Hamiltonian Field Theory Close to the Wave Equation: From Fermi-Pasta-Ulam to Water Waves

Matteo Gallone and Antonio Ponno

Abstract In the present work we analyse the structure of the Hamiltonian field theory in the neighbourhood of the wave equation $q_{tt} = q_{xx}$. We show that, restricting to "graded" polynomial perturbations in q_x, p and their space derivatives of higher order, the local field theory is equivalent, in the sense of the Hamiltonian normal form, to that of the Korteweg-de Vries hierarchy of second order. Within this framework, we explain the connection between the theory of water waves and the Fermi-Pasta-Ulam system.

1 Introduction

The present work aims to treat the perturbations of a linear string in the framework of classical Hamiltonian field theory. The unperturbed base model we have in mind, the linear string, is described by the one-dimensional wave equation

$$q_{tt} = c^2 q_{xx}, \tag{1}$$

where $q : \mathbb{R} \times D \to \mathbb{R} : (t, x) \to q(t, x)$ is the unknown, real-valued field, and c is a real, positive parameter, the *speed of the wave*. As usual, partial derivatives are denoted by subscripts, i.e. $q_t = \partial_t q$, $q_x = \partial_x q$ and so on. Concerning the space domain D and the boundary conditions of the field q, we here focus on the 1-periodic case, namely $D = \mathbb{T} := \mathbb{R}/\mathbb{Z}$ (the L-periodic case, with $D = \mathbb{R}/(L\mathbb{Z})$, can always be reduced to the case $L = 1$ by rescaling both the independent variables to $x' = x/L$, $t' = t/L$).

M. Gallone (✉)
Dipartimento di Matematica "F. Enriques", Università degli Studi di Milano, Milano, Italy
e-mail: matteo.gallone@unimi.it

A. Ponno
Dipartimento di Matematica "T. Levi-Civita", Università degli Studi di Padova, Padova, Italy
e-mail: ponno@math.unipd.it

Solving Eq. (1), for any initial condition $q(0, x)$, $q_t(0, x)$ defined on $\mathbb{T}$ and regular enough, is a standard exercise in Fourier analysis. Indeed, substituting $q(t, x) = \sum_{k \in \mathbb{Z}} \hat{q}_k(t) e^{\iota 2\pi kx}$ (ι is the imaginary unit) into (1), one gets

$$\frac{d^2 \hat{q}_k}{dt^2} = -4\pi^2 c^2 k^2 \, \hat{q}_k \, ,$$

which implies $\hat{q}_k(t) = a_k e^{\iota \omega_k t} + \bar{a}_{-k} e^{-\iota \omega_k t}$, where the a_k are complex constants (the bar denoting complex conjugation), and

$$\omega_k := 2\pi c |k| \, ; \quad k \in \mathbb{Z} \, . \tag{2}$$

Observe that $\omega_{-k} = \omega_k$, which implies $\overline{\hat{q}}_k = \hat{q}_{-k}$, i.e. q is real. Relation (2) defines the dispersion relation of the wave equation. A given space periodic system, characterised by a certain dispersion relation $k \to \omega_k$, is said to be *non dispersive* if $\omega_{k+1} - \omega_k$ is piecewise constant, i.e. if ω_k is piecewise linear in k and this is clearly the case for the wave equation. One can check that the solution $q(t, x)$ of the problem is time periodic for all initial conditions, the period being $2\pi/\omega_1 = 1/c$.

It is almost impossible to give a complete account of physical phenomena that, to the first linear approximation, are described by the wave equation. Let us just mention, to have in mind concrete examples that we are going to analyse later, wave propagation in fluids and long-wavelength vibrations of interacting particle chains. In all these problems, the need to go beyond the first approximation arises, in order to take into account the effects of both nonlinearity and dispersion, typically determining whether some interesting form of energy localisation may take place, as opposed to a fast energy spreading among the degrees of freedom of the system. One is thus led to look for a general treatment of the possible perturbations of Eq. (1) regardless of the specific physical problem giving rise to it. This in turn calls for the restriction to a mathematical context where the possible perturbations constitute a well-defined ordered class of objects. We do this within the framework of Hamiltonian field theory, at the price to exclude, among others, all the dissipative effects from the theory (no claim is made here about their irrelevance: the other way around. See, for example, the enlightening discussion made by Nekhoroshev in [33]). Moreover, we consider nonlinear and dispersive perturbations depending on q_x, p and their higher order derivatives, but not on q. Indeed, all systems made of interacting particles, such as solids, fluids and gasses, in absence of external forces, and on a sufficiently large space scale, are described by a certain wave equation at the linear level, with perturbations depending, in principle, only by the space derivatives of the field (and its momentum, possibly). This is due to the fact that interactions in matter depend on *differences* of coordinates, which in the continuum approximation corresponds to *derivatives*.

On the other hand, considering smooth perturbations of the wave equation depending on q (not only through derivatives) would be interesting as well. For example, as shown by Bambusi and Nekhoroshev and by Nekhoroshev [6, 7, 33], the smooth perturbations of the wave equation depending on q only (no derivatives)

give rise to very nice, long-lasting localisation phenomena. Whether be possible to include such a class of problems in our treatment, drawing meaningful conclusions, looks unclear, at present.

Although we decided to focus on one-dimensional systems, it is worth mentioning that the techniques presented here can be generalised to study problems in higher space dimension. In this case one can predict, for example, energy localisation for a certain class of anisotropic rectangular lattices [22].

The paper is organised as follows. In Sect. 2 we introduce the Hamiltonian formalism of classical field theory, at the end of which we provide an informal presentation of the main results. Section 3 contains the elements of perturbation theory framed in the more general context of Poisson systems, which is the one appropriate to our purposes. Section 4 contains the formal statements and proofs of the results. The application of such results to the FPU problem and to the water wave problem is treated in Sect. 5. Finally, a short list of open problems is provided in Sect. 6.

2 Outline of the Method and Results

2.1 Hamiltonian Field Theory

For the sake of completeness, we report here a short review on what is meant by *Hamiltonian field theory*. The reader is referred to the monographs [16, 25], and [32], for details and/or a more extensive treatment of the subject.

In Hamiltonian field theory the dynamical variables (e.g. coordinates and conjugate momenta) are points in a certain function space, the phase space of the system, and the observables, including the Hamiltonian, are functionals, admitting a density, defined on the phase space.

In order to specify the notations used below, let us first consider the space of smooth functions, or fields $u : \mathbb{T} \to \mathbb{R}$. A functional $F[u]$, with density $\mathscr{F}$ depending on x and on $u(x)$ and its derivatives up to a given order, is defined as

$$F[u] = \oint \mathscr{F}(x, u, u_x, u_{xx}, \dots) \, dx \, , \tag{3}$$

where here and in the sequel we make use of the short hand notation $\oint := \int_{\mathbb{T}}$. The functional derivative (or variational derivative) of F with respect to u, denoted by $\delta F / \delta u$, is defined by the relation

$$\delta F[u, \delta u] := \frac{d}{d\epsilon} F[u + \epsilon \delta u]\big|_{\epsilon=0} = \oint \frac{\delta F}{\delta u} \delta u \, dx \, , \tag{4}$$

for any smooth finite increment δu defined on $\mathbb{T}$. Through repeated integrations by parts and erasing the boundary terms one finds

$$\frac{\delta F}{\delta u} = \sum_{j \geq 0} (-1)^j \frac{d^j}{dx^j} \frac{\partial \mathscr{F}}{\partial (\partial_x^j u)} = \frac{\partial \mathscr{F}}{\partial u} - \frac{d}{dx} \frac{\partial \mathscr{F}}{\partial u_x} + \frac{d^2}{dx^2} \frac{\partial \mathscr{F}}{\partial u_{xx}} + \cdots , \qquad (5)$$

the sum above being finite if $\mathscr{F}$ is a polynomial in u and its derivatives up to a given finite order (as will be in our case). Relation (4) defines the Gateaux, or weak differential of the functional F at u with increment δu, which under further requirements coincide with the Fréchet, or strong differential of F; see e.g. [39]. The functional derivative is also referred to, in the mathematical literature, as the L_2-gradient of F with respect to u. Indeed, in the Hilbert space $L_2(\mathbb{T})$ of square integrable functions on $\mathbb{T}$, endowed with the usual scalar product $\langle f, g \rangle := \oint fg \, dx$, one can rewrite (4) as $\delta F = \langle \delta F / \delta u, \delta u \rangle := \langle \nabla F, \delta u \rangle$, identical in form to its finite-dimensional counterpart.

In the *Hamiltonian field theory* considered in the present paper, the phase space Γ of the system is the space of two components, smooth, real-valued fields $(q(x), p(x))$ defined on $\mathbb{T}$. The observables of the theory are the functionals $F : \Gamma \to \mathbb{R}$ admitting a density $\mathscr{F}$ which is a polynomial in $q(x)$, $p(x)$ and their space derivatives up to a finite order, with coefficients possibly depending on x. One then selects, among the observables, the Hamiltonian defining the given system, namely

$$H[q, p] := \oint \mathscr{H}(x, q, p, q_x, p_x, \dots) \, dx . \qquad (6)$$

The motion of the system, a certain curve $\gamma : [t_1, t_2] \ni t \mapsto (q, p)(t) \in \Gamma$, is then specified by a stationary action principle, as in the finite-dimensional case. Indeed, defining the *action functional $S[q, p]$* as

$$S[q, p] := \int_{t_1}^{t_2} \left[\langle p, q_t \rangle - H \right] dt = \int_{t_1}^{t_2} \oint [pq_t - \mathscr{H}] \, dt \, dx , \qquad (7)$$

one defines the actual motion of the system as the critical point of S in the space of smooth curves $(q(t, x), p(t, x))$ in Γ with fixed ends on the first component: $q(t_1, x) := q_1(x)$, $q(t_2, x) := q_2(x)$, q_1 and q_2 being two assigned fields on $\mathbb{T}$. The smooth increment curves $(\delta q, \delta p)(t)$ must then satisfy the condition $\delta q(t_1, x) = \delta q(t_2, x) = 0$. With the notation just introduced, and performing simple integrations by parts, one gets the differential δS of the action S, namely

$$\delta S = \int_{t_1}^{t_2} \oint \left[\left(q_t - \frac{\delta H}{\delta p} \right) \delta p - \left(p_t + \frac{\delta H}{\delta q} \right) \delta q \right] dt \, dx . \qquad (8)$$

This is zero for any increment $(\delta q, \delta p)(t)$ if and only if the following Hamilton equations hold:

$$q_t = \frac{\delta H}{\delta p} \; ; \; p_t = -\frac{\delta H}{\delta q} \; . \tag{9}$$

This is the Hamilton principle of stationary action in classical field theory.

In this work, we restrict our attention to scalar fields q and p defined on the (flat) unit circle $\mathbb{T}$. However, all the above construction and most of the results presented below can be extended to vector fields defined on any multi-dimensional space domain (not necessarily a torus).

Consider now a functional $F[q, p] := \oint \mathcal{F}(x, q, p, q_x, p_x, \ldots)dx$. Its time derivative along the solutions of the Hamilton equations (9) associated with H is computed by means of repeated integrations by parts with respect to x. The result can be written as $dF/dt = \{F, H\}_{q,p}$, where

$$\{F, H\}_{q,p} := \oint \left(\frac{\delta F}{\delta q} \frac{\delta H}{\delta p} - \frac{\delta F}{\delta p} \frac{\delta H}{\delta q} \right) dx := \langle \nabla F, \mathsf{J}_2 \nabla H \rangle \tag{10}$$

is the *Poisson bracket* of the functionals F and H. In the second definition above, $\mathsf{J}_2 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is the standard 2×2 symplectic matrix, $\nabla F = \begin{pmatrix} \delta F/\delta q \\ \delta F/\delta p \end{pmatrix}$ and the same for H. The product $\xi^T \mathsf{J}_2 \eta = \xi_1 \eta_2 - \xi_2 \eta_1$, for any pair of vectors $\xi, \eta \in \mathbb{R}^2$, defines the symplectic 2-form. The Poisson bracket (10) defines a bilinear, skew-symmetric product on the algebra of functionals defined on Γ, and it satisfies the Jacobi identity $\{\{F, G\}_{q,p}, H\}_{q,p} + \{\{G, H\}_{q,p}, F\}_{q,p} + \{\{H, F\}_{q,p}, G\}_{q,p} \equiv 0$ and the Leibniz rule $\{FG, H\}_{q,p} = F\{G, H\}_{q,p} + \{F, H\}_{q,p}G$ for any triple of functionals F, G, H. The algebra of functionals on Γ endowed with the Poisson bracket becomes a Poisson algebra and is typically referred to as the algebra of observables.

Remark 1 Given any skew-symmetric bilinear product on an algebra, the Jacobi identity characterises it as a Lie bracket. The latter, by further assuming the Leibniz rule, becomes a Poisson bracket (by definition). Thus, a Poisson algebra is a Lie algebra of Leibniz type.

The fundamental Poisson brackets of the Hamiltonian field theory on $\mathbb{T}$ are

$$\{q(x), p(y)\}_{q,p} = \delta(x - y) \; ; \; \{q(x), q(y)\}_{q,p} = \{p(x), p(y)\}_{q,p} = 0 \, , \tag{11}$$

where $\delta(x)$ is the Dirac delta distribution on $\mathbb{T}$. This is proved by considering the identity $\oint \delta(x - y)f(y)dy = f(x)$, valid for any continuous function on $\mathbb{T}$, from which $\delta f(x)/\delta f(y) = \delta(x - y)$ follows. As a consequence, the Hamilton equations (9) can be written in the form

$$q_t = \{q, H\}_{q,p} \; ; \; p_t = \{p, H\}_{q,p} \; . \tag{12}$$

2.2 Results: Informal Presentation

Within the Hamiltonian formalism just introduced, we study a well-defined class of problems, defined as follows. We introduce a "bookkeeping parameter" λ and give a weight λ^2 to both q_x and p, weighting any successive derivative ∂_x of them by λ. Defining $r := q_x$, this amounts to assume a "grading" (perturbative ordering of the dynamical variables and their derivatives) $r \sim p \ll r_x \sim p_x \ll r_{xx} \sim p_{xx} \dots$, and $(r_x)^2 \sim r^3$, where, in a loose notation, $\sim$ and $\ll$ mean "of the same order of" and "of an order smaller than", respectively. For the sake of simplicity, we assume the smooth density $\mathscr{H}$ of H to be a function of q_x, p and their space derivatives up to order four. Such a limitation is due to the fact that, in the present paper, we do not consider λ-expansions of the Hamiltonian H to degree higher than four, and with the chosen grading, derivatives of q_x and p of order higher than four enter the perturbative problem from degree five on (in λ). The parameter λ is formal: it is necessary to define the grading and to trace the perturbative ordering, and it can be set to one at the end of the computations.

Definition 1 The class of problems considered in the present work is defined by the family of Hamiltonians of the form

$$H_\lambda := \frac{1}{\lambda^4} \oint \mathscr{H}(\lambda^2 q_x, \lambda^2 p, \lambda^3 q_{xx}, \lambda^3 p_x, \dots, \lambda^6 q_{xxxxx}, \lambda^6 p_{xxxx})\, dx\,, \qquad (13)$$

with the condition

$$\left(\frac{\partial^2 \mathscr{H}}{\partial q_x^2}\bigg|_{\lambda=0}\right)\left(\frac{\partial^2 \mathscr{H}}{\partial p^2}\bigg|_{\lambda=0}\right) > 0\,. \qquad (14)$$

By Taylor expanding $\mathscr{H}$ in powers of λ, close to $\lambda = 0$, and assuming without loss of generality that $\mathscr{H}|_{(q,p)=0} = 0$, one gets a perturbative ordering of the Hamiltonian of the form

$$H_\lambda = H_0 + \lambda H_1 + \lambda^2 H_2 + \lambda^3 H_3 + \lambda^4 H_4 + \cdots\,. \qquad (15)$$

We here observe that the absence of a term proportional to $1/\lambda^2$ in the latter expansion is due to the conservation of the total momentum $\oint p\, dx$, which can be always set to zero.

The main results are now presented in an informal way, their precise statements and proofs being provided below. The condition (14), which characterises the elliptic nature of the fixed point $q = p = 0$, implies that there exists a canonical transformation bringing the unperturbed Hamiltonian H_0 into the standard wave form

$$K_0 := \oint \frac{p^2 + (q_x)^2}{2}\, dx\,, \qquad (16)$$

and leaving the perturbative expansion (15) unaltered. The equations of motion associated with the latter Hamiltonian are $q_t = p$, $p_t = q_{xx}$, i.e. in second-order form, the wave equation $q_{tt} = q_{xx}$.

Now, in terms of the variables $r := q_x$ and p, the expanded Hamiltonian (15) reads $K_0 + \lambda H_1 + \lambda^2 H_2 + \cdots$, where $K_0 = \frac{1}{2} \oint (p^2 + r^2) dx$, and the H_j are functionals whose density is a homogeneous polynomial of "grade" j in r, p and their derivatives. One then conveniently performs the change of field variables $(r, p) \mapsto (u, v)$ defined by $u = (r + p)/\sqrt{2}$, $v = (r - p)/\sqrt{2}$, in terms of which $K_0 = \frac{1}{2} \oint (u^2 + v^2) dx$, and its flow separates the left from right wave: $u_t = u_x$, $v_t = -v_x$, so that u and v are simply the left and right translation of the corresponding initial datum, respectively.

The key idea is now to decouple the left from the right dynamics to higher orders. To such an end, we build up an explicit transformation of the field variables

$$\mathscr{T}_\lambda : (u, v) \mapsto (\tilde{u}, \tilde{v}) ,$$

λ-close to the identity, which sets the Hamiltonian $H = K_0 + \lambda H_1 + \lambda^2 H_2 + \cdots$ (expressed in the (u, v) variables) into normal form to order $1 \leq s \leq 4$ with respect to K_0. This means, by definition, that $H \circ \mathscr{T}_\lambda^{-1} = K_0 + \lambda Z_1 + \lambda^2 Z_2 + \cdots$ is such that the Z_j are first integrals of K_0, for $1 \leq j \leq s$.

The results proved below are the following. In the general case, i.e. no further hypotheses being added to the Definition 1, we show that the normal form Hamiltonian to order $s = 2$ has the form $K_0 + \lambda^2 Z_2 + \cdots$, and the corresponding dynamics of the variables $\tilde{u}$, $\tilde{v}$ reads

$$\begin{cases} \tilde{u}_t = c_l \tilde{u}_x + a_l \kappa_3(\tilde{u}) + \cdots \\ \tilde{v}_t = -c_r \tilde{v}_x - a_r \kappa_3(\tilde{v}) + \cdots \end{cases} . \tag{17}$$

On the other hand, in certain relevant cases, such as the "mechanical" one, where $\mathscr{H} = p^2/2 + \mathscr{U}(q_x, q_{xx}, \ldots, q_{xxxxx})$, or that of the water waves, one has $H_1 = H_3 \equiv 0$. In such situations the normal form Hamiltonian to order $s = 4$ has the form $K_0 + \lambda^2 Z_2 + \lambda^4 Z_4 + \cdots$, whose associated dynamics reads

$$\begin{cases} \tilde{u}_t = c_l \tilde{u}_x + a_l \kappa_3(\tilde{u}) + b_l \kappa_5(\tilde{u}) + \cdots \\ \tilde{v}_t = -c_r \tilde{v}_x - a_r \kappa_3(\tilde{v}) - b_r \kappa_5(\tilde{v}) + \cdots \end{cases} . \tag{18}$$

In systems (17) and (18) $a_{l/r}$, $b_{l/r}$ and $c_{l/r}$ are certain constants (depending on the model, on the parameter λ and on the initial condition), whereas κ_3 and κ_5 are the vector fields of the first and second integral in the KdV hierarchy [1], namely

$$\kappa_3(w) = \gamma w w_x + w_{xxx} = \partial_x \frac{\delta I_3}{\delta w} , \tag{19}$$

$$\kappa_5(w) = \frac{5}{6} \gamma^2 w^2 w_x + \frac{10}{3} \gamma w_x w_{xx} + \frac{5}{3} \gamma w w_{xxx} + u_{xxxxx} = \partial_x \frac{\delta I_5}{\delta w} . \tag{20}$$

Here $\gamma \in \mathbb{R}$ is a parameter, whose value is explicitly determined by the first order normal form transformation, whereas the first two integrals I_3 and I_5 of the KdV hierarchy are given by

$$I_3 = \oint \left(\frac{\gamma}{6} w^3 - \frac{1}{2} (w_x)^2 \right) dx , \qquad (21)$$

$$I_5 = \oint \left(\frac{5\gamma^2}{72} w^4 + \frac{5\gamma}{12} w^2 w_{xx} + \frac{1}{2} (w_{xx})^2 \right) dx . \qquad (22)$$

The conclusion is that *both in the general and in the special case, the dynamics of the perturbed wave equation is integrable in the KdV hierarchy sense to the second perturbative order included.*

Remark 2 The standard Hamiltonian normal form construction to leading order always leads to (17). On the other hand, in order to get (18), the second step of Hamiltonian normalisation is not enough, in general. With the aid of Hamiltonian transformations, we generally succeed in decoupling equations of motion for the two independent variables to higher orders but, in general, this is not enough to conjugate the equations of motion to those of the KdV integrable hierarchy. It is remarkable that, at this point, each of the two decoupled equations of motion falls in a class that was analysed by Kodama [26, 29–31] (and whose results have been extended to equations on the torus in [23]). Without entering the details, which could deserve an entire work, the idea is the following. One starts from a PDE of the form

$$u_t = F(u) := F_0(u) + \lambda F_1(u) + \lambda^2 F_2(u) + O(\lambda^3) \qquad (23)$$

and one considers the effect of a change of variables $u \mapsto u + \lambda G(u)$. Denoting with $[\cdot, \cdot]$ the commutator of two vector fields, the effect of the transformation on the RHS of the PDE (23) is

$$F(u) \mapsto e^{\lambda[G, \cdot]} F(u) = F_0(u) + \lambda \big(F_1(u) + [G, F_0](u) \big)$$
$$+ \lambda^2 \left(F_2(u) + [G, F_1](u) + \frac{1}{2} [G, [G, F_0]](u) \right) + O(\lambda^3) . \qquad (24)$$

The latter conjugation of the vector field F holds in general, i.e. for any G. The Kodama transformation consists in making use of the natural grading of the KdV equation in order to choose a G consisting of a finite sum of monomials and satisfying two fundamental requirements. The first one is $[G, F_0] = 0$, which allows to leave F_1 in the KdV hierarchy, as it is given by the normal form construction. The second one consists just in "forcing" $F_2 + [G, F_1]$ to fit the KdV hierarchy, even though F_2 does not. This part of the theory is only sketched in the present review and we refer to [23, 26] for details.

Remark 3 The treatment of the general case to orders $s = 3$ and $s = 4$ requires three and four perturbative steps, respectively, and is currently in progress.

3 Abstract Setting: Perturbation Theory in Poisson Systems

In order to treat our problem, we need to frame our Hamiltonian field theory in the more general context of Poisson systems [32, 36]. Such a short digression is adapted to our present purposes and does not aim at any generality.

3.1 Poisson Formalism

Definition 2 Let Γ be the phase space of the system and let $\mathscr{A}(\Gamma)$ be the algebra of real-valued smooth functions defined on Γ. A binary application, or product, $\{\cdot, \cdot\} : \mathscr{A}(\Gamma) \times \mathscr{A}(\Gamma) \to \mathscr{A}(\Gamma)$ is called a *Poisson bracket* on Γ if it satisfies the following properties

(i)   Skew-symmetry: $\{F, G\} = -\{G, F\}$;
(ii) Left-linearity: $\{\alpha F + \beta G, H\} = \alpha\{F, H\} + \beta\{G, H\}$;
(iii) Jacobi identity: $\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0$;
(iv) Leibniz rule: $\{FG, H\} = F\{G, H\} + \{F, H\}G$,

$\forall F, G, H \in \mathscr{A}(\Gamma)$ and $\alpha, \beta \in \mathbb{R}$. The pair $(\mathscr{A}, \{\,,\,\})$ is called *Poisson algebra*.

Remark 4 The bracket $\{\cdot, \cdot\}_{q,p}$ defined in (10) satisfies axioms (i)-(iv) in the above definition. Thus, the axiomatic definition above contains both the usual Hamiltonian mechanics and the field theory (as well as quantum mechanics).

For the sake of concreteness, let us consider the case where Γ is the space of two components, smooth, real-valued fields $u(x) = (u_1(x), u_2(x))^T$ defined on $\mathbb{T}$ (what we show can be exported to the case of n components, complex-valued fields on a d-dimensional domain D).

By analogy with the standard case (10), a bilinear, skew-symmetric, Leibniz bracket on such a space is defined by the formula

$$\{F, G\}_J := \langle \nabla F, J\nabla G \rangle := \oint \sum_{i,j=1}^{2} \frac{\delta F[u]}{\delta u_i} J_{ij}[u] \frac{\delta G[u]}{\delta u_j} \, dx \,, \tag{25}$$

where $J_{ij}[u]$ is a tensor valued operator, skew-symmetric with respect to the L_2 scalar product $\langle\,,\,\rangle$, functionally dependent on u. Notice that with the choice $J = \mathsf{J}_2$, and denoting $u_1 = q$, $u_2 = p$, (25) coincides with (10). On the other hand, the bracket (25) does not satisfy the Jacobi identity (hypothesis (iii) above), in general. We state without proof the following Proposition [32], which characterises the Poisson brackets of the form (25).

Proposition 1 *The bracket (25) satisfies the Jacobi identity, so that it is a Poisson bracket, if and only if the skew-symmetric tensor $J[u]$ satisfies the Schouten identity*

$$\sum_{s=1}^{2} \left(J_{is} D_{u_s} J_{jk} + J_{js} D_{u_s} J_{ki} + J_{ks} D_{u_s} J_{ij} \right) = 0 \tag{26}$$

for all u and all $i, j, k = 1, 2$.

Here D_{u_s} denotes the weak partial derivative with respect to u_s, defined in the usual way:

$$\left(D_{u_s} f \right) h := \frac{d}{d\epsilon} f[u_s + \epsilon h]) \Big|_{\epsilon=0} , \tag{27}$$

for any f functionally dependent on u. Observe that, for example, $D_{u_1} u_1 = 1$, $D_{u_2} \partial_x u_2 = \partial_x$ and so on. Thus, any skew-symmetric tensor $J[u]$ satisfying the identity (26) is a *Poisson tensor*, i.e. it defines through (25) a Poisson bracket. An obvious but fundamental consequence of Proposition 1 is the following

Corollary 1 *Any skew-symmetric tensor J independent of u (i.e. constant on the phase space) is a Poisson tensor.*

Remark 5 One does not require $J[u]$ to be non-degenerate, so that J is allowed to have a nontrivial kernel. The functionals F such that $J\nabla F = 0$ are called *Casimir invariants* of the given Poisson structure and represent constants of motion for all Hamiltonian systems: $\{H, F\} = 0$ for any $H \in \mathscr{A}(\Gamma)$.

Within this framework, fixing a Hamiltonian $H[u]$ in the given Poisson algebra, the associated dynamics is defined in the usual way, namely

$$u_t = \{u, H\}_J = J\nabla_u H , \tag{28}$$

to be read by components, $\nabla_u H$ being the functional gradient of $H[u]$. Of course, any functional F evolves along the solutions of (28) according to $F_t = \{F, H\}_J$. Hamiltonian dynamical systems, in the generalised Poisson sense, have the form (28), which includes the standard (symplectic) case.

The fundamental feature of generalised Hamiltonian systems is their invariant character under any change of variables.

Proposition 2 *Any smooth change of variables $f : u \mapsto \tilde{u} = f[u]$ maps the Hamiltonian system $u_t = J\nabla_u H$ into the Hamiltonian system $\tilde{u}_t = \tilde{J}\nabla_{\tilde{u}}\tilde{H}$, where $\tilde{H} = H \circ f^{-1}$, whereas the transformed Poisson tensor $\tilde{J}$ is given by*

$$\tilde{J}[\tilde{u}] := (D_u f) J (D_u f)^T \Big|_{u=f^{-1}[\tilde{u}]} . \tag{29}$$

The corresponding Poisson brackets are related, for any $F, G \in \mathscr{A}(\Gamma)$, by

$$\{F, G\}_J \circ f^{-1} = \{F \circ f^{-1}, G \circ f^{-1}\}_{\tilde{J}} . \tag{30}$$

In the latter formula, D_u denotes the weak Jacobian of u, as defined in (27). The proof of the above Proposition is direct and not reported. The important point is the following: if J is a Poisson tensor, its transformed $\tilde{J}$ under any f is a Poisson tensor. Of course, the Hamilton equations are not invariant in form under f, which happens if and only if $\tilde{J} = J$. *Canonical transformations* are then defined as those transformations f leaving the Poisson tensor invariant. In order to check the canonicity of a transformation f, it is easier to make use of (30) which, with $J = \tilde{J}$, yields $\{F, G\}_J \circ f^{-1} = \{F \circ f^{-1}, G \circ f^{-1}\}_J$.

Remark 6 If $J = \mathsf{J}_2$, the transformation law (29), together with the canonicity condition $\tilde{J} = J$, yields the requirement that the Jacobian $D_u f$ be symplectic.

The equation of motion (28) can be rewritten as $u_t = \mathscr{L}_H u$, where the operator $\mathscr{L}_H \cdot = \{\cdot, G\}_J$, such that $\mathscr{L}_H F = \{F, H\}_J$ for any F, is the Lie derivative of F in the direction of the Hamiltonian vector field $J \nabla H$. One can then formally solve the equation by exponentiation, which defines the flow Φ_H^t of the system, namely

$$u(t) = e^{t\mathscr{L}_H} w := \Phi_H^t(w) , \tag{31}$$

where $w = u(0)$ is an arbitrary initial condition. Of course the exponential operator above is defined, as usual, by its formal series

$$e^{t\mathscr{L}_H} = 1 + t\mathscr{L}_H + \frac{t^2}{2}\mathscr{L}_H^2 + O(t^3) . \tag{32}$$

Now, since the evolution equation $F_t = \{F, H\}_J = \mathscr{L}_H F$ of any functional F is solved by $e^{t\mathscr{L}_H} F(w)$, which must equal $F[u(t)] = F[\Phi_H^t(w)]$ for any initial condition w, one gets the useful relation

$$e^{t\mathscr{L}_H} F = F \circ \Phi_H^t , \tag{33}$$

which is known as the exchange Lemma; we will make use of it below.

The Hamiltonian flow $\Phi_H^t : \Gamma \to \Gamma$ represents a one-parameter family of canonical transformations of Γ into itself (the family is a group if the flow is global).

Proposition 3 *For any t such that Φ_H^t exists, and any pair of functionals F and G, one has*

$$\{F, G\}_J \circ \Phi_H^t = \{F \circ \Phi_H^t, H \circ \Phi_H^t\}_J . \tag{34}$$

Proof Define $\Delta(t)$ the difference between the left and the right-hand side of (34), and observe that $\Delta(0) \equiv 0$. Making use of relation (33), and of the Jacobi identity, one gets $d\Delta(t)/dt = \{\Delta(t), H\}_J = \mathscr{L}_H \Delta(t)$, whose solution is $\Delta(t) = e^{t\mathscr{L}_H} \Delta(0) \equiv 0$.   □

Remark 7 In the above treatment, the Hamiltonian H is arbitrary. It follows that any functional G, regarded as a Hamiltonian, generates a one-parameter family of canonical transformations, which is given by its flow $\Phi_G^s = e^{s\mathscr{L}_G}$, where $\mathscr{L}_G = \{\ , G\}_J$. In the jargon, G is called the generating Hamiltonian, and $\mathscr{L}_G = d\Phi_G^s/ds|_{s=0}$ the generator of the transformation.

As a final point of this section, we state a simple version of the Nöther theorem in the Poisson framework.

Theorem 1 *If the Hamiltonian $H[u]$ is invariant with respect to the flow $e^{s\mathscr{L}_K}$ of generator $\mathscr{L}_K = \{\ , K\}_J$, i.e. $e^{s\mathscr{L}_K} H = H$ for any s close to zero, then $\{H, K\}_J = 0$.*

Proof The derivative of $e^{s\mathscr{L}_K} H = H$ with respect to s, at $s = 0$, gives the result.
                                                                                                   □

In the practice, one usually "sees" a certain symmetry of H, i.e. one is able to write down a certain transformation Ψ^s such that $\Psi^0 = 1$ and $H \circ \Psi^s = H$ for any s around zero. Then, if Ψ^s is a Hamiltonian flow, its generating Hamiltonian K is a constant of motion of the given system.

3.2 Perturbation Theory

The target of Hamiltonian perturbation theory, which goes back to Poincaré and Birkhoff, is the following. Given a Hamiltonian

$$H = H_0 + \lambda H_1 + \lambda^2 H_2 + O(\lambda^3) , \qquad (35)$$

formally ordered with respect to the small parameter λ, one looks for a canonical transformation, λ-close to the identity, erasing completely or in part the perturbation terms $H_{j\geq 1}$ up to a given order (possibly infinite, as in the KAM theory). As is well known, the complete removal of the perturbation terms, even to the first few orders, is not possible, in general. The best one can do is instead to find a canonical transformation setting H in normal form, according to the following definition.

Definition 3 The Hamiltonian $H_0 + \lambda Z_1 + \cdots + \lambda^n Z_n + O(\lambda^{n+1})$ is said to be in normal form to order $n \geq 1$ with respect to H_0 if $\mathscr{L}_{H_0} Z_j = \{Z_j, H_0\} = 0$ for any $j = 1, \ldots, n$.

Observe that $Z_j \equiv 0$ fits the normal form requirement, which means that the definition includes the possibility of complete removal of some perturbation terms.

The canonical transformation bringing the Hamiltonian (35) into normal form with respect to H_0, to order λ^2 included, is given by composing the flows of two unknown Hamiltonians G_1 and G_2, namely

$$u \mapsto \tilde{u} = e^{-\lambda^2 \mathscr{L}_2} e^{-\lambda \mathscr{L}_1} u \,, \tag{36}$$

where $\mathscr{L}_j := \mathscr{L}_{G_j}$, $j = 1, 2$. The *inverse* transformation maps the Hamiltonian (35) into

$$
\begin{aligned}
\tilde{H} = e^{\lambda^2 \mathscr{L}_2} e^{\lambda \mathscr{L}_1} H &= H_0 + \lambda \left(\mathscr{L}_1 H_0 + H_1 \right) + \\
&+ \lambda^2 \left(\mathscr{L}_2 H_0 + \mathscr{L}_1 H_1 + \frac{1}{2} \mathscr{L}_1^2 H_0 + H_2 \right) + O(\lambda^3) \,,
\end{aligned}
\tag{37}
$$

which is obtained by expanding the exponentials. The two generating Hamiltonians are then found by imposing that, according to the Definition 3, the quantities

$$
\begin{aligned}
Z_1 &:= H_1 + \mathscr{L}_1 H_0 \,, \\
Z_2 &:= \mathscr{L}_2 H_0 + \mathscr{L}_1 H_1 + \frac{1}{2} \mathscr{L}_1^2 H_0 + H_2
\end{aligned}
\tag{38}
$$

be first integrals of H_0. Observing that $\mathscr{L}_j H_0 = -\mathscr{L}_{H_0} G_j$, the latter two equations for the four unknowns Z_j and G_j, can be rewritten in the form

$$
\begin{aligned}
\mathscr{L}_{H_0} G_1 &:= H_1 - Z_1 \,, \\
\mathscr{L}_{H_0} G_2 &:= \mathscr{L}_1 H_1 + \frac{1}{2} \mathscr{L}_1^2 H_0 + H_2 - Z_2 \,.
\end{aligned}
\tag{39}
$$

These equations have one and the same structure, namely

$$\mathscr{L}_{H_0} G_j = S_j - Z_j \,, \quad (j = 1, 2) \tag{40}$$

with obvious definitions of the S_j.

Remark 8 Looking for a transformation to an arbitrary order n, one finds at any order $j = 1, \ldots, n$ an equation of the form (40), where S_j is a known quantity if all the equations up to order $j - 1$ have been solved.

Equation (40) is known as the *homological equation* of order j, which has to be solved determining the unknowns Z_j and G_j under the condition $\mathscr{L}_{H_0} Z_j = 0$.

In what follows we suppose that the flow $\Phi_{H_0}^s$ of H_0 is global (i.e. it exists for all $s \in \mathbb{R}$) and uniformly bounded with respect to s.

Definition 4 The time average of any F along the unperturbed flow of H_0 is denoted by

$$\langle F \rangle_0 := \lim_{t \to \infty} \frac{1}{t} \int_0^t F \circ \Phi_{H_0}^s \, ds \, . \tag{41}$$

If the flow of H_0 is τ-periodic, i.e. $\Phi_{H_0}^\tau = 1$, then $\langle F \rangle_0 = \frac{1}{\tau} \int_0^\tau F \circ \Phi_{H_0}^s ds$.

Lemma 1

$$\mathscr{L}_{H_0} \langle F \rangle_0 = 0 \, . \tag{42}$$

Proof Composing the left and right-hand side of (41) with the flow $\Phi_{H_0}^r$, one gets, on the right-hand side, $\lim \frac{1}{t} \int_0^t F \circ \Phi_{H_0}^{s+r} \, ds = \lim \frac{1}{t} \left(\int_r^0 + \int_0^t + \int_t^{t+r} \right) F \circ \Phi_{H_0}^a \, da = \lim \frac{1}{t} \int_0^t F \circ \Phi_{H_0}^a \, da$. Thus $\langle F \rangle_0 \circ \Phi_{H_0}^r = F$, which implies (42), and vice versa. $\qquad\square$

Lemma 2 *The solution of the homological equation (40) is given by*

$$Z_j = \langle S_j \rangle_0 \; ; \; G_j = \langle G_j \rangle_0 + \lim_{t \to \infty} \frac{1}{t} \int_0^t (s - t) e^{s \mathscr{L}_{H_0}} \left(S_j - \langle S_j \rangle_0 \right) ds \, . \tag{43}$$

If the flow of H_0 is τ-periodic, $G_j = \langle G_j \rangle_0 + \frac{1}{\tau} \int_0^\tau s \, e^{s \mathscr{L}_{H_0}} \left(S_j - \langle S_j \rangle_0 \right) ds$.

Proof Applying $e^{s \mathscr{L}_{H_0}}$ to Eq. (40), taking into account the invariance of Z_j (by the definition of normal form), and taking the time average, one gets the first of (43) in the limit. By the latter result, the homological equation becomes $\mathscr{L}_{H_0} G_j = S_j - \langle S_j \rangle_0$. Applying $(s - t) e^{s \mathscr{L}_{H_0}}$ to the latter equation and time averaging, one gets the second of (43) in the limit. $\qquad\square$

Remark 9 The generating Hamiltonians G_j solving the homological equation are defined up to their average along the flow of H_0, i.e. up to an arbitrary constant of motion of H_0. Thus, both the normal form Hamiltonian and the transformation bringing to it are not unique. In the sequel, we make the choice $\langle G_j \rangle_0 \equiv 0$.

Theorem 2 (Averaging Principle) *The canonical transformation*

$$u \mapsto \tilde{u} = e^{-\lambda^2 \mathscr{L}_2} e^{-\lambda \mathscr{L}_1} u \, ,$$

generated by

$$G_1 = \lim_{t\to\infty} \frac{1}{t} \int_0^t (s-t)e^{s\mathcal{L}_0} \left(H_1 - \langle H_1\rangle_0\right) ds \; ;$$

$$G_2 = \lim_{t\to\infty} \frac{1}{t} \int_0^t (s-t)e^{s\mathcal{L}_0} \left(S_2 - \langle S_2\rangle_0\right) ds \; ; \qquad (44)$$

$$S_2 := H_2 + \frac{1}{2}\{H_1, G_1\} + \frac{1}{2}\{\langle H_1\rangle_0, G_1\} \; ,$$

maps the perturbed Hamiltonian $H = H_0 + \lambda H_1 + \lambda^2 H_2 + O(\lambda^3)$ into the normal form $\tilde{H} = e^{\lambda^2 \mathcal{L}_2} e^{\lambda \mathcal{L}_1} H = H_0 + \lambda Z_1 + \lambda^2 Z_2 + O(\lambda^3)$, explicitly given by

$$\tilde{H} = H_0 + \lambda\langle H_1\rangle_0 + \lambda^2 \left(\langle H_2\rangle_0 + \frac{1}{2}\langle\{H_1, G_1\}\rangle_0 \right) + O(\lambda^3) \; . \qquad (45)$$

Proof By Lemma 2, solving the first of the homological equations (39) yields Z_1 and G_1. By substituting $\mathcal{L}_1 H_0 = Z_1 - H_1 = \langle H_1\rangle_0 - H_1$ into the right-hand side of the second of the homological equations (39), one gets the latter in the form $\mathcal{L}_{H_0} G_2 = S_2 - Z_2$, with S_2 as in (44). Solving by Lemma 2 again yields Z_2 and G_2. □

Remark 10 As a matter of fact, in order to get the normal form Hamiltonian (45), one does not need to compute G_2. This is a general fact: Z_{j+1} depends on $G_1, \ldots, G_j$.

4 Hamiltonian Field Theory Close to $q_{tt} = q_{xx}$

We now come back to our problem and solve it by applying all the tools introduced in the previous section.

Let us start by considering a Hamiltonian $H = \oint \mathcal{H} dx$, whose density $\mathcal{H}$ does not depend explicitly on t and x and is an analytic function of q_x, p and their spatial derivatives up to a certain finite order, in the neighbourhood of the origin. Since $\mathcal{H}$ is invariant under time, space and q translations, Theorem 1 (Nöther) applies.

Proposition 4 *$H = \oint \mathcal{H} dx$, $I = \oint q_x p \, dx$ and $P = \oint p \, dx$ are the three first integrals corresponding to the symmetries $t \to t + s$, $x \to x + s$ and $q \to q + s$, respectively. Moreover, $\{I, P\} = 0$, so that the three first integrals are in involution.*

Proof The conservation of H is obvious. The Hamilton equations for I at time s are: $q_s = q_x$ and $p_s = p_x$, whose solution is $q(t, x+s)$ and $p(t, x+s)$, clearly corresponding to the x-translation. The Hamilton equations for P are $q_s = 1$, $p_s = 0$, solved by $q(t, x) + s$ and $p(t, x)$, corresponding to the q-translation. Finally, observe that $\{I, P\}_{q,p} = \oint(\delta I/\delta q)(\delta P/\delta p)dx = -\oint p_x \, dx = 0$. □

Remark 11 One can always restrict the dynamics to the submanifold $P = \oint p \, dx = 0$ by the canonical transformation $q = q'$, $p = P + p'$.

For the sake of convenience, we repeat below the definition of the class of Hamiltonian functionals considered, with the appropriate grading.

Definition 5 The perturbative ordering of the Hamiltonian H is defined by the following scaling:

$$H_\lambda := \frac{1}{\lambda^4} \oint \mathcal{H}(\lambda^2 q_x, \lambda^2 p, \lambda^3 q_{xx}, \lambda^3 p_x, \ldots, \lambda^6 q_{xxxxx}, \lambda^6 p_{xxxx}) \, dx \, . \tag{46}$$

By Taylor expanding in powers of λ, close to $\lambda = 0$, assuming without loss of generality that $\mathcal{H}|_{(q,p)=0} = 0$, and taking into account Remark 11, one gets

$$H_\lambda = H_0 + \lambda H_1 + \lambda^2 H_2 + \lambda^3 H_3 + \lambda^4 H_4 + \cdots , \tag{47}$$

where

$$H_0 = \oint \frac{ap^2 + b(q_x)^2}{2} dx + cI \tag{48}$$

with a, b and c some constants and $I = \oint q_x p \, dx$;

$$H_1 = \oint d_1 q_x p_x \, dx \, ; \tag{49}$$

$$H_2 = \oint \left[e_1(q_x)^3 + e_2 p^3 + e_3(q_x)^2 p + e_4 q_x p^2 + e_5(q_{xx})^2 + \right.$$
$$\left. + \; e_6(p_x)^2 + e_7 q_{xx} p_x \right] dx \, ; \tag{50}$$

$$H_3 = \oint \left[f_1(q_x)^2 p_x + f_2 q_{xx} p^2 + f_3 q_{xx} p_{xx} \right] dx \, ; \tag{51}$$

$$H_4 = \oint \left[g_1(q_x)^4 + g_2 p^4 + g_3(q_x)^2 p^2 + g_4(q_x)^3 p + g_5 q_x p^3 + \right.$$
$$+ \; g_6(q_{xx})^2 q_x + g_7(q_{xx})^2 p + g_8(p_x)^2 q_x + g_9(p_x)^2 p +$$
$$+ \; g_{10} q_{xxx} p^2 + g_{11}(q_x)^2 p_{xx} + g_{12}(q_{xxx})^2 + g_{13}(p_{xx})^2 +$$
$$\left. + \; g_{14} q_{xxxx} p_x \right] dx \, , \tag{52}$$

and so on. Here $d_1, e_1, \ldots, g_{14}$ are given constants.

Remark 12 Since $\mathscr{H}$ is independent of x, the density of each H_j is independent of x. It follows that $\{I, H_j\} = 0$ for any $j \geq 0$.

Proposition 5 *If the constants $a := \partial^2 \mathscr{H}/\partial p^2|_0$ and $b := \partial^2 \mathscr{H}/\partial (q_x)^2|_0$ appearing in (48) are different from zero and have the same sign, there exists a time-dependent canonical transformation which brings the Hamiltonian H_0 in the canonical wave equation form $K_0 = \frac{1}{2}\oint[p^2 + (q_x)^2]dx$ and preserves the structure of the perturbations H_j to any order $j \geq 0$.*

Proof Let $a = \sigma|a|$ and $b = \sigma|b|$, with $\sigma = \pm 1$. One first performs the canonical rescaling $q = \sqrt{|a|}\, q'$, $p = \sqrt{|b|}\, p'$, $H = \sigma|ab|H'$, $t = \sigma t'$, which brings H_0 into $K_0 + c'I$, where $c' = \sigma c/\sqrt{|ab|}$. Then one performs the transformation $(q', p') = \Phi_{c'I}^t(q'', p'') = \Phi_I^{c't}(q'', p'')$, where Φ_I^t denotes the flow of $I = \oint q_x p\, dx$. The latter transformation is canonical and erases $c'I$. Clearly, both transformations do not change the structure of any H_j nor the value of the coefficients of the Hamiltonians $H_1, \ldots, H_4$. Observe that the flow of I is the left translation of (q, p), so that it is global and preserves the regularity of the initial condition. $\qquad\square$

Remark 13 Consider $K_0 + \lambda H_1 = \frac{1}{2}\oint[p^2 + (q_x)^2 + 2\lambda d_1 q_x p_x]dx$. Its Hamilton equations read

$$q_t = p - \lambda d_1 q_{xx} \ ; \ \ p_t = q_{xx} + \lambda d_1 p_{xx} \ .$$

Both q and p satisfy the linear Boussinesq equation

$$u_{tt} = u_{xx} + (\lambda d_1)^2 u_{xxxx} \ .$$

The condition on a and b in the Proposition 5 above identifies the elliptic fixed points in the given class of Hamiltonians. One is then left with the problem of simplifying the dynamics of $K_0 + \lambda H_1 + \lambda^2 H_2 + \cdots$. The perturbations to various order have the structure listed above and no further simplification can be made, in general. However, there is a relevant class of Hamiltonians that display a much simpler structure, namely the class of *mechanical Hamiltonians* of the form $\mathscr{H} = p^2/2 + \mathscr{U}$, where $\mathscr{U}$ depends only on q_x and its derivatives. Such Hamiltonians usually arise as the continuum limit of some lattice system, the notable case being just that of the vibrating string.

Proposition 6 *Suppose that $\mathscr{H} = p^2/2 + \mathscr{U}(q_x, q_{xx}, \ldots, q_{xxxxx})$. Then, if the condition $b := \partial^2 \mathscr{U}/\partial(q_x)^2|_0 > 0$ holds, H_0 can be brought in the canonical wave form K_0, $H_1 = H_3 \equiv 0$, and*

$$H_2 = \oint \left[\alpha_1(q_x)^3 + \alpha_2(q_{xx})^2\right] dx \ ;$$

$$H_4 = \oint \left[\beta_1 (q_x)^4 + \beta_2 (q_{xx})^2 q_x + \beta_3 (q_{xxx})^2 \right] dx .$$

Proof The momentum p cannot appear out of H_0, by definition. Notice that in this case there is no term proportional to I in H_0. $\square$

In the latter significant case one can obviously rename $H_2 \to H_1$ and $H_4 \to H_2$, $\lambda^2 \to \lambda$.

4.1 Traveling Waves

The equations of motion associated with $K_0 = \oint \frac{p^2 + (q_x^2)}{2} dx$ reduce to the wave equation for the field q:

$$q_t = p \; ; \; p_t = q_{xx} , \qquad \Longleftrightarrow \qquad q_{tt} = q_{xx} . \tag{53}$$

In order to simplify the analysis of perturbations of the wave equation, it is convenient to perform a change of variables that maps the functions (q, p) into the *Riemann invariants* (u, v):

$$u = \frac{q_x + p}{\sqrt{2}} \; ; \; v = \frac{q_x - p}{\sqrt{2}} . \tag{54}$$

The equations of motion for u and v are the left and right translation equation, respectively:

$$\begin{cases} u_t = u_x \\ v_t = -v_x \end{cases} . \tag{55}$$

Indeed, the solution of the above system corresponding to the initial condition $(u_0(x), v_0(x))$ is $(u_0(x + t), v_0(x - t))$, i.e. a rigid translation of the initial profiles. The flow of the wave equation, that is used to compute normal forms, is particularly manageable in these new variables, being a left translation for u and a right translation for v (at positive times).

The change of variables (54) is not canonical and it maps the standard Poisson tensor J_2 into the Gardner tensor [24]

$$J = \begin{pmatrix} \partial_x & 0 \\ 0 & -\partial_x \end{pmatrix} . \tag{56}$$

In particular, as can be checked, formula (29) for the transformation (54) reads

$$D_{q,p}(u,v) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} D^T_{q,p}(u,v) = \begin{pmatrix} \partial_x & 0 \\ 0 & -\partial_x \end{pmatrix}.$$

The Hamiltonian K_0, expressed in terms of (u,v), reads $K_0 = \oint \frac{u^2+v^2}{2}\,dx$, so that the translation equations for u and v are the Hamilton equations associated with K_0 in the Gardner structure.

The explicit expression of the Hamiltonians (48)–(52) in the (u,v) variables is:

$$K_0 = \oint \frac{u^2+v^2}{2}\,dx\,; \tag{57}$$

$$H_1 = \oint \frac{d_1}{\sqrt{|ab|}}uv_x\,dx\,; \tag{58}$$

$$
\begin{aligned}
H_2 = \oint \Bigg\{ \frac{1}{2^{3/2}} \Big[&\Big(\frac{e_1}{|b|^{3/2}} + \frac{e_2}{|a|^{3/2}} + \frac{e_3}{|b|\sqrt{|a|}} + \frac{e_4}{|a|\sqrt{|b|}} \Big)u^3 \\
+ &\Big(\frac{e_1}{|b|^{3/2}} - \frac{e_2}{|a|^{3/2}} - \frac{e_3}{|b|\sqrt{|a|}} + \frac{e_4}{|a|\sqrt{|b|}} \Big)v^3 \\
+ &\Big(\frac{3e_1}{|b|^{3/2}} - \frac{3e_2}{|a|^{3/2}} + \frac{e_3}{|b|\sqrt{|a|}} - \frac{e_4}{|a|\sqrt{|b|}} \Big)u^2 v \\
+ &\Big(\frac{3e_1}{|b|^{3/2}} + \frac{3e_2}{|a|^{3/2}} - \frac{e_3}{|b|\sqrt{|a|}} - \frac{e_4}{|a|\sqrt{|b|}} \Big)uv^2 \Big] + \\
+ \frac{1}{2}\Big[&\Big(\frac{e_5}{|b|} + \frac{e_6}{|a|} + \frac{e_7}{\sqrt{|ab|}} \Big)u_x^2 + \\
+ &\Big(\frac{e_5}{|b|} + \frac{e_6}{|a|} - \frac{e_7}{\sqrt{|ab|}} \Big)v_x^2 \Big] \Bigg\}\,dx\,.
\end{aligned}
\tag{59}
$$

4.2 The Generic Case

In order to perform a canonical transformation as stated in Proposition 2, one has to compute time averages, as required in Theorem 2. General formulas applying to the case of an unperturbed flow consisting of left/right translations are provided in the next lemma.

Lemma 3 *Suppose that f and g are continuous functions on $\mathbb{T}$. Then*

$$\oint \oint f(x \pm s)\,dx\,ds = \oint f(x)\,dx\,; \tag{60}$$

$$\oint \oint f(x \pm s)g(x \mp s)\,dx\,ds = \oint f(x)\,dx \oint g(y)\,dy\,; \tag{61}$$

$$\int_0^1 \oint s\, f(x \pm s) g(x \mp s)\, dx\, ds =$$

$$= \frac{1}{2} \oint f(x)\, dx \oint g(y)\, dy \pm \frac{1}{2} \oint g(x)\, \partial_x^{-1} f(x)\, dx \, , \quad (62)$$

where $\partial_x^{-1} f(x)$ denotes the unique primitive of f with zero average on $\mathbb{T}$.

Proof All these proofs consist of straightforward computations in Fourier space. First, we prove (61):

$$\oint \oint f(x \pm s) g(x \mp s)\, dx\, ds = \int_0^1 \int_0^1 \sum_{k,k' \in \mathbb{Z}} \hat{f}_k \hat{g}_{k'} e^{2\pi\iota k(x \pm s)} e^{2\pi\iota k'(x \mp s)}\, dx\, ds$$

$$= \sum_{k,k' \in \mathbb{Z}} \hat{f}_k \hat{g}_{k'} \delta_{k+k',0} \delta_{k-k',0} = \hat{f}_0 \hat{g}_0 \, .$$

From here, (60) follows by choosing $g = 1$. In order to prove (62), we Fourier transform the LHS:

$$\int_0^1 \oint s\, f(x \pm s) g(x \mp s)\, dx\, ds = \sum_{k \in \mathbb{Z}} \hat{f}_k \hat{g}_{-k} \int_0^1 s e^{\pm 4\pi\iota ks}\, ds \, .$$

It remains to notice that

$$\int_0^1 s e^{\pm 4\pi\iota ks}\, ds = \delta_{k,0} \int_0^1 s\, ds + (1 - \delta_{k,0}) \int_0^1 s e^{\pm 4\pi\iota ks}\, ds = \frac{1}{2} \delta_{k,0} \pm \frac{1}{2} \frac{1}{2\pi\iota k} (1 - \delta_{k,0})$$

and to recognise that $1/(2\pi\iota k)$ is the Fourier-multiplier corresponding to the operator ∂_x^{-1}. $\square$

Proposition 7 *There exists a (formal) near-to-identity, canonical transformation $(u, v) \mapsto (\tilde{u}, \tilde{v})$ mapping H_λ into*

$$\tilde{H}_\lambda = K_0 + \lambda^2 Z_2 + O(\lambda^3) \, , \quad (63)$$

where

$$K_0 = \oint \frac{\tilde{u}^2 + \tilde{v}^2}{2}\, dx \, ; \quad (64)$$

$$Z_2 = \oint \left\{ \frac{1}{2^{3/2}} \left[\left(\frac{e_1}{|b|^{3/2}} + \frac{e_2}{|a|^{3/2}} + \frac{e_3}{|b|\sqrt{|a|}} + \frac{e_4}{|a|\sqrt{|b|}} \right) \tilde{u}^3 + \right. \right.$$

$$+ \left(\frac{e_1}{|b|^{3/2}} - \frac{e_2}{|a|^{3/2}} - \frac{e_3}{|b|\sqrt{|a|}} + \frac{e_4}{|a|\sqrt{|b|}} \right) \tilde{v}^3 \right] +$$

$$+ \frac{1}{2} \left[\left(\frac{e_5}{|b|} + \frac{e_6}{|a|} + \frac{e_7}{\sqrt{|ab|}} - \frac{d_1^2}{2|ab|} \right) \tilde{u}_x^2 + \right.$$

$$\left. \left. + \left(\frac{e_5}{|b|} + \frac{e_6}{|a|} - \frac{e_7}{\sqrt{|ab|}} - \frac{d_1^2}{2|ab|} \right) \tilde{v}_x^2 \right] \right\} dx \,. \tag{65}$$

Proof *First perturbative step*: Using (45) and (61) one has $Z_1 = 0$:

$$Z_1 = \int_0^1 e^{s\mathcal{L}_{H_0}} H_1 \, ds$$

$$= \int_0^1 \oint \frac{d_1}{\sqrt{|ab|}} u(x+s) v_x(x-s) \, dx \, ds$$

$$\overset{(61)}{=} \frac{d_1}{\sqrt{|ab|}} \oint u(x) \, dx \oint v_y(y) \, dy = 0,$$

where in the last step we used that the v_y has zero average.

Additional term at second order: We need the expression of G_1 to compute Z_2. Using (45) and (62) we have

$$G_1 = \int_0^1 s e^{s\mathcal{L}_{H_0}} H_1 \, ds$$

$$= \frac{d_1}{\sqrt{|ab|}} \int_0^1 \oint s \, u(x+s) v_x(x-s) \, dx \, ds$$

$$\overset{(62)}{=} -\frac{d_1}{2\sqrt{|ab|}} \oint uv \, dx \,.$$

The computation of functional derivatives yields:

$$\frac{\delta G_1}{\delta u} = -\frac{d_1}{2\sqrt{|ab|}} v \,; \qquad \frac{\delta G_1}{\delta v} = -\frac{d_1}{2\sqrt{|ab|}} u \,;$$

$$\frac{\delta H_1}{\delta u} = \frac{d_1}{\sqrt{|ab|}} v_x \,; \qquad \frac{\delta H_1}{\delta v} = -\frac{d_1}{\sqrt{|ab|}} u_x \,,$$

and one finally obtains

$$\{H_1, G_1\} = \oint \left(\frac{\delta H_1}{\delta u} \partial_x \frac{\delta G_1}{\delta u} - \frac{\delta H_1}{\delta v} \partial_x \frac{\delta G_1}{\delta v} \right) dx$$

$$= -\frac{d_1^2}{2|ab|} \oint \left(v_x^2 + u_x^2 \right) dx .$$

Computation of the second-order normal form: Using (45), one has to time average (with respect to the unperturbed flow of K_0) the following expression:

$$H_2 + \frac{1}{2}\{H_1 - Z_1, G_1\} = \oint \left\{ \frac{1}{2^{3/2}} \left[\left(\frac{e_1}{|b|^{3/2}} + \frac{e_2}{|a|^{3/2}} + \frac{e_3}{|b|\sqrt{|a|}} + \frac{e_4}{|a|\sqrt{|b|}} \right) u^3 \right. \right.$$

$$+ \left(\frac{e_1}{|b|^{3/2}} - \frac{e_2}{|a|^{3/2}} - \frac{e_3}{|b|\sqrt{|a|}} + \frac{e_4}{|a|\sqrt{|b|}} \right) v^3 + \left(\frac{3e_1}{|b|^{3/2}} - \frac{3e_2}{|a|^{3/2}} + \frac{e_3}{|b|\sqrt{|a|}} - \frac{e_4}{|a|\sqrt{|b|}} \right) u^2 v$$

$$+ \left. \left(\frac{3e_1}{|b|^{3/2}} + \frac{3e_2}{|a|^{3/2}} - \frac{e_3}{|b|\sqrt{|a|}} - \frac{e_4}{|a|\sqrt{|b|}} \right) u v^2 \right] +$$

$$+ \left. \frac{1}{2} \left[\left(\frac{e_5}{|b|} + \frac{e_6}{|a|} + \frac{e_7}{\sqrt{|ab|}} - \frac{d_1^2}{2|ab|} \right) u_x^2 + \left(\frac{e_5}{|b|} + \frac{e_6}{|a|} - \frac{e_7}{\sqrt{|ab|}} - \frac{d_1^2}{2|ab|} \right) v_x^2 \right] \right\} dx .$$

As a consequence of (61) and under the assumption of $\oint u \, dx = \oint v \, dx = 0$:

$$\int_0^1 \oint u^2(x+s) v(x-s) \, dx \, ds = \left(\oint u^2(x) \, dx \right) \left(\oint v(x) \, dx \right) = 0 ;$$

$$\int_0^1 \oint u(x+s) v^2(x-s) \, dx \, ds = \left(\oint u(x) \, dx \right) \left(\oint v^2(x) \, dx \right) = 0 .$$

Moreover

$$\int_0^1 \oint u^3(x+s) \, dx \, ds = \oint u^3(x) \, dx ;$$

$$\int_0^1 \oint v^3(x+s) \, dx \, ds = \oint v^3(x) \, dx ;$$

$$\int_0^1 \oint u_x^2(x+s) \, dx \, ds = \oint u_x^2(x) \, dx ;$$

$$\int_0^1 \oint v_x^2(x+s) \, dx \, ds = \oint v_x^2(x) \, dx ,$$

and this completes the proof. $\square$

Remark 14 $\tilde{H}_\lambda$ is *always* the Hamiltonian of a pair of counter-propagating Korteweg-de Vries equations (up to a small remainder), i.e. its vector field $J\nabla\tilde{H}_\lambda$ is of the form (17). Such a result is somehow expected from, and in agreement with the existing results treating particular cases in the literature, among which those concerning the FPU problem (starting with the seminal work of Zabusky and Kruskal [37]) and the propagation of surface water waves (where the first deduction of the KdV equation goes back to Boussinesq [13]).

4.3 The Mechanical Case

For mechanical Hamiltonians of the form $\mathscr{H} = p^2/2 + \mathscr{U}$, where $\mathscr{U}$ depends on q_x and its derivatives, starting from Proposition 6 and repeating the analysis made in the general case, we perform the change of variables $(q, p) \mapsto (u, v)$, which yields

$$K_0 = \oint \frac{u^2 + v^2}{2} \, dx \,, \tag{66}$$

$$H_2 = \oint \left[\frac{\alpha_1}{2^{3/2}} \left(u^3 + 3u^2 v + 3uv^2 + v^3 \right) + \frac{\alpha_2}{2} \left((u_x)^2 + 2u_x v_x + (v_x)^2 \right) \right] dx \,, \tag{67}$$

$$H_4 = \oint \left\{ \beta_1 \left[\frac{u^4 + 4u^3 v + 6u^2 v^2 + 4uv^3 + v^4}{4} \right] \right.$$
$$+ \beta_2 \left[\frac{(u_x)^2 + 2u_x v_x + (v_x)^2}{2} \right] \frac{u+v}{\sqrt{2}} \tag{68}$$
$$\left. + \frac{\beta_3}{2} [(u_{xx})^2 + 2u_{xx} v_{xx} + (v_{xx})^2] \right\} dx \,.$$

Proposition 8 *There exists a (formal) near-to-identity, canonical transformation* $(u, v) \mapsto (\tilde{u}, \tilde{v})$ *mapping H_λ into*

$$\tilde{H}_\lambda = K_0 + \lambda^2 Z_2 + \lambda^4 Z_4 + O(\lambda^6) \,, \tag{69}$$

where

$$K_0 = \oint \frac{\tilde{u}^2 + \tilde{v}^2}{2} \, dx \,, \tag{70}$$

$$Z_2 = \oint \left[\frac{\alpha_1}{2^{3/2}} (\tilde{u}^3 + \tilde{v}^3) + \frac{\alpha_2}{2} (\tilde{u}_x^2 + \tilde{v}_x^2) \right] dx \,, \tag{71}$$

$$Z_4 = \oint \left\{ \left(\frac{\beta_1}{4} - \frac{9\alpha_1^2}{16} \right) (\tilde{u}^4 + \tilde{v}^4) + \left(\frac{\beta_2}{2^{3/2}} - \frac{3\alpha_1\alpha_2}{\sqrt{2}} \right) [\tilde{u}(\tilde{u}_x)^2 + \tilde{v}(\tilde{v}_x)^2] + \right.$$

$$+ \left(\frac{\beta_3}{2} - \frac{\alpha_2^2}{2} \right) [(\tilde{u}_{xx})^2 + (\tilde{v}_{xx})^2] \right\} dx + \left(\frac{3\beta_1}{2} - \frac{9\alpha_1^2}{2} \right) \langle \tilde{u}^2 \rangle \langle \tilde{v}^2 \rangle +$$

$$+ \frac{9\alpha_1^2}{16} \left(\langle \tilde{u}^2 \rangle^2 + \langle \tilde{v}^2 \rangle^2 \right).$$

$$(72)$$

Proof *First perturbative step:* using Proposition 2 we have

$$Z_2 = \int_0^1 e^{s\mathscr{L}_{H_0}} H_2 \, ds$$

$$\overset{(61)}{=} \oint \left\{ \frac{\alpha_1}{2^{3/2}} (u^3 + v^3) + \frac{\alpha_2}{2} [(u_x)^2 + (v_x)^2] \right\} dx$$

$$+ \frac{3\alpha_1}{2^{3/2}} \left(\langle u^2 \rangle \langle v \rangle + \langle u \rangle \langle v^2 \rangle \right);$$

here the last term vanishes because $\langle u \rangle = \oint u \, dx = 0$ and $\langle v \rangle = \oint v \, dx = 0$.
Generator of the first order transformation:

$$G_2 = \int_0^1 s e^{s\mathscr{L}_{H_0}} (H_2 - Z_2) \, ds$$

$$= \int_0^1 \oint s \left\{ \frac{3\alpha_1}{2^{3/2}} [u^2(x+s)v(x-s) + u(x+s)v^2(x-s)] \right.$$

$$\left. + \alpha_2 u_x(x+s)v_x(x-s) \right\} dx \, ds$$

$$\overset{(62)}{=} \frac{3\alpha_1}{2^{5/3}} \left(\langle u^2 \rangle \langle v \rangle + \langle u \rangle \langle v^2 \rangle \right) + \frac{3\alpha_1}{2^{5/2}} \left(\oint v^2 \partial_x^{-1} u \, dx + \oint v \partial_x^{-1} u^2 \, dx \right)$$

$$+ \frac{\alpha_2}{2} \oint u v_x \, dx$$

$$= \frac{3\alpha_1}{2^{5/2}} \left(\oint v^2 \partial_x^{-1} u \, dx + \oint v \partial_x^{-1} u^2 \, dx \right) + \frac{\alpha_2}{2} \oint u v_x \, dx,$$

where in the last step we used $\langle u \rangle = 0$ and $\langle v \rangle = 0$. Making use of the functional derivatives

$$\frac{\delta G_1}{\delta u} = \frac{3\alpha_1}{2^{3/2}} \left[-u \partial_x^{-1} v - \frac{1}{2} \partial_x^{-1} v^2 \right] + \frac{\alpha_2}{2} v_x;$$

$$\frac{\delta G_1}{\delta v} = \frac{3\alpha_1}{2^{3/2}} \left[\frac{1}{2} \partial_x^{-1} u^2 + v \partial_x^{-1} u \right] - \frac{\alpha_2}{2} u_x;$$

$$\frac{\delta(H_2 - Z_2)}{\delta u} = \frac{3\alpha_1}{2^{3/2}}\left[2uv + v^2\right] - \alpha_2 v_{xx};$$

$$\frac{\delta(H_2 - Z_2)}{\delta v} = \frac{3\alpha_1}{2^{3/2}}\left[u^2 + 2uv\right] - \alpha_2 u_{xx},$$

we can compute the Poisson bracket

$$\{H_2 - Z_2, G_2\} = \oint\left[\frac{\delta(H_2 - Z_2)}{\delta u}\partial_x\frac{\delta G_2}{\delta u} - \frac{\delta(H_2 - Z_2)}{\delta v}\partial_x\frac{\delta G_2}{\delta v}\right]dx.$$

Since we do not need its full expression, we can use (61) to simplify computations and consider only those terms that do not vanish after taking the average with respect to the flow of K_0. We obtain

$$\left\langle\{H_2 - Z_2, G_2\}\right\rangle_0 = \oint\left[-\frac{9\alpha_1^2}{16}(u^4 + v^4) - \frac{3\alpha_1\alpha_2}{\sqrt{2}}(u_x^2 u + v_x^2 v)+\right.$$
$$\left. -\frac{\alpha_2^2}{2}\left((u_{xx})^2 + (v_{xx})^2\right)\right]dx + \frac{9\alpha_1^2}{16}\left(\langle u^2\rangle^2 + \langle v^2\rangle^2\right)$$
$$- \frac{9\alpha_1^2}{2}\langle u^2\rangle\langle v^2\rangle,$$

whereas

$$\langle H_4\rangle_0 = \oint\left\{\frac{\beta_1}{4}(u^4 + v^4) + \frac{\beta_2}{2^{3/2}}\left[u(u_x)^2 + v(v_x)^2\right] + \frac{\beta_3}{2}\left[(u_{xx})^2 + (v_{xx})^2\right]\right\}dx$$
$$+ \frac{3\beta_1}{2}\langle u^2\rangle\langle v^2\rangle.$$

Summing the right-hand sides of the two previous equations we get

$$Z_4 = \oint\left\{\left(\frac{\beta_1}{4} - \frac{9\alpha_1^2}{16}\right)(u^4 + v^4) + \left(\frac{\beta_2}{2^{3/2}} - \frac{3\alpha_1\alpha_2}{\sqrt{2}}\right)\left[u(u_x)^2 + v(v_x)^2\right]+\right.$$
$$\left.+\left(\frac{\beta_3}{2} - \frac{\alpha_2^2}{2}\right)\left[(u_{xx})^2 + (v_{xx})^2\right]\right\}dx + \left(\frac{3\beta_1}{2} - \frac{9\alpha_1^2}{2}\right)\langle u^2\rangle\langle v^2\rangle$$
$$+ \frac{9\alpha_1^2}{16}\left(\langle u^2\rangle^2 + \langle v^2\rangle^2\right).$$

<div align="right">□</div>

Here, as in the generic case, Z_2 is in the KdV hierarchy, i.e. the vector field $J\nabla(\tilde{K}_0 + \lambda^2 Z_2)$ has the form of the right-hand side of (17). On the other hand, Z_4 is not, in general, in the KdV hierarchy: the two components of its vector field $J\nabla Z_4$ are not proportional to κ_5 (as defined in (20)), which is due to the impossibility to fit

all the required constraints on its parameters, in general. However, it is still possible to get a dynamics within the KdV hierarchy to order λ^4 by applying the Kodama normalisation procedure to the vector field $J\nabla(\tilde{K}_0 + \lambda^2 Z_2 + \lambda^4 Z_4)$. Although such a normalisation is noncanonical, in principle, it actually yields a system of equations in the form (18). Neglecting the remainder, these equations turn out to be Hamiltonian a fortiori, with the correct Gardner-Poisson tensor (56). The deep reason behind this fact is far from being deeply understood, at present.

Concrete examples are discussed in the next Sect. 5, where we also provide an explicit example of Kodama transformation.

5 Applications

5.1 The Fermi-Pasta-Ulam Problem

The Fermi-Pasta-Ulam (FPU) chain consists of N identical (unit) masses connected by nonlinear springs to their nearest neighbours. The dynamics is generated by the Hamiltonian

$$H = \sum_{j\in\mathbb{Z}_N} \left[\frac{p_j^2}{2} + \phi(q_{j+1} - q_j) \right] , \tag{73}$$

where $\mathbb{Z}_N := \mathbb{Z}/(N\mathbb{Z})$, and ϕ is the potential

$$\phi(z) := \frac{z^2}{2} + \alpha\frac{z^3}{3} + \beta\frac{z^4}{4} + O(z^5) , \tag{74}$$

and $\alpha, \beta, \ldots$ are the parameters measuring the strength of the nonlinear terms. One usually refers to the *α-model* if α is the only non-zero parameter; to the *β-model* if β is the only non-zero parameter; to the *$\alpha + \beta$-model* if both α and β are non-zero, and to the *generalised FPU model* if the lowest degree of the nonlinearity is greater than or equal 5.

When all the parameters in the nonlinearity are set to zero, the Hamiltonian (73) reduces to that of a *harmonic chain*, where particles interact through linear forces only. The latter system is integrable in the sense of Liouville, and the Hamiltonian is diagonalised by the (discrete) Fourier transform

$$p_j = \frac{1}{\sqrt{2N}} \sum_{k=-N}^{N} \hat{p}_k e^{\imath\pi\frac{jk}{N}} , \tag{75}$$

and similarly for q_j. The integrals of motion are the energies of the Fourier modes

$$E_k = \frac{|\hat{p}_k|^2 + \omega_k^2|\hat{q}_k|^2}{2} , \qquad k = -N, \ldots, N-1 , \tag{76}$$

where $\omega_k := 2\left|\sin\left(\frac{k\pi}{2N}\right)\right|$ are the proper frequencies of oscillation. Observe that $E_k = E_{-k}$, for all k.

The nonlinear model (73) was introduced by Fermi, Pasta and Ulam (FPU), supported by Tsingou [18], with the purpose of analysing its thermalisation process. The authors expected that the interaction between the Fourier modes due to the nonlinear terms, and the consequent energy sharing between them, would have brought the system to reach the thermal equilibrium on a short time scale. In particular, as a detector of thermal equilibrium, they expected to observe the "equipartition of energy", i.e. a final state of the system where, on time average, all Fourier energies have almost the same value, i.e. $E_k \simeq E/N$, where E is the total energy. Their numerical simulations showed instead a completely different scenario: by initially exciting the lowest frequency mode ($k = 1$), within their available computation time, energy sharing was observed to effectively take place only among the first few modes and, instead of a continuous trend to equipartition, the dynamics showed an almost recurrent behaviour. The first explanation of the latter phenomenon goes back to Zabusky and Kruskal [37], who approximated the traveling wave dynamics of the system by the KdV equation, and based their argument on the recurrent behaviour of its solitons. On the Hamiltonian side, the first correct computation of the resonant normal form of the lattice system, in action angle-variables, is due to [35]. Such a construction was only later recognised to include that of Zabusky and Kruskal [8, 10].

Nowadays, it is well known that a key role in the explanation of the FPU phenomenon, or paradox, is played by the *integrability* of the resonant normal form either of the lattice system or of its infinite-dimensional approximation (we refer to [4, 19] and the references therein). Indeed, the KdV equation admits a complete set of (infinitely many) integrals of motion, whose conservation prevents a fast energy sharing among the Fourier modes. Moreover, the preservation of the analyticity of the initial condition causes an exponential decay of its Fourier energies [28]. These two aspects resemble very much the observations in the FPU experiment.

In fact, the connection FPU-KdV can be made rigorous using the normal form construction of Theorem 2, as follows. As a preliminary step, we perform the canonical change of variables $(q, p) \mapsto (s, r)$ defined by the generating function

$$F(q, s) = \sum_{j \in \mathbb{Z}_N} s_j (q_j - q_{j+1}), \qquad (77)$$

which gives

$$
\begin{aligned}
r_j &= -\frac{\partial F}{\partial s_j} = q_{j+1} - q_j, \\
p_j &= \frac{\partial F}{\partial q_j} = s_j - s_{j-1}.
\end{aligned}
\qquad (78)
$$

In terms of the new variables (s, r) the Hamiltonian (73) reads

$$H = \sum_{j \in \mathbb{Z}_N} \left[\frac{(s_{j+1} - s_j)^2}{2} + \phi(r_j) \right] , \tag{79}$$

whose equations of motion are

$$
\begin{aligned}
\dot{s}_j &= \frac{\partial H}{\partial r_j} = \phi'(r_j) , \\
\dot{r}_j &= -\frac{\partial H}{\partial s_j} = s_{j+1} + s_{j-1} - 2s_j .
\end{aligned}
\tag{80}
$$

Remark 15 The periodicity of the q_j implies $\sum_{j \in \mathbb{Z}_N} r_j = 0$, whereas the periodicity of the s_j implies $\sum_{j \in \mathbb{Z}_N} p_j = 0$.

We now assume the existence of a pair of analytic functions $R, S : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ such that

$$
\begin{aligned}
s_j(t) &= \frac{\sqrt{\varepsilon}}{h} S(x, \tau)|_{x=hj, \tau=ht} \\
r_j(t) &= \sqrt{\varepsilon} R(x, \tau)|_{x=hj, \tau=ht} .
\end{aligned}
\qquad h := \frac{1}{N} .
\tag{81}
$$

Notice that the choice of the functions R and S is *not* unique. For example, one can add to them any linear combination of the form $\sum_{m \in \mathbb{Z}} c_m \sin(\pi m x / h)$, which vanishes at the lattice sites $x = hj$. Having in mind long-wavelength initial conditions, a natural choice consists in restricting R and S to the Fourier polynomials supported on the first few harmonics at $\tau = 0$, and in regarding the discrete system as a sampling of the continuous one at any $\tau > 0$. This is allowed by the following proposition.

Proposition 9 *Consider the Hamiltonian functional*

$$\mathcal{H}[S, R] = \oint \left[\frac{1}{\varepsilon} \phi(\sqrt{\varepsilon} R(x, \tau)) - \frac{1}{2} S(x, \tau) \Delta_h S(x, \tau) \right] dx , \tag{82}$$

where

$$\Delta_h := \frac{4}{h^2} \sinh^2 \left(\frac{h}{2} \partial_x \right) = \partial_x^2 + \frac{h^2}{12} \partial_x^4 + O(h^4) \tag{83}$$

is the discrete Laplacian. Then, its Hamilton equations restricted to the lattice coincide with the FPU equations (80).

Proof Considering (S, R) as a canonical pair coordinate-momentum, one has

$$S_\tau = \frac{\delta \mathscr{H}}{\delta R} = \frac{1}{\sqrt{\varepsilon}} \phi'(\sqrt{\varepsilon} R) \,,$$

$$R_\tau = -\frac{\delta \mathscr{H}}{\delta S} = \Delta_h S \,. \tag{84}$$

The latter equations, restricted to the lattice, i.e. to the points $x = hj$, coincide with those obtained by substituting (81) into (80). □

One thus embeds the dynamics of the FPU lattice (80) within that of the infinite-dimensional Hamiltonian system (84). The latter consists of a system of nonlinear dispersive Hamiltonian PDEs for any expansion to finite order of the discrete Laplacian (83). Moreover, making use of the latter expansion and of the explicit expression (74) of ϕ, one observes that the Hamiltonian (82) has the grading of Definition 5 with

$$\lambda \sim \sqrt{\varepsilon} \sim h^2 \,. \tag{85}$$

Let us see in which sense KdV equation allows us to explain rigorously, in the case of the α-chain, the FPU phenomenon, namely the fact that, if one low-frequency mode is initially excited, then the energy quickly flows to a small packet of modes whose energy, on time average, decreases exponentially with the mode index. The main result is conveniently formulated in terms of the quantities

$$\kappa := \frac{k}{N} \;\; ; \;\; \mathscr{E}_\kappa := \frac{E_k}{N} \,, \tag{86}$$

denoting the specific mode index (or wave number) and the corresponding specific energy, respectively. We are interested in the evolution of initial data supported on one harmonic mode of long wavelength, i.e. specific index $\kappa_0 = k_0/N \ll 1$.

Theorem 3 (Bambusi-Ponno [8]) *Consider an initial condition of the form*

$$\mathscr{E}_{\kappa_0}(0) = C_0 \mu^4, \qquad \mathscr{E}_\kappa(0) = 0, \qquad \forall \kappa \neq \kappa_0 \,, \tag{87}$$

where C_0 is any fixed constant and $\mu := \kappa_0 := k_0/N \ll 1$.

Then, for any fixed time T_f there exist positive constants μ^*, σ, C_1 and C_2 (dependent on C_0 and T_f) such that, for all κ, $\mu < \mu^*$ and $|t| \leq T_f/\mu^3$

(i)

$$\mathscr{E}_\kappa(t) \leq \mu^4 C_1 e^{-\sigma \kappa/\mu} \,; \tag{88}$$

(ii) *there exists a sequence of almost periodic functions $\{F_n(t)\}_{n\in\mathbb{N}}$ and an associated specific sequence*

$$F_\kappa = \mu^4 F_n \,, \quad if \quad \kappa = n\kappa_0 \,; \quad F_\kappa = 0 \quad otherwise \,, \tag{89}$$

such that

$$|\mathscr{E}_\kappa(t) - F_\kappa(t)| \;\leq\; C_2\mu^5 \,. \tag{90}$$

The proof of this theorem is based on the fact that a solution of the KdV equation with an analytic initial datum on the torus remains analytic for all times [28]. In particular, analyticity implies the exponential decay of Fourier coefficients, which in turn implies the exponential decay of the Fourier coefficients for the FPU system.

On the other hand, technical difficulties arise when comparing the dynamics of the discrete system with the dynamics of the continuous one, due to the contribution of the *singular* remainder of the discrete Laplacian that contains higher-order derivatives. The latter problem is overcome by a combined use of the analyticity property of the KdV flow, closeness to the identity of the canonical transformation and Grönwall lemma [8].

However, when comparing the above result with the numerical simulations and with the recent results on relating the FPU dynamics to that of the Toda lattice [5], one realises that it is not optimal: the time scale of closeness to the KdV dynamics numerically observed turns out to be longer than $t \sim \mu^{-3} \sim \varepsilon^{-3/4}$. In fact, there is an actual hope to improve the latter result which rests on the fact that the normal form of the FPU problem is in the *KdV hierarchy* not only to the first but also to the second perturbative order. Then, an extension of Theorem 3 could work with a second-order normal form transformation yielding the (presumably) optimal result of localisation of the Fourier spectrum on time scales $\sim \mu^5 \sim \varepsilon^{-5/4}$.

Within this context, we present below the normal form construction of the FPU problem, including the Kodama transformation.

Proposition 10 *The Hamiltonian* (82) *can be mapped into the normal form*

$$\tilde{H} = K_0 + Z_1 + Z_2 + \dots \,, \tag{91}$$

with

$$K_0 = \oint \frac{\tilde{u}^2 + \tilde{v}^2}{2}\, dx \tag{92}$$

$$Z_1 = \frac{h^2}{4!2} \oint \left[\frac{4\alpha\sqrt{2\varepsilon}}{h^2}(\tilde{u}^3 + \tilde{v}^3) + \tilde{u}\tilde{u}_{xx} + \tilde{v}\tilde{v}_{xx} \right] dx \tag{93}$$

$$Z_2 = \frac{3}{20}\frac{h^4}{(4!)^2} \oint \left[\left(\frac{\beta}{\alpha^2} - \frac{1}{2}\right)\frac{240\alpha^2\varepsilon}{h^4}(\tilde{u}^4 + \tilde{v}^4) + \frac{20\alpha\sqrt{2\varepsilon}}{h^2}(\tilde{u}^2\tilde{u}_{xx} + \tilde{v}^2\tilde{v}_{xx}) \right.$$

$$+(\tilde{u}_{xx})^2 + (\tilde{v}_{xx})^2\Big] dx + \Big(\tfrac{3\beta\varepsilon}{8} - \tfrac{\alpha^2\varepsilon}{4}\Big)\Big(\oint \tilde{u}^2\,dx\Big)\Big(\oint \tilde{v}^2\,dx\Big) +$$

$$+\frac{\alpha^2\varepsilon}{32}\Big(\langle\tilde{u}^2\rangle^2 + \langle\tilde{v}^2\rangle^2\Big) \tag{94}$$

Proof By introducing the Riemann variables

$$u := \frac{S_x + R}{\sqrt{2}} \;\; ; \;\; v := \frac{S_x - R}{\sqrt{2}} \,, \tag{95}$$

the result is actually a Corollary of Proposition 8, with the substitutions $\alpha_1 = \frac{\alpha\sqrt{\varepsilon}}{6\sqrt{2}}$, $\alpha_2 = -\frac{h^2}{4!2}$ and $\beta_1 = \frac{\beta\epsilon}{4}$. □

Remark 16 The equations of motion of $K_0 + Z_1$ are those of two counter-propagating KdV equations, i.e. of the form (17), for any α. On the other hand, the equations of motion of $K_0 + Z_1 + Z_2$ are not in the KdV hierarchy, i.e. in the form (18), unless the special condition $\beta = 5\alpha^2/6$ holds.

In order to bring the continuous FPU equations of motion into the KdV hierarchy form (18), one must look for a suitable Kodama transformation, as sketched in Remark 2 [20].

Proposition 11 *The Kodama transformation*

$$\tilde{u} = w + g(w) \;\; ; \;\; \tilde{v} = z + g(z) \,, \tag{96}$$

where

$$g(w) := \frac{h^2}{4!}\Big(\frac{7}{2} - \frac{9}{2}\frac{\beta}{\alpha^2}\Big)w_{xx} + \frac{\alpha\sqrt{\varepsilon}}{\sqrt{2}}\Big(\frac{13}{12} - \frac{3}{2}\frac{\beta}{\alpha^2}\Big)\Big(w^2 - \oint w^2\,dx\Big) +$$

$$-\frac{1}{6}\Big(w_x\partial_x^{-1}w - \oint w^2\,dx\Big) \,, \tag{97}$$

maps the equations of motion of the Hamiltonian normal form (91) into the integrable KdV form (18).

Proof The proof consists in a long, though direct computation. Details can be found in [20]. Observe that, according to the grading (85), $g \sim \lambda$, which does not affect the first order normal form. □

A natural question arises now, namely whether it is possible to construct a normal form transformation, including the Kodama procedure, conjugating the continuous FPU equations to those of the KdV hierarchy to perturbative orders higher than the second one. This is an open problem for initial data generically supported on lower

modes, but it has recently been addressed for initial data close to the traveling wave. In [23] it is proved that for almost-traveling waves, the conjugation to the third-order works only if the parameters correspond to a curve in the space of parameters containing the Toda lattice.

In general, it is expected that the FPU normal form is in the KdV hierarchy to a finite perturbative order, depending on the model. This is easily seen by considering the family of generalised FPU-systems [9] defined by a Hamiltonian of the form (73) with

$$\phi(z) = \frac{z^2}{2} + \frac{z^p}{p}, \qquad p \geq 3 .$$
(98)

Instead of fixing a model and going on with the perturbative order, we here consider how the first order normal form depends on the exponent p. The Hamiltonian (82) with potential (98) reads

$$H = \oint \left[\frac{R^2}{2} + \gamma \varepsilon^{\frac{p-2}{2}} \frac{R^p}{p} + \frac{1}{2}(S_x)^2 - \frac{h^2}{12}(S_{xx})^2 \right] dx + O(h^4) .$$
(99)

Passing to the (u, v) variables (95), one gets $H = K_0 + H_1$, where

$$K_0 = \oint \frac{u^2 + v^2}{2} dx ,$$
(100)

$$H_1 = \oint \left[\gamma \varepsilon^{\frac{p-2}{2}} \frac{(u - v)^p}{2^{p/2} p} - \frac{h^2}{24} \big((u_x)^2 + 2u_x v_x + (v_x)^2 \big) \right] dx .$$
(101)

Averaging H_1 (using (61)) one computes the normal form $\tilde{H} = K_0 + Z_1 + \cdots$ of the system, where

$$
\begin{aligned}
Z_1 = \langle H_1 \rangle = {} & \oint \frac{\gamma \varepsilon^{\frac{p-2}{2}}}{2^{p/2} p} \big(u^p + (-v)^p \big) - \frac{h^2}{24} \big((u_x)^2 + (v_x)^2 \big) \, dx \\
& + \frac{\gamma \varepsilon^{\frac{p-2}{2}}}{2^{p/2} p} \sum_{j=1}^{p-1} (-1)^j \binom{p}{j} \left(\oint u^{p-j} \, dx \right) \left(\oint v^j \, dx \right) .
\end{aligned}
$$
(102)

For $p = 3$ one finds that $K_0 + Z_1$ is the Hamiltonian of two uncoupled KdV equations, as expected. For $p = 4$, the so-called β-model, $K_0 + Z_1$ is the Hamiltonian of two uncoupled modified KdV (mKdV) equations. Thus, the first order normal form is integrable for $p = 3, 4$. On the other hand, for $p \geq 5$ the first order normal form Hamiltonian is that of two generalised, nonintegrable KdV equations, that are also nonlinearly coupled for $p \geq 6$. For this class of models the integrability of the normal form, and the consequent FPU phenomenon of energy localisation due to closeness to integrability, are lost to first order if the degree of

nonlinearity is high enough ($p \geq 5$). More than this, in [9] it is suggested that the blow-up of solutions characterising the nonintegrable KdV equations might play a relevant role in the problem.

As a last point, we stress that the method of infinite-dimensional perturbation theory allows to analyse the FPU system, treated in Proposition 10, in the singular limit $h \to 0$ with fixed, small specific energy ε. Such a limit is justified on the short term, where dispersion is expected to play a minor role with respect to nonlinearity, which explains why the normal modes start to effectively share their energy. Taking the limit $h \to 0$, at fixed ε, of the FPU terms (92), (93) and (94), one finds

$$H = K_0 + Z_1 + Z_2 + \dots \tag{103}$$

with

$$K_0 = \oint \frac{u^2 + v^2}{2}\, dx\,, \tag{104}$$

$$Z_1 = \frac{\alpha\sqrt{\varepsilon}}{2\sqrt{2}} \oint \frac{u^3 + v^3}{3}\, dx\,, \tag{105}$$

$$Z_2 = \left(\frac{\beta}{\alpha^2} - \frac{1}{2}\right)\frac{\alpha^2\varepsilon}{4} \oint \frac{u^4 + v^4}{4}\, dx\,. \tag{106}$$

The equations of motion associated with this normal form Hamiltonian consist of a pair of uncoupled, generalised Burgers equations, whose solution displays a gradient catastrophe at a finite shock time t_s. It has recently been proved that the Fourier energy spectrum of such a system displays a power law decay characterised by the universal exponent $-8/3$ exactly at t_s. Such a prediction fits very well the numerical spectrum of the FPU system [21]. Of course, the dynamics on times longer than t_s cannot be described in this limit and dispersive effects must be re-included, in agreement with the grading (85).

5.2 Water Waves

Consider an ideal fluid occupying, at rest, the domain

$$\Omega_{0,L} := \left\{(x, z) \in [0, L] \times \mathbb{R} : -h < z < 0\right\}, \tag{107}$$

with $L > 0$. We study the evolution of the free surface under the action of gravity, in the irrotational regime. Thus, given a periodic function $\eta : [0, L] \to \mathbb{R}$, we define the domain

$$\Omega_{\eta,L} := \left\{(x, z) \in [0, L] \times \mathbb{R} : -h < z < \eta(x)\right\}. \tag{108}$$

Irrotationality makes it possible to describe the velocity of the fluid u as gradient of a function called *velocity potential* by $u = \nabla\phi$. This problem admits a Hamiltonian formulation [14, 15, 38] and the conjugated variables are the wave profile $\eta(x)$ and the trace of the velocity potential at the free surface:

$$\psi(x) := \phi(x, \eta(x)). \tag{109}$$

The Hamiltonian of the system is

$$H(\eta, \psi) = \oint \left(\frac{1}{2}g\eta^2 + \frac{1}{2}\psi G(\eta)\psi\right) dx, \tag{110}$$

where $G(\eta)$ is the *Dirichlet-to-Neumann operator* defined as follows. Given a function $\psi(x)$ and consider the boundary value problem

$$\Delta\phi = 0, \qquad (x, z) \in \Omega_{\eta, L} \tag{111}$$

$$\phi_z\Big|_{z=-h} = 0 \tag{112}$$

$$\phi(0) = \phi(L) \tag{113}$$

$$\phi\Big|_{z=\eta(x)} = \psi \tag{114}$$

and let ϕ be its solution. Then

$$G(\eta)\psi := \sqrt{1 + \eta_x^2}\,\partial_n\phi\Big|_{z=\eta(x)} = (\phi_z - \eta_x\phi_x)\Big|_{z=\eta(x)}, \tag{115}$$

where ∂_n denotes the derivative in the direction normal to $z = \eta(x)$.

We are interested in solutions of the form

$$\eta(x) = \mu^2 h^3 \sqrt{2}\tilde{\eta}(\mu x), \qquad \psi(x) = \mu\sqrt{2ghh^2}\tilde{\psi}(\mu x), \qquad \mu = 1/L \ll 1, \tag{116}$$

that corresponds to a canonical transformation when rescaling time to

$$\tilde{t} = \frac{t}{\mu\sqrt{gh}} \tag{117}$$

and the physical space becomes the torus of unitary length.

Note that the dependence on η of the Dirichlet-to-Neumann operator causes the Hamiltonian (110) not to fall within the class of mechanical Hamiltonians of Sect. 4.3.

The small parameter of the theory is $\lambda = (h\mu)^2$. Expanding the Hamiltonian in λ one gets[1]

$$H = H_0 + \lambda H_1 + \lambda^2 H_2 + O(\lambda^3) \tag{118}$$

with H_0 being in the same form of (92) but with renamed variables:

$$H_0 = \oint \frac{\tilde{\eta}^2 + \tilde{\psi}_y^2}{2} \, dy \,, \tag{119}$$

$$H_1 = \frac{1}{2} \oint \left(-\frac{1}{3}\tilde{\psi}_{yy}^2 + \sqrt{2}\tilde{\eta}\tilde{\psi}_y^2 \right) dy \,, \tag{120}$$

$$H_2 = \frac{1}{2} \oint \left(\frac{2}{15}\tilde{\psi}_{yyy}^2 - \sqrt{2}\tilde{\eta}\tilde{\psi}_{yy}^2 \right) dy \,. \tag{121}$$

Note that, the Hamiltonian contains terms with the product of $\tilde{\eta}$ and $\tilde{\psi}$ and thus does not fit the definition of *mechanical Hamiltonian* given above. Anyway, as for the FPU problem, it is convenient to use characteristic variables (u, v) defined as

$$\tilde{\eta}(y, t) = \frac{u(y, t) + v(y, t)}{\sqrt{2}} \,, \tag{122}$$

$$\tilde{\psi}_y(y, t) = \frac{u(y, t) - v(y, t)}{\sqrt{2}} \tag{123}$$

we then obtain

$$K_0 = \oint \frac{u^2 + v^2}{2} \, dy \,, \tag{124}$$

$$H_1 = \oint \left(-\frac{1}{12}(u_y^2 + v_y^2) + \frac{u^3 + v^3}{4} + \frac{u_y v_y}{6} - \frac{u^2 v + u v^2}{4} \right) dy \,, \tag{125}$$

$$H_2 = \oint \left(\frac{1}{2}\frac{u_{yy}^2 + v_{yy}^2}{15} - \frac{1}{4}(u u_y^2 + v v_y^2) - \frac{1}{15}u_{yy}v_{yy} \right. \tag{126}$$

$$\left. -\frac{1}{4}(u v_y^2 - 2u u_y v_y + v u_y^2 - 2v u_y v_y) \right) dy \,. \tag{127}$$

---

[1] This step is far from being a trivial Taylor expansion as it involves the asymptotic expansion of the Dirichlet-to-Neumann operator (see [3] for details).

Applying the techniques of Theorem 2 one has

Proposition 12 *Within the normal form procedure outlined above, Hamiltonian* (118) *can be mapped into the normal form*

$$\tilde{H} = \tilde{K}_0 + \lambda Z_1 + \lambda^2 Z_2 + \dots \tag{128}$$

with

$$Z_1 = \oint \left[\frac{\tilde{u}^3 + \tilde{v}^3}{4} - \frac{1}{12}(\tilde{u}_y^2 + \tilde{v}_y^2) \right] dy, \tag{129}$$

$$Z_2 = \oint \left[\frac{1}{64}(\tilde{u}^4 + \tilde{v}^4) + \frac{7}{48}(\tilde{u}^2 \tilde{u}_{yy} + \tilde{v}^2 \tilde{v}_{yy}) + \frac{29}{720}(\tilde{u}_{yy}^2 + \tilde{v}_{yy}^2) \right] dy$$

$$+ \frac{1}{8}\langle \tilde{u}^2 \rangle \langle \tilde{v}^2 \rangle. \tag{130}$$

Proof This result is proved computing normal form Proposition 8.

First perturbative step: We use (61) to average H_1 along the flow of H_0 obtaining the expression for Z_1 in (129). The Hamiltonian generating the canonical transformation can be computed using (62):

$$G_1 = -\oint \left[\frac{1}{12} v_y u + \frac{1}{8} u^2 \partial_y^{-1} v - \frac{1}{8} v^2 \partial_y^{-1} u \right] dy.$$

We can therefore compute the L_2-gradient of G_1 and of $H_1 - Z_1$ obtaining

$$\frac{\delta G_1}{\delta u} = -\frac{1}{12} v_y - \frac{1}{4} u \partial_y^{-1} v - \frac{1}{8} \partial_y^{-1} v^2,$$

$$\frac{\delta G_1}{\delta v} = \frac{1}{12} u_y + \frac{1}{8} \partial_y^{-1} u^2 + \frac{1}{4} v \partial_y^{-1} u,$$

$$\frac{\delta (H_1 - Z_1)}{\delta u} = -\frac{1}{6} v_{yy} - \frac{1}{2} uv - \frac{1}{4} v^2,$$

$$\frac{\delta (H_1 - Z_1)}{\delta v} = -\frac{1}{6} u_{yy} - \frac{1}{2} uv - \frac{1}{4} u^2.$$

Second perturbative step: We use (61) to average H_2 and $\{H_1 - Z_1, G_1\}$ obtaining:

$$\langle \{H_1, G_1\} \rangle_0 = \frac{1}{8} \oint \left[\frac{1}{9}(u_{yy}^2 + v_{yy}^2) + \frac{1}{3}(u^2 u_{yy} + v^2 v_{yy}) + \frac{1}{4}(u^4 + v^4) \right] dy$$

$$\frac{1}{4}\langle u^2 \rangle \langle v^2 \rangle,$$

$$\langle H_2 \rangle_0 \;=\; \oint \left[\frac{u_{yy}^2 + v_{yy}^2}{30} + \frac{1}{8}\left(u^2 u_{yy} + v^2 v_{yy}\right) \right] dy \,.$$

We obtain $Z_2 = \langle H_2 \rangle_0 + \frac{1}{2}\langle\{H_1, G_1\}\rangle_0$ that is precisely (130). □

As for equations of the FPU lattice, these Hamiltonians are *not* in the Korteweg-de Vries hierarchy. Exactly as in the previous case, Kodama's theory solves the problem and with a close-to-identity change of variables maps the Hamiltonian into:

$$H_{\text{NF}}(u, v) \;=\; K_0(u) + \lambda K_1(u) + \lambda^2 c_2 K_2(u) + K_0(v) + \lambda K_1(v) + \lambda^2 c_2 K_2(v) \tag{131}$$

with c_2 being some explicit constant.

In case μ is a small free parameter not related to L and the water waves are studied on the whole real line (that is, $x \in \mathbb{R}$ and thus, imposing $\lim_{x \to \infty} \phi(x) = 0$ instead of $\phi(0) = \phi(L)$ in (113)), the following result holds

Theorem 4 (Bambusi [3]) *For any s', there exists $\lambda^* > 0$ and s, s'', s.t., if $0 < \lambda < \lambda^*$, then there exists a map $T_\lambda : B_1^s \to W^{s'',1} \times W^{s'',1}$, with the following properties*

(i) *$\sup_{(u,v) \in B_1^s} \|T_\lambda(u, v) - (u, v)\|_{W^{s'',1} \times W^{s'',1}} \leq C\lambda,$*

(ii) *Let I_λ be an interval containing the origin and $z(\cdot) = (u(\cdot), v(\cdot)) \in C^1(I_\lambda; B_1^s)$ be a solution of the Hamiltonian system (131) with $c_2 = \frac{299}{389}$ define*

$$z_a = (u_a, v_a) \;:=\; T_\lambda(u, v) \,. \tag{132}$$

Then there exists $R \in C^1(I_\lambda, W^{s',2} \times W^{s',2})$ s.t. one has

$$\dot{z}_a \;=\; J\nabla H(z_a(t)) + \lambda^3 R(t) \qquad \forall t \in I_\lambda \,, \tag{133}$$

where H is the Hamiltonian of water waves problem in the variables u and v.

An interesting non-trivial dynamical information one can obtain from this Theorem concerns the goodness of the approximation of the normal form dynamics. That is, for smooth enough initial data, it is possible to go back to the original non-scaled variables and to get the estimate on the wave profile

$$\sup_{|t| \leq T^*/\mu^3 \sqrt{gh}} \|\eta(t) - \eta_a(t)\|_{L^\infty} \leq C\mu^6 \,. \tag{134}$$

Note that the difference between wave profiles can be proved to be small only for times in which the second perturbative correction is negligible. Thus, as for the FPU system, an interesting open problem is the understanding which results can hold for larger time scales.

We are confident that these two results can be proved also in the periodic setting presented above.

6   Conclusions and Open Problems

In the framework of Hamiltonian field theory, the continuum limit of the FPU chain for long-wavelength excitations and the Hamiltonian of water waves belong to the same wider class of perturbations of the wave equation. This is not the case of other lattice models, such as the Klein-Gordon, for which one has to take into account the presence of the mass term.

Recently, the analysis of lattice model using the machinery of water waves has received a certain interest especially for systems in two spatial dimensions [22, 27] or for the analysis of higher-order normal forms for one-dimensional systems [23].

As a comparison, water waves are now a hot topic in research. The main goals in the field are results on well-posedness as well as regularity result for solutions or existence of quasi periodic or traveling wave solutions (see e.g. the recent results [2, 11, 12, 17]).

In this sense, many open questions remain open and can hopefully be addressed in the next future:

- The analysis at second order performed in Subsec. 5.1 does not allow us to conclude that the dynamics of the integrable system is close to the dynamics of the original system. Actually, it is known how to obtain a result on the dynamics, but only over times over which the effects of the second-order term is invisible. One of the open major problems is to understand how to go beyond the time scale of Theorem 3.
- From the point of view of statistical physics, the regime on which Theorem 3 is proved is not significant as the specific energy of the system $\varepsilon \sim 1/N^4$. The thermodynamic limit would require ε to be constant and independent of the size of the system. This is read, in terms of the normal form construction, as a zero-dispersion limit of the Korteweg-de Vries equation. It would be interesting to study the effect of this limit.
- Last, small attention has been given to the analysis of the FPU model when the dispersion is neglected (see [34]). An interesting question to address would be if Eqs. (103)–(106) can be used to explain some properties of the dynamics, especially for short time scales, low Fourier modes or in the regime of high specific energy.

References

1. Arnold, V.: Mathematical Methods of Classical Mechanics. Graduate text in Mathematics. Springer, Berlin (1997)
2. Baldi, P., Berti, M., Haus, E., Montalto, R.: Time quasi-periodic gravity water waves in finite depth. Invent. Math. **214**(2), 739–911 (2018)
3. Bambusi, D.: Hamiltonian studies on counter-propagating water waves. Water Waves **3**, 49–83 (2021)
4. Bambusi, D., Carati, A., Maiocchi, A., Maspero, A.: Some analytic results on the FPU paradox. Fields Inst. Commun. **75**, 235–254 (2015)
5. Bambusi, D. Maspero, A.: Birkhoff coordinates for the Toda Lattice in the limit of infinitely many particles with an application to FPU. J. Funct. Anal. **270**(5), 1818–1887 (2016)
6. Bambusi, D., Nekhoroshev, N.N.: A property of exponential stability in nonlinear wave equations near the fundamental linear mode. Phys. D **122**, 73–104 (1998)
7. Bambusi, D., Nekhoroshev, N.N.: Long time stability in perturbations of completely resonant PDE's. Acta Appl. Math. **70**, 1–22 (2002)
8. Bambusi, D., Ponno, A.: Commun. Math. Phys. **264**, 539–561 (2006)
9. Bambusi, D., Ponno, A.: Resonance, Metastability and Blow-up in FPU. In: Gallavotti, G. (ed.), The Fermi-Pasta-Ulam Problem. Springer Lecture Notes in Physics, vol. 728, 191–205 (2008)
10. Benettin, G., Livi, R., Ponno, A.: The Fermi-Pasta-Ulam problem: scaling laws vs. initial conditions. J. Stat. Phys. **135**, 873–893 (2009)
11. Berti, M., Franzoi, L., Maspero, A.: Pure gravity traveling quasi-periodic water waves with constant vorticity. arXiv:2101.12006
12. Berti, M., Franzoi, L., Maspero, A.: Traveling quasi-periodic water waves with constant vorticity. ARMA **240**(1), 99–202 (2021)
13. Boussinesq, J.: Essai sur la theorie des eaux courantes, Memoires presentes par divers savants l'Acad. des Sci. Inst. Nat. France, XXIII, 1877
14. Craig, W., Groves, M.D.: Hamiltonian long-wave approximations to the water-wave problem. Wave Motion **19**(4), 367–389 (1994)
15. Craig, W., Sulem, C.: Numerical simulation of gravity waves. J. Comput. Phys. **108**(1), 73–83 (1993)
16. Dubrovin, B.A., Novikov, S.P., Fomenko, A.T.: Modern Geometry—Methods and Applications, Part I. Springer, Berlin (1992)
17. Feola, R., Giuliani, F.: Quasi-periodic traveling waves on an infinitely deep fluid under gravity. Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. **31**(4), 901–916 (2020)
18. Fermi, E., Pasta, J., Ulam, S.: Studies of non linear problems Los-Alamos Internal Report, 1955 Document LA-1940 first published. In: Enrico Fermi Collected Papers, vol II, pp 977–988. The University of Chicago Press, Chicago, and Accademia Nazionale dei Lincei, Roma (1965)
19. Gallavotti, G. (ed.): The Fermi–Pasta–Ulam Problem: A Status Report. Lecture Notes in Physics, vol. 728. Springer, Berlin (2008)
20. Gallone, M.: Hydrodynamics of the Fermi-Pasta-Ulam model and its integrable aspects. Master Thesis (2015)
21. Gallone, M., Marian, M., Ponno, A., Ruffo, S.: Burgers turbulence in the Fermi-Pasta-Ulam-Tsingou chain. Phys. Rev. Lett. **129**, 114101 (2022)
22. Gallone, M., Pasquali, S.: Metastability phenomena in two-dimensional rectangular lattices with nearest-neighbour interaction. Nonlinearity **34** 4983 (2021)
23. Gallone, M., Ponno, A., Rink, B.: Korteweg-de Vries and Fermi-Pasta-Ulam-Tsingou: asymptotic integrability of quasi unidirectional waves. J. of Phys. A: Math Theor. **54**, 305701/1-29 (2021)
24. Gardner, C.S.: Korteweg-de Vries equation and generalizations. IV. The Korteweg-de Vries equation as a hamiltonian system. J. Math. Phys. **12**, 1548–1551 (1971)

25. Gelfand, I.M., Fomin, S.V.: Calculus of Variations. Dover, New York (2000)
26. Hiraoka, Y., Kodama, Y.: Normal form and solitons. In: Mikhailov, A.V. (ed.), Integrability. LNP, vol. 767, pp. 175–214. Springer, Berlin (2009)
27. Hristov, N., Pelinovsky, D.E.: Justification of the KP-II approximation in dynamics of two-dimensional FPU systems. Zeitschrift für angewandte Mathematik un Physik **73**, 213 (2022)
28. Kappeler, T., Pöschel, J.: On the periodic KdV equation in weighted Sobolev spaces. Ann. I. H. Poincaré – AN **26**, 841–853 (2009)
29. Kodama, Y.: Normal forms for weakly dispersive wave equations. Phys. Lett. A **112**(5), 193–196 (1985)
30. Kodama, Y.: Normal form and solitons. In: Topics in Soliton Theory and Exactly Solvable Nonlinear Equations (Oberwolfach, 1986), pp. 319–340. World Sci. Publishing, Singapore (1987)
31. Kodama, Y.: On solitary-wave interaction. Phys. Lett. A **123**(6), 276–282 (1987)
32. Marsden, J.E., Ratiu, T.S.: Introduction to Mechanics and Symmetry, 2nd edn. Springer, Berlin (1999)
33. Nekhoroshev, N.N.: Strong Stability of the Approximate Fundamental Mode of the Nonlinear String Equation, pp. 151–217. Trans. Moscow Math. Soc. (2002)
34. Poggi, P., Ruffo, S., Kantz, H.: Shock waves and time scales to reach equipartition in the Fermi-Pasta-Ulam model. Phys. Rev. E**52**, 307 (1995)
35. Shepelyansky, D.L.: Nonlinearity **10**, 1331 (1997)
36. Vaisman, I.: Lecture on the Geometry of Poisson Manifolds. Progress in Mathematics, vol. 118. Birkhäuser Verlag, Basel (1994)
37. Zabusky, N.J., Kruskal, M.D.: Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. Phys. Rev. Lett. **15**, 240–243 (1965)
38. Zakharov, V.E.: Stability of periodic waves of finite amplitude on the surface of a deep fluid. J. Appl. Mech. Tech. Phys. **9**, 190–194 (1968)
39. Zorn, M.: Derivatives and Fréchet differentials. Bull. Am. Math. Soc. **52**, 133–137 (1946)