{\mathcal{L}}$ is sufficient for (25).

Example 9. Let $Q = L^2(\mathbb{R}\times T^1;\, \mathrm{d}\eta\mathrm{d}\phi)$ be the Lebesgue space of functions x depending on $\eta \in \mathbb{R}$ and a periodic phase variable $\phi \in T^1 \cong [0,\, 2\pi]$, and let the unitary operator $\mathcal{T}_{\rm con}$ be defined by $(\mathcal{T}_{\rm con}\, x)(\eta,\, \phi) = x(\eta,\, \phi + s_0\eta)$ for some s_0. The Lagrangian $\mathcal{L}$ of the embedded Klein–Gordon equation, cf. (11), is *not* invariant under the action of $\mathcal{T}_{\rm con}$ as the differential operator ∂_η transforms into $\partial_\eta + s_0\partial_\phi$. However, the condition (25) is satisfied.

2.2.3 Groups of symmetry transformations

The concept of symmetry groups is well established in mechanics and mathematics and plays a fundamental role in the analysis of Hamiltonian systems. Here we summarize the definitions and basic properties.

Definition 10. A *(weak) symmetry group* (with respect to the Lagrangian $\mathcal{L}$) is a one-parameter family of $s \mapsto \mathcal{T}_{\rm con}(s)$, $s \in \mathbb{R}$, of (weak) symmetry transformations that satisfies the following properties:

1. The family is a group of unitary transformations, i.e., $\mathcal{T}_{\rm con}(0) = \mathrm{Id}_{Q\to Q}$ and
$$\mathcal{T}_{\rm con}(s+\tilde{s}) = \mathcal{T}_{\rm con}(s)\mathcal{T}_{\rm con}(\tilde{s}), \qquad \tilde{\mathcal{T}}_{\rm con}(s) = (\mathcal{T}_{\rm con}(s))^{-1} = (\mathcal{T}_{\rm con}(s))' = \mathcal{T}_{\rm con}(-s)$$
for all $s, \tilde{s} \in \mathbb{R}$.

2. The generator $\mathcal{A}_{\rm con}$ with $\mathcal{A}_{\rm con}x = \lim_{s\to 0} s^{-1}(\mathcal{T}_{\rm con}(s)x - x)$ is defined on a dense subset of Q. Consequently, the group $s \mapsto \mathcal{T}_{\rm vel}(s)$ is generated by $\mathcal{A}_{\rm vel} = \mathcal{A}_{\rm con} \times \mathcal{A}_{\rm con}$.

Remark 11. If $\mathcal{L}$ is invariant under the action of a symmetry group Noether's Theorem provides the integral of motion

$$\mathcal{I}(x,\, x_t) = \langle \pi(x,\, x_t),\, \mathcal{A}_{\rm con}x\rangle, \quad \pi(x,\, x_t) = \partial_{x_t}\mathcal{L}(x,\, x_t), \tag{26}$$

i.e., $\mathcal{I}$ is conserved for any solution to the Hamiltonian equation (20).

Example 12. Let Q and $\mathcal{L}$ be as in Example 9, and for fixed $\eta_0 \in \mathbb{R}$, $\phi_0 \in \mathbb{R}$ and all $s \in \mathbb{R}$ let $(\mathcal{T}_{\rm con}(s)\, x)(\eta,\, \phi) = x(\eta + s\eta_0,\, \phi + s\phi_0)$. Then, $s \mapsto \mathcal{T}_{\rm con}(s)$ is a symmetry group with generator $\mathcal{A}_{\rm con} = \eta_0\partial_\eta + \phi_0\partial_\phi$ and integral of motion $\mathcal{I} = \eta_0\mathcal{I}_{\rm space} + \phi_0\mathcal{I}_{\rm phase}$, where $\mathcal{I}_{\rm space}$ and $\mathcal{I}_{\rm phase}$ are given by (12).

Lemma 13. *Each (weak) symmetry group satisfies* $\sigma|_z(\mathcal{A}_{\rm vel}z,\, \cdot) = \mathrm{d}\mathcal{I}|_z(\cdot)$ *for all* $z \in TQ$ *and* $\mathcal{I}$ *from* (26).

Proof. For given $z = (x,\, x_t) \in TQ$ consider the curve $t \mapsto z(t) = \mathcal{T}_{\rm vel}(t)z \in TQ$, and its image under $\mathfrak{F}\mathcal{L}$, that is $t \mapsto \overline{z}(t) = (x(t),\, \pi(t))$ with $x(t) = \mathcal{T}_{\rm con}(t)x$ and $\pi(t) = \pi(z(t))$. Moreover, let $\acute{\overline{z}} = (\acute{x},\, \acute{\pi})$ be an arbitrary tangent vector in $T|_{\overline{z}(0)}T^*Q$. Condition (25) implies $\langle \pi(t),\, \acute{x}\rangle = \langle \pi(\mathcal{T}_{\rm vel}(t)z(0)),\, \acute{x}\rangle = \langle \pi(0),\, \mathcal{T}_{\rm con}(-t)\, \acute{x}\rangle$ and differentiation and evaluation for $t = 0$ yield $\langle \pi_t(0),\, \acute{x}\rangle = -\langle \pi(0),\, \mathcal{A}_{\rm con}\acute{x}\rangle$. This identity and the definition of $\overline{\sigma}_{\rm can}$ provide

$$\begin{aligned}\overline{\sigma}_{\rm can}|_{\overline{z}(0)}(\overline{z}_t(0),\, \acute{\overline{z}}) &= \overline{\sigma}_{\rm can}|_{\overline{z}_0}\Big((\mathcal{A}_{\rm con}x(0),\, \pi_t(0)),\, (\acute{x},\, \acute{\pi})\Big)\\ &= \langle \acute{\pi},\, \mathcal{A}_{\rm con}x(0)\rangle - \langle \pi_t(0),\, \mathcal{A}_{\rm con}\acute{x}\rangle = \langle \acute{\pi},\, \mathcal{A}_{\rm con}x(0)\rangle + \langle \pi(0),\, \mathcal{A}_{\rm con}\acute{x}\rangle.\end{aligned}$$

Moreover, for $\overline{\mathcal{I}}(x,\, \pi) = \langle \pi,\, \mathcal{A}_{\rm con}x\rangle$ we find $\mathrm{d}\overline{\mathcal{I}}|_{\overline{z}(0)}\big(\acute{\overline{z}}\big) = \langle \acute{\pi},\, \mathcal{A}_{\rm con}\, x(0)\rangle + \langle \pi(0),\, \mathcal{A}_{\rm con}\, \acute{x}\rangle$ and, hence, $\overline{\sigma}_{\rm can}|_{\overline{z}(0)}(\overline{z}_t(0),\, \cdot) = \mathrm{d}\overline{\mathcal{I}}|_{\overline{z}(0)}(\cdot)$. Finally, pulling back this identity via $\mathfrak{F}\mathcal{L}$ and using $z_t(0) = \mathcal{A}_{\rm vel}z(0)$ completes the proof . □

2.2.4 Moving frames

In this section we consider a time-parametrized family of invertible transformations of the configuration space $\mathcal{M}_{\rm con}(t) : Q \to Q$ and denote the family of inverse transformations by $\tilde{\mathcal{M}}_{\rm con}(t)$. Taking into account the time dependence we shall lift this transformation to the tangent bundle as follows: Each time-parametrized curve $t \mapsto x(t)$ in Q provides a lifted curve $t \mapsto (x(t),\, x_t(t))$ in TQ, where $x_t(t)$ denotes the tangent vector at time t, i.e. $x_t(t) = \frac{\mathrm{d}}{\mathrm{d}t}x(t)$. Consequently, the lift of the transformed curve $t \mapsto \tilde{x}(t) = \mathcal{M}_{\rm con}(t)\, x(t)$ is given by

$$t \mapsto (\tilde{x}(t),\, \tilde{x}_t(t)) = \big(\tilde{x}(t),\, \tfrac{\mathrm{d}}{\mathrm{d}t}\tilde{x}_t(t)\big) = \big(\mathcal{M}_{\rm con}(t)\, x(t),\, \mathcal{M}_{\rm con}(t)\, x_t(t) + \big(\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}_{\rm con}(t)\big)\, x(t)\big),$$

and we read-off the definition of $\mathcal{M}_{\rm vel}(t)$, that is

$$\mathcal{M}_{\rm vel}(t) : (x,\, x_t) \mapsto (\tilde{x},\, \tilde{x}_t) = \big(\mathcal{M}_{\rm con}(t)\, x,\, \mathcal{M}_{\rm con}(t)\, x_t + \big(\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}_{\rm con}(t)\big)\, x\big). \tag{27}$$

The transformation of a Hamiltonian structure under a time-dependent transformation is in general quite complicated. Therefore we solely discuss time-dependent transformations that are related to *moving frames*.

Definition 14. The transformation $\mathcal{M}_{\rm con}(t)$ is called a *moving-frame transformation* (with respect to the Lagrangian $\mathcal{L}$) if it is related to a symmetry group $s \mapsto \mathcal{T}_{\rm con}(s)$ via $\mathcal{M}_{\rm con}(t) = \mathcal{T}_{\rm con}(t)$. This implies $\mathcal{M}_{\rm con}(0) = \mathrm{Id}_{Q\to Q}$ and $\tilde{\mathcal{M}}_{\rm con}(t) = \mathcal{M}_{\rm con}(-t)$ for all t.

Example 15. Let Q be as in Example 9, and let $\mathcal{L}_{\rm met}$ be the metric Lagrangian from Example 2. Obviously, $\mathcal{L}_{\rm met}$ is invariant under Galilean transformations $(t,\,\eta,\,\phi) \mapsto (t,\,\tilde\eta,\,\phi)$, where $\tilde\eta = \eta - ct$ denotes the spatial coordinate in the moving frame. The corresponding time-dependent transformations $\mathcal{M}_{\rm con}(t) : x \mapsto \tilde x$ and $\mathcal{M}_{\rm vel}(t) : (x,\,x_t) \mapsto (\tilde x,\,\tilde x_t)$ can be read-off from the identification $x(t,\,\eta,\,\phi) = \tilde x(t,\,\eta - ct,\,\phi)$ and are given by

$$\tilde x(\eta,\,\phi) = x(\eta + ct,\,\phi), \quad \tilde x_t(\eta,\,\phi) = x_t(\eta + ct,\,\phi) + c\,x_\eta(\eta + ct,\,\phi),$$

where x_η abbreviates the derivative of x with respect to η. The underlying symmetry group $(\mathcal{T}_{\rm con}(s)x)(\eta,\,\phi) = x(\eta + cs,\,\phi)$ has the generator $\mathcal{A}_{\rm con} = c\partial_\eta$ and the conserved quantity $\mathcal{I}(x,\,x_t) = c\int_{\mathbb{R}} x_t\,x_\eta\,\mathrm{d}\eta \in \mathbb{R}$. The lifted transformations $\mathcal{M}_{\rm vel}(t)$ and $\mathcal{T}_{\rm vel}(t)$ are really different because of $\mathcal{T}_{\rm vel}(t) = \mathcal{T}_{\rm con}(t) \times \mathcal{T}_{\rm con}(t)$.

For moving-frame transformations we can decompose the lifted map as follows: Definition 14 implies $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{M}_{\rm con}(t) = \mathcal{A}_{\rm con}\mathcal{T}_{\rm con}(t)$, and using (27) we conclude that

$$\begin{aligned} \mathcal{M}_{\rm vel}(t) &= \mathcal{R}_{\rm vel} \circ \mathcal{T}_{\rm vel}(t) \quad \text{with} \quad \mathcal{R}_{\rm vel} : (x,\,x_t) \mapsto (x,\,x_t + \mathcal{A}_{\rm con}\,x), \\ \tilde{\mathcal{M}}_{\rm vel}(t) &= \tilde{\mathcal{T}}_{\rm vel}(t) \circ \tilde{\mathcal{R}}_{\rm vel} \quad \text{with} \quad \tilde{\mathcal{R}}_{\rm vel} : (\tilde x,\,\tilde x_t) \mapsto (\tilde x,\,\tilde x_t - \mathcal{A}_{\rm con}\,\tilde x). \end{aligned}$$

In what follows we denote by $\tilde{\mathcal{L}}$ the transformed Lagrangian, i.e. $\tilde{\mathcal{L}}(t) = \mathcal{L} \circ \tilde{\mathcal{M}}_{\rm vel}(t)$, and with $(\tilde{\mathcal{H}},\,\tilde\sigma)$ the Hamiltonian structure corresponding to $\tilde{\mathcal{L}}$. However, since the Legendre transformation does not commute with $\mathcal{M}_{\rm vel}(t)$ we can not expect $\tilde{\mathcal{H}}$, that is the Legendre transform of $\tilde{\mathcal{L}}$, to equal the transformed Hamiltonian. For this reason we identify $\mathcal{H}$ with $\mathcal{E}$, and define $\tilde{\mathcal{E}} = \mathcal{E} \circ \tilde{\mathcal{M}}_{\rm vel}(t)$. This notation is motivated by normal systems, see (18), for which the Hamiltonian $\mathcal{H}$ equals the total energy $\mathcal{E} = \mathcal{K} + \mathcal{V}$. Finally, we write $\tilde{\mathcal{I}} = \mathcal{I} \circ \tilde{\mathcal{M}}_{\rm vel}(t)$, where $\mathcal{I}(z) = \langle \pi(z),\,\mathcal{A}_{\rm con}x\rangle$ is the integral of motion associated to the symmetry group.

Next we prove that all these quantities do not depend on time, as it is already indicated by the notation, and derive the transformation rules for the Hamiltonian structure.

Theorem 16. *Moving-frame transformations satisfy*

$$\tilde{\mathcal{L}} = \mathcal{L} \circ \tilde{\mathcal{R}}_{\rm vel}, \quad \tilde{\mathcal{E}} = \mathcal{E} \circ \tilde{\mathcal{R}}_{\rm vel}, \quad \tilde{\mathcal{I}} = \mathcal{I} \circ \tilde{\mathcal{R}}_{\rm vel}.$$

Moreover, we have

$$\tilde{\mathcal{H}} = \tilde{\mathcal{E}} + \tilde{\mathcal{I}}, \quad \tilde\sigma = (\tilde{\mathcal{R}}_{\rm vel})^*\sigma,$$

where $(\tilde H,\,\tilde\sigma)$ is the Hamiltonian structure associated to $\tilde{\mathcal{L}}$.

Proof. Let t be fixed, and for arbitrary $z = (x,\, x_t)$ let $\tilde{z} = (\tilde{x},\, \tilde{x}_t) = \mathcal{M}_{\mathrm{vel}}(t)z$. Due to the invariance of $\mathcal{L}$ under $\mathcal{T}_{\mathrm{vel}}(t)$ we have $\tilde{\mathcal{L}} = (\mathcal{L} \circ \mathcal{T}_{\mathrm{vel}}(-t)) \circ \tilde{\mathcal{R}}_{\mathrm{vel}} = \mathcal{L} \circ \tilde{\mathcal{R}}_{\mathrm{vel}}$, and this implies

$$\tilde{\pi}(\tilde{z}) = \partial_{\tilde{x}_t} \tilde{\mathcal{L}}(\tilde{z}) = \partial_{\tilde{x}_t}(\mathcal{L}(\tilde{x},\, \tilde{x}_t - \mathcal{A}_{\mathrm{con}}\tilde{x}_t)) = \pi(\tilde{x},\, \tilde{x}_t - \mathcal{A}_{\mathrm{con}}\tilde{x})$$

so that $\tilde{\pi} = \pi \circ \tilde{\mathcal{R}}_{\mathrm{vel}}$. We conclude that $\mathfrak{F}\tilde{\mathcal{L}} = \mathfrak{F}\mathcal{L} \circ \tilde{\mathcal{R}}_{\mathrm{vel}} = (\tilde{\mathcal{R}}_{\mathrm{vel}})^* \mathfrak{F}\mathcal{L}$ and hence $\tilde{\sigma} = (\tilde{\mathcal{R}}_{\mathrm{vel}})^* \sigma$. The unitarity of $\tilde{\mathcal{T}}_{\mathrm{con}}(t)$, the identity $\mathcal{A}_{\mathrm{con}}\tilde{\mathcal{T}}_{\mathrm{con}}(t) = \tilde{\mathcal{T}}_{\mathrm{con}}(t)\mathcal{A}_{\mathrm{con}}$ and Formula (25) yield

$$\begin{aligned}
\tilde{\mathcal{I}}(\tilde{z}) &= \mathcal{I}(\tilde{\mathcal{M}}_{\mathrm{vel}}(t)\tilde{z}) = \langle \pi(\tilde{\mathcal{M}}_{\mathrm{vel}}(t)\tilde{z}),\, \mathcal{A}_{\mathrm{con}}\tilde{\mathcal{T}}_{\mathrm{con}}(t)\tilde{x} \rangle = \langle \pi(\tilde{\mathcal{M}}_{\mathrm{vel}}(t)\tilde{z}),\, \tilde{\mathcal{T}}_{\mathrm{con}}(t)\mathcal{A}_{\mathrm{con}}\tilde{x} \rangle \\
&= \langle \pi(\mathcal{T}_{\mathrm{vel}}(t)\tilde{\mathcal{M}}_{\mathrm{vel}}(t)\tilde{z}),\, \mathcal{A}_{\mathrm{con}}\tilde{x} \rangle = \langle \pi(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),\, \mathcal{A}_{\mathrm{con}}\tilde{x} \rangle = \mathcal{I}(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),
\end{aligned}$$

the desired result for $\mathcal{I}$. Analogously, with $(\mathcal{E} + \mathcal{L})(z) = \langle \pi(z),\, x \rangle$ we find

$$\begin{aligned}
(\tilde{\mathcal{E}} + \tilde{\mathcal{L}})(\tilde{z}) &= (\mathcal{E} + \mathcal{L})(\tilde{\mathcal{M}}_{\mathrm{vel}}(t)\tilde{z}) = \langle \pi(\tilde{\mathcal{T}}_{\mathrm{vel}}(t)\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),\, \tilde{\mathcal{T}}_{\mathrm{con}}(t)(\tilde{x}_t - \mathcal{A}_{\mathrm{con}}\tilde{x}) \rangle \\
&= \langle \pi(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),\, \tilde{x}_t - \mathcal{A}_{\mathrm{con}}\tilde{x} \rangle = (\mathcal{E} + \mathcal{L})(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),
\end{aligned}$$

which implies the formula for $\mathcal{E}$. Finally,

$$\begin{aligned}
\tilde{\mathcal{H}}(\tilde{z}) &= \langle \tilde{\pi}(\tilde{z}),\, \tilde{x}_t \rangle - \tilde{\mathcal{L}}(\tilde{z}) = \langle \pi(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),\, \tilde{x}_t \rangle - \tilde{\mathcal{L}}(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}) \\
&= \langle \pi(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),\, \tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{x}_t \rangle + \langle \pi(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),\, \mathcal{A}_{\mathrm{con}}\tilde{x} \rangle - \tilde{\mathcal{L}}(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}) = \mathcal{E}(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}) + \mathcal{I}(\tilde{\mathcal{R}}_{\mathrm{vel}}\tilde{z}),
\end{aligned}$$

and the proof is finished. $\square$

$$\begin{tikzcd}
\mathcal{L} \arrow[d, "\mathcal{H} = \langle \partial_{x_t}\mathcal{L},\, x_t \rangle - \mathcal{L}"] & & \tilde{\mathcal{L}} \arrow[ll, "\tilde{\mathcal{M}}_{\mathrm{vel}}(t)"'] \arrow[d, "\tilde{\mathcal{H}} = \langle \partial_{\tilde{x}_t}\tilde{\mathcal{L}},\, \tilde{x}_t \rangle - \tilde{\mathcal{L}}"'] \\
\mathcal{H} = \mathcal{E} & \tilde{\mathcal{E}} \arrow[l, "\tilde{\mathcal{M}}_{\mathrm{vel}}(t)"'] \arrow[r, "\tilde{\mathcal{H}} = \tilde{\mathcal{E}} + \tilde{\mathcal{I}}"] & \tilde{\mathcal{H}}
\end{tikzcd}$$

Figure 2: $\mathcal{L}$, $\mathcal{E}$, $\mathcal{H}$ and their transformed counterparts in a moving frame.

The results of Theorem 16 can be reinterpreted as the *transformation rule* for Hamiltonian structures, see Figure 2. In fact, $\tilde{\sigma}$ equals the pull-back of σ, and to obtain $\tilde{\mathcal{H}}$ we pull back the sum of the Hamiltonian $\mathcal{E} = \mathcal{H}$ and conserved quantity $\mathcal{I}$. As a consequence we gain the following result.

Corollary 17. *The following equivalences are satisfied.*

1. *A curve $t \mapsto x(t) \in Q$ solves the Lagrangian equation to $\mathcal{L}$ if and only if the transformed curve $t \mapsto \tilde{x}(t) = \mathcal{M}_{\mathrm{con}}(t)x(t) \in Q$ solves the Lagrangian equation to $\tilde{\mathcal{L}}$.*

2. *A curve $t \mapsto z(t) \in TQ$ solves the Hamiltonian equation to $(\mathcal{H}, \sigma)$ if and only if the transformed curve $t \mapsto \tilde{z}(t) = \mathcal{M}_{\rm vel}(t) z(t) \in TQ$ solves the Hamiltonian equation to $(\tilde{\mathcal{H}}, \tilde{\sigma})$.*

Proof. Since $(\tilde{\mathcal{H}}, \tilde{\sigma})$ is the Hamiltonian structure associated to $\tilde{\mathcal{L}}$ it is sufficient to prove the equivalence in the Hamiltonian framework. Let $\hat{z}(t) = \tilde{\mathcal{R}}_{\rm vel} \tilde{z}(t)$ such that $z(t) = \tilde{\mathcal{M}}_{\rm vel}(t) \tilde{z}(t)$ satisfies $\hat{z}(t) = \mathcal{T}_{\rm vel}(t) z(t)$. Now suppose that $t \mapsto \tilde{z}(t)$ solves the Hamiltonian equation to $(\tilde{\mathcal{H}}, \tilde{\sigma})$. This means
$\tilde{\sigma}|_{\hat{z}(t_0)}(\tilde{z}_t(t)_0, \cdot) = \mathrm{d}\tilde{\mathcal{H}}|_{\tilde{z}(t_0)}(\cdot)$ for arbitrary but fixed t_0, and Theorem 16 provides

$$\sigma|_{\hat{z}(t_0)}(\hat{z}_t(t_0), \cdot) = \mathrm{d}\mathcal{E}|_{\hat{z}(t_0)}(\cdot) + \mathrm{d}\mathcal{I}|_{\hat{z}(t_0)}(\cdot).$$

By construction we have $\hat{z}_t(t_0) = \mathcal{T}_{\rm vel}(t_0) z_t(t_0) + \mathcal{A}_{\rm vel} z(t_0)$, and exploiting Lemma 13 we find

$$\sigma|_{\mathcal{T}_{\rm vel}(t_0) z(t_0)}(\mathcal{T}_{\rm vel}(t_0) z_t(t_0), \cdot) = \mathrm{d}\mathcal{E}|_{\mathcal{T}_{\rm vel}(t_0) z(t_0)}(\cdot) = \mathrm{d}\mathcal{H}|_{\mathcal{T}_{\rm vel}(t_0) z(t_0)}(\cdot),$$

and the invariance of $\mathcal{L}$, $\mathcal{H}$ and σ under $\mathcal{T}_{\rm vel}(t_0)$ (cf. Remark 8) shows that $t \mapsto z(t)$ solves the Hamiltonian equation to $(\mathcal{H}, \sigma)$. Finally, in order to establish the equivalence we argue in the reverse direction. $\square$

2.2.5 Scaling transformations

The two-scale problems considered in §3 involve suitable scalings of space and time variables. We always suppose that there exist positive constants β and γ such that $\tau = \varepsilon^\beta t$ and $y = \varepsilon^\gamma \eta$. In particular, $\frac{\mathrm{d}\tau}{\mathrm{d}t} = \varepsilon^\beta$ and $\frac{\mathrm{d}y}{\mathrm{d}\eta} = \varepsilon^\gamma$ are the scaling constants for time and space, respectively.

The spatial scaling can be encoded in a linear and invertible *scaling transformation* $\mathcal{S}_{\rm con} : Q \to P$ that maps the microscopic configuration space Q into P, the space of all macroscopic configurations. In what follows P is always a Hilbert space, usually some L^2–space, with inner product $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ and $\mathbb{S}_{\rm con} : P \to Q$ is the inverse to $\mathcal{S}_{\rm con}$. The elements of P are denoted by X and are functions of the macroscopic space variable y.

Definition 18. A *scaling transformation* is a scaled isometry $\mathcal{S}_{\rm con} : Q \to P$, i.e.,

$$\langle x, \tilde{x} \rangle = \varepsilon^\mu \langle\!\langle \mathcal{S}_{\rm con}\, x, \mathcal{S}_{\rm con}\, \tilde{x} \rangle\!\rangle$$

for some exponent μ and all $x, \tilde{x} \in Q$.

Notice that P does not depend on the scaling parameter ε, whereas the transformations $\mathcal{S}_{\rm con}$ and $\mathbb{S}_{\rm con}$ as well as the scaled Lagrangian and Hamiltonian structures will strongly depend on ε. Nevertheless, for the moment ε is an arbitrary but fixed parameter and hence we do not denote explicitly the dependence on ε.

Example 19. Let $Q = L^2(\mathbb{R};\,\mathrm{d}\eta)$ and $P = L^2(\mathbb{R};\,\mathrm{d}y)$, and consider the two-scale ansatz $x(t,\,\eta) = \varepsilon^\alpha X\big(\varepsilon^\beta t,\,\varepsilon^\gamma\eta\big)$. In this case we have

$$\big(\mathcal{S}_{\rm con}\,x\big)(y) = \varepsilon^{-\alpha}x\big(\varepsilon^{-\gamma}y\big),\quad \big(\mathbb{S}_{\rm con}\,X\big)(\eta) = \varepsilon^{\alpha}X(\varepsilon^{\gamma}\eta),$$

providing

$$\langle x,\,\tilde{x}\rangle = \int\limits_{\mathbb{R}} x(\eta)\tilde{x}(\eta)\,\mathrm{d}\eta = \varepsilon^{2\alpha}\int\limits_{\mathbb{R}} X(\varepsilon^\gamma\eta)\tilde{X}(\varepsilon^\gamma\eta)\,\mathrm{d}\eta = \varepsilon^{2\alpha-\gamma}\langle\!\langle X,\,\tilde{X}\rangle\!\rangle.$$

The transformation $\mathcal{S}_{\rm con} : Q \to P$ does not take into account the time scaling, since this is related to *reparametrization of curves* as follows: Let $t \mapsto x(t)$ be any curve in Q with tangent vectors $t \mapsto x_t(t)$, and let $t \mapsto \tilde{X}(t) = \mathcal{S}_{\rm con}\,x(t)$ be the transformed curve in P, which has tangent vectors $t \mapsto \tilde{X}_t(t) = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{X}(t) = \mathcal{S}_{\rm con}\,x_t(t)$. In view of the time scaling we are not interested in $\tilde{X}(t)$, but refer to the reparametrized curve

$$\tau \mapsto X(\tau) = \tilde{X}(t(\tau)) = \mathcal{S}_{\rm con}\,x(t(\tau))$$

with rescaled tangent vectors

$$\tau \mapsto X_\tau(\tau) = \tfrac{\mathrm{d}}{\mathrm{d}\tau}X(\tau) = \tfrac{\mathrm{d}t}{\mathrm{d}\tau}\,\mathcal{S}_{\rm con}\,x_t(t(\tau)). \tag{28}$$

For this reason we denote elements of $TP \cong P{\times}P$ by $(X,\,X_\tau)$ instead of $(X,\,X_t)$. Moreover, we must take into account this reparametrization when defining $\mathcal{S}_{\rm vel}$, i.e. the lift of $\mathcal{S}_{\rm con}$ to a map $TQ \to TP$. In fact, using (28) we find

$$\mathcal{S}_{\rm vel} : (x,\,x_t) \mapsto (X,\,X_\tau) = \big(\mathcal{S}_{\rm con}\,x,\,\varepsilon^{-\beta}\mathcal{S}_{\rm con}\,x_t\big).$$

Example 20. Using the notations from Example 19 we obtain

$$\big(\mathcal{S}_{\rm vel}(x,\,x_t)\big)(y) = \big(\varepsilon^{-\alpha}x,\,\varepsilon^{-\alpha-\beta}x_t\big)\big(\varepsilon^{-\gamma}y\big),\quad \big(\mathbb{S}_{\rm vel}(X,\,X_\tau)\big)(\eta) = \big(\varepsilon^{\alpha}X,\,\varepsilon^{\alpha+\beta}X_\tau\big)(\varepsilon^\gamma\eta).$$

Following the proof of Theorem 5, we derive the transformation rules for the Lagrangian and Hamiltonian structures. To this end, let $\mathbb{L} = \mathcal{L}\circ\mathbb{S}_{\rm vel}$ be the transformed Lagrangian and $(\mathbb{H},\,\sigma)$ the associated Hamiltonian structure on TP, i.e., $\mathbb{H}$ is the Legendre transform of $\mathbb{L}$ and $\sigma = (\mathfrak{F}\mathbb{L})^*\overline{\sigma}_{\rm can}$, where $\overline{\sigma}_{\rm can}$ is the canonical symplectic structure on T^*P.

Theorem 21. *We have $\mathbb{H} = \mathcal{H}\circ\mathbb{S}_{\rm vel}$ and $\sigma = \frac{\mathrm{d}\tau}{\mathrm{d}t}\hat{\sigma}$, where $\hat{\sigma} = (\mathbb{S}_{\rm vel})^*\sigma$.*

Proof. The definition of $\mathbb{L}$ implies $\langle\!\langle\Pi(Z),\,\cdot\rangle\!\rangle = \langle\pi(\mathbb{S}_{\rm vel}Z),\,\frac{\mathrm{d}\tau}{\mathrm{d}t}\mathbb{S}_{\rm con}\cdot\rangle$, where $Z = (X,\,X_\tau)$ and $\Pi = \partial_{X_\tau}\mathbb{L}$. We conclude that

$$\mathbb{H}(Z) = \langle\!\langle\Pi(Z),\,X_\tau\rangle\!\rangle - \mathbb{L}(Z) = \langle\pi(\mathbb{S}_{\rm vel}Z),\,\tfrac{\mathrm{d}\tau}{\mathrm{d}t}\mathbb{S}_{\rm con}Z_\tau\rangle - \mathcal{L}(\mathbb{S}_{\rm vel}(Z)) = \mathcal{H}(\mathbb{S}_{\rm vel}Z)$$

and $\langle\!\langle D\Pi|_Z(\dot{Z}),\, \cdot\rangle\!\rangle = \langle D\pi|_{\mathbb{S}_{\rm vel}Z}(\mathbb{S}_{\rm vel}\dot{Z}),\, \frac{{\rm d}t}{{\rm d}\tau}\mathbb{S}_{\rm con}\cdot\rangle$ for all $\dot{Z} \in T|_Z TP$. Inserting this identity into the definition of σ, compare (19), we obtain

$$\begin{aligned}\sigma|_Z(\dot{Z},\, \acute{Z}) &= \langle\!\langle D\Pi|_Z(\acute{Z}),\, \dot{X}\rangle\!\rangle - \langle\!\langle D\Pi|_Z(\dot{Z}),\, \acute{Z}\rangle\!\rangle\\ &= \langle D\pi|_{\mathbb{S}_{\rm vel}Z}(\mathbb{S}_{\rm vel}\acute{Z}),\, \tfrac{{\rm d}\tau}{{\rm d}t}\mathbb{S}_{\rm con}\dot{X}\rangle - \langle D\pi|_{\mathbb{S}_{\rm vel}Z}(\mathbb{S}_{\rm vel}\dot{Z}),\, \tfrac{{\rm d}\tau}{{\rm d}t}\mathbb{S}_{\rm con}\acute{X}\rangle\\ &= \tfrac{{\rm d}\tau}{{\rm d}t}\langle D\pi|_{\mathbb{S}_{\rm vel}Z}(\mathbb{S}_{\rm vel}\acute{Z}),\, \mathbb{S}_{\rm con}\dot{X}\rangle - \tfrac{{\rm d}\tau}{{\rm d}t}\langle D\pi|_{\mathbb{S}_{\rm vel}Z}(\mathbb{S}_{\rm vel}\dot{Z}),\, \mathbb{S}_{\rm con}\acute{X}\rangle\\ &= \tfrac{{\rm d}\tau}{{\rm d}t}\sigma|_{\mathbb{S}_{\rm vel}Z}\Big(\mathbb{S}_{\rm vel}\dot{Z},\, \mathbb{S}_{\rm vel}\acute{Z}\Big) = \tfrac{{\rm d}\tau}{{\rm d}t}\hat{\sigma}|_Z\Big(\dot{Z},\, \acute{Z}\Big),\end{aligned}$$

which is the desired result for σ. □

The additional scaling parameter in the formula for σ appears naturally due to the reparametrization of curves. More precisely, the microscopic Hamiltonian equation is equivalent to

$$\hat{\sigma}|_Z\big(\tfrac{\rm d}{{\rm d}t}Z,\, \cdot\big) = {\rm d}\mathbb{H}|_z(\cdot),$$

but since here the solution still depends on t we reparametrize via $\frac{\rm d}{{\rm d}t} = \frac{{\rm d}\tau}{{\rm d}t}\frac{\rm d}{{\rm d}\tau}$.

Example 22. Let $\mathcal{L}$ be a normal Lagrangian, cf. (18), and consider a simple *time scaling* $t \rightsquigarrow \tau = \varepsilon t$ with $P = Q$ and the two-scale ansatz $x(t) = X(\varepsilon t)$. Then

$$\mathbb{S}_{\rm vel} = \begin{pmatrix} 1 & 0\\ 0 & \varepsilon\end{pmatrix}, \quad \mathbb{L}(X,\, X_\tau) = \tfrac{1}{2}\varepsilon^2\langle X_\tau,\, MX_\tau\rangle - \mathcal{V}(X) \quad \text{with} \quad \Pi(X,\, X_\tau) = \varepsilon^2 MX_\tau,$$

and a simple calculation yields

$$\mathbb{H}(X,\, X_\tau) = \tfrac{1}{2}\varepsilon^2\langle X_\tau,\, MX_\tau\rangle + \mathcal{V}(X), \quad \sigma = (\mathfrak{F}\mathbb{L})^*\overline{\sigma}_{\rm can} \simeq \varepsilon^2 M\begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix}.$$

The pull-back of σ via $\mathbb{S}_{\rm vel}$ is given by

$$\hat{\sigma} = (\mathbb{S}_{\rm vel})^*\sigma \simeq \mathbb{S}_{\rm vel}^T\begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix}\mathbb{S}_{\rm vel} = \varepsilon\begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix},$$

and differs from σ by the factor ε^{-1}.

In what follows we refer to $\mathbb{L}$ and $(\mathbb{H},\, \sigma)$ as the *macroscopic* Lagrangian and Hamiltonian structures, but we recall that microscopic and macroscopic structures are completely equivalent as long as ε is a fixed but positive parameter. Consequently, we find the following transformation rules for solutions.

Corollary 23. *The following equivalences are satisfied.*

1. *A curve $t \mapsto x(t) \in Q$ solves the microscopic Lagrangian equation to $\mathcal{L}$ if and only if the transformed and reparametrized curve $\tau \mapsto X(\tau) = \mathcal{S}_{\rm con}\, x(t(\tau)) \in P$ solves the macroscopic Lagrangian equation to $\mathbb{L}$.*
2. *A curve $t \mapsto z(t) \in TQ$ solves the microscopic Hamiltonian system to $(\mathcal{H},\, \sigma)$ if and only if the transformed and reparametrized curve $\tau \mapsto Z(\tau) = \mathcal{S}_{\rm vel}\, z(t(\tau)) \in TP$ solves the macroscopic Hamiltonian equation to $(\mathbb{H},\, \sigma)$.*

2.2.6 Two-scale transformations

Since space-time scalings depend on the parameter ε, from now on we denote a scaling transformation and its inverse by $\mathcal{S}_{\mathrm{con}}(\varepsilon)$ and $\mathbb{S}_{\mathrm{con}}(\varepsilon)$, respectively. Consequently, both the macroscopic Lagrangian and Hamiltonian structures will depend on ε, and thus we write $\mathbb{L} = \mathbb{L}(\varepsilon)$, $\mathbb{H} = \mathbb{H}(\varepsilon)$ and $\sigma = \sigma(\varepsilon)$. However, we always choose the macroscopic configuration space P as independent of ε.

The *two-scale transformations* considered in §3 are compositions of a scaling transformation $\mathcal{S}_{\mathrm{con}}(\varepsilon) : Q \to P$, a moving-frame transformation $\mathcal{M}_{\mathrm{con}}(t) : Q \to Q$ and a symmetry transformation $\mathcal{T}_{\mathrm{con}} : Q \to Q$. More precisely, a general exact two-scale transformation $\mathcal{T}_{\mathrm{con}}(\varepsilon,\,\tau) : Q \to P$ and its inverse $\mathbb{T}_{\mathrm{con}}(\varepsilon,\,t) : P \to Q$ are given by

$$\mathcal{T}_{\mathrm{con}}(\varepsilon,\,\tau) = \mathcal{S}_{\mathrm{con}}(\varepsilon) \circ \mathcal{M}_{\mathrm{con}}\big(t(\varepsilon,\,\tau)\big) \circ \mathcal{T}_{\mathrm{con}}, \quad \mathbb{T}_{\mathrm{con}}(\varepsilon,\,t) = \mathcal{T}_{\mathrm{con}}^{-1} \circ \mathcal{M}_{\mathrm{con}}(-t) \circ \mathbb{S}_{\mathrm{con}}(\varepsilon).$$

For convenience we parametrize forward and backward transformations by τ and t, respectively, i.e.,

$$\mathcal{T}_{\mathrm{con}}(\varepsilon,\,\tau(\varepsilon,\,t)) \circ \mathbb{T}_{\mathrm{con}}(\varepsilon,\,t) = \mathrm{Id}_{P\to P}, \quad \mathbb{T}_{\mathrm{con}}(\varepsilon,\,t(\varepsilon,\,\tau)) \circ \mathcal{T}_{\mathrm{con}}(\varepsilon,\,\tau) = \mathrm{Id}_{Q\to Q}.$$

Moreover, the lifted transformations are given by

$$\mathcal{T}_{\mathrm{vel}}(\varepsilon,\,\tau) = \mathcal{S}_{\mathrm{vel}}(\varepsilon) \circ \mathcal{M}_{\mathrm{vel}}\big(t(\varepsilon,\,\tau)\big) \circ \mathcal{T}_{\mathrm{vel}}, \quad \mathbb{T}_{\mathrm{vel}}(\varepsilon,\,t) = \mathcal{T}_{\mathrm{vel}}^{-1} \circ \mathcal{M}_{\mathrm{vel}}(-t) \circ \mathbb{S}_{\mathrm{vel}}(\varepsilon).$$

Example 24. The KdV reduction relies on the scaling $\tau = \varepsilon^3 t$, $y = \varepsilon\eta$ and the two-scale ansatz

$$x(t,\,\eta) = \varepsilon X\big(\varepsilon^3 t,\,\varepsilon\eta + \varepsilon c t\big),$$

where $x \in Q = L^2(\mathbb{R};\,\mathrm{d}\eta)$ and $X \in P = L^2(\mathbb{R};\,\mathrm{d}y)$. From this ansatz we can read-off directly the inverse two-scale transformation $\mathbb{T}_{\mathrm{con}}(\varepsilon,\,t) = \mathcal{M}_{\mathrm{con}}(-t) \circ \mathbb{S}_{\mathrm{con}}(\varepsilon)$, which consist of the inverse scaling transformation $(\mathbb{S}_{\mathrm{con}}(\varepsilon)\,X)(\eta) = \varepsilon X(\varepsilon\eta)$ and the inverse of the moving-frame transformation

$$(\mathcal{M}_{\mathrm{con}}(t)\,x)(\eta) = x(\eta - c\,t).$$

Moreover, $\mathcal{I}(x,\,x_t) = -c\,\mathcal{I}_{\mathrm{space}}(x,\,x_t) = -c\int_{\mathbb{R}} x_t(\eta)\,x_\eta(\eta)\,\mathrm{d}\eta$ is the integral of motion associated to $\mathcal{M}_{\mathrm{con}}(t)$.

For later purposes we prove two auxiliary results. The first lemma describes how to restrict Lagrangian and Hamiltonian structures to subspaces of P, and the second one allows us to compute Σ from Σ, the matrix-valued maps corresponding to σ and σ, respectively.

Lemma 25. *Let $\tilde{P} \subset P$ be a closed subspace of P, embedded via a linear and continuous operator $\tilde{\mathbb{J}}_{\mathrm{con}} : \tilde{P} \hookrightarrow P$ with canonical lift $\tilde{\mathbb{J}}_{\mathrm{vel}} = \tilde{\mathbb{J}}_{\mathrm{con}} \times \tilde{\mathbb{J}}_{\mathrm{con}} : T\tilde{P} \hookrightarrow TP$. Moreover, let $\tilde{\mathbb{L}} = \mathbb{L} \circ \tilde{\mathbb{J}}_{\mathrm{vel}}$ be the restricted Lagrangian on $T\tilde{P}$ and $(\tilde{\mathbb{H}},\,\tilde{\sigma})$ the associated Hamiltonian structure. Then, $\tilde{\mathbb{H}} = \mathbb{H} \circ \tilde{\mathbb{J}}_{\mathrm{vel}}$ and $\tilde{\sigma} = (\tilde{\mathbb{J}}_{\mathrm{vel}})^{*}\sigma$.*

Proof. The projector corresponding to $\tilde{\mathbb{J}}_{\text{vel}}$ is denoted by $\mathbb{J}_{\text{vel}} : TP \twoheadrightarrow T\tilde{P}$ and equals the adjoint of $\tilde{\mathbb{J}}_{\text{vel}}$. Notice that $\mathbb{J}_{\text{vel}} \circ \tilde{\mathbb{J}}_{\text{vel}} = \text{Id}_{T\tilde{P}}$ but $\ker(\tilde{\mathbb{J}}_{\text{vel}} \circ \mathbb{J}_{\text{vel}}) \supsetneq \{0\}$ for $\tilde{P} \subsetneq P$. Besides this modification the proof is entire similar to that of Theorem 5. $\square$

Lemma 26. *Let ε and t be fixed and suppose there exist two linear and invertible transformations $\boldsymbol{S} : TP \to TQ$ with $\langle \boldsymbol{S}\, Z,\, \boldsymbol{S}\, \tilde{Z}\rangle = \varepsilon^{\mu}\langle\langle Z,\, \tilde{Z}\rangle\rangle$ for some μ and $\boldsymbol{T} : TP \to TP$ such that $\mathbb{T}_{\text{vel}}(\varepsilon,\, t) = \boldsymbol{S} \circ \boldsymbol{T}$. Then,*

$$\Sigma|_Z = \tfrac{\mathrm{d}\tau}{\mathrm{d}t}\varepsilon^{\mu}\boldsymbol{T}'\boldsymbol{S}^{-1}\Sigma|_{\boldsymbol{ST}Z}\boldsymbol{ST}, \tag{29}$$

where $\boldsymbol{T}'$ is the adjoint to $\boldsymbol{T}$.

Proof. Let $Z \in TP$ be fixed and choose two arbitrary tangent vectors $\dot{Z},\, \acute{Z} \in T|_Z TP$. Moreover, set $z = \boldsymbol{ST}Z$ and $\dot{z} = \boldsymbol{ST}\dot{Z}$, $\acute{z} = \boldsymbol{ST}\acute{Z}$. The definition of σ and the linearity of $\mathbb{T}_{\text{vel}}(\varepsilon,\, t) = \boldsymbol{S} \circ \boldsymbol{T}$ imply $\sigma|_Z(\dot{Z},\, \acute{Z}) = \frac{\mathrm{d}\tau}{\mathrm{d}t}\, \sigma|_z(\dot{z},\, \acute{z})$, and this gives

$$\begin{aligned}\langle\langle \Sigma|_z\dot{Z},\, \acute{Z}\rangle\rangle &= \tfrac{\mathrm{d}\tau}{\mathrm{d}t}\, \langle \Sigma|_z\dot{z},\, \acute{z}\rangle = \tfrac{\mathrm{d}\tau}{\mathrm{d}t}\, \langle \boldsymbol{SS}^{-1}\Sigma|_z\boldsymbol{ST}\dot{Z},\, \boldsymbol{ST}\acute{Z}\rangle \\ &= \tfrac{\mathrm{d}\tau}{\mathrm{d}t}\, \varepsilon^{\mu}\langle\langle \boldsymbol{S}^{-1}\Sigma|_z\boldsymbol{ST}\dot{Z},\, \boldsymbol{T}\acute{Z}\rangle\rangle = \tfrac{\mathrm{d}\tau}{\mathrm{d}t}\, \varepsilon^{\mu}\langle\langle \boldsymbol{T}'\boldsymbol{S}^{-1}\Sigma|_z\boldsymbol{ST}\dot{Z},\, \acute{Z}\rangle\rangle,\end{aligned}$$

the desired result. $\square$

Remark 27. For $\Sigma = \Sigma_{\text{met}}$ and $\boldsymbol{S} = \tilde{\boldsymbol{S}} \times \tilde{\boldsymbol{S}}$ with $\tilde{\boldsymbol{S}} : P \to Q$ we have $\boldsymbol{S}^{-1}\Sigma\boldsymbol{S} = \Sigma_{\text{met}}$, where Σ_{met} and Σ_{met} correspond to the metric symplectic forms on TQ and TP, respectively, see Example 2. In this case (29) becomes $\Sigma = \frac{\mathrm{d}\tau}{\mathrm{d}t}\varepsilon^{\mu}\boldsymbol{T}'\Sigma_{\text{met}}\boldsymbol{T}$.

2.3 Reduction principles

In this section we suppose that an exact two-scale transformation has already transformed the original microscopic system into a macroscopic one on TP, where P is a Hilbert space with inner product $\langle\langle\cdot,\, \cdot\rangle\rangle$. As before, the macroscopic system has Lagrangian $\mathbb{L}(\varepsilon)$, and the associated Hamiltonian structure on TP is given by $(\mathbb{H}(\varepsilon), \sigma(\varepsilon))$. In the previous section we have shown how $\mathbb{H}(\varepsilon)$ and $\sigma(\varepsilon)$ can be computed directly from their microscopic counterparts, but $\mathbb{H}(\varepsilon)$ is always the macroscopic Legendre transform of $\mathbb{L}(\varepsilon)$, and $\sigma(\varepsilon)$ equals $(\mathfrak{F}\mathbb{L}(\varepsilon))^{*}\overline{\sigma}_{\text{can}}$.

In what follows we describe how the explicit dependence on ε allows for a consistent model reduction. As illustrated in §3, a typical two-scale transformation provides an expansion of the macroscopic Lagrangian in powers of the scaling parameter ε, i.e., we have

$$\mathbb{L}(\varepsilon) = \varepsilon^{\kappa}\left(\mathbb{L}_0 + \varepsilon\mathbb{L}_1 + \varepsilon^2\mathbb{L}_2 + \ldots\right) \tag{30}$$

at least on a formal level, where κ can be positive or even negative depending on the underlying two-scale ansatz. Recall that such an expansion is not available for the original microscopic system.

Since we deal only with Hamiltonian structures on tangent bundles we benefit from Principle 3. In particular, the expansions

$$\mathbb{H}(\varepsilon) = \varepsilon^\kappa \left(\mathbb{H}_0 + \varepsilon\mathbb{H}_1 + \varepsilon^2\mathbb{H}_2 + \ldots\right), \quad \sigma(\varepsilon) = \varepsilon^\kappa \left(\sigma_0 + \varepsilon\sigma_1 + \varepsilon^2\sigma_2 + \ldots\right) \tag{31}$$

are consistent with (30), i.e., $\mathbb{H}_i$ is the Legendre transform of $\mathbb{L}_i$, and we have $\sigma_i = (\mathfrak{F}\mathbb{L}_i)^*\overline{\sigma}_{\mathrm{can}}$.

In the simplest case the reduced model is obtained by considering the leading order terms for $\mathbb{L}$ and $(\mathbb{H}, \sigma)$, and ignoring all terms that contribute to higher orders in ε. However, depending on the two-scale ansatz the leading order system can be degenerate. For this reason we distinguish the following cases:

Case A: The symplectic form σ_0 is non-degenerate, i.e. there is no $(Z, Z_\tau) \in TTP$ with $\sigma_0|_Z(Z_\tau, \cdot) \equiv 0$.

Case B: σ_0 is degenerate, but $\mathbb{L}_0$ depends on X_τ.

Case C: The leading order Lagrangian $\mathbb{L}_0$ is quasi-stationary, this means independent of X_τ, and this yields $\mathbb{H}_0 = -\mathbb{L}_0$ and $\sigma_0 = 0$.

Reduction in Case A Whenever we end up with Case A, the formal reduction provides a non-degenerate macroscopic Hamiltonian system and thus we have established already a (formal) micro-macro transition. In particular, the reduced Lagrangian reads $\mathbb{L}^{\mathrm{red}} = \mathbb{L}_0$ and the associated Hamiltonian structure is given by $(\mathbb{H}^{\mathrm{red}}, \sigma^{\mathrm{red}}) = (\mathbb{H}_0, \sigma_0)$, so that the macroscopic Hamiltonian equation on TP is given by

$$\sigma^{\mathrm{red}}|_Z(Z_\tau, \cdot) = \mathrm{d}\mathbb{H}^{\mathrm{red}}|_Z(\cdot). \tag{32}$$

Recall, that we can neglect the pre-factor ε^κ as it drops out in both the Lagrangian and Hamiltonian equations on TP.

Since we have derived the reduced macroscopic structures by means of formal expansions with respect to ε, we are confronted with the *justification problem.* More precisely, it is not obvious that solutions to (32) provide (approximate) solutions to the microscopic system. Of course, any curve $\tau \mapsto Z(\tau) \in TP$ that solves $\sigma|_Z(Z_\tau, \cdot) = \mathrm{d}\mathbb{H}(\cdot)$ and that obeys an expansion in powers of ε, must satisfy (32) to leading order, but the existence of such an expansion for the solution $Z(\tau)$ must be proven. This problem is very subtle and cannot be addressed here. Rigorous justification results for linear and some (weakly) nonlinear systems are given in [Mie08]. For a brief discussion of the difficulties that arise in the case of strong nonlinearities we refer to §3.2, which shows that such an ε-expansion can be valid only under additional assumptions concerning the initial data, the macroscopic time-interval under consideration and, finally, the regularity properties of the macroscopic equation.

Reduction in Case B In contrast to Case A, the Cases B and C allow for further reduction steps, which we explain next. We start with Case B and refer to the KdV reduction in §3.3 as a typical example. For simplicity we suppose that $\mathbb{L}_0$ depends linearly on X_τ, i.e., we assume that the momentum $\Pi_0 = \partial_{X_\tau}\mathbb{L}_0$ is a function of X but not of X_τ. As a consequence, the associated Hamiltonian structure lives on P, this means $\mathbb{H}_0$ is a function on P and σ is a symplectic form on P. In fact, $\mathbb{H}_0 = \langle\langle \Pi_0(X),\, X_\tau\rangle\rangle - \mathbb{L}_0(X,\, X_\tau)$ provides $\partial_{X_\tau}\mathbb{H}_0 = 0$. This implies that the right hand side in

$$D\Pi_0|_{(X,\,X_\tau)}(\dot{X},\,\dot{X}_\tau) = \partial_X\Pi_0|_X(\dot{X})$$

is independent of both X_τ and $\dot{X}_\tau$, and due to (19) the form σ_0 lives actually on P. Thus we end up with the following macroscopic model. The reduced Lagrangian $\mathbb{L}^{\mathrm{red}} = \mathbb{L}_0$ lives on TP and has a consistent Hamiltonian structure on P given by $\mathbb{H}^{\mathrm{red}} = \mathbb{H}_0|_P$ and $\sigma^{\mathrm{red}} = \sigma_0|_P$.

Reduction in Case C: Restriction to sub-spaces In some cases the leading order reduction turns out to be quasi-stationary, i.e., $\mathbb{L}_0$ does not depend on X_τ, and this implies $\mathbb{H}_0 = -\mathbb{L}_0$ and $\sigma_0 = 0$. Whenever this happens, we obtain a reduced macroscopic model as follows. We restrict the macroscopic configurations to

$$P_0 = \{X_0 \in P\ :\ 0 = \partial_X\mathbb{L}_0(X_0) = \partial_X\mathbb{H}_0(X_0)\ \}\,,$$

and determine the reduced Lagrangian and Hamiltonian structures by restricting the next-leading order terms $\mathbb{L}_i$ and $(\mathbb{H}_i,\,\sigma_i)$ to P_0. However, in general we shall expand additionally the solution X in powers of ε, and this may produce correction terms in the expansions (30) and (31). This problem will be discussed now, where for our purposes we can assume that P_0 is a closed linear subspace of P.

In order to identify suitable correction terms we start with the ansatz

$$Z = (X,\,X_\tau) = Z_0 + \varepsilon Z_1 = (X_0,\,X_{0\tau}) + \varepsilon(X_1,\,X_{1\tau}),$$

and study the Lagrangian $\tilde{\mathbb{L}}(\varepsilon,\,Z_0,\,Z_1) = \mathbb{L}(\varepsilon,\,Z_0 + \varepsilon Z_1)$ defined on $T\widetilde{P}$ with $\widetilde{P} = P_0 \times P$. Exploiting $\partial_{X_\tau}\mathbb{L}_0 \equiv 0$ and $\partial_X\mathbb{L}_0(X_0) = 0$ for all $X_0 \in P_0$ we find

$$\begin{aligned}\widetilde{\mathbb{L}}(\varepsilon,\,Z_0,\,Z_1) &= \varepsilon^\kappa\Big(\mathbb{L}_0(X_0 + \varepsilon X_1) + \varepsilon\mathbb{L}_1(Z_0 + \varepsilon Z_1) + \varepsilon^2\mathbb{L}_2(Z_0 + \varepsilon Z_1) + \ldots\Big)\\ &= \varepsilon^\kappa\Big(\mathbb{L}_0(X_0) + \varepsilon\mathbb{L}_1(Z_0) + \varepsilon^2\Big(\mathbb{L}_2(Z_0) + \widehat{\mathbb{L}}_2(Z_0,\,Z_1)\Big) + \ldots\Big),\end{aligned}$$

with first correction term

$$\widehat{\mathbb{L}}_2(Z_0,\,Z_1) = \tfrac{1}{2}\langle\langle\partial_X^2\mathbb{L}_0(X_0)X_1,\,X_1\rangle\rangle + \langle\langle\partial_X\mathbb{L}_1(Z_0),\,X_1\rangle\rangle + \langle\langle\partial_{X_\tau}\mathbb{L}_1(Z_0),\,X_{1\tau}\rangle\rangle.$$

In particular, any possible correction εZ_1 effects $\mathbb{L}_2$ but neither $\mathbb{L}_0$ nor $\mathbb{L}_1$.

Case C1: Reduced model via $\mathbb{L}_1$ If the next-leading order Lagrangian $\mathbb{L}_1$ depends on X_τ, then the reduced Lagrangian is given by $\mathbb{L}^{\rm red} = \mathbb{L}_1|_{TP_0}$, and in this case we can ignore the correction term εZ_1. Moreover, according to Lemma 25 the corresponding Hamiltonian structure is given by $(\mathbb{H}_1,\, \sigma_1)|_{TP_0}$. An example for this case is the three-wave-interaction discussed in §3.5.

As before, the reduction to TP_0 is formal and must be justified rigorously. In the simplest case the space P_0 is an *invariant manifold* for $(\mathbb{H},\, \sigma)$. This means that for all initial data chosen from TP_0 the solution $\tau \mapsto Z(\tau)$ to the original problem belongs to TP_0 for all times $\tau > 0$. In general, we expect that the restriction to TP_0 provides a reasonable reduced model if P_0 is an *approximate invariant manifold*, so that solutions to $(\mathbb{H}^{\rm red},\, \sigma^{\rm red})$ are approximate solutions to $(\mathbb{H},\, \sigma)$. For the *justification* in this case one has to prove that for all initial data chosen from TP_0 the real trajectory stays close to TP_0 (up to higher orders in ε) at least for sufficiently small *macroscopic* times, see for instance [SW00, GM04, GM06, GMS07].

Example 28. Let M be an integer, $Q = L^2([0,\, M];\, \mathrm{d}\eta)$ the Lebesgue space of all *periodic* functions on the interval $[0,\, M]$, and let $\mathcal{L}(x,\, x_t) = \mathcal{K}(x_t) - \mathcal{V}(x)$ be defined by

$$\mathcal{K}(x_t) = \tfrac{1}{2}\langle x_t,\, x_t\rangle, \quad \mathcal{V}(x) = -\tfrac{1}{2}\langle \triangle x,\, x\rangle, \quad \langle x,\, \tilde{x}\rangle = \int\limits_0^M x\,\tilde{x}\,\mathrm{d}\eta,$$

with discrete Laplacian $(\triangle x)(\eta) = x(\eta+1) + x(\eta-1) - 2x(\eta)$, so that the microscopic law of motion is the discrete wave equation $x_{tt} = \triangle x$. Moreover, consider the time scaling from Example 22, this means $\tau = \varepsilon t$, $x = X$, $P = Q$, $y = \eta$. Then, $\mathbb{L}$ obeys an (exact) expansion in powers of ε^2 via

$$\mathbb{L}\big(\varepsilon^2,\, X,\, X_\tau\big) = \mathbb{L}_0(X) + \varepsilon^2\mathbb{L}_1(X_\tau), \quad \mathbb{L}_0(X) = -\mathcal{V}(X), \quad \mathbb{L}_1(X_\tau) = \mathcal{K}(X_\tau).$$

The leading order Lagrangian and Hamiltonian equations read $\triangle X = 0$ and provide

$$P_0 = \{X_0 \in P \;:\; X_0(y+1) = X_0(y)\,\}.$$

Exploiting the next-leading order terms corresponding to $\mathbb{L}_1$ we find

$$\mathbb{L}^{\rm red} = \mathbb{H}^{\rm red} = \tfrac{1}{2}\langle X_{0\tau},\, X_{0\tau}\rangle, \quad \sigma^{\rm red} \simeq \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

so the macroscopic evolution is governed by $X_{0\tau\tau} = 0$. Moreover, the reduction is exact as both microscopic and reduced dynamics are equivalent for all initial data $(X(0),\, X_\tau(0)) \in TP_0$.

Case C2: Reduced model via $\mathbb{L}_2$ It may happen that even the next-leading Lagrangian $\mathbb{L}_1|_{TP_0}$ is quasi-stationary, i.e., $\mathbb{L}_1(Z_0) = \mathbb{L}_1(X_0)$ for all $Z_0 = (X_0,\, X_{0\tau}) \in$

TP_0. Then the general reduction procedure depends on the particular properties of $\mathbb{L}_1$. Here we restrict to the case we meet in §3.4 (nlS equation), where the two-scale transformation implies

$$\mathbb{L}_1|_{P_0} = -\mathbb{H}_1|_{P_0} = 0. \tag{33}$$

For $\mathbb{L}_1|_{P_0} \neq \text{const}$ we would restrict X_0 further by imposing additionally $\partial_{X_0}(\mathbb{L}_1|_{P_0}) = 0$.

Notice that (33) does *not* necessarily imply $\partial_X \mathbb{L}_1(X_0) = 0 \in \mathrm{Lin}(P, \mathbb{R})$ for all $X_0 \in P_0$ and therefore we proceed as follows. Our strategy is to choose X_1 in such a way that it is a stationary point of $\widehat{\mathbb{L}}_2(X_0, X_1)$. This means we seek $X_1 = X_1(X_0)$ as solution to the affine equation

$$\partial_X^2 \mathbb{L}_0(X_0) X_1 + \partial_X \mathbb{L}_1(X_0) = 0.$$

Provided this is possible, our reduced Lagrangian on TP_0 is given by

$$\mathbb{L}^{\mathrm{red}}(Z_0) = \mathbb{L}_2(Z_0) + \widehat{\mathbb{L}}_2(X_0,\, X_1(X_0)),$$

and since the term $\widehat{\mathbb{L}}_2$ does not contribute to the fiber derivative $\mathfrak{F}\mathbb{L}^{\mathrm{red}}$, one can show (similarly to Lemma 25) that

$$\mathbb{H}^{\mathrm{red}}(Z_0) = \mathbb{H}_2(Z_0) + \widehat{\mathbb{H}}_2(X_0,\, X_1(X_0)), \quad \sigma^{\mathrm{red}} = \sigma_2|_{TP_0},$$

where $\widehat{\mathbb{H}}_2 = -\widehat{\mathbb{L}}_2$ is the corresponding Hamiltonian structure on TP_0.

3 Two-scale reductions for the atomic chain

The abstract framework developed in the previous section shall now be applied to the examples mentioned in the introduction. The microscopic system will be either the Fermi–Pasta–Ulam (FPU) chain

$$\ddot{x}_j(t) = \Phi_1'\big(x_{j+1}(t) - x_j(t)\big) - \Phi_1'\big(x_j(t) - x_{j-1}(t)\big) \tag{34}$$

or the Klein–Gordon (KG) chain

$$\ddot{x}_j(t) = \alpha\big(x_{j+1}(t) + x_{j-1}(t) - 2x_j(t)\big) - \Phi_0'(x_j(t)) \tag{35}$$

with harmonic constant $\alpha = \Phi_1''(0) \in \mathbb{R}$. Without loss of generality we always assume $0 = \Phi_0(0) = \Phi_0'(0) = \Phi_1(0) = \Phi_1'(0)$, and restrict our considerations to infinite chains. In the case that the two-scale ansatz refers to small amplitudes, the linearized atomic chain

$$\ddot{x}_j(t) = \alpha\big(x_{j+1}(t) + x_{j-1}(t) - 2x_j(t)\big) - \Phi_0''(0)x_j$$

becomes important. This linearized chain allows for propagating plane wave solutions $\mathtt{e}^{\mathtt{i}(\omega t+\theta j)}$, provided that the *frequency* ω and the *wave number* θ satisfy the *dispersion relation*

$$\omega^2 = \Omega^2(\theta) = 2\alpha\big(1-\cos\theta\big) + \Phi_0''(0). \tag{36}$$

The atomic chain falls into the class of normal Hamiltonian systems, see (18), with configuration space $Q_{\mathrm{discr}} = \ell^2(\mathbb{Z})$. The Lagrangian reads $\mathcal{L}_{\mathrm{discr}}(x,\,\dot{x}) = \mathcal{K}_{\mathrm{discr}}(\dot{x}) - \mathcal{V}_{\mathrm{discr}}(x)$ with kinetic and potential energy given by

$$\mathcal{K}_{\mathrm{discr}}(\dot{x}) = \tfrac{1}{2}\sum_{j\in\mathbb{Z}} \dot{x}_j^2, \quad \mathcal{V}_{\mathrm{discr}}(x) = \sum_{j\in\mathbb{Z}} \Big(\Phi_1(x_{j+1}-x_j) + \Phi_0(x_j)\Big), \tag{37}$$

and Newton's equations (13) equal the Euler-Lagrange equations to $\mathcal{L}_{\mathrm{discr}}$ on TQ_{discr}. Moreover, the Hamiltonian is given by $\mathcal{H}_{\mathrm{discr}}(x,\,\dot{x}) = \mathcal{K}_{\mathrm{discr}}(\dot{x}) + \mathcal{V}_{\mathrm{discr}}(x)$, so that Newton's equations are equivalent to

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} = \begin{pmatrix} \partial_x \mathcal{H}_{\mathrm{discr}} \\ \partial_{\dot{x}} \mathcal{H}_{\mathrm{discr}} \end{pmatrix},$$

which is a Hamiltonian ODE on TQ_{discr} with metric symplectic form, i.e., we have $\sigma = \sigma_{\mathrm{met}}$ in the sense of Example 2.

3.1 The embedded atomic chain

In order to derive effective models we start with a suitable embedding of the atomic chain. At first we replace the discrete lattice index $j \in \mathbb{Z}$ by a continuous variable $\eta \in \mathbb{R}$. In addition, if the two-scale ansatz involves oscillatory microstructure, we consider k additional *phase variables* $\phi = (\phi_1,\,..,\,\phi_k)$, which are supposed to take values in the k-dimensional torus T^k. This embedding gives rise to the formal identification

$$x_j(t) = x(t,\,j,\,0), \quad \dot{x}_j(t) = x_t(t,\,j,\,0),$$

where the instantaneous configuration $x(t,\,\cdot,\,\cdot)$ is for each t a function in η and ϕ.

Next, we identify the Lagrangian $\mathcal{L}$ of the embedded system. To this end, we replace all sums over j in (37) by integrals with respect to η and ϕ. This yields

$$\mathcal{L}(x,\,x_t) = \mathcal{K}(x_t) - \mathcal{V}(x) \tag{38}$$

with

$$\mathcal{K}(x_t) = \int\limits_{\mathbb{R}\times T^k} \tfrac{1}{2} x_t^2 \,\mathrm{d}\eta\mathrm{d}\phi, \quad \mathcal{V}(x) = \int\limits_{\mathbb{R}\times T^k} \Big(\Phi_1\big(\nabla_{1,0}^{+}x\big) + \Phi_0(x)\Big)\,\mathrm{d}\eta\mathrm{d}\phi, \tag{39}$$

where $\nabla^{+}_{1,0}$ is a discrete differential operator, see Remark 29 below. Notice that the Euler-Lagrange equation for $\mathcal{L}$, i.e.

$$x_{tt} = \nabla^{-}_{1,0}\Phi_1'\big(\nabla^{+}_{1,0}x\big) - \Phi_0'\big(x\big), \quad x = x(t,\, \eta,\, \phi),$$

is still fully equivalent to (an uncountable number of uncoupled copies of) Newton's equations (13). However, the embedding gives rise to additional symmetries, and thus we gain new integrals of motion. In fact, the Lagrangian (38) is invariant under the continuous groups of space shifts $\eta \rightsquigarrow \eta + \eta_0$ and phase shifts $\phi \rightsquigarrow \phi + \phi_0$, and Noether's theorem provides that

$$\mathcal{I}_{\mathrm{space}}(x,\, x_t) = \int\limits_{\mathbb{R}\times T^k} x_t\, x_\eta\, \mathrm{d}\eta \mathrm{d}\phi \in \mathbb{R}, \quad \mathcal{I}_{\mathrm{phase}}(x,\, x_t) = \int\limits_{\mathbb{R}\times T^k} x_t\, x_\phi\, \mathrm{d}\eta \mathrm{d}\phi \in \mathbb{R}^k$$

are conserved for any solution of the microscopic system. Recall that $x_\eta \in \mathbb{R}$ and $x_\phi \in \mathbb{R}^k$ denote the derivatives of x with respect to η and ϕ, respectively. These conservation laws have no counterpart within the classical mechanics of mass points as they are a byproduct of the embedding.

Remark 29. For given $\delta \in \mathbb{R}$ and $\theta \in T^k$ let

$$(\nabla^{+}_{\delta,\theta}x)(\eta,\, \phi) = x(\eta + \delta,\, \phi + \theta) - x(\eta,\, \phi), \quad (\nabla^{-}_{\delta,\theta}x)(\eta,\, \phi) = x(\eta,\, \phi) - x(\eta - \delta,\, \phi - \theta),$$

and $\triangle_{\delta,\theta} = \nabla^{+}_{\delta,\theta} - \nabla^{-}_{\delta,\theta}$. These definitions imply

$$\nabla^{\pm}_{-\delta,-\theta} = -\nabla^{\mp}_{\delta,\theta}, \quad \big(\nabla^{\pm}_{\delta,\theta}\big)^* = -\nabla^{\mp}_{\delta,\theta}, \quad (\triangle_{\delta,\theta})^* = \triangle_{\delta,\theta},$$

where $*$ denotes the adjoint operator with respect to the L^2–inner product.

3.2 From FPU to the wave equation

Here we derive the quasi-linear wave equation from Newton's equation for FPU chains. Recall that the underlying two-scale ansatz is given by (14), and involves neither a microstructure nor a moving frame. For the embedded system this ansatz reads

$$x(t,\, \eta) = \varepsilon^{-1}X(\varepsilon t,\, \varepsilon\eta), \tag{40}$$

and Example 19 provides $Q = L^2(\mathbb{R};\, \mathrm{d}\eta)$ and $P = L^2(\mathbb{R};\, \mathrm{d}y)$ as well as the (lifted) inverse two-scale transformation

$$\mathbb{T}_{\mathrm{vel}}(\varepsilon) : (X,\, X_\tau)(y) \mapsto (x,\, x_t)(\eta) = \big(\varepsilon^{-1}X,\, X_\tau\big)(\varepsilon\eta). \tag{41}$$

Lemma 30. *The two-scale transformation* (41) *yields* $\mathbb{L} = \mathbb{K} - \mathbb{V}$ *and* $\mathbb{E} = \mathbb{K} + \mathbb{V}$ *with*

$$\mathbb{K}(\varepsilon,\, X_\tau) = \varepsilon^{-1}\, \mathbb{K}_0(X_\tau), \quad \mathbb{K}_0(X_\tau) = \int_{\mathbb{R}} \tfrac{1}{2} X_\tau^2 \mathrm{d}y, \quad \mathbb{V}(\varepsilon,\, X) = \varepsilon^{-1} \int_{\mathbb{R}} \Phi_1\big(\varepsilon^{-1}\nabla_\varepsilon^+ X\big)\mathrm{d}y.$$

Moreover, we have $\mathbb{H} = \mathbb{E}$ *and* $\sigma = \varepsilon^{-1}\sigma_{\mathrm{met}}$*, where* σ_{met} *is the metric symplectic form on* TP*, see Example 2.*

Proof. All assertions are direct consequences of (41) and the abstract results from §2.2, see Theorem 21. In particular, $\mathbb{H} = \mathbb{E}$ holds, since the two-scale transformation does not involve a moving frame. □

Next we identify the leading order terms in the expansion with respect to ε. Using formal Taylor expansion

$$\big(\varepsilon^{-1}\nabla_\varepsilon^+ X\big)(y) = \varepsilon^{-1}(X(y+\varepsilon) - X(y)) = X_y(y) + \tfrac{1}{2}\varepsilon X_{yy}(y) + \mathcal{O}\big(\varepsilon^2\big),$$

we find $\mathbb{V}(\varepsilon,\, X) = \varepsilon^{-1}\mathbb{V}_0(X) + \mathcal{O}(1)$ with $\mathbb{V}_0(X) = \int_{\mathbb{R}} \Phi_1(X_y)\,\mathrm{d}y$ and conclude that

$$\mathbb{L}^{\mathrm{red}}(X,\, X_\tau) = \mathbb{K}_0(X_\tau) - \mathbb{V}_0(X), \quad \mathbb{H}^{\mathrm{red}}(X,\, X_\tau) = \mathbb{K}_0(X_\tau) + \mathbb{V}_0(X).$$

Notice that $\mathbb{V}_0$ is defined only on $H^1(\mathbb{R};\, \mathrm{d}y)$, which is dense in P. Finally, $\sigma^{\mathrm{red}} = \sigma_{\mathrm{met}}$ completes the leading order reduction and we end up with the following macroscopic model:

Theorem 31. *Both the formally reduced Lagrangian and Hamiltonian equations are equivalent to*

$$X_{\tau\tau} - \Phi_1'(X_y)_y = 0. \tag{42}$$

Remark 32. We claimed in the introduction that the Hamiltonian two-scale reduction is always related to the Hamiltonian structure on TP but fails if we use the canonical structure on T^*P. In this example we clearly see the reason for this: The canonical momentum corresponding to $\mathbb{L}(\varepsilon)$ is given by $\Pi = \varepsilon^{-1}X_\tau$ and, replacing X_τ by Π, we find

$$\overline{\mathbb{H}}(\varepsilon,\, X,\, \Pi) = \langle\!\langle \partial_{X_\tau}\mathbb{L},\, X_\tau \rangle\!\rangle - \mathbb{L} = \varepsilon \int_{\mathbb{R}} \tfrac{1}{2}\Pi^2\,\mathrm{d}y + \varepsilon^{-1}\int_{\mathbb{R}} \Phi_1\big(\varepsilon^{-1}\nabla_\varepsilon^+ X\big)\,\mathrm{d}y,$$

the Hamiltonian on T^*P. As long as we fix $\varepsilon > 0$, the canonical equations $X_\tau = \varepsilon\Pi$ and $\Pi_\tau = \varepsilon^{-2}\nabla_\varepsilon^- \Phi_1'(\varepsilon^{-1}\nabla_\varepsilon^+ X)$ are fully equivalent to the Hamiltonian equations on TP. However, formal expansion of $\overline{\mathbb{H}}$ with respect to ε yields, to leading order ε^{-1}, the reduced Hamiltonian $\overline{\mathbb{H}}^{\mathrm{red}}(X,\, \Pi) = \varepsilon^{-1}\mathbb{V}_0(X)$ and the corresponding canonical equations $X_\tau = 0$, $\Pi_\tau = \varepsilon^{-1}\Phi_1'(X_y)_y$ are apparently different from the wave equation (42). Of course, here we can overcome this problem by multiplying $\mathbb{L}$ with ε, but this is not always possible as the KdV reduction in §3.3 shows.

To conclude this section we discuss some aspects of (42) which are closely related to the justification problem. In particular, it comes out that (42) can provide a reasonable macroscopic model for the FPU chain only under additional assumptions and this shows that the formal expansions from the reduction step truly need to be justified rigorously. We introduce new variables $W = X_\tau$ and $R = X_y$ and rewrite equation (42) in the form

$$\partial_\tau R - \partial_y W = 0, \quad \partial_\tau W - \partial_y \Phi_1'(R) = 0. \tag{43}$$

This is a first order system of macroscopic conservation laws with characteristic speeds $\lambda_\pm = \pm\sqrt{\Phi_1''(R)}$ and is called the p-system (with $p = -\Phi_1'$), see [Daf00]. These equations formally imply the conservation of energy, i.e., any *smooth* solution to (43) satisfies $\partial_\tau E - \partial_y(W\Phi_1'(R)) = 0$ with $E = \frac{1}{2}W^2 + \Phi_1(R)$.

Now suppose that Φ_1 is concave or, more general, restrict R to the region of concavity of Φ_1. In this case, the system (43) is elliptic and its initial value problem is ill-posed. For this reason the microscopic system behaves as follows: Even if we initialize the chain with data satisfying $x_j(0) = \varepsilon^{-1}X_{\mathrm{ini}}(\varepsilon j)$ and $\dot{x}_j(0) = W_{\mathrm{ini}}(\varepsilon j)$, where X_{ini} and W_{ini} are infinitely smooth macroscopic functions, the atomic data will immediately start to oscillate on the microscopic scale, see [Her05, DH07] for numerical simulations. Therefore, the two-scale ansatz *cannot* be satisfied for any $\tau > 0$ and we conclude that any rigorous justification of (42) must exclude non-convex Φ_1.

Next suppose that Φ_1 is strictly convex, which provides the strict hyperbolicity of the p-system, and assume for simplicity that Φ_1' is also strictly convex, so that all eigenvalues are genuinely nonlinear. However, even in this case there exist limitations for the validity of (43). In fact, it is well known that the nonlinearity of Φ_1 causes the following generic situation: Given smooth initial data for (43), there exists a critical time $0 < \tau_0 < \infty$ at which the first macroscopic shock is formed. In particular, there exists a smooth solution for $0 < \tau < \tau_0$, and for these times we can expect that (40) provides an approximate solution of the microscopic system. However, for $\tau > \tau_0$ the macroscopic energy E is not conserved anymore and thus the p-system can not be related to the macroscopic dynamics of the chain, since the chain conserves the energy exactly. This phenomenon is usually called the *shock problem* and appears analogously in all zero dispersion limits, compare for instance the surveys in [Lax86, Lax91, LLV93]. For the FPU chain the macroscopic dynamics beyond the shock can be understood by Whitham's modulation theory with periodic travelling waves, see [FV99, DHM06, DHR06, DH07] and [HFM81, DM98, El05] for the complete integrable Toda chain. Moreover, for harmonic lattices the macroscopic limit under the hyperbolic scaling can be established rigorously by means of weak convergence methods (cf. [Mie06, Mie08]). The transport of energies can be studied via Wigner-Husimi measures, see [Mie06].

3.3 From FPU to KdV

To derive the KdV equation for FPU chains we rely on the two-scale ansatz

$$x(t,\,\eta) = \varepsilon X\big(\varepsilon^3 t,\,\varepsilon\eta+\varepsilon ct\big), \tag{44}$$

which is related to a moving frame with drift velocity c. Example 24 provides

$$\mathbb{T}_{\rm vel}(\varepsilon,\,t) : (X,\,X_\tau)(y) \mapsto (x,\,x_t)(\eta) = \big(\varepsilon X,\,\varepsilon^4 X_\tau + \varepsilon^2 c X_y\big)(\varepsilon\eta+\varepsilon ct) \tag{45}$$

with P and Q as in §3.2. The transformation $\mathbb{T}_{\rm vel}$ is defined only on $H^1(\mathbb{R};\,{\rm d}y)\times L^2(\mathbb{R};\,{\rm d}y)$, a dense subset of TP, but in order to focus on the basic features of the reduction procedure we do not stress out this explicitly.

Lemma 33. *Under the exact two-scale transformation* (45) *the energies $\mathcal{K}$ and $\mathcal{V}$ transform into their ε-parametrized counterparts*

$$\mathbb{K}(\varepsilon,\,X,\,X_\tau) = \varepsilon^3\int_{\mathbb{R}} \tfrac{1}{2}\big(\varepsilon^2 X_\tau + c\,X_y\big)^2{\rm d}y,\quad \mathbb{V}(\varepsilon,\,X) = \varepsilon^{-1}\int_{\mathbb{R}}\Phi_1\big(\varepsilon\nabla_\varepsilon^+ X\big){\rm d}y \tag{46}$$

and the matrix Σ corresponding to the symplectic form σ is given by $\Sigma = \varepsilon^5\Sigma_1 + \varepsilon^7\Sigma_2$ with

$$\Sigma_1 = \begin{pmatrix} -2\,c\,\partial_y & 0 \\ 0 & 0 \end{pmatrix},\quad \Sigma_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{47}$$

Moreover, due to the time dependence of the two-scale transformation, $\mathbb{H}$ differs from $\mathbb{E}$ and satisfies $\mathbb{H} = \mathbb{E} + \mathbb{I}$ with

$$\mathbb{I}(\varepsilon,\,X,\,X_\tau) = -\varepsilon^3 c\int_{\mathbb{R}}\big(\varepsilon^2\,X_\tau + cX_y\big)X_y\,{\rm d}y. \tag{48}$$

Proof. For the proof of (46) and (48) we insert the ansatz (44) into the definitions of $\mathcal{V}$, $\mathcal{K}$ and $\mathcal{I}$, cf. Formula (39) and Example 24, and replace $\varepsilon^3 t + \varepsilon c\eta$ by y in the arising integrals. Moreover, the identity $\mathbb{H} = \mathbb{E} + \mathbb{I}$ is provided by Theorems 16 and 21. Finally, the linear two-scale transformation $\mathbb{T}_{\rm vel}(\varepsilon,\,t)$ can be identified with $\boldsymbol{S}\circ\boldsymbol{T}$, where $\boldsymbol{S} : TP \to TQ$ is given by $(\boldsymbol{S}\,Z)(\eta) = Z(\varepsilon\eta+\varepsilon ct)$ and $\boldsymbol{T}$ abbreviates the operator-valued matrix

$$\boldsymbol{T} = \begin{pmatrix} \varepsilon & 0 \\ \varepsilon^2 c\partial_y & \varepsilon^4 \end{pmatrix}.$$

Due to Lemma 26 and Remark 27 we find $\Sigma = \frac{{\rm d}\tau}{{\rm d}t}\frac{{\rm d}\eta}{{\rm d}y}\boldsymbol{T}^*\,\Sigma_{\rm met}\,\boldsymbol{T}$, which yields (47) after a short computation. □

Leading order reduction At first we expand the various energies with respect to ε up to $\mathcal{O}(\varepsilon^6)$. To this end, we define $v_i = \Phi_1^{(i)}(0)$ so that the Taylor polynomial of Φ_1 reads $\Phi_1(x) = \frac{v_2}{2}x^2 + \frac{v_3}{6}x^3 + \text{h.o.t.}$

Lemma 34. *The transformed energies $\mathbb{K}$, $\mathbb{V}$ and $\mathbb{I}$ satisfy*

$$\mathbb{K} = \varepsilon^3\mathbb{K}_0 + \varepsilon^5\mathbb{K}_1 + \mathcal{O}\big(\varepsilon^6\big), \quad \mathbb{V} = \varepsilon^3\mathbb{V}_0 + \varepsilon^5\mathbb{V}_1 + \mathcal{O}\big(\varepsilon^6\big), \quad \mathbb{I} = \varepsilon^3\mathbb{I}_0 + \varepsilon^5\mathbb{I}_1 + \mathcal{O}\big(\varepsilon^6\big),$$

where

$$\begin{aligned}
\mathbb{K}_0(X) &= \tfrac{c^2}{2}\int_{\mathbb{R}} X_y^2\,\mathrm{d}y, \quad \mathbb{K}_1(X,\,X_\tau) = c\int_{\mathbb{R}} X_\tau X_y\,\mathrm{d}y,\\
\mathbb{V}_0(X) &= \tfrac{v_2}{2}\int_{\mathbb{R}} X_y^2\,\mathrm{d}y, \quad \mathbb{V}_1(X) = -\tfrac{v_2}{24}\int_{\mathbb{R}} X_{yy}^2\,\mathrm{d}y + \tfrac{v_3}{6}\int_{\mathbb{R}} X_y^3\,\mathrm{d}y,\\
\mathbb{I}_0(X) &= -c^2\int_{\mathbb{R}} X_y^2\,\mathrm{d}y, \quad \mathbb{I}_1(X,\,X_\tau) = -c\int_{\mathbb{R}} X_\tau X_y\,\mathrm{d}y.
\end{aligned}$$

Proof. The expansions for $\mathbb{K}$ and $\mathbb{I}$ follow immediately from Lemma 33. To prove the remaining assertions we start with $\varepsilon\nabla_\varepsilon^+ X = \varepsilon^2 X_y + \varepsilon^3\frac{1}{2}X_{yy} + \varepsilon^4\frac{1}{6}X_{yyy} + \mathcal{O}(\varepsilon^5)$ and obtain

$$\begin{aligned}
\Phi_1\big(\varepsilon\nabla_\varepsilon^+ X\big) &= \tfrac{v_2}{2}\big(\varepsilon\nabla_\varepsilon^+ X\big)^2 + \tfrac{v_3}{6}\big(\varepsilon\nabla_\varepsilon^+ X\big)^3 + \mathcal{O}\big(|\varepsilon\nabla_\varepsilon^+ X|^4\big)\\
&= \tfrac{v_2}{2}\big(\varepsilon^4 X_y^2 + \varepsilon^5 X_y X_{yy} + \varepsilon^6\tfrac{1}{4}X_{yy}^2 + \varepsilon^6\tfrac{1}{3}X_y X_{yyy}\big) + \varepsilon^6\tfrac{v_3}{6}X_y^3 + \mathcal{O}\big(\varepsilon^7\big).
\end{aligned}$$

We insert this expression into the formula for $\mathbb{V}(\varepsilon,\,X)$ and due to $\int_{\mathbb{R}} X_y X_{yy}\mathrm{d}y = 0$ and $\int_{\mathbb{R}} X_y X_{yyy}\mathrm{d}y = -\int_{\mathbb{R}} X_{yy}^2\mathrm{d}y$, we obtain the asserted expansion for $\mathbb{V}$. □

In the next step we can read-off the leading order terms of $\mathbb{L} = \mathbb{K}-\mathbb{V}$, $\mathbb{H} = \mathbb{K}+\mathbb{V}+\mathbb{I}$ and Σ. However, the order of ε at which we find the reduced Lagrangian and Hamiltonian structures depends on the choice of the moving-frame speed c. Let us start with the case $c^2 \neq v_2$. Under this assumption the leading order terms correspond to ε^3. More precisely, we obtain $\sigma^{\text{red}} = 0$ and

$$\mathbb{L}^{\text{red}}(X) = -\mathbb{H}^{\text{red}}(X) = \mathbb{K}_0(X) - \mathbb{V}_0(X) = \tfrac{1}{2}\big(c^2 - v_2\big)\int_{\mathbb{R}} X_y^2\,\mathrm{d}y.$$

In particular, both the reduced Lagrangian and Hamiltonian equations turn out to be equivalent to $X_{yy} = 0$ and have no non-trivial solutions at all. Thus, we assume

$$c^2 = v_2 = \Phi_1''(0), \tag{49}$$

i.e., the moving-frame speed equals the sound velocity of the linearized FPU chain. In this case we find $\mathbb{K}_0 = \mathbb{V}_0 = -\frac{1}{2}\mathbb{I}_0$ and this leads to *cancelations* in $\mathbb{L}$ and $\mathbb{H}$. Consequently, the leading order terms in the Lagrangian and Hamiltonian structure correspond to ε^5 and we end up with the following macroscopic model:

Theorem 35. *With* (49) *the reduced Lagrangian and Hamiltonian structures are given by*

$$\mathbb{L}^{\mathrm{red}}(X,\, X_\tau) = \mathbb{K}_1(X,\, X_\tau) - \mathbb{V}_1(X), \quad \mathbb{H}^{\mathrm{red}}(X) = \mathbb{V}_1(X), \quad \Sigma^{\mathrm{red}} = \Sigma_1. \tag{50}$$

In particular, both the reduced Lagrangian and Hamiltonian equations are equivalent to

$$2\, cX_{\tau y} - \tfrac{1}{12}\, v_2\, X_{yyyy} - v_3 X_y X_{yy} = 0, \tag{51}$$

which is a KdV equation for X_y.

Proof. The identities (50) can be read-off from Lemma 34 and (51) follows by a direct calculation. □

The KdV reduction with (49) is an example for Case B from Section §2.3, i.e., the reduced Hamiltonian structure lives on P and not on TP. Moreover, the term $\mathbb{I}_0$, which produces the cancelation in $\mathbb{L}_0$ and $\mathbb{H}_0$, equals up to the sign the macroscopic integral of motion $\mathbb{I}^{\mathrm{red}}(X) = \int_{\mathbb{R}} X_y^2 \mathrm{d}y$, associated to the invariance under shift in the y-direction.

Remark 36. As before, the formal two-scale reduction relies on the Hamiltonian structure on TP but fails if we use the canonical structure on T^*P. Even worse, here we cannot overcome this problem by a simple rescaling of $\mathbb{L}$. To understand this, we consider the rescaled Lagrangian (for $c^2 = v_2$)

$$\tilde{\mathbb{L}}(\varepsilon,\, X,\, X_\tau) = \varepsilon^{-5}\mathbb{L}(\varepsilon,\, X,\, X_\tau) = \mathbb{K}_1(X,\, X_\tau) + \varepsilon^2\mathbb{K}_2(X_\tau) + \tilde{\mathbb{V}}(\varepsilon,\, X),$$

where

$$\mathbb{K}_2(X_\tau) = c\int_{\mathbb{R}} X_\tau X_y\, \mathrm{d}y, \quad \tilde{\mathbb{V}}(\varepsilon,\, X) = \varepsilon^{-5}\big(\mathbb{V}(\varepsilon,\, X) - \varepsilon^3\mathbb{V}_0(X)\big) \approx \mathbb{V}_1(X) = \mathcal{O}(1).$$

The canonical momenta are given by $\tilde{\Pi} = cX_y + \varepsilon^2 X_\tau$ and computing the associated Hamiltonian on T^*P we find

$$\tilde{\tilde{H}}(\varepsilon,\, X,\, \Pi) = \varepsilon^{-2}\tilde{\tilde{H}}_0(X,\, \Pi) \;+\; \tilde{\mathbb{V}}(\varepsilon,\, X), \quad \tilde{\tilde{H}}_0(X,\, \Pi) = \int_{\mathbb{R}} \tfrac{1}{2}\big(\Pi^2 - c^2\, X_y^2\big)\, \mathrm{d}y.$$

In particular, the canonical equations corresponding to the leading order Hamiltonian $\tilde{\tilde{\mathbb{H}}}_0$ do not equal (51).

Finally, we mention that rigorous justification results for the KdV reduction can be found in [SW00, FP99]. However, these results do not use the reduced Lagrangian or Hamiltonian structures, but work on the equation of motion directly.

3.4 From KG to nlS

We start with the two-scale ansatz (15) for a modulated pulse in the KG chain (35) with $\alpha = 1$ and aim to show that the complex amplitude A satisfies the nlS equation. Recall that the plane waves appearing in (15) model a microstructure of harmonic oscillations, and thus we can regard the nlS equation as a macroscopic modulation equation.

In contrast to the previous examples, here the two-scale ansatz does not provide immediately an exact two-scale transformation, but we can setup the problem as follows: We embed the discrete lattice $\mathbb{Z}$ into the cylinder $\mathbb{R}\times T^1$ and identify each microscopic configuration with a function $x \in Q = L^2(\mathbb{R}\times T^1;\, \mathrm{d}\eta\mathrm{d}\phi)$ depending on the microscopic continuous space variable η and a periodic phase variable $\phi \in T^1 \cong [0,\, 2\pi]$. Moreover, in accordance to the scaling, we choose $P = L^2(\mathbb{R}\times T^1;\, \mathrm{d}y\mathrm{d}\phi)$ and make the two-scale ansatz

$$x(t,\, \eta,\, \phi) = \varepsilon X\big(\varepsilon^2 t,\, \varepsilon\eta - \varepsilon ct,\, \phi + \omega t + \theta\eta\big), \tag{52}$$

which gives rise to the inverse two-scale transformation

$$\Big(\mathbb{T}_{\mathrm{vel}}(\varepsilon,\, t)(X,\, X_\tau)\Big)(\eta,\, \phi) = \big(\varepsilon X,\, \varepsilon^3 X_\tau - \varepsilon^2 c X_y + \varepsilon\omega X_\phi\big)(\varepsilon\eta - \varepsilon ct,\, \phi + \omega t + \theta\eta). \tag{53}$$

In this section we show that this transformation implies both the particular form of the microstructure and the nlS equation.

Remark 37. From (52) we read-off the identity $\mathbb{T}_{\mathrm{con}}(\varepsilon,\, t) = \mathcal{T}_{\mathrm{con}}^{-1}\circ\mathcal{M}_{\mathrm{con}}(-t)\circ\mathbb{S}_{\mathrm{con}}(\varepsilon)$, where $(\mathbb{S}_{\mathrm{con}}(\varepsilon)X)(\eta,\, \phi) = \varepsilon X(\varepsilon\eta,\, \phi)$ denotes the inverse scaling transformation. Moreover, $(\mathcal{M}_{\mathrm{con}}(t)\, x)(\eta,\, \phi) = x(\eta + c\,t,\, \phi - \omega t)$ is a moving frame transformation with associated integral of motion $\mathcal{I}(x,\, x_t) = \int_{\mathbb{R}\times T^1} x_t(cx_\eta - \omega x_\phi)\mathrm{d}\eta\mathrm{d}\phi$ and $(\mathcal{T}_{\mathrm{con}}\, x)(\eta,\, \phi) = x(\eta,\, \phi - \theta\eta)$ corresponds to a weak symmetry transformation.

Lemma 38. *Under* (53) *the transformed energies* $\mathbb{K}$, $\mathbb{V}$ *and* $\mathbb{I}$ *take the form*

$$\mathbb{K}(\varepsilon,\, X,\, X_\tau) = \varepsilon^{-1}\int\limits_{\mathbb{R}\times T^1} \tfrac{1}{2}\big(\varepsilon^3 X_\tau - \varepsilon^2 c\, X_y + \varepsilon\omega X_\phi\big)^2\, \mathrm{d}y\mathrm{d}\phi,$$

$$\mathbb{V}(\varepsilon,\, X) = \varepsilon^{-1}\int\limits_{\mathbb{R}\times T^1} \big(-\varepsilon^2\tfrac{1}{2}X\triangle_{\varepsilon,\theta}X + \Phi_0(\varepsilon X)\big)\, \mathrm{d}y\mathrm{d}\phi,$$

$$\mathbb{I}(\varepsilon,\, X,\, X_\tau) = \varepsilon^{-1}\int\limits_{\mathbb{R}\times T^1} \big(\varepsilon^3 X_\tau - \varepsilon^2 c\, X_y + \varepsilon\omega X_\phi\big)\big(\varepsilon^2 c\, X_y - \varepsilon\omega X_\phi\big)\, \mathrm{d}y\mathrm{d}\phi,$$

and we have $\mathbb{L} = \mathbb{K} - \mathbb{V}$, $\mathbb{E} = \mathbb{K} + \mathbb{V}$ *and* $\mathbb{H} = \mathbb{E} + \mathbb{I}$. *Moreover, the matrix* Σ *corresponding to the symplectic form* σ *satisfies* $\Sigma = \varepsilon^3\Sigma_2 + \varepsilon^4\Sigma_3 + \varepsilon^5\Sigma_4$, *with*

$$\Sigma_2 = \begin{pmatrix} -2\,\omega\,\partial_\phi & 0 \\ 0 & 0 \end{pmatrix}, \quad \Sigma_3 = \begin{pmatrix} 2\,c\,\partial_y & 0 \\ 0 & 0 \end{pmatrix}, \quad \Sigma_4 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

and is non-degenerate due to Σ_4.

Proof. At first we study the time dependent transformation $\mathcal{M}_{\rm con}(t)\circ\mathcal{T}_{\rm con}:Q\to Q$, see Remark 37, and write $\hat{x}=\mathcal{T}_{\rm con}x$ and $\tilde{x}=\mathcal{M}_{\rm con}(t)\hat{x}$. According to Theorem 5, the transformation $\mathcal{T}_{\rm con}$ transforms $(\mathcal{H},\,\sigma)$ into $(\hat{\mathcal{H}},\,\hat{\sigma})$ with $\hat{\mathcal{H}}=\mathcal{H}\circ\mathcal{T}_{\rm vel}^{-1}$ and $\hat{\sigma}=(\mathcal{T}_{\rm vel}^{-1})^*\sigma$. Then we apply the moving frame transformation $\mathcal{M}_{\rm vel}(t)$ and Theorem 16 provides the Hamiltonian structure $(\tilde{\mathcal{H}},\,\tilde{\sigma})$ with $\tilde{\sigma}=(\mathcal{M}_{\rm vel}(-t))^*\hat{\sigma}$ and $\tilde{\mathcal{H}}=(\hat{\mathcal{E}}+\hat{\mathcal{I}})\circ\mathcal{M}_{\rm vel}(-t)$ with $\hat{\mathcal{E}}=\hat{\mathcal{H}}$ and

$$\hat{\mathcal{I}}(\hat{x},\,\hat{x}_t)=\int\limits_{\mathbb{R}\times T^1}\hat{x}_t(c\hat{x}_\eta-\omega\hat{x}_\phi)\mathrm{d}\eta\mathrm{d}\phi. \tag{54}$$

Moreover, exploiting Theorem 21 for the scaling transformation $\mathcal{S}_{\rm con}(\varepsilon)=\mathbb{S}_{\rm con}^{-1}(\varepsilon)$, we find

$$\mathbb{H}=(\mathcal{E}+\mathcal{I})\circ\mathcal{T}_{\rm vel}^{-1}\circ\mathcal{M}_{\rm vel}(-t)\circ\mathbb{S}_{\rm vel}(\varepsilon),\quad \sigma=\varepsilon^2(\mathcal{T}_{\rm vel}^{-1}\circ\mathcal{M}_{\rm vel}(-t)\circ\mathbb{S}_{\rm vel}(\varepsilon))^*\sigma$$

with $\mathcal{I}=\hat{\mathcal{I}}\circ\mathcal{T}_{\rm vel}$. According to (39) and (54), the microscopic energies are given by $\mathcal{K}(x_t)=\frac{1}{2}\int_{\mathbb{R}\times T^1}x_t^2\,\mathrm{d}\eta\mathrm{d}\phi$ and

$$\mathcal{V}(x)=\int\limits_{\mathbb{R}\times T^1}\Big(-\tfrac{1}{2}x\triangle_{1,\,0}x+\Phi_0(x)\Big)\,\mathrm{d}\eta\mathrm{d}\phi,\quad \mathcal{I}(x,\,x_t)=\int\limits_{\mathbb{R}\times T^1}x_t(cx_\eta-c\theta x_\phi-\omega x_\phi)\mathrm{d}\eta\mathrm{d}\phi,$$

where the discrete operators ∇ and $\triangle$ are defined in Remark 29. The expressions for $\mathbb{K}$, $\mathbb{V}$, $\mathbb{L}$, $\mathbb{E}$ and $\mathbb{I}$ now follow by inserting (52) into the formulas for $\mathcal{K}$, $\mathcal{V}$ and $\mathcal{I}$. For the computation of Σ we identify the linear two-scale transformation $\mathbb{T}_{\rm vel}(\varepsilon,\,t)$ with $\boldsymbol{S}\circ\boldsymbol{T}$, where $\boldsymbol{S}:TP\to TQ$ is given by $(\boldsymbol{S}Z)(\eta,\,\phi)=Z(\varepsilon\eta-\varepsilon ct,\,\phi+\omega t+\theta\eta)$ and $\boldsymbol{T}$ abbreviates the operator-valued matrix

$$\boldsymbol{T}=\begin{pmatrix}\varepsilon & 0\\ -\varepsilon^2\,c\,\partial_y+\varepsilon\,\omega\,\partial_\phi & \varepsilon^3\end{pmatrix}$$

with components in $\mathrm{Lin}(P,\,P)$. Finally, Remark 27 yields $\Sigma=\frac{\mathrm{d}\tau}{\mathrm{d}t}\frac{\mathrm{d}\eta}{\mathrm{d}y}\boldsymbol{T}'\,\Sigma_{\rm met}\,\boldsymbol{T}$ and this implies the desired result. □

Leading order reduction Next we derive the formal expansions with respect to ε. To this end we introduce the constants $v_i=\Phi_0^{(i)}(0)$ and find, due to $v_0=v_1=0$,

$$\Phi_0(\varepsilon X)=\varepsilon^2\tfrac{v_2}{2}X^2+\varepsilon^3\tfrac{v_3}{6}X^3+\varepsilon^4\tfrac{v_4}{24}X^4+\mathcal{O}\big(\varepsilon^5\big). \tag{55}$$

Lemma 39. *The transformed energies satisfy*

1. $\mathbb{I}=\varepsilon\mathbb{I}_0+\varepsilon^2\mathbb{I}_1+\varepsilon^3\mathbb{I}_2+\mathcal{O}(\varepsilon^4)$ *with* $\mathbb{I}_0(X)=-\omega^2\int\limits_{\mathbb{R}\times T^1}X_\phi^2\,\mathrm{d}y\mathrm{d}\phi$,

$$\mathbb{I}_1(X)=2\omega\,c\int\limits_{\mathbb{R}\times T^1}X_yX_\phi\,\mathrm{d}y\mathrm{d}\phi,\quad \mathbb{I}_2(X,\,X_\tau)=-c^2\int\limits_{\mathbb{R}\times T^1}X_y^2\,\mathrm{d}y\mathrm{d}\phi-\omega\int\limits_{\mathbb{R}\times T^1}X_\tau X_\phi\,\mathrm{d}y\mathrm{d}\phi,$$

2. $\mathbb{K} = \varepsilon\mathbb{K}_0 + \varepsilon^2\mathbb{K}_1 + \varepsilon^3\mathbb{K}_2 + \mathcal{O}(\varepsilon^4)$ *with* $\mathbb{K}_0(X) = -\frac{1}{2}\,\mathbb{I}_0(X)$*,* $\mathbb{K}_1(X) = -\frac{1}{2}\,\mathbb{I}_1(X)$ *and*

$$\mathbb{K}_2(X,\,X_\tau) = -\tfrac{1}{2}\,\mathbb{I}_2(X) + \tfrac{1}{2}\,\omega \int\limits_{\mathbb{R}\times T^1} X_\tau X_\phi \,\mathrm{d}y\mathrm{d}\phi,$$

3. $\mathbb{V} = \varepsilon\mathbb{V}_0 + \varepsilon^2\mathbb{V}_1 + \varepsilon^3\mathbb{V}_2 + \mathcal{O}(\varepsilon^4)$ *with*

$$\begin{aligned}
\mathbb{V}_0(X) &= -\tfrac{1}{2}\int\limits_{\mathbb{R}\times T^1} X\triangle_{0,\,\theta}X \,\mathrm{d}y\mathrm{d}\phi, &+\ \tfrac{v_2}{2}\int\limits_{\mathbb{R}\times T^1} X^2 \,\mathrm{d}y\mathrm{d}\phi,\\
\mathbb{V}_1(X) &= -\tfrac{1}{2}\int\limits_{\mathbb{R}\times T^1} X\big(\nabla^+_{0,\theta}X_y + \nabla^-_{0,\theta}X_y\big) \,\mathrm{d}y\mathrm{d}\phi &+\ \tfrac{v_3}{6}\int\limits_{\mathbb{R}\times T^1} X^3 \,\mathrm{d}y\mathrm{d}\phi,\\
\mathbb{V}_2(X) &= -\tfrac{1}{4}\int\limits_{\mathbb{R}\times T^1} X\big(\nabla^+_{0,\theta} - \nabla^-_{0,\theta} + 2\,\mathrm{Id}\big)X_{yy} \,\mathrm{d}y\mathrm{d}\phi &+\ \tfrac{v_4}{24}\int\limits_{\mathbb{R}\times T^1} X^4 \,\mathrm{d}y\mathrm{d}\phi.
\end{aligned} \tag{56}$$

Proof. The expansions for $\mathbb{K}$ and $\mathbb{I}$ follow directly from Lemma 38. Moreover, Taylor expansion with respect to ε yields

$$\triangle_{\varepsilon,\,\theta}X = \triangle_{0,\,\theta}X + \varepsilon\big(\nabla^+_{0,\theta} + \nabla^-_{0,\theta}\big)X_y + \varepsilon^2\,\tfrac{1}{2}\big(\nabla^+_{0,\theta} - \nabla^-_{0,\theta} + 2\mathrm{Id}\big)X_{yy} + \mathcal{O}\big(\varepsilon^3\big),$$

and this gives rise to the first kind of integrals in (56). Finally, inserting (55) into $\int_{\mathbb{R}\times T^1}\Phi_0(\varepsilon X)\,\mathrm{d}\eta\mathrm{d}\phi$ completes the proof. □

According to Lemma 39 the leading order Lagrangian and Hamiltonian equations are given by

$$-\omega^2 X_{\phi\phi} + \triangle_{0,\,\theta}X - v_2 X = 0, \tag{57}$$

and, using Fourier transform with respect to ϕ, we conclude that this equation has nontrivial solutions if and only if ω and θ satisfy $m^2\omega^2 = \Omega^2(m\theta)$ for some integer m, where Ω is the dispersion relation for the linearized chain with $\alpha = 1$, that is

$$\Omega^2(\theta) = v_2 + 2(1 - \cos\theta). \tag{58}$$

In what follows we always assume $\omega^2 = \Omega(\theta)^2$ as well as the *non-resonance* condition

$$m^2\omega^2 \neq \Omega(m\theta)^2 \text{ for } m \in \mathbb{Z}\setminus\{-1, 1\}, \tag{59}$$

which imply that the solution space to (57) in $L^2(T^1; \mathrm{d}\phi)$ is spanned by $\cos\phi$ and $\sin\phi$.

Theorem 40. *Suppose* $\omega^2 = \Omega^2(\theta)$ *and* (59)*. Then, the leading order Lagrangian and Hamiltonian equations* (57) *are quasi-stationary (corresponding to* $\Sigma_0 = 0$*) and have solutions*

$$X_0(y,\,\phi) = \pi^{-1/2}\Big(B_1(y)\cos(\phi) + B_2(y)\sin(\phi)\Big) \tag{60}$$

with $B_1,\,B_2 \in L^2(\mathbb{R}; \mathrm{d}y)$ *arbitrary. Moreover,* (60) *implies* $0 = \mathbb{L}_0(X_0) = \mathbb{H}_0(X_0)$.

Proof. All results follow from Lemma 39. □

Introducing a complex valued amplitude A by $2\sqrt{\pi}A = B_1 - \mathtt{i}B_2$, Equation (60) transforms into

$$X_0(y,\, \phi) = 2\,\mathrm{Re}\,\big(A(y)\mathtt{e}^{\mathtt{i}\phi}\big), \tag{61}$$

and is hence equivalent the original two-scale ansatz (15).

Elimination of the microstructure The leading order reduction determines the structure of the microscopic oscillations together with the dispersion relation. As discussed in Case C2 of §2.3, this allows for a further reduction step that yields the macroscopic modulation equation for the amplitudes B_1 and B_2, or, equivalently, for the complex-valued amplitude A. Let P_0 be the L^2–space of complex-valued functions depending on y, i.e.,

$$P_0 = \big\{(B_1,\, B_2) \in L^2(\mathbb{R};\mathrm{d}y)\times L^2(\mathbb{R};\mathrm{d}y)\big\} \cong \big\{A \in L^2(\mathbb{R};\,\mathbb{C})\big\}\,,$$

which can be viewed as a closed and proper subset of P due to (60). By construction, each element of P_0 satisfies the leading order equations exactly, and thus we can use the next-leading order terms in order to derive the effective macroscopic dynamics on TP_0.

Below we choose the moving frame speed c appropriately, and this yields $\mathbb{L}_1|_{P_0} \equiv 0$ as well as $\mathbb{H}_1|_{P_0} \equiv 0$ due to cancelations. Consequently, the reduced structures are related to $\mathbb{L}_2$, and hence we must take care of the correction terms $\widehat{\mathbb{L}}_2$ and $\widehat{\mathbb{H}}_2$ coming from the ansatz $X = X_0 + \varepsilon X_1$, see Case C2 in §2.3.

Lemma 41. *With $X = X_0 + \varepsilon X_1$ we have*

$$\widehat{\mathbb{L}}_2(X_0,\, X_1) = \mathbb{L}_0(X_1) + \int\limits_{\mathbb{R}\times T^1} X_1\big(2\,\omega\, cX_{0y\phi} + \big(\nabla^+_{0,\theta}X_{0y} + \nabla^-_{0,\theta}X_{0y}\big) - \tfrac{v_3}{2}X_0^2\big)\,\mathrm{d}y\mathrm{d}\phi$$

and $\widehat{\mathbb{H}}_2(X_0,\, X_1) = -\widehat{\mathbb{L}}_2(X_0,\, X_1)$. Moreover, all corrections to the symplectic structure are of order ε^3 and do not contribute to Σ_2.

Proof. The correction terms for $\mathbb{I}$, $\mathbb{K}$, and $\mathbb{V}$ can be read-off from Lemma 39. More precisely, we find

1. $\mathbb{I}_i(X_0 + \varepsilon\, X_1) = \mathbb{I}_i(X_0) + \widehat{\mathbb{I}}_i(X_0,\, X_1)$ with $\widehat{\mathbb{I}}_0(X_0,\, X_1) = 0$ and

$$\widehat{\mathbb{I}}_1(X_0,\, X_1) = 2\omega^2 \int\limits_{\mathbb{R}\times T^1} X_1 X_{0\phi\phi}\,\mathrm{d}y\mathrm{d}\phi,$$

$$\widehat{\mathbb{I}}_2(X_0,\, X_1) = \mathbb{I}_0(X_1) - 4\omega\, c \int\limits_{\mathbb{R}\times T^1} X_1 X_{0y\phi}\,\mathrm{d}y\mathrm{d}\phi,$$

2. $\mathbb{K}_i(X_0 + \varepsilon X_1) = \mathbb{K}_i(X_0) + \widehat{\mathbb{K}}_i(X_0, X_1)$ with $\widehat{\mathbb{K}}_0(X_0, X_1) = 0$ and

$$\widehat{\mathbb{K}}_1(X_0, X_1) = -\tfrac{1}{2}\widehat{\mathbb{I}}_1(X_0, X_1), \quad \widehat{\mathbb{K}}_2(X_0, X_1) = -\tfrac{1}{2}\widehat{\mathbb{I}}_2(X_0, X_1),$$

3. $\mathbb{V}_i(X_0 + \varepsilon X_1) = \mathbb{V}_i(X_0) + \widehat{\mathbb{V}}_i(X_0, X_1)$ with $\widehat{\mathbb{V}}_0(X_0, X_1) = 0$ and

$$\widehat{\mathbb{V}}_1(X_0, X_1) = -\int\limits_{\mathbb{R}\times T^1} X_1 \triangle_{0,\theta} X_0 \, \mathrm{d}y\mathrm{d}\phi + v_2 \int\limits_{\mathbb{R}\times T^1} X_1 X_0 \, \mathrm{d}y\mathrm{d}\phi,$$

$$\widehat{\mathbb{V}}_2(X_0, X_1) = \mathbb{V}_0(X_1) - \int\limits_{\mathbb{R}\times T^1} X_1 \big(\nabla^+_{0,\theta} X_{0y} + \nabla^-_{0,\theta} X_{0y}\big) \, \mathrm{d}y\mathrm{d}\phi + \tfrac{v_3}{2} \int\limits_{\mathbb{R}\times T^1} X_1 X_0^2 \, \mathrm{d}y\mathrm{d}\phi.$$

Finally, due to $\mathbb{L} = \mathbb{K} - \mathbb{V}$ and $\mathbb{H} = \mathbb{K} + \mathbb{V} + \mathbb{I}$ all assertions are direct consequences of these identities. □

Lemma 42. *If the moving frame speed c is given by*

$$c\,\omega = -\Omega'(\theta)\Omega(\theta) = -\sin\theta \tag{62}$$

then $\mathbb{L}_1|_{P_0} = -\mathbb{H}_1|_{P_0} = 0$. Otherwise the Lagrangian and Hamiltonian equations to $\mathbb{L}_1$ and $\mathbb{H}_1$ have no non-trivial solution at all.

Proof. A direct calculation shows

$$\mathbb{I}_1(X_0) = -2\,\mathbb{K}_1(X_0) = 2\,\omega\,c \int\limits_{\mathbb{R}} \big(B_{1y}B_2 - B_1 B_{2y}\big)\,\mathrm{d}y = 4\,\omega\,c \int\limits_{\mathbb{R}} B_{1y}B_2\,\mathrm{d}y,$$

and using

$$X\big(\nabla^+_{0,\theta} X_y + \nabla^-_{0,\theta} X_y\big) = 2\sin\theta(B_1\cos\phi + B_2\sin\phi)\big(-B_{1y}\sin\phi + B_{2y}\cos\phi\big)$$

as well as $0 = \int\limits_{T^1} (B_1\cos\phi + B_2\sin\phi)^3 \mathrm{d}\phi$ we find

$$\mathbb{V}_1(X_0) = -\sin\theta \int\limits_{\mathbb{R}} \big(B_1 B_{2y} - B_{1y}B_2\big)\,\mathrm{d}y = 2\sin\theta \int\limits_{\mathbb{R}} B_{1y}B_2\,\mathrm{d}y,$$

so that $\mathbb{L}_1|_{P_0} = -\mathbb{H}_1|_{P_0} = 0$ for (62). Finally, for other values of c the Lagrangian equations for $\mathbb{L}_1|_{P_0}$ equal $B_{1y} = B_{2y} = 0$. □

Condition (62) implies that the moving frame moves with the negative *group velocity* associated to the dispersion relation (58) (the negative sign appears since our phase definition is $\phi = \omega t + \theta\eta$). Compare this with the case $c^2 \neq \Phi_1''(0)$ from §3.3.

Next we prove that the non-resonance condition (59) provides the higher order correction X_1 in dependence of the first order solution X_0.

Lemma 43. *Suppose $\omega^2 = \Omega^2(\theta)$, $c\,\omega = -\Omega'(\theta)\Omega(\theta)$, and the non-resonance condition* (59)*, and let X_0 be fixed. Then, each solution X_1 to the equation*

$$\partial_{X_1}\widehat{\mathbb{L}}_2(X_0,\, X_1) = 0$$

satisfies $X_1 - \widehat{X}_1(X_0) \in P_0$, where the special solution $\widehat{X}_1(X_0)$ is given in the proof. Moreover, for each $\widetilde{X}_0 \in P_0$ we have

$$\widehat{\mathbb{L}}_2\Big(X_0,\, \widehat{X}_1(X_0) + \widetilde{X}_0\Big) = -\widehat{\mathbb{H}}_2\Big(X_0,\, \widehat{X}_1(X_0) + \widetilde{X}_0\Big) = -\overline{\mathbb{V}}_2(X_0)$$

with

$$\overline{\mathbb{V}}_2(X_0) = C\int\limits_{\mathbb{R}\times T^1} \big(B_1^2 + B_2^2\big)^2\,\mathrm{d}y\mathrm{d}\phi, \qquad C = \frac{v_3^2}{8\pi}\bigg(\frac{1}{4(4\omega^2 - \Omega^2(2\theta))} - \frac{1}{2v_2}\bigg).$$

Proof. The choice of c implies

$$\widehat{\mathbb{L}}_2(X_0,\, X_1) = \mathbb{L}_0(X_1) - \tfrac{v_3}{2}\int\limits_{\mathbb{R}\times T^1} X_1 X_0^2\,\mathrm{d}y\mathrm{d}\phi,$$

and hence the equation for X_1 becomes

$$\omega^2 X_{1\phi\phi} - \triangle_{0,\theta}X_1 + v_2 X_1 = -\tfrac{v_3}{2}X_0^2 = -\tfrac{v_3}{2}\Big(\tfrac{B_1^2+B_2^2}{2\pi} + \tfrac{B_1^2-B_2^2}{2\pi}\cos 2\phi + \tfrac{B_1B_2}{\pi}\sin 2\phi\Big). \tag{63}$$

This equation can be solved explicitly by Fourier transform with respect to ϕ and noting that the operator $\omega^2\partial_{\phi\phi} - \triangle_{0,\theta} + v_2$ is symmetric with kernel orthogonal to X_0^2. Some elementary analysis shows that each solution (63) can be written as $X_1 = \widehat{X}_1(X_0) + \widetilde{X}_0$, where $\widetilde{X}_0 \in P_0$ and

$$\widehat{X}_1(X_0) = C_1X_0^2 + C_2, \quad C_1 = \frac{v_3}{2(4\omega^2 - \Omega^2(2\theta))}, \quad C_2 = -\bigg(C_1 + \frac{v_3}{2v_2}\bigg)\frac{B_1^2 + B_2^2}{2\pi}.$$

Multiplying (63) by X_1 and integrating over $\mathbb{R}\times T^1$ gives

$$2\,\mathbb{L}_0\Big(\widehat{X}_1(X_0) + \widetilde{X}_0\Big) = \tfrac{v_3}{2}\int\limits_{\mathbb{R}\times T^1}\Big(\widehat{X}_1(X_0) + \widetilde{X}_0\Big)X_0^2\,\mathrm{d}y\mathrm{d}\phi = \tfrac{v_3}{2}\int\limits_{\mathbb{R}\times T^1}\widehat{X}_1(X_0)X_0^2\,\mathrm{d}y\mathrm{d}\phi$$

for all $\widetilde{X}_0 \in P_0$, and hence we find

$$\widehat{\mathbb{L}}_2\Big(X_0,\, \widehat{X}_1(X_0) + \widetilde{X}_0\Big) = \widehat{\mathbb{L}}_2\Big(X_0,\, \widehat{X}_1(X_0)\Big) = -\tfrac{v_3}{4}\int\limits_{\mathbb{R}\times T^1}\big(C_1X_0^2 + C_2\big)X_0^2\,\mathrm{d}y\mathrm{d}\phi$$

which implies the desired result. □

Finally, we combine all results and obtain the macroscopic model on TP_0.

Theorem 44. *Under the assumptions made in Lemma 43 the reduced Lagrangian and Hamiltonian are given by*

$$\mathbb{L}^{\rm red}(A,\,A_\tau) = \mathbb{K}^{\rm red}(A,\,A_\tau) - \mathbb{V}^{\rm red}(A), \quad \mathbb{H}^{\rm red}(A,\,A_\tau) = \mathbb{V}^{\rm red}(A),$$

with

$$\mathbb{K}^{\rm red}(A,\,A_\tau) = \mathtt{i}2\pi\omega \int\limits_{\mathbb{R}} \left(A\overline{A}_\tau - A_\tau\overline{A}\right) \mathrm{d}y,$$

$$\mathbb{V}^{\rm red}(A) = 2\pi\rho_1 \int\limits_{\mathbb{R}} |A_y|^2 \,\mathrm{d}y + 2\pi\rho_2 \int\limits_{\mathbb{R}} |A|^4 \,\mathrm{d}y,$$

where the constants ρ_1 and ρ_2 are given in (67)*, and $A = (B_1 - \mathtt{i}B_2)/(2\sqrt{\pi})$ is the complex-valued amplitude. Moreover, in terms of $(A,\,A_\tau)$ the reduced symplectic matrix $\Sigma^{\rm red}$ is given by*

$$\Sigma^{\rm red} = 4\,\pi\omega\mathtt{i}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{64}$$

and both the reduced Lagrangian and Hamiltonian equations are equivalent to

$$\mathtt{i}2\,\omega A_\tau = \rho_1 A_{yy} - 2\rho_2 \,|A|^2\, A, \tag{65}$$

which is a nonlinear Schrödinger equation.

Proof. Using the results from Lemma 39 and Lemma 41 we end up with

$$\widetilde{\mathbb{L}}(X_0,\,X_1) = \varepsilon^3(\mathbb{K}_2(X_0) - \mathbb{V}_2(X_0) + \widehat{\mathbb{L}}_2(X_0,\,X_1)) + \mathcal{O}\big(\varepsilon^4\big),$$
$$\widetilde{\mathbb{H}}(X_0,\,X_1) = \varepsilon^3(\mathbb{K}_2(X_0) + \mathbb{V}_2(X_0) + \mathbb{I}_2(X_0) + \widehat{\mathbb{H}}_2(X_0,\,X_1)) + \mathcal{O}\big(\varepsilon^4\big),$$

where we have used that $\mathbb{L}_i(X_0) = \mathbb{H}_i(X_0) = 0$ for $i = 0, 1$ and $X_0 \in P_0$, compare Theorem 40 and Lemma 42. Moreover, due to Lemma 43 we can eliminate X_1, and this yields

$$\begin{aligned}\mathbb{L}^{\rm red}(X_0) &= \varepsilon^3(\mathbb{K}_2(X_0) - \mathbb{V}_2(X_0) - \overline{\mathbb{V}}_2(X_0)),\\ \mathbb{H}^{\rm red}(X_0) &= \varepsilon^3(\mathbb{K}_2(X_0) + \mathbb{V}_2(X_0) + \mathbb{I}_2(X_0) + \overline{\mathbb{V}}_2(X_0)),\end{aligned} \tag{66}$$

compare Case C2 in §2.3. Inserting (60) into the formulas from Lemma 39 gives

$$\mathbb{I}_2(X_0) = -c^2 \int\limits_{\mathbb{R}} \left(\big(B_{1y}\big)^2 + \big(B_{2y}\big)^2\right) \mathrm{d}y - \omega \int\limits_{\mathbb{R}} (B_{1\tau}B_2 - B_1B_{2\tau})\,\mathrm{d}y,$$

$$\mathbb{K}_2(X_0) = \tfrac{1}{2}c^2 \int\limits_{\mathbb{R}} \left(\big(B_{1y}\big)^2 + \big(B_{2y}\big)^2\right) \mathrm{d}y + \omega \int\limits_{\mathbb{R}} (B_{1\tau}B_2 - B_1B_{2\tau})\,\mathrm{d}y,$$

$$\mathbb{V}_2(X_0) = \tfrac{1}{2}\,\cos\theta \int\limits_{\mathbb{R}} \left(\big(B_{1y}\big)^2 + \big(B_{2y}\big)^2\right) \mathrm{d}y + \tfrac{v_4}{32\pi} \int\limits_{\mathbb{R}} \big(B_1^2 + B_2^2\big)^2 \,\mathrm{d}y.$$

By construction, B_1 and B_2 satisfy $B_1 = \sqrt{\pi}\big(A+\overline{A}\big)$, $B_2 = \mathtt{i}\sqrt{\pi}\big(A-\overline{A}\big)$, and thus we find $B_{1\tau}B_2 - B_1B_{2\tau} = \mathtt{i}2\pi\left(A\overline{A}_\tau - A_\tau\overline{A}\right)$ as well as

$$\big(B_{1y}\big)^2 + \big(B_{2y}\big)^2 = 4\pi\, A_y\,\overline{A}_y = 4\pi\,|A_y|^2, \quad \big({B_1}^2 + {B_2}^2\big)^2 = 16\pi^2\left(A\,\overline{A}\right)^2 = 16\pi^2\,|A|^4.$$

We define

$$\rho_1 := \Omega(\theta)\Omega''(\theta) = \cos\theta - c^2, \quad \rho_2 := \frac{v_4}{4} - \frac{v_3^2}{2v_2} + \frac{v_3^2}{4(4\omega^2 - \Omega^2(2\theta))} \tag{67}$$

and

$$\begin{aligned}
\mathbb{K}^{\mathrm{red}} &:= \omega\int_{\mathbb{R}} \left(B_{1\tau}B_2 - B_1B_{2\tau}\right)\mathrm{d}y,\\
\mathbb{V}^{\mathrm{red}} &:= \tfrac{\rho_1}{2}\int_{\mathbb{R}} \Big(\big(B_{1y}\big)^2 + \big(B_{2y}\big)^2\Big)\,\mathrm{d}y + \tfrac{\rho_2}{8\pi}\int_{\mathbb{R}} \big(B_1^2 + B_2^2\big)^2\,\mathrm{d}y,
\end{aligned}$$

and this implies the formulas for $\mathbb{L}^{\mathrm{red}}$ and $\mathbb{H}^{\mathrm{red}}$. To compute Σ^{red}, recall $\Sigma = \varepsilon^3\Sigma_2 + \mathcal{O}(\varepsilon^4)$ independent of X_1, and notice that the ansatz (60) can be written as

$$\begin{pmatrix} X \\ X_\tau \end{pmatrix} = \boldsymbol{T}_0 \begin{pmatrix} B_1 \\ B_2 \\ B_{1\tau} \\ B_{2\tau} \end{pmatrix}, \quad \boldsymbol{T}_0 = \frac{1}{\sqrt{\pi}}\begin{pmatrix} \cos\phi & \sin\phi & 0 & 0 \\ 0 & 0 & \cos\phi & \sin\phi \end{pmatrix}$$

with $\boldsymbol{T}_0 : TP_0 \to TP$. The adjoint operator $\boldsymbol{T}_0' : TP \to TP_0$ reads

$$\boldsymbol{T}_0'\begin{pmatrix} X \\ X_\tau \end{pmatrix} = \frac{1}{\sqrt{\pi}}\left(\int_{T^1} X\cos\phi\,\mathrm{d}\phi,\ \int_{T^1} X\sin\phi\,\mathrm{d}\phi,\ \int_{T^1} X_\tau\cos\phi\,\mathrm{d}\phi,\ \int_{T^1} X_\tau\sin\phi\,\mathrm{d}\phi\right)^T$$

and with respect to the variables B_1, B_2, $B_{1\tau}$, $B_{2\tau}$ we find

$$\Sigma^{\mathrm{red}} = \boldsymbol{T}_0'\begin{pmatrix} 2\omega\partial_\phi & 0 \\ 0 & 0 \end{pmatrix}\boldsymbol{T} = 2\,\omega\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which implies (64). From this and (66) we conclude that both the reduced Lagrangian and Hamiltonian equations on P_0 read

$$\begin{aligned}
-2\,\omega B_{2\tau} &= -\rho_1 B_{1yy} + \tfrac{1}{2\pi}\rho_2\big(B_1^2 + B_2^2\big)B_1,\\
+2\,\omega B_{1\tau} &= -\rho_1 B_{2yy} + \tfrac{1}{2\pi}\rho_2\big(B_1^2 + B_2^2\big)B_2
\end{aligned}$$

and rewriting this in terms of A we find (65). □

As before, Theorem 44 concerns a reduced macroscopic model on P_0 that is obtained by means of *formal expansions*. In particular, it is not obvious that the nlS equation (65) combined with the modulation ansatz (61) yields approximate solutions for the KG chain. However, the careful residual analysis from [GM04, GM06] provides rigorous justification results, and thus we can regard P_0 as an approximate invariant manifold.

Remark 45. Like for the KdV example, the terms $\mathbb{I}_0$ and $\mathbb{I}_1$, which cause the cancelations in $\mathbb{L}_0$ and $\mathbb{L}_1$, provide macroscopic conservation laws. In fact, with some calculations we find

$$\mathbb{I}_0 \sim \int_{\mathbb{R}} |A|^2 \, \mathrm{d}y \quad \text{and} \quad \mathbb{I}_1 \sim \int_{\mathbb{R}} \mathrm{Im}(A_y \overline{A}) \, \mathrm{d}y,$$

which equal the macroscopic integrals of motion associated with the symmetries under phase shifts $A \mapsto \mathtt{e}^{\mathtt{i}s} A$, and shifts in the y-direction, respectively.

3.5 Three-wave-interaction for the KG chain

Here we discuss the interaction of three pulses in the KG chain, see (16). More precisely, we consider three pulses p_1, p_2, and p_3, and aim to understand how the resulting microstructure is modulated on the hyperbolic scale for space and time.

Pulses in the KG chain We briefly summarize some aspects of *pulses*, and refer to [GM04, Gia08, GMS07] for more details. A plane wave is a solution to the linearized chain

$$x_j(t) = A\mathtt{e}^{\mathtt{i}(\omega t + \theta j)} + \text{c.c.} = A\mathtt{e}^{+\mathtt{i}(\omega t + \theta j)} + \overline{A}\mathtt{e}^{-\mathtt{i}(\omega t + \theta j)}, \quad j \in \mathbb{Z},$$

with complex amplitude A, frequency ω, and wave number θ. Notice that $(-\theta, -\omega)$ gives the same pulse as (θ, ω), whereas $(\theta, -\omega)$ is the pulse that travels in opposite direction. Obviously, each plane wave must satisfy the dispersion relation

$$\omega^2 = \Omega^2(\theta) = v_2 + 2\alpha(1 - \cos\theta) \quad \text{with } v_2 = \Phi_0''(0).$$

Since the amplitude A can always be chosen arbitrarily, we can identify each plane wave with a point in

$$\mathcal{P} = \{(\theta, \omega) \,:\, \omega^2 = \Omega^2(\theta)\} \subset T^1 \times \mathbb{R}.$$

In what follows we assume the stability condition

$$\min\{4\alpha + v_2, \, v_2\} = \min_{\theta \in [0, 2\pi)} \Omega^2(\theta) > 0, \tag{68}$$

so that each single plane wave is a stable solution to the linearized chain.

A simple *pulse* is a modulation of a plane wave by a slowly varying amplitude

$$x_j^{(k)}(t) = \varepsilon \, A_k(\varepsilon t, \varepsilon j) \, \mathtt{e}^{\mathtt{i}(\omega_k t + \theta_k j)} + \text{c.c.}.$$

On the hyperbolic scale $\tau = \varepsilon t$ and $y = \varepsilon j$ a pulse will simply travel with group velocity $c_k = \Omega'(\theta_k)$. However, if different pulses associated with $p_k \in \mathcal{P}$ meet each other they interact in case their frequencies and wave vectors are in resonance. Three plane waves $p_1, p_2, p_3 \in \mathcal{P}$ are called in *three-wave resonance* if there exists a choice of three signs $m_i \in \{-1, +1\}$ such that $m_1p_1 + m_2p_2 + m_3p_3 = (0, 0) \in T^1 \times \mathbb{R}$. By using complex conjugates and replacing p_k by $-p_k$ if necessary, we can always assume that

$$p_1 + p_2 + p_3 = 0, \quad \text{i.e.,} \quad \begin{cases} \theta_1 + \theta_2 + \theta_3 = 0 & \in T^1, \\ \omega_1 + \omega_2 + \omega_3 = 0 & \in \mathbb{R}. \end{cases} \tag{69}$$

This *resonance condition* arises naturally as it is equivalent to the *cancelation of oscillations* via

$$\mathtt{e}^{\mathtt{i}\,(\omega_1 t+\theta_1 j)}\mathtt{e}^{\mathtt{i}\,(\omega_2 t+\theta_2 j)}\mathtt{e}^{\mathtt{i}\,(\omega_3 t+\theta_3 j)} = 1 \quad \text{for all } t \in \mathbb{R} \text{ and } j \in \mathbb{Z},$$

and guarantees that the product of two pulses contains oscillatory terms that appear also in the third pulse.

Of course, the KG chain allows for resonances between more than three pulses, but in our context these can be ignored for the following reason. According to (16), the pulse amplitudes scale with ε, so that three-pulse resonances, which are related to quadratic products such as x_1x_2, correspond to the power ε^2. Interactions of more than three pulses, however, contribute to order $\mathcal{O}(\varepsilon^3)$, and are thus not relevant on the hyperbolic scale.

However, to make the presentation as simple as possible we now assume that the three plane waves p_1, p_2, and p_3 are chosen such that except for (69) there are no further resonances. More precisely, we define

$$\mathcal{Z} = \{\, (k_1, k_2) \in \mathbb{Z}^2 \;:\; (\omega_1 k_1 + \omega_2 k_2)^2 = \Omega^2(\theta_1 k_1 + \theta_2 k_2) \,\},$$

and make the following assumption.

Assumption 46. The vectors $p_1, p_2 \in \mathcal{P}$ are chosen such that

$$\mathcal{Z} = \{\, (1, 0),\; (-1, 0),\; (0, 1),\; (0, -1),\; (1, 1),\; (-1, -1) \,\}.$$

Obviously, we always have $(\pm 1, 0) \in \mathcal{Z}$ by $p_1 \in \mathcal{P}$ and similarly $(0, \pm 1) \in \mathcal{Z}$ by $p_2 \in \mathcal{P}$. If additionally $p_3 \in \mathcal{P}$ satisfies the three-wave resonance condition (69), then $\pm(1, 1)$ also lies in $\mathcal{Z}$. Thus, Assumption 46 already implies (69) and additionally excludes any further resonances involving these three plane waves.

Remark 47. (i) According to [Gia08], the resonance condition (69) is equivalent to

$$\alpha^2\mu_1\mu_2(\mu_1+\mu_2) + (2\alpha\mu_1\mu_2+\delta)\sqrt{(\alpha\mu_1+\delta)(\alpha\mu_2+\delta)} + \delta\alpha(\mu_1\mu_2+\mu_1+\mu_2) + \tfrac{5}{4}\delta^2 = 0$$

with $\mu_i = (1-\cos\theta_i)/2$ and $\delta = v_2/4$. Hence, for $\alpha > 0$ (attractive nearest-neighbour interactions) the resonance condition cannot be satisfied as the stability condition (68) implies $\delta > 0$. However, for $\alpha \in (-\frac{1}{4}v_2, -\frac{3}{16}v_2)$ (repulsive case) the stability condition is still satisfied, but now there exists a one-parameter family of solutions

$(\mu_1,\, \mu_2)$.
(ii) In general, it is not easy to check the *non-resonance conditions* implied by Assumption 46, i.e., to prove that no further plane waves are contained in $\mathcal{Z}$. The mapping $\mathbb{Z}^2 \ni k \mapsto (k\cdot(\theta_1, \theta_2), k\cdot(\omega_1, \omega_2)) \in T^1 \times \mathbb{R}$ may have a dense image and hence comes close to the set $\mathcal{P}$ very often, giving rise to a small divisor problem. However, by varying also α and v_2, it is possible to choose θ_1, θ_2 as rational multiples of π and to make ω_1/ω_2 rational as well. Then, the image of the above mapping hits every bounded set in finitely many points. Then, Assumption 46 appears very reasonable.
(iii) In Remark 53 below we provide a weaker variant of Assumption 46.

Invertible two-scale ansatz The resonance and non-resonance conditions imposed by Assumption 46 imply that there exist exactly two independent phase variables. Therefore, concerning the embedding of the microscopic system, it is necessary and sufficient to introduce a two-dimensional phase variable $\phi = (\phi_1,\, \phi_2) \in T^2$, i.e.,

$$Q = L^2\big(\mathbb{R}{\times}T^2;\, \mathrm{d}\eta\mathrm{d}\phi\big), \quad P = L^2\big(\mathbb{R}{\times}T^2;\, \mathrm{d}y\mathrm{d}\phi\big).$$

Similarly to the nlS example we start with the invertible two-scale ansatz

$$x(t,\, \eta,\, \phi) = \varepsilon X(\varepsilon t,\, \varepsilon\eta,\, \phi + \omega t + \theta\eta), \tag{70}$$

so that the corresponding inverse two-scale transformation $\mathbb{T}_{\mathrm{vel}}(\varepsilon,\, t) : TP \to TQ$ reads

$$\Big(\mathbb{T}_{\mathrm{vel}}(\varepsilon,\, t)(X,\, X_\tau)\Big)(\eta,\, \phi) = \big(\varepsilon X,\, \varepsilon^2 X_\tau + \varepsilon\omega \cdot X_\phi\big)(\varepsilon\eta,\, \phi + \omega t + \theta\eta) \tag{71}$$

with $\theta = (\theta_1,\, \theta_2)$, $\omega = (\omega_1,\, \omega_2)$, and $\partial_\phi = (\partial_{\phi_1},\, \partial_{\phi_2})$.

Remark 48. The ansatz (70) provides $\mathbb{T}_{\mathrm{con}}(\varepsilon,\, t) = \mathcal{T}_{\mathrm{con}}^{-1}\circ\mathcal{M}_{\mathrm{con}}(-t)\circ\mathbb{S}_{\mathrm{con}}(\varepsilon) : P \to Q$ with inverse scaling transformation $(\mathbb{S}_{\mathrm{con}}(\varepsilon)X)(\eta,\, \phi) = \varepsilon X(\varepsilon\eta,\, \phi)$, weak symmetry transformation $(\mathcal{T}_{\mathrm{con}}\, x)(\eta,\, \phi) = x(\eta,\, \phi - \theta\eta)$, and moving frame transformation $(\mathcal{M}_{\mathrm{con}}(t)\, x)(\eta,\, \phi) = x(\eta,\, \phi - \omega t)$ associated to the integral of motion $\mathcal{I}(x,\, x_t) = -\omega \cdot \mathcal{I}_{\mathrm{phase}}(x,\, x_t) = -\int_{\mathbb{R}\times T^2} x_t(\omega \cdot x_\phi)\mathrm{d}\eta\mathrm{d}\phi$.

Leading order reduction We start with the computation of the transformed structures.

Lemma 49. *The transformation* (71) *provides* $\mathbb{L} = \mathbb{K}{-}\mathbb{V}$, $\mathbb{E} = \mathbb{K}{+}\mathbb{V}$, *and* $\mathbb{H} = \mathbb{E}{+}\mathbb{I}$, *as well as the following expansions:*

1. $\mathbb{I} = \varepsilon\mathbb{I}_0 + \varepsilon^2\mathbb{I}_1 + \mathcal{O}(\varepsilon^3)$ *with*

$$\mathbb{I}_0(X) = -\int_{\mathbb{R}\times T^2} (\omega \cdot X_\phi)^2\, \mathrm{d}y\mathrm{d}\phi, \quad \mathbb{I}_1(X,\, X_\tau) = -\int_{\mathbb{R}\times T^2} X_\tau(\omega \cdot X_\phi)\, \mathrm{d}y\mathrm{d}\phi,$$

2. $\mathbb{K} = \varepsilon\mathbb{K}_0 + \varepsilon^2\mathbb{K}_1 + \mathcal{O}(\varepsilon^3)$ *with* $\mathbb{K}_0 = -\frac{1}{2}\,\mathbb{I}_0$ *and* $\mathbb{K}_1 = -\mathbb{I}_1$,

3. $\mathbb{V} = \varepsilon\mathbb{V}_0 + \varepsilon^2\mathbb{V}_1 + \mathcal{O}(\varepsilon^3)$ *with*

$$\begin{aligned}\mathbb{V}_0(X) &= -\tfrac{\alpha}{2}\int\limits_{\mathbb{R}\times T^2} X\triangle_{0,\theta}X\,\mathrm{d}y\mathrm{d}\phi &+\ \tfrac{v_2}{2}\int\limits_{\mathbb{R}\times T^2} X^2\,\mathrm{d}y\mathrm{d}\phi,\\ \mathbb{V}_1(X) &= -\tfrac{\alpha}{2}\int\limits_{\mathbb{R}\times T^2} X\big(\nabla^+_{0,\theta} + \nabla^-_{0,\theta}\big)X_y\,\mathrm{d}y\mathrm{d}\phi &+\ \tfrac{v_3}{6}\int\limits_{\mathbb{R}\times T^2} X^3\,\mathrm{d}y\mathrm{d}\phi,\end{aligned}$$

where $\alpha = \Phi_1''(0)$, $v_2 = \Phi_0''(0)$, *and* $v_3 = \Phi_0'''(0)$.

Moreover, the matrix Σ *corresponding to* σ *obeys the exact expansion*

$$\Sigma = \varepsilon^2\Sigma_1 + \varepsilon^3\Sigma_2, \quad \Sigma_1 = \begin{pmatrix} -2\,\omega\cdot\partial_\phi & 0\\ 0 & 0\end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix},$$

so that Σ *is non-degenerate due to* Σ_2.

Proof. Analogously to the proof of Lemma 38 we find the equations for $\mathbb{L}$ and $\mathbb{H}$ along with

$$\begin{aligned}\mathbb{I}(\varepsilon,\,X,\,X_\tau) = \mathcal{I}\circ\mathbb{T}_{\mathrm{vel}}(\varepsilon,\,t) &= \varepsilon^{-1}\int\limits_{\mathbb{R}\times T^2}\big(\varepsilon^2X_\tau + \varepsilon\omega\cdot X_\phi\big)(-\varepsilon\omega\cdot X_\phi)\,\mathrm{d}y\mathrm{d}\phi,\\ \mathbb{K}(\varepsilon,\,X,\,X_\tau) = \mathcal{K}\circ\mathbb{T}_{\mathrm{vel}}(\varepsilon,\,t) &= \varepsilon^{-1}\int\limits_{\mathbb{R}\times T^2}\tfrac{1}{2}\big(\varepsilon^2X_\tau + \varepsilon\omega\cdot\partial_\phi X\big)^2\,\mathrm{d}y\mathrm{d}\phi,\\ \mathbb{V}(\varepsilon,\,X) = \mathcal{V}\circ\mathbb{T}_{\mathrm{vel}}(\varepsilon,\,t) &= \varepsilon^{-1}\int\limits_{\mathbb{R}\times T^2}\big(-\varepsilon^2\tfrac{\alpha}{2}X\triangle_{\varepsilon,\theta}X + \Phi_0(\varepsilon X)\big)\,\mathrm{d}y\mathrm{d}\phi.\end{aligned}$$

Moreover, the expansions with respect to ε follow from direct calculations, and using

$$\boldsymbol{T} = \begin{pmatrix} \varepsilon & 0\\ \varepsilon\,\omega\cdot\partial_\phi & \varepsilon^2\end{pmatrix}$$

the matrix Σ can be calculated by means of Remark 27. ☐

As a consequence of Lemma 49 we obtain $\Sigma_0 = 0$ and $\mathbb{L}_0(X) = -\mathbb{H}_0(X)$, and the leading order equation

$$(\omega\cdot\partial_\phi)^2X - \alpha\triangle_{0,\theta}X + v_2X = 0.$$

is again quasi-stationary. Applying Fourier transformation with respect to ϕ, a general function X has the form $X(y,\phi) = \sum_{k\in\mathbb{Z}^2}F_k(y)\mathrm{e}^{\mathrm{i}k\cdot\phi}$ and solves the above equation if and only if $F_k = 0$ for all $k \notin \mathcal{Z}$.

Lemma 50. *Under Assumption 46 the leading order Lagrangian and Hamiltonian equations are quasi-stationary, and all solutions are given by*

$$X_0(y,\,\phi) = \sum_{n=1}^{3} A_n(y)\mathtt{e}^{\mathtt{i}\phi_n} + \text{c.c.}, \tag{72}$$

with $\phi_3 = -\phi_1 - \phi_2$ and arbitrary $A_n \in L^2(\mathbb{R};\mathbb{C})$, $n = 1, 2, 3$. Moreover, we have $\mathbb{L}_0(X_0) = \mathbb{H}_0(X_0) = 0$ for all X_0 with (72).

Elimination of the microstructure As outlined in §2.3, we derive the reduced macroscopic model by restricting the next-leading order terms to the space

$$P_0 = \{X_0 \in P \;:\; \partial_X\mathbb{L}_0(X_0) = \partial_X\mathbb{H}_0(X_0) = 0\} \cong \{A = (A_1, A_2, A_3) \in (L^2(\mathbb{R};\,\mathbb{C}))^3\}.$$

Notice that, in contrast to the nlS example from §3.4, here $\mathbb{L}_1|_{TP_0}$ and $\mathbb{H}_1|_{TP_0}$ do not vanish, and provide the reduced Lagrangian and Hamiltonian. In particular, we need not care for the correction terms coming from $X = X_0 + \varepsilon X_1$.

Theorem 51. *Under Assumption 46 the reduced Lagrangian and Hamiltonian are given by*

$$\mathbb{L}^{\text{red}}(A,\,A_\tau) = \mathbb{K}^{\text{red}}(A,\,A_\tau) - \mathbb{V}^{\text{red}}(A), \quad \mathbb{H}^{\text{red}}(A,\,A_\tau) = \mathbb{V}^{\text{red}}(A),$$

with

$$\mathbb{K}^{\text{red}}(A,\,A_\tau) = \mathtt{i}\sum_{n=1}^{3}\omega_n\int_{\mathbb{R}} A_n\overline{A}_{n\tau}\,\mathrm{d}y + \text{c.c.}$$

$$\mathbb{V}^{\text{red}}(A) = v_3\int_{\mathbb{R}} A_1A_2A_3\,\mathrm{d}y + \mathtt{i}\sum_{n=1}^{3}\omega_n\omega_n'\int_{\mathbb{R}} A_n\overline{A_{ny}}\,\mathrm{d}y + \text{c.c.},$$

where $\omega_n\omega_n' = \Omega(\theta_n)\Omega'(\theta_n)$. Moreover,

$$\Sigma^{\text{red}} = -2\mathtt{i}\begin{pmatrix}\mathbf{\Omega} & 0\\ 0 & 0\end{pmatrix}, \qquad \mathbf{\Omega} = \begin{pmatrix}\omega_1 & 0 & 0\\ 0 & \omega_2 & 0\\ 0 & 0 & \omega_3\end{pmatrix},$$

is the reduced symplectic matrix, and the reduced Lagrangian and Hamiltonian equations are equivalent to the three-wave-interaction equations (17).

Proof. According to §2.3 we have

$$\mathbb{L}^{\text{red}} = c\mathbb{K}_1|_{TP_0} - c\mathbb{V}_1|_{TP_0}, \quad \mathbb{H}^{\text{red}} = c\mathbb{K}_1|_{TP_0} + c\mathbb{V}_1|_{TP_0} + c\mathbb{I}_1|_{TP_0}, \quad \sigma^{\text{red}} = c\sigma_1|_{TP_0},$$

where for convenience we introduced a trivial scaling by $c = 1/(4\pi^2)$. Inserting (72) into the formulas from Lemma 49, and exploiting Assumption 46 we obtain

$$c\,\mathbb{K}_1(A,\,A_\tau) = -c\mathbb{I}_1(A,\,A_\tau) = \sum_{n=1}^{3}\int_{\mathbb{R}} \mathtt{i}\omega_n A_n \overline{A_{n\tau}}\,\mathrm{d}y + \text{c.c.},$$

$$c\mathbb{V}_1(A) = \Big(\sum_{n=1}^{3}\int_{\mathbb{R}} \mathtt{i}\omega_n\omega_n' A_n\overline{A_{ny}}\,\mathrm{d}y + v_3\int_{\mathbb{R}} A_1A_2A_3\,\mathrm{d}y\Big) + \text{c.c.},$$

where we used $\alpha\sin\theta_n = \omega_n\omega_n'$ and the properties of ∇ and $\triangle$, see Remark 29. Concerning Σ^{red} we observe that the ansatz (72) can be written as

$$\begin{pmatrix} X \\ X_\tau \end{pmatrix} = \boldsymbol{T}_0 \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_{1\tau} \\ A_{2\tau} \\ A_{3\tau} \end{pmatrix} + \text{c.c.}, \quad \boldsymbol{T}_0 = \begin{pmatrix} \mathtt{e}^{\mathtt{i}\phi_1} & \mathtt{e}^{\mathtt{i}\phi_2} & \mathtt{e}^{-\mathtt{i}(\phi_1+\phi_2)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathtt{e}^{\mathtt{i}\phi_1} & \mathtt{e}^{\mathtt{i}\phi_2} & \mathtt{e}^{-\mathtt{i}(\phi_1+\phi_2)} \end{pmatrix}$$

with $\boldsymbol{T}_0 : TP_0 \to TP$. The adjoint operator $\boldsymbol{T}_0' : TP \to TP_0$ reads

$$\boldsymbol{T}_0'\begin{pmatrix} X \\ X_\tau \end{pmatrix} = \int_{T^2} \begin{pmatrix} \mathtt{e}^{-\mathtt{i}\phi_1} & 0 \\ \mathtt{e}^{-\mathtt{i}\phi_2} & 0 \\ \mathtt{e}^{\mathtt{i}(\phi_1+\phi_2)} & 0 \\ 0 & \mathtt{e}^{-\mathtt{i}\phi_1} \\ 0 & \mathtt{e}^{-\mathtt{i}\phi_2} \\ 0 & \mathtt{e}^{\mathtt{i}(\phi_1+\phi_2)} \end{pmatrix} \begin{pmatrix} X \\ X_\tau \end{pmatrix}\,\mathrm{d}\phi,$$

and we find

$$\Sigma^{\mathrm{red}} = c\,\boldsymbol{T}_0'\Sigma_1\boldsymbol{T}_0 = c\,\boldsymbol{T}_0'\begin{pmatrix} -2\omega\cdot\partial_\phi & 0 \\ 0 & 0 \end{pmatrix}\boldsymbol{T}_0 = -2\mathtt{i}\begin{pmatrix} \boldsymbol{\Omega} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

Finally, the Lagrangian equations to $\mathbb{L}^{\mathrm{red}}$ are given by

$$\partial_\tau\left(\partial_{\overline{A}_\tau}\mathbb{L}^{\mathrm{red}}(A,\,A_\tau)\right) - \partial_{\overline{A}}\mathbb{L}^{\mathrm{red}}(A,\,A_\tau) = 0$$

and equal

$$-2\,\mathtt{i}\boldsymbol{\Omega}A_\tau = \partial_{\overline{A}}\mathbb{H}^{\mathrm{red}}(A),$$

which is the Hamiltonian equation to $(\mathbb{H}^{\mathrm{red}}, \sigma^{\mathrm{red}})$. Finally, both equations coincide with (17). □

As mentioned in the introduction, one can obtain the macroscopic equations (17) also by inserting the two-scale ansatz (16) into the Klein–Gordon chain (35) and requiring the coefficients of the terms $\varepsilon^2\mathtt{e}^{\mathtt{i}(\omega_n t+\theta_n j)}$ to vanish. Based on this formal expansion one can then justify the validity of (17), see [Gia06, Gia08] and §7.2 in [GHM06].

Remark 52. Inserting (72) into the formulas from Lemma 49 and exploiting the resonance condition, we obtain

$$c\mathbb{K}_0|_{P_0} = -\tfrac{1}{2}c\mathbb{I}_0|_{P_0} = c\mathbb{V}_0|_{P_0} = \tfrac{1}{2}c\mathbb{E}_0|_{P_0} = \sum_{n=1}^{3}\omega_n^2 \int_{\mathbb{R}} |A_n|^2\,\mathrm{d}y.$$

These equalities reflect the cancelation in $\mathbb{L}_0$ and manifest the equipartition of energy for plane-wave solutions. Moreover, the total energy $\mathbb{E}_0$ is the first integral associated to the invariance under phase shifts.

Remark 53. Assumption 46, which excludes all other possible resonances except for $p_1+p_2+p_3 = 0$, can be weakened as follows. As already mentioned, on the hyperbolic scale we can ignore resonances of more than three pulses. We shall, however, exclude the possibility that further pulses are created via three-pulse resonance, because otherwise we expect the three-pulse solution that involves p_1, p_2, and p_3 to be unstable on the hyperbolic scale. This gives rise to the non-resonance conditions

$$2p_1,\, 2p_2,\, 2p_3,\, p_1 - p_2,\, p_1 - p_3,\, p_2 - p_3 \notin \mathcal{P} \setminus \{\pm p_1,\, \pm p_2,\, \pm p_3\}.$$

Assuming this, it can happen that there exist further degenerate three-pulse resonances between p_1, p_2, and p_3, as for instance $2p_1 + p_3 = 0$ or $2p_1 - p_2 = 0$. In this case we still obtain a stable three-pulse solution, but the reduction procedure provides a different modulation equation. In fact, such degenerate resonances give rise to further cubic coupling terms in the formula for $\mathbb{V}^{\mathrm{red}}$, as for instance $\frac{v_3}{2}\int_{\mathbb{R}}(A_1^2 + A_2^2)A_3\mathrm{d}y + \text{c.c.}$ or $\frac{v_3}{2}\int_{\mathbb{R}} A_1^2\overline{A_2}\mathrm{d}y + \text{c.c.}$, respectively. Altogether, in order to guarantee that (17) is a reasonable macroscopic model it is sufficient to assume the resonance condition

$$(1,\, 0),\, (0,\, 1),\, (1,\, 1) \in \mathcal{Z}$$

and the non-resonance conditions

$$(0,\, 2),\, (2,\, 0),\, (1,\, -1),\, (2,\, 1),\, (1,\, 2),\, (2,\, 2) \notin \mathcal{Z}.$$

3.6 Outlook to further examples

Finally, we give a brief overview on two other classes of micro-macro transitions that can also be studied with respect to Hamiltonian and Lagrangian reductions. However, since these examples lead to additional problems, their investigation is left for a forthcoming study.

Coupled systems describe the interactions between modulated pulses and waves with long wave-length. The interesting feature here is that the corresponding two-scale ansatz

$$x(t,\, \eta,\, \phi) = \varepsilon^{\alpha}X(\varepsilon t,\, \varepsilon j) + \varepsilon^{\beta}A(\varepsilon t,\, \varepsilon j)\mathtt{e}^{+\mathtt{i}(\omega t+\theta j)} + \text{c.c.} \tag{73}$$

combines contributions with different orders of magnitude. For instance, if we derive the effective macroscopic model for $\alpha = 0$ and $\beta = 1$ by inserting (73) into the microscopic equation of motion, we find

$$\partial_{\tau\tau} X = c_{\mathrm{m}}^2\,\partial_{yy} X, \quad \mathtt{i}\,\partial_\tau A = \mathtt{i}\,c_{\mathrm{gr}}\,\partial_y A - \rho_0\,\partial_y X\,A. \tag{74}$$

However, the asymmetric coupling between both equations prevents (74) from being the Euler-Lagrange equation of a suitable chosen macroscopic Lagrangian with variables X and A, and we conclude that the reduction of Lagrangian and Hamiltonian structures yields a different reduced model.

Whitham's modulation theory is another example postponed to our forthcoming paper. This theory was originally developed in the context of PDEs, see [Whi74, Kam06], but can also be applied to discrete systems, see for instance [HLM94, FV99]. The main building block for Whitham's modulation theory are *periodic travelling waves*. These are exact solutions to (13) satisfying $x_j(t) = \mathfrak{X}(\omega t + \theta j)$ with $\phi = \omega t + \theta j$. For the atomic chain the profile $\mathfrak{X}$ must fulfil the following advance-delay differential equation

$$\omega^2\,\mathfrak{X}_{\phi\phi}(\phi) = \Phi_1'\Big(\mathfrak{X}(\phi+\theta) - \mathfrak{X}(\phi)\Big) - \Phi_1'\Big(\mathfrak{X}(\phi) - \mathfrak{X}(\phi-\theta)\Big) - \Phi_0'(\mathfrak{X}(\phi)).$$

In case that both Φ_0' and Φ_1' are linear, we can solve this equation by means of Fourier transformation, and will recover plane waves with (36), but for nonlinear potentials more sophisticated methods are necessary, compare for instance [DHM06] and references therein. The basic ideas behind Whitham's modulation theory can be summarized as follows: We consider the KG chain and start with the following two-scale ansatz

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

Here, Θ is the *modulated phase* and provides the fields of wave number and frequency via $\omega(\tau,\, y) = \partial_\tau\Theta(\tau,\, y)$ and $\theta(\tau,\, y) = \partial_y\Theta(\tau,\, y)$, and for each $(\tau,\, y)$ the function $\phi \mapsto \mathfrak{X}(\tau,\, y,\, \phi)$ is assumed to be a periodic travelling wave. Whitham's approach to the Lagrangian reduction allows to derive easily the corresponding macroscopic model. For the KG chain we find two nonlinear conservation laws

$$\partial_\tau\theta(\tau,\, y) - \partial_y\omega(\tau,\, y) = 0, \qquad \partial_\tau S(\tau,\, y) + \partial_y g(\tau,\, y) = 0, \tag{75}$$

which are closed by the Gibbs equation $\mathrm{d}L = S\,\mathrm{d}\omega + g\,\mathrm{d}\theta$ and the equation of state $L = L(\theta,\, \omega)$, which provides the action of a travelling wave as a function of ω and θ. Moreover, it can be shown that (75) is a system of Hamiltonian PDEs.

The new feature appearing in this example is that the corresponding two-scale transformation depends on the modulated phase Θ, which in turn depends on the solution to the macroscopic equation. In other words, within Whitham's modulation theory we do not know the two-scale transformations a priori and this complicates the reduction of Lagrangian and Hamiltonian structures. Finally, the modulation theory for FPU chains leads to further complications, since the Galilean invariance of (34) causes a coupling between macroscopic waves and modulated oscillations, see [FV99, Her05, GHM06, DHR06, DH07].

References

[AM78] R. Abraham and J.E. Marsden, *Foundations of Mechanics*, 2. ed., Perseus books, Cambridge Massachusetts, 1978, Updated 1985 Printing.

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

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

[BG06] ______, *From discrete systems to continuous variational problems: an introduction*, Topics on concentration phenomena and problems with multiple scales (A. Braides and V.C. Piat, eds.), Springer, 2006, pp. 3–78.

[BLM06] T. Blesgen, S. Luckhaus, and L. Mugnai, *Discrete free energy functionals for elastic materials*, Analysis, Modeling and Simulation of Multiscale Problems (A. Mielke, ed.), Springer, 2006.

[CJLL06] D. Cohen, T. Jahnke, K. Lorenz, and Ch. Lubich, *Integrators for highly oscillatory Hamiltonian systems*, Analysis, Modeling and Simulation of Multiscale Problems (A. Mielke, ed.), Springer, 2006.

[Daf00] C.M. Dafermos, *Hyperbolic conservation laws in continuum physics*, Grundlehren d. mathem. Wissenschaften, vol. 325, Springer, Berlin, 2000.

[DH07] W. Dreyer and M. Herrmann, *Numerical experiments on the modulation theory for the nonlinear atomic chain*, Physica D (2007), in press, see `doi:10.1016/j.physd.2007.09.003`.

[DHM06] W. Dreyer, M. Herrmann, and A. Mielke, *Micro-macro transition for the atomic chain via Whitham's modulation equation*, Nonlinearity **19** (2006), no. 2, 471–500.

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

[DM98] P. Deift and T.-R. McLaughlin, *A continuum limit of the Toda lattice*, Mem. Americ. Math. Soc., vol. 131/624, American Mathematical Society, 1998.

[El05] G.A. El, *Resolution of a shock in hyperbolic systems modified by weak dispersion*, Chaos **15** (2005), 037103.

[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** (2000), no. 6-7, 1519–1540.

[FP99] G. Friesecke and R.L. Pego, *Solitary waves on FPU lattices. I. Qualitative properties, renormalization and continuum limit*, Comm. Math. Phys. **12** (1999), no. 6, 1601–1627.

[FT02] G. Friesecke and F. Theil, *Validity and failure of the Cauchy–Born hypothesis in a two–dimensional mass–spring lattice*, J. Nonlin. Sci. **12** (2002), 445–478.

[FV99] A.-M. Filip and S. Venakides, *Existence and modulation of traveling waves in particle chains*, Comm. Pure Appl. Math. **51** (1999), no. 6, 693–735.

[GHM06] J. Giannoulis, M. Herrmann, and A. Mielke, *Continuum descriptions for the dynamics in discrete lattices: derivation and justification*, Analysis, Modeling and Simulation of Multiscale Problems (A. Mielke, ed.), Springer, 2006.

[Gia06] J. Giannoulis, *Three-wave interaction in discrete lattices*, PAMM **6** (2006), 475–476.

[Gia08] ———, *Interaction of modulated pulses in nonlinear lattices*, in preparation, 2008.

[GKM96] Zh. Ge, H.-P. Kruse, and J.E. Marsden, *The limits of Hamiltonian structures in the three dimensional elasticity, shells, and rods*, J. Nonlinear Sci. **6** (1996), no. 1, 19–57.

[GKMS95] Zh. Ge, H.-P. Kruse, J.E. Marsden, and C. Scovel, *The convergence of Hamiltonian structures in the shallow water approximation*, Canad. Appl. Math. Quart. **3** (1995), no. 3, 277–302.

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

[GM06] ———, *Dispersive evolution of pulses in oscillator chains with general interaction potentials*, Discr. Cont. Dynam. Systems Ser. B **6** (2006), 493–523.

[GMS07] J. Giannoulis, A. Mielke, and Ch. Sparber, *Interaction of modulated pulses in the nonlinear Schrödinger equation with periodic potential*, WIAS preprint 1221, 2007.

[Her05] M. Herrmann, *Ein Mikro-Makro-Übergang für die nichtlineare atomare Kette mit Temperatur*, Phd thesis, Humboldt-Universität zu Berlin, 2005.

[HFM81] B.L. Holian, H. Flaschka, and D.W. McLaughlin, *Shock waves in the toda lattice: Analysis*, Phys. Rev. A **24** (1981), no. 5, 2595–2623.

[HLM94] M.H. Hays, C.D. Levermore, and P.D. Miller, *Macroscopic lattice dynamics*, Physica D **79** (1994), no. 1, 1–15.

[HLW02] E. Hairer, Ch. Lubich, and G. Wanner, *Geometric Numerical Integration*, Springer Series in Comp. Mathem., vol. 31, Springer, Berlin, 2002.

[Jar93] C. Jaryzynski, *Multiple-time-scale approach to ergodic adiabatic systems: Another look*, Phys. Rev. Lett. **71** (1993), no. 6, 839–842.

[Kam06] A.M. Kamchatnov, *Nonlinear periodic waves and their modulations*, World Scientific, Singapore, 2006.

[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** (1992), 85–91.

[Lax86] P.D. Lax, *On dispersive difference schemes*, Physica D **18** (11986), 250–254.

[Lax91] ———, *The zero dispersion limit, a deterministic analogue of turbulence*, Comm. Pure Appl. Math. **44** (1991), 1047–1056.

[LLV93] P.D. Lax, C.D. Levermore, and S. Venakides, *The generation and propagation of oscillations in dispersive initial value problems and their limiting behavior*, Important developments in soliton theory (A.S. Fokas and V.E. Zakharov, eds.), Springer, 1993, pp. 205–241.

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

[Mie06] ———, *Macroscopic behavior of microscopic oscillations in harmonic lattices via Wigner–Husimi transforms*, Arch. Ration. Mech. Anal. **181** (2006), no. 3, 401–448.

[Mie08] ———, *Weak-convergence methods for Hamiltonian multiscale problems*, Discr. Cont. Dynam. Systems Series A **20** (2008), no. 1, 53–79.

[NV05] A.I. Neishtadt and A.A. Vasiliev, *Destruction of adiabatic invariance at resonances in slow-fast hamiltonian systems*, preprint, 2005.

[Sch06] B. Schmidt, *A derivation of continuum nonlinear plate theory from atomistic models*, Multiscale Model. Simul. **5** (2006), no. 2, 664–694.

[SU07] G. Schneider and H. Uecker, *The mathematics of light pulses in dispersive media*, Jahresber. Deutsch. Math.-Verein. **109** (2007), no. 3, 139–161.

[SW00] G. Schneider and C.E. Wayne, *Counter-propagating waves on fluid surfaces and the continuum limit for the Fermi-Pasta-Ulam model*, International Conference on Differential Equations (K. Gröger, B. Fiedler, and J. Sprekels, eds.), World Scientific, 2000, pp. 390–403.

[The06] F. Theil, *A proof of crystallization in two dimensions*, Comm. Math. Phys. **262** (2006), no. 1, 209–236.

[TR99] N.P. Tretiakov and J.N. Teixeira Rabelo, *Fast driving: Effective equations of motion for classical systems*, Europhys. Lett. **48** (1999), no. 2, 143–149.

[Whi74] G.B. Whitham, *Linear and Nonlinear Waves*, Pure And Applied Mathematics, vol. 1237, Wiley Interscience, New York, 1974.