Università degli studi di Padova

Dipartimento di Fisica e Astronomia “Galileo Galilei”
Dipartimento di Matematica

Tesi di Laurea magistrale in Fisica

Hydrodynamics of the FPU Problem and its Integrable Aspects

Relatore: Prof. Antonio Ponno

Laureando: Matteo Gallone

Anno Accademico 2014/2015

Contents

Introduction **iii**

1 The Fermi-Pasta-Ulam system **1**
1.1 The Fermi-Pasta-Ulam experiment 1
1.2 Connection with Toda and KdV equations 2
1.3 Physical meaning of the model 4

2 Hamiltonian systems **9**
2.1 Poisson structures 9
2.1.1 Canonical transformations 12
2.1.2 Integrability: Liouville 18
2.2 Isospectral method and Lax pairs 20
2.3 Some tools from functional analysis 23
2.3.1 Normed spaces 23
2.3.2 Differentiation on linear spaces 24
2.3.3 Anti-derivative operators 25
2.4 The Korteweg-de Vries equation 26
2.4.1 Conservation laws and integrals of the motion 27
2.4.2 Bi-Hamiltonian structure of KdV 30
2.4.3 KdV hierarchy 32
2.4.4 Generalized KdV equations 34
2.4.5 mKdV hierarchy 34

3 Perturbation theory **37**
3.1 Classical theory of perturbations 37
3.1.1 A first perturbative step: Lie method 38
3.2 Infinite-dimensional Hamiltonian perturbation theory 42
3.2.1 Hamiltonian normal form 42
3.2.2 Some semplifications for the second order 44
3.3 Normal form of vector fields 45
3.3.1 Transformation rule for vector fields 46

3.3.2 Continuous case . . . 48
3.3.3 Relation between normal form of vector field and Hamiltonian normal form . . . 49

4 Normal form of the $\alpha + \beta$-FPU system **51**
4.1 Extension to continuum . . . 51
4.2 Hamiltonian normal form of $\alpha + \beta$-FPU . . . 55
4.2.1 First order normal form . . . 58
4.2.2 Analyticity of the interpolating function . . . 59
4.2.3 Hamiltonian generating canonical transformation to first order . . . 61
4.2.4 Second order . . . 63
4.3 Integrability of the FPU hierarchy . . . 63
4.3.1 First order . . . 63
4.3.2 Second order . . . 64
4.4 Normalization of vector fields . . . 65
4.4.1 First perturbative step . . . 65
4.4.2 Relation between vector fields and Hamiltonian transformations . . . 67
4.5 Second order homological equation . . . 67
4.6 Canonicity of last transformation . . . 70

5 Normal form of the β-FPU system **73**
5.1 Hamiltonian first order normal form . . . 73
5.2 Hamiltonian second order normal form . . . 75
5.3 Integrability at second order . . . 75

A Symmetries of the continuous FPU **81**
A.1 Nöther theorem for Hamiltonian systems . . . 81
A.2 Translational invariance . . . 81

B Perturbative computations for the $\alpha + \beta$-model **83**
B.1 Exactness of the generating Hamiltonian at first order . . . 83
B.2 Normal form of second order perturbation . . . 85
B.3 Canonicity at first order: a direct calculation . . . 87

Bibliography **90**

Introduction

The physics we deal with in our everyday life consists mostly in thermodynamics, classical mechanics and electromagnetism. Even though from a fundamental point of view electromagnetism plays the most important role, from a phenomenological point of view one is more familiar with classical mechanics and thermodynamics. With these two theories one can describe a lot of phenomena we are very familiar with: falling bodies, collisions, heating or cooling…

All this picture works perfectly until one recognizes that every "big" system is formed by particles whose dynamics is described by classical mechanics. One thus expects that large systems described with this theory show thermodynamical behaviours. This is not true in general. For example integrable systems are very faraway from such a behaviour.

Some physicists in the first half of the XX century thought that generic small perturbations added to an integrable system could lead to a thermodynamical behaviour. In this direction, the Fermi-Pasta-Ulam work was the first numerical experiment aimed to check whether this hypothesis was true or not.

The outcome of the experiment was quite surprising and, to the eyes of Enrico Fermi, seemed a paradox to solve in order to build up a strong bridge between classical mechanics and thermodynamics.

Looking at the literature on the Fermi-Pasta-Ulam problem, one sees that a wide part of the works tries to explain the phenomenology or to determine the equipartition time. In this thesis we approached the problem from a different point of view. We tried to check how long the Fermi-Pasta-Ulam system displays an integrable dynamics because, up to the time-scale this is true, one cannot expect energy sharing between all the degrees of freedom. This point of view gives us a lower bound to the equipartition time because we do not know what happens during the time the system no longer behaves in an integrable way.[(1)] Actually one can expect both that the integrable behaviour lasts up equilibrium (if any), or that it ends up earlier and just a few integrals of motion survive up to the time the system thermalizes. We have some indications on what happens on longer time-scales from numerical experiments to which the result obtained in this thesis, agrees perfectly.

[(1)] We do not even know, actually, if such an equipartition time exists.

In the first chapter of this thesis we briefly resume the history of the FPU system, its connection between the two closest integrable systems (Toda and KdV) and we analyze the physical meaning of the model. In the second chapter we sum up some aspects of classical mechanics we will use in the following chapters.

In the third chapter we present the Hamiltonian theory of perturbations and the standard way to compute normal form of vector fields. The latter is the non-canonical and the most general way to construct normal forms which will be necessary in the last two chapters in order to map the FPU normal form into the KdV hierarchy.

The last two chapters are devoted to the computation of normal forms for the $\alpha + \beta$ and pure β models to second order. In these chapters we start from the system introduced in chapter 1 and we first compute the first order normal form to get the KdV (or the mKdV in the β model) and then we go beyond computing the normal form at second order. We then try to find another transformation to map our system into the KdV hierarchy. We succeeded in it for the $\alpha + \beta$ model but we failed for the pure β model. This is a proof of the integrability to second order for the $\alpha + \beta$ model and a signal for a possible integrability breaking in the pure β model.

CHAPTER 1

The Fermi-Pasta-Ulam system

1.1 The Fermi-Pasta-Ulam experiment

In the 1950s Enrico Fermi, John Pasta and Stanislaw Ulam conceived a series of numerical experiments aimed to enlighten the dynamical foundations of statistical mechanics. Following a very diffused idea, according to which every non-linearity in a physical system would cause an ergodic behaviour, they planned to simulate with a computer the dynamics of discrete systems with increasing complexity.

Starting from the simplest possible one, which they recognized to be a one dimensional chain of N oscillators, they looked at the time evolution of energies associated to Fourier modes. This is an indicator of the approach to thermal equilibrium since such a state is characterized by the equipartition of energy between them.

In a Hamiltonian framework we say that the system used in the experiment was a system of N particles with unitary mass and Hamiltonian

$$H = \sum_{n=1}^{N-1} \frac{p_n^2}{2} + \sum_{n=1}^{N} \left[\frac{1}{2}(q_{n-1} - q_n)^2 + \frac{\alpha}{3}(q_{n-1} - q_n)^3 + \frac{\beta}{4}(q_{n-1} - q_n)^4 \right]. \tag{1.1}$$

This system is today called the *Fermi-Pasta-Ulam (FPU) system.* Since the first numerical experiment ran with only the first Fourier mode initially excited, we refer today to *FPU initial data* as the ones on which only long wavelength Fourier modes are excited.

Fermi and co-workers expected to see Fourier modes exchanging energy until all of them had reached the same value (state of equipartition of energy), which is a signal of the thermodynamical equilibrium. The outcome of the numerical experiment wasn't this and, later, people referred to this as the *FPU paradox.*

In fact, instead of a gradual drift to the equipartition of energy, they saw a nearly periodic exchange of energy between the first few Fourier modes. In figure 1.1 one of their numerical results is reported.

The impact of the FPU work, 10 years later, stimulated a work of Zabusky and Kruskal [23] in which they explained the FPU paradox in terms of interaction of solitons

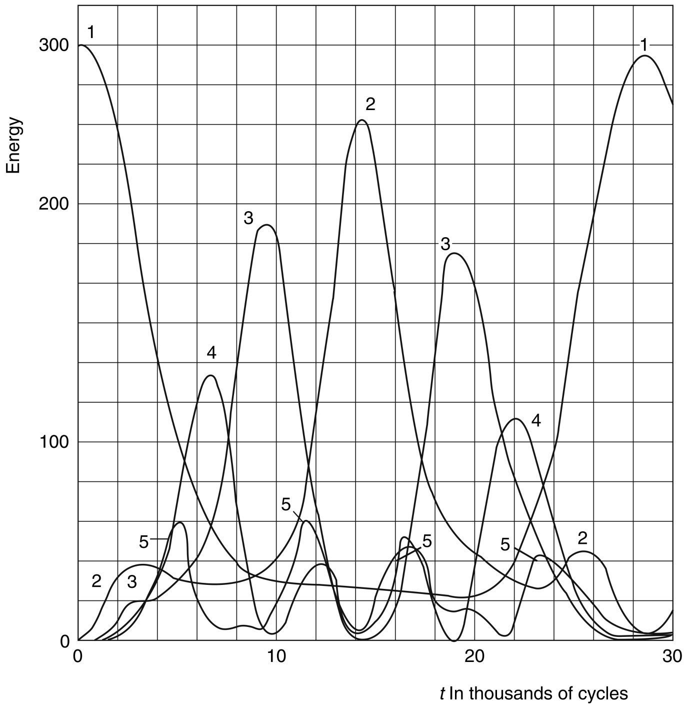

Figure 1.1: *Time dependence of energy per Fourier mode for the first Fourier mode excited as initial state. In this numerical experiment $N = 32$, $\alpha = 1/4$ and $\beta = 0$. (From [12]).*

of the KdV equation. In this work they conclude that if the system approaches the thermodynamical equilibrium, it happens on a time-scale longer than the one observed by FPU. Later it has been shown in some numerical experiments that the FPU system approaches the thermodynamical equilibrium on such a longer time. Nowadays, the phase of the dynamics during which the interaction between solitons takes place is usually referred to as the "metastable state", and such a scenario is usually referred to as the "metastable scenario".

The golden year for FPU problem can be considered today the 1982. Two works of this year are really important cornerstones for the understanding of the FPU paradox. The first one by Fucito et al. [14] introduced the so-called metastable scenario; the second one, by Ferguson et al. [11], explained the long time approach to the equilibrium of FPU showing that for the studies concerning the approach to equilibrium is better to regard the FPU model as a perturbed Toda chain instead of a perturbed harmonic chain.

Further developments are more recent and in [3, 6, 21] it is shown what is the relation between KdV and FPU proving that, when long wavelength modes are initially excited, the first order normal form of the FPU chain is precisely the KdV.

1.2 Connection with Toda and KdV equations

The Toda model is a model of one-dimensional crystal with non-linear nearest-neighbour interactions. In the case of periodic boundary conditions, denoting with $\mathbb{Z}_N = \mathbb{Z}/N$, its

Hamiltonian has the form

$$H(q,p) = \sum_{k\in\mathbb{Z}_N} \left(\frac{p_k^2}{2} + V_0(e^{\Lambda(q_{k+1}-q_k)} - 1 - \Lambda(q_{k+1}-q_k)) \right), \tag{1.2}$$

with V_0 and Λ free parameters of the potential. If one expands the potential in Taylor series, one gets

$$H(q,p) = \sum_{k\in\mathbb{Z}_N} \left[\frac{p_k^2}{2} + V_0\Lambda^2 \left(\frac{1}{2}(q_{k+1}-q_k)^2 + \frac{\Lambda}{6}(q_{k+1}-q_k)^3 + \dots \right) \right]. \tag{1.3}$$

One then sees that with the choice $\Lambda^2 = 2\alpha$ and $V_0 = \Lambda^{-2}$, one gets FPU system up to the cubic order. In other words the Toda Hamiltonian is tangent to $\alpha+\beta$-FPU Hamiltonian to third order. The Toda system is integrable and then, for the time FPU remains tangent to Toda, we can't expect FPU to show an approach to the thermodynamical equilibrium. In the metastable state what one observes is then the formation of the actions related to the tangent Toda system, the so-called "Toda packet".

Numerical experiments like [5, 7] show that, calling $\beta_T = 2\alpha^2/3$ the Taylor coefficient of the Toda potential expansion involving the fourth power in q, if $\beta \neq \beta_T$ the dynamics proceeds as follows. Fixed, for example $\varepsilon \sim 10^{-4}$, in a short time (say $\sim 10^3$) one observes the formation of the metastable state which coincides with the state of the pure Toda system. In a longer time (say $\sim 10^9$) it is possible to see energy to be shared between all the normal modes.

If $\beta = \beta_T$ the Toda packet lasts for more time. This is interpreted as indication of the stronger tangency between Toda and FPU.

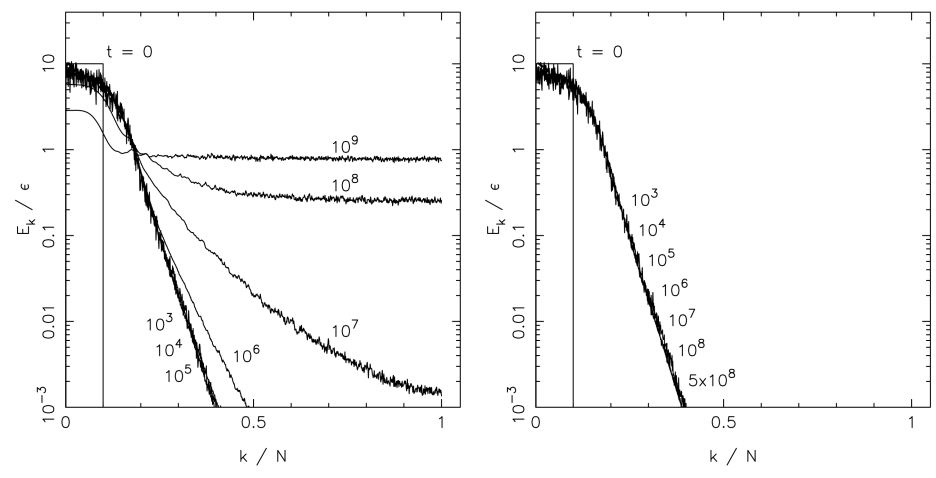

Figure 1.2: *Comparison between time evolution of energy per Fourier modes for Toda and FPU systems. On the left (FPU) one sees the formation of the Toda Packet for $t \sim 10^3$ and the freezing of Toda actions until $t \sim 10^5$. One thus sees the slow walk to equipartition. On the right (Toda) one sees the freezing of actions for Toda model standing, as a numerical proof of its integrable dynamics. (From [7])*

The connection with the Korteweg-de Vries (KdV) equation was first established by Zabusky and Kruskal looking at the recurrence of initial states in the soliton dynamics of the KdV. Looking at Figure 1.1 one can see that after nearly 29 thousands cycles the status of the FPU system resembles the initial one.

Solitons in the KdV have a similar behaviour and, after some time, they form back the initial status. Several years later a work by Ponno [20] explains the metastable packet in terms of the KdV. In particular, in that paper, it is explained its width ($\varepsilon^{1/4}$) and its formation time ($\varepsilon^{-3/4}$). Moreover, in two works by Bambusi and Ponno [3, 21] it is shown that the normal form of the FPU system consists in a couple of KdV equations.

1.3 Physical meaning of the model

The Hamiltonian of a wide part of physical systems can be written as a sum of a Kinetic part and a Potential one. Thus its general form is

$$\mathcal{H}(x, v) = K(v) + V(x), \tag{1.4}$$

where x are the coordinates and v the conjugate momenta. For a one dimensional system of N interacting particles with mass m, nearest neighbours interaction and periodic boundary conditions (1.4) takes the form

$$\mathcal{H}(x_n, v_n) = \sum_{n\in\mathbb{Z}_N} \left(\frac{v_n^2}{2m} + V(x_{n+1} - x_n) \right). \tag{1.5}$$

Our goal, in this section, is to show how it is possible to perform a *canonical* coordinate transformation[(1)] which maps the Hamiltonian $\mathcal{H}$ of equation (1.5) in the Hamiltonian H_{FPU} of the Fermi-Pasta-Ulam system,[(2)] which is

$$H_{FPU}(q_n, p_n) = \sum_{n\in\mathbb{Z}_N} \left(\frac{p_n^2}{2} + \phi_a(q_{n+1} - q_n) \right). \tag{1.6}$$

As notation we will use the (x_n, v_n) coordinates for the physical dimensional coordinates and the (q_n, p_n) for the dimensionless FPU coordinates.

To reach our aim we will use the following proposition which will be proved in the second chapter.

Proposition 1.3.1. *A coordinate transformation*

$$(q, p, H, T) \mapsto (x, v, \mathcal{H}, t) = (\alpha q, \beta p, \gamma H, \delta T) \tag{1.7}$$

is canonical, and then it preserves the Hamiltonian structure, if $\alpha\beta = \gamma\delta$.

We want to build this transformation and, moreover, we want that after the transformation, the new variabiles are dimensionless.

Let us say that the system described by the Hamiltonian $\mathcal{H}$ in (1.5) is in equilibrium when $x_n = x_{n,eq}$ and let's call a the quantity $a = x_{n,eq} - x_{n-1,eq}$. It is obvious that sufficient conditions for the system in the $\{x_{n,eq}, p_{n,eq}\}$ configuration is that

$$\begin{cases} p_{n,eq} = 0 & \forall n \in \mathbb{Z}_N \\ V'(x_{n+1,eq} - x_{n,eq}) - V'(x_{n,eq} - x_{n-1,eq}) = 0 & \forall n \in \mathbb{Z}_N. \end{cases} \tag{1.8}$$

(1) For the definition of canonical coordinate transformation see section 2.1.1.

(2) Connection with equation (1.1) is discussed below.

So if $\forall n \in \mathbb{Z}_N$, $x_{n+1} - x_n = a$, the system is in equilibrium. It is easy to calculate the $x_{n,eq}$ for a system of N identical particles with periodic boundary conditions described by Hamiltonian (1.5). These are

$$x_{n,eq} = n\frac{L}{N} \tag{1.9}$$

where L represents the length of the system and N the total number of particles. The relation (1.9) together with the definition of a implies $a = L/N$. If one thus fix the lenght L, a is the inverse of the total number of particles in the chain. A way to create a dimensionless variable is to set

$$x_n = aq_n + x_{n,eq} \tag{1.10}$$

so the q_n has the physical meaning of *displacement from equilibrium* of the n-th particle in units of a and a is the length scale of oscillations.

To get a dimensionless momentum we set

$$v_n = \frac{ma}{\tau}p_n \tag{1.11}$$

where τ is a parameter with the physical dimension of a time.

To get a dimensionless time T it is natural to set

$$t = \tau T. \tag{1.12}$$

Now, since we want a canonical transformation, (1.7) must hold, and then

$$\gamma = \frac{ma^2}{\tau^2}. \tag{1.13}$$

We thus get, for the Hamiltonian transformation,

$$\mathcal{H}(x_n(q), v_n(p)) = \frac{ma^2}{\tau^2}H(q_n, p_n). \tag{1.14}$$

Substituting equations (1.9)-(1.12) into (1.5) we get

$$\mathcal{H}(q_n, p_n) = \frac{ma^2}{\tau^2}\sum_{n\in\mathbb{Z}_N}\left(\frac{p_n^2}{2} + \frac{\tau^2}{ma^2}V(a + a(q_{n+1} - q_n))\right) \tag{1.15}$$

The transformed Hamiltonian emerges from the previous equation and it is

$$H(q_n, p_n) = \sum_{n\in\mathbb{Z}_N}\left(\frac{p_n^2}{2} + \frac{\tau^2}{ma^2}V(a + a(q_{n+1} - q_n))\right) \tag{1.16}$$

which is clearly in the form of (1.6).

We want to show that without any further assumption the power series in the neighbourhood of $q_{n+1} - q_n = 0$ of (1.16) is, up to fourth order, the one of the classical FPU model, which is

$$H_{FPU}(q_n, p_n) = \sum_{n\in\mathbb{Z}_N}\left(\frac{p_n^2}{2} + \frac{1}{2}(q_{n+1} - q_n)^2 + \frac{\alpha}{3}(q_{n+1} - q_n)^3 + \frac{\beta}{4}(q_{n+1} - q_n)^4\right). \tag{1.17}$$

This calculations will provide us also relations between the FPU coefficients of the series and the physical potential.

The series expansion of $V(a + a(q_{n+1} - q_n))$ around $q_{n+1} - 1_n = 0$ is

$$\begin{aligned} V(a + a(q_{n+1} - q_n)) = V(a) + aV'(a)(q_{n+1} - q_n) + \frac{a^2}{2}V''(a)(q_{n+1} - q_n)^2 + \\ + \frac{a^3}{3!}V'''(a)(q_{n+1} - q_n)^3 + \frac{a^4}{4!}V''''(a)(q_{n+1} - q_n)^4 + \dots \end{aligned} \tag{1.18}$$

Substituting (1.18) in (1.16) we get

$$\begin{aligned} H(q_n, p_n) = \sum_{n\in\mathbb{Z}_N} \Bigg[\frac{p_n^2}{2} + \frac{\tau^2}{ma^2} \Bigg(V(a) + aV'(a)(q_{n+1} - q_n) + \frac{a^2}{2}V''(a)(q_{n+1} - q_n)^2 + \\ + \frac{a^3}{3!}V'''(a)(q_{n+1} - q_n)^3 + \frac{a^4}{4!}V''''(a)(q_{n+1} - q_n)^4 + \dots \Bigg) \Bigg] . \end{aligned} \tag{1.19}$$

Notice that $V(a), V'(a), V''(a), \dots$ are constants.We can thus rescale the energies and put $V(a) = 0$ without changing Hamilton's equations. We note then

$$\sum_{n\in\mathbb{Z}_N} (q_{n+1} - q_n) = 0 \tag{1.20}$$

and so the only important terms of the Hamiltonian are

$$\begin{aligned} H(q_n, p_n) = \sum_{n\in\mathbb{Z}_N} \Bigg(\frac{p_n^2}{2} + \frac{1}{2}\frac{\tau^2 V''(a)}{m}(q_{n+1} - q_n)^2 + \\ + \frac{1}{3!}\frac{V'''(a)a\tau^2}{m}(q_{n+1} - q_n)^3 + \frac{1}{4!}\frac{V''''(a)\tau^2 a^2}{m}(q_{n+1} - q_n)^4 \Bigg) . \end{aligned} \tag{1.21}$$

We can now set the free parameter τ to get

$$\frac{\tau^2 V''(a)}{m} = 1, \tag{1.22}$$

thus getting, as natural time-scale for the physical problem, the quantity

$$\tau = \sqrt{\frac{m}{V''(a)}} \tag{1.23}$$

which is the very well known relation between the period of the small oscillations around the equilibrium in a and the potential energy V[(3)]. Using the relation (1.23) in (1.21) we note that if we set

$$\alpha = \frac{1}{2}\frac{V'''(a)a}{V''(a)} \qquad \beta = \frac{1}{6}\frac{V''''(a)a^2}{V''(a)} \tag{1.24}$$

we put the Hamiltonian in the form

$$H_{FPU}(q_n, p_n) = \sum_{n\in\mathbb{Z}_N} \left(\frac{p_n^2}{2} + \frac{1}{2}(q_{n+1} - q_n)^2 + \frac{1}{3}\alpha(q_{n+1} - q_n)^3 + \frac{1}{4}\beta(q_{n+1} - q_n)^4 \right) \tag{1.25}$$

(3) See for example [1] chap 5 par 22.

which is the classical Hamiltonian of the FPU problem[(4)].

If we transform (1.5) we get for $\phi_a(\xi)$ the expression

$$\phi_a(\xi) = \frac{\tau^2}{ma^2}\left(V(a + a\xi) - V(a) - V'(a)a\xi\right). \tag{1.26}$$

We conclude this section noting that all the calculations we will perform will be in the (q_n, p_n) variables. If we want to know something about the real physical problem we have to map back our results to the physical coordinates with (1.9) - (1.12), (1.14) where the parameter τ in (1.12) is given by the (1.23).

Potential	Expression	τ	α	β
Harmonic	$V = \frac{1}{2}\omega^2 x^2$			
Lennard-Jones	$V(x) = V_0\left[\left(\frac{a}{x}\right)^{12} - 2\left(\frac{a}{x}\right)^6\right]$	$\sim 3 \times 10^{-13}\,\mathrm{s}$	-10.5	61.83
Morse	$V = V_0\left(e^{-2(x/a-1)} - 2e^{-(x/a-1)}\right)$	$\sim 10^{-13}\,\mathrm{s}$	-1.5	1.17

Table 1.1: *Computation of the values for the $\alpha + \beta$-FPU parameters for common interaction potentials. The parameters in the potentials are chosen in order to have the deep of the potential well equal to V_0 and the equilibrium point at $x = a$. For the numerical evaluation of times we chose $a \sim 10^{-10}\,\mathrm{m}$, $m \sim 10^{-26}\,\mathrm{kg}$, $V_0 \sim 10^{-19}\,\mathrm{J}$ which are typical order of magnitude values for a crystal (see [16]).*

[(4)] But it is different from (1.1) because of the different boundary conditions.

CHAPTER 2

Hamiltonian systems

In this chapter we briefly recall some classical definitions and properties of Hamiltonian systems with a particular attention to the *Poisson environment* and its role in the studies of infinite dimensional systems. Poisson environment is a generalization of the classical symplectic environment in Hamiltonian mechanics.

Last section is devoted entirely to the Korteweg-de Vries equations which are integrable Hamiltonian systems involved in our study of FPU problem.

2.1 Poisson structures

Let us consider a manifold Γ as phase space of our physical system.(1) We then introduce a vector field on Γ, i.e. a section of its tangent bundle, $X(x) \in T\Gamma$ connected to the time evolution of our physical system by the differential equation

$$\dot{x} = X(x). \tag{2.1}$$

The difference between general dynamical systems and Hamiltonian systems is given by Poisson structures. This is a request on the phase space which, for the latters, is more rich than a purely naked differential manifold.

The same is true even if we are considering infinite-dimensional physical systems like wave propagation. In those cases the phase space Γ is an infinite-dimensional space of functions (which most of the time is a Hilbert space) and its time evolution is not given by an ODE but by a PDE:

$$u_t = X(u), \tag{2.2}$$

where we denote with pedices the partial differentiations.

We have now to state some definitions in order to be more precise on the ideas introduced above.

(1) Usually the phase space of a finite-dimensional classical mechanical system with configuration space on a manifold M is its cotangent bundle $\Gamma = T^*M$.

Definition 2.1.1 (Poisson bracket)**.** Let Γ be a (finite or infinite dimensional) manifold and $\mathscr{A}(\Gamma)$ the algebra of real valued smooth functions[(2)] defined on Γ. A function $\{\cdot,\cdot\}:\mathscr{A}(\Gamma)\times\mathscr{A}(\Gamma)\to\mathscr{A}(\Gamma)$ is called a *Poisson bracket* on Γ if it satisfies the following properties:

i) $\{F,G\}=-\{G,F\}$ (skew-symmetry);

ii) $\{\alpha F+\beta G,H\}=\alpha\{F,H\}+\beta\{G,H\}$ (left linearity);

iii) $\{F,\{G,H\}\}+\{G,\{H,F\}\}+\{H,\{F,G\}\}=0$ (Jacobi identity);

iv) $\{FG,H\}=F\{G,H\}+\{F,H\}G$ (left Leibniz rule);

$\forall F,G,H\in\mathscr{A}(\Gamma)$ and $\alpha,\beta\in\mathbb{R}$.

Observe that i) and ii) implies right linearity which means that actually Poisson brackets are bi-linear. Observe also that i) and iv) implies the right Leibniz rule.

Definition 2.1.2 (Poisson algebra)**.** Given a phase space Γ, the algebra of real valued smooth functions on Γ, $\mathscr{A}(\Gamma)$, and a Poisson bracket $\{\cdot,\cdot\}$ on $\mathscr{A}(\Gamma)$, we say that the pair $(\mathscr{A}(\Gamma),\{\cdot,\cdot\})$ is a *Poisson algebra* on Γ.

It is then straightforward to define a Poisson manifold as follows

Definition 2.1.3 (Poisson manifold)**.** Given a phase space Γ and a Poisson algebra on it $(\mathscr{A}(\Gamma),\{\cdot,\cdot\})$, we say that the structure $(\Gamma,\mathscr{A}(\Gamma),\{\cdot,\cdot\})$ is a *Poisson manifold.*

In the following proposition we show that the choice of a bi-linear Poisson bracket is equivalent to the choice of a skew-symmetric tensor on Γ. This tensor is called *Poisson tensor.*[(3)]

Proposition 2.1.4. *Given a Poisson algebra on a finite dimensional phase space, there exists a skew-symmetric tensor J such that*

$$\{F,G\}=\langle\nabla F,J\nabla G\rangle \tag{2.3}$$

and

$$\sum_s\left(J^{is}\frac{\partial J^{jk}}{\partial x^s}+J^{js}\frac{\partial J^{ki}}{\partial x^s}+J^{ks}\frac{\partial J^{ij}}{\partial x^s}\right)=0. \tag{2.4}$$

where we denoted with $\langle\cdot,\cdot\rangle$ the Euclidean scalar product.

The same proposition can be stated for infinite dimensional spaces. Here we focus on L_2 where it is enough to substitute the gradient with the L_2-gradient, if one wants the analogous of (2.3) and (2.4), the Euclidean scalar product with the L_2-scalar product and the last condition is

$$\sum_s\left(J^{is}\frac{DJ^{jk}}{DF^s}+J^{js}\frac{DJ^{ki}}{DF^s}+J^{ks}\frac{DJ^{ij}}{DF^s}\right)=0 \tag{2.5}$$

where D/DF is the weak derivative defined in subsection 2.3.2.

[(2)] Sometimes it is denoted by $\mathcal{C}^\infty(M,\mathbb{R})$.

[(3)] Sometimes it is also referred as *Poisson bivector.*

Remark 2.1.5. *Because of this duality between skew-symmetric tensors and Poisson brackets sometimes one refers to a Poisson manifold as a structure* $(\Gamma, \mathscr{A}(\Gamma), J)$ *where* J *is the Poisson tensor related to* $\{\cdot,\cdot\}$ *via* **`Proposition 2.1.4`**.

Remark 2.1.6. *Sometimes we deal with different Poisson structures on the same manifold. To each Poisson bracket is related a (different) Poisson tensor. If* J *and* $\tilde{J}$ *are two different Poisson tensors we refer to the related brackets as* $\{\cdot,\cdot\}_J$ *and* $\{\cdot,\cdot\}_{\tilde{J}}$. *We will use this when dealing with the bi-Hamiltonian structures in subsection 2.4.2.*

We are then ready to define what a Hamiltonian system is.

Definition 2.1.7 (Hamiltonian system)**.** Given a phase space Γ endowed with a Poisson structure (i.e. a Poisson manifold $(\Gamma, \mathscr{A}(\Gamma), \{\,\cdot\,,\,\cdot\,\})$), a *Hamiltonian system* on Γ is a dynamical system described by a differential equation (ODE or PDE) whose vector field has the form

$$X(x) = X_H(x) = \{x, H\} \tag{2.6}$$

where $H \in \mathscr{A}(\Gamma)$ is called *Hamiltonian* of the system.

In the definition above the Poisson bracket $\{x, H\}$ is defined by means of its components. More precisely we shall write

$$X^k(x) = \{x^k, H\} \tag{2.7}$$

where k isn't necessary a discrete index.

As a consequence of the proposition above we can replace Poisson brackets by an expression involving Poisson tensor:[(4)]

$$\begin{aligned} X_H(x) &= J(x)\nabla_x H(x) \qquad \text{finite-dimensional case} \\ X_H(u) &= J(u)\nabla_{L_2} H(u) \qquad \text{infinite-dimensional case} \end{aligned} \tag{2.8}$$

Poisson brackets are useful also in writing the time derivative of a function $f : \Gamma\times\mathbb{R} \to \mathbb{R}$ along a trajectory of a Hamiltonian system as the following proposition shows

Proposition 2.1.8. *Let* $f \in \mathcal{C}^\infty(\Gamma\times\mathbb{R}, \mathbb{R})$ *be a function, let* $(\Gamma, \mathscr{A}(\Gamma), \{\cdot,\cdot\})$ *be a Poisson manifold and* $H \in \mathscr{A}(\Gamma)$ *be a Hamiltonian function. Then if* $\gamma : \mathbb{R} \to \Gamma$ *is a trajectory of the Hamiltonian system with Hamiltonian* H *we have*

$$\frac{df(t,\gamma(t))}{dt} = \frac{\partial f}{\partial t}(t,\gamma(t)) + \{f(t,\gamma(t)), H(\gamma(t))\} \tag{2.9}$$

Proof. Using the chain rule we get[(5)]

$$\begin{aligned} \frac{df(t,\gamma(t))}{dt} &= \frac{\partial f}{\partial t}(t,\gamma(t)) + \langle \nabla f(t,\gamma(t)), \dot{\gamma}(t)\rangle \\ &= \frac{\partial f}{\partial t}(t,\gamma(t)) + \langle \nabla f(t,\gamma(t)), X_H(\gamma(t))\rangle = \\ &= \frac{\partial f}{\partial t}(t,\gamma(t)) + \langle \nabla f(t,\gamma(t)), J(\gamma(t))\nabla H(\gamma(t)))\rangle = \\ &= \frac{\partial f}{\partial t}(t,\gamma(t)) + \{f(t,\gamma(t)), H(\gamma(t))\} \end{aligned}$$

[(4)]For the definition of L_2-gradient, ∇_{L_2} see subsection 2.3.2.
[(5)]We use ∇ in order to be very general and to not specify if we are using L_2-gradient or x-gradient.

which is precisely our thesis. We used, in order: the fact that γ is a solution of (2.1) with $X(x)$ given by (2.7), equation (2.8) and then `Proposition 2.1.4`. □

Differently from symplectic environment where the Poisson tensor J is required to be non degenerate, in the Poisson environment Poisson tensor can be degenerate. In other words there can be some functions $C \in \mathscr{A}(\Gamma)$ whose gradient is in the kernel of J:

$$J(x)\nabla C(x) = 0 \tag{2.10}$$

where we denoted with ∇ the gradient without specifying whether it is the L_2 or the usual one. These functions (or functionals) are called *Casimirs* associated to the given Poisson tensor J.

Proposition 2.1.9. *Let $C \in \mathscr{A}(\Gamma)$ be a Casimir invariant associated to the Poisson tensor J, then C is a constant of motion for any Hamiltonian system with vector field X_H associated to J.*

Proof. To show that C is a first integral it is sufficient to prove that $\frac{dC}{dt} = 0$ along the trajectory of a Hamiltonian system with Hamiltonian H. From `Proposition 2.1.8` we have

$$\frac{dC}{dt} = \frac{\partial C}{\partial t} + \{C, H\}.$$

Recalling now `Proposition 2.1.4` we can write the Poisson bracket using the Poisson tensor

$$\{C, H\} = \langle \nabla C, J\nabla H\rangle = -\langle J\nabla C, \nabla H\rangle = 0$$

where we used the skew-symmetry of Poisson tensor and the fact that $\nabla C \in \ker J$. Thus we get

$$\frac{dC}{dt} = \frac{\partial C}{\partial t} = 0$$

because C is a function on Γ and therefore cannot depend explicitly from t. Since the proof is given for a generic Hamiltonian H the result is valid for every Hamiltonian $H \in \mathscr{A}(\Gamma)$. □

2.1.1 Canonical transformations

In the previous section we defined a Hamiltonian system in an intrinsic way. This means that the Hamiltonian character of a dynamical system is independent of the coordinates on which it is described. An immediate consequence of this fact is that mapping a Hamiltonian system into another dynamical system through a diffeomorphism gives us another Hamiltonian system. Among all possible changes of variables, an important role is played by the ones which preserve the Hamiltonian structure. These are called canonical for this reason.

We begin this section presenting how the Hamiltonian structure changes under diffeomorphisms.

Proposition 2.1.10. *Let $(\Gamma, \mathscr{A}(\Gamma), J)$ a Hamiltonian system with Hamiltonian H and let $f : \tilde{\Gamma} \mapsto \Gamma$ be a diffeomorphism. Let $\tilde{H} = H \circ f$ be the transformed Hamiltonian. The* *dynamical system on* $(\tilde{\Gamma}, \mathscr{A}(\tilde{\Gamma}), \tilde{J})$ *is an Hamiltonian system with Hamiltonian* $\tilde{H}$ *and Poisson tensor*

$$\tilde{J} = (Df)^{-1} J (Df)^{-T} \tag{2.11}$$

where T *denotes transposition.*

Proof. As a first step we find the transformation rule for the Poisson tensor. For the sake of simplicity we present the proof for a finite dimensional system. For the same reason we will use Einstein sum convention. We will denote with x coordinates on $\tilde{\Gamma}$ and with y coordinates on Γ. Starting from Hamilton equations on Γ:

$$\dot{y}^i = J^{ij}(y) \frac{\partial H(y)}{\partial y^j}.$$

We now compose both sides of the equation with $y^i = f^i(x)$. On the left hand side we get

$$\dot{y}^i = \frac{\partial f^i(x)}{\partial x^k} \dot{x}^k,$$

while on the right hand side

$$\frac{\partial H(f(x))}{\partial y^j} = \left. \frac{\partial \tilde{H}(f^{-1}(y))}{\partial y^j} \right|_{y=f(x)} = \left. \frac{\partial \tilde{H}(f^{-1}(y))}{\partial x^l} \right|_{y=f(x)} \left. \frac{\partial {f^{-1}}^l(y)}{\partial y^j} \right|_{y=f(x)}.$$

Substituting these two expressions in the differential equation above we get

$$\frac{\partial f^i(x)}{\partial x^k} \dot{x}^k = J^{ij}(f(x)) \frac{\partial \tilde{H}(x)}{\partial x^l} \left. \frac{\partial {f^{-1}}^l(y)}{\partial y^j} \right|_{y=f(x)}.$$

In a matricial form this reads

$$\frac{\partial f(x)}{\partial x} \dot{x} = J(f(x)) \left(\frac{\partial f^{-1}(y)}{\partial y} \right)^T_{y=f(x)} \frac{\partial \tilde{H}(x)}{\partial x},$$

which is exactly

$$\dot{x} = \left(\frac{\partial f(x)}{\partial x} \right)^{-1} J(f(x)) \left(\frac{\partial f(x)}{\partial x} \right)^{-T} \frac{\partial \tilde{H}(x)}{\partial x}.$$

To show that the transformed system is still Hamiltonian we have to prove that is a Poisson tensor. Namely we have to prove that that $\{\cdot,\cdot\}_{\tilde{J}}$ is a Poisson bracket. First of all we prove the skew-symmetry of $\tilde{J}$ under the hypothesis that $J^T = -J$:

$$\tilde{J}^T = \left[\left(\frac{\partial f(x)}{\partial x} \right)^{-1} J(f(x)) \left(\frac{\partial f(x)}{\partial x} \right)^{-T} \right]^T = \left(\frac{\partial f(x)}{\partial x} \right)^{-1} J^T(f(x)) \left(\frac{\partial f(x)}{\partial x} \right)^{-T} = -\tilde{J}.$$

We then have that $\tilde{J}$ is obviously linear since it is a composition of linear operators.

To verify Jacobi identity we prove that[(6)]

$$\epsilon_{ijk} \tilde{J}^{is} \frac{\partial \tilde{J}^{jk}}{\partial x^s} = 0$$

[(6)] It is an equivalent statement of (2.4).

where ϵ_{ijk} is the three dimensional Levi-Civita symbol. Substituting the expression above we get (calling $g = f^{-1}$)

$$\begin{aligned}\epsilon_{ijk}\tilde{J}^{is}\frac{\partial \tilde{J}^{jk}}{\partial x^s} &= \epsilon_{ijk}\frac{\partial g^i}{\partial y^m}J^{mn}\frac{\partial g^s}{\partial y^n}\frac{\partial}{\partial x^s}\left(\frac{\partial g^j}{\partial y^o}J^{op}\frac{\partial g^k}{\partial y^p}\right) = \\ &= \epsilon_{ijk}\frac{\partial g^i}{\partial y^m}J^{mn}\frac{\partial g^s}{\partial y^n}\left(\frac{\partial^2 g^j}{\partial x^s \partial y^o}J^{op}\frac{\partial g^k}{\partial y^p} + \frac{\partial g^j}{\partial y^o}\frac{\partial J^{op}}{\partial y^q}\frac{\partial f^q}{\partial x^s}\frac{\partial g^k}{\partial y^p} + \frac{\partial g^j}{\partial y^o}J^{op}\frac{\partial^2 g^k}{\partial x^s \partial y^p}\right).\end{aligned}$$

The first and the third term cancel each other because of the skew-symmetry of J. We are left with

$$\begin{aligned}\epsilon_{ijk}\frac{\partial g^i}{\partial y^m}\frac{\partial g^s}{\partial y^n}\frac{\partial f^q}{\partial x^s}\frac{\partial g^j}{\partial y^o}\frac{\partial g^k}{\partial y^p}J^{mn}\frac{\partial J^{op}}{\partial y^q} &= \epsilon_{ijk}\frac{\partial g^i}{\partial y^m}\delta^q_n\frac{\partial g^j}{\partial y^o}\frac{\partial g^k}{\partial y^p}J^{mn}\frac{\partial J^{op}}{\partial y^q} = \\ = \epsilon_{ijk}\frac{\partial g^i}{\partial y^m}\frac{\partial g^j}{\partial y^o}\frac{\partial g^k}{\partial y^p}J^{mn}\frac{\partial J^{op}}{\partial y^n} &= \det\left(\frac{\partial g}{\partial y}\right)\epsilon_{mop}J^{mn}\frac{\partial J^{op}}{\partial y^n}.\end{aligned}$$

Where we used the following property of the Levi-civita symbol

$$\epsilon_{ijk}A^i_m A^j_n A^k_o = \det A\, \epsilon_{mno}.$$

Now since f is a diffeomorphism also g is a diffeomorphism and then the determinant of its jacobian is not vanishing. We then obtain that $\tilde{J}$ satisfies Jacobi if and only if J satisfies Jacobi.

Last we have to show that $\{AB, C\}_{\tilde{J}} = \{A, C\}_{\tilde{J}}B + A\{B, C\}_{\tilde{J}}$. It is easy since

$$\{AB, C\}_{\tilde{J}} = \partial_i(AB)\tilde{J}^{ij}\partial_j C = \partial_i A\tilde{J}^{ij}\partial_j C\, B + A\partial_i B\tilde{J}^{ij}\partial_j C = \{A, C\}_{\tilde{J}}B + A\{B, C\}_{\tilde{J}}$$

which is precisely what we had to prove. □

From the above proposition we can state the definition of canonical transformation.

Definition 2.1.11 (Canonical transformation)**.** Let $f : \tilde{\Gamma} \to \Gamma$ be a diffeomorphism. We say that f is a *canonical transformation* if

$$\tilde{J} = J \tag{2.12}$$

or, in other words, if it leaves unchanged the Poisson tensor.

Remark 2.1.12. *Directly from* `Proposition 2.1.10` *and* `Definition 2.1.11` *we get that if Df is the jacobian of the diffeomorphism this is a canonical transformation if*

$$(Df)J(Df)^T = J. \tag{2.13}$$

We can look at the canonical transformations as the ones which leave unchanged the equation of motions. Starting from the action functional one can show the following proposition to hold.

Proposition 2.1.13. *Let (q, p) be the coordinate on the cotangent bundle of a configuration manifold M endowed with the standard symplectic structure. The transformation $(q, p, H, t) \mapsto (Q, P, K, T)$ is canonical if*

$$pdq - Hdt = c(PdQ - KdT) + dF(q, Q, t) \tag{2.14}$$

for some constant c and some function F.

The constant c is often referred as the *valence* of the canonical transformation. In the first chapter we used the following corollary:

Corollary 2.1.14. *The rescaling*

$$(q, p, H, t) \mapsto (Q, P, K, T) = (\alpha q, \beta p, \gamma H, \delta t) \tag{2.15}$$

is canonical if

$$\alpha\beta = \gamma\delta. \tag{2.16}$$

Proof. We will make use of `Proposition 2.1.13`. To do so we calculate RHS of (2.14) for the rescaling. We get:

$$P = \beta p \qquad dQ = \alpha dq \qquad K = \gamma H \qquad dT = \delta dt.$$

Substituting in (2.14) we get the following equality

$$pdq - Hdt = c(\alpha\beta pdq - \gamma\delta Hdt) + dF$$

which holds if $F = 0$ and $\alpha\beta = \gamma\delta$. □

Here we propose another method to perform a canonical transformation. It is using a function which depend in a mixed way from new and old variables. Under the hypothesis of non-degenerate second derivatives one can use such a function to build a canonical transformation. Such a function is called *generating function* and, to be more precise, the following proposition holds:

Proposition 2.1.15. *Let F be a function of $2n$ variables define in an open set of $\mathbb{R}^{2n}$, $F(q, Q)$ and, in such a domain, we suppose*

$$\det\left(\frac{\partial^2 F}{\partial q \partial Q}\right) \neq 0. \tag{2.17}$$

Equations

$$p = \frac{\partial F}{\partial q}(q, Q), \qquad\qquad P = -\frac{\partial F}{\partial Q}(q, Q) \tag{2.18}$$

defines implicitly a local transformation of coordinates $(p, q) = f(Q, P)$ from a neighbourhood $\tilde{U} \subset \mathbb{R}^{2n}$ to its image $U \subset \mathbb{R}^{2n}$ which is canonical.

Proof. The definition of the transformation is well posed since the condition on the determinant ensures the local invertibility of (2.18). So that f is locally invertible. To show the canonicity we see that it is preserved the Liouville 1-form:

$$d\mathscr{F} = \frac{\partial F}{\partial q}(q(Q, P), Q)dq + \frac{\partial F}{\partial Q}(q(Q, P), Q)dQ = pdq - PdQ$$

which states exactly `Proposition 2.1.13` with $\mathscr{F}(Q, P) = F(q(Q, P), Q)$. □

Among the transformations defined by a function F with the properties listed above, the identity map misses. This is an important lack if one wants to perform perturbation theory with generating functions. By the way this is not a real problem since the following result holds.

Proposition 2.1.16. *Let $S(P,q)$ be a given function which, in its domain, has*

$$\det\left(\frac{\partial^2 S}{\partial P \partial q}\right) \neq 0. \tag{2.19}$$

Then equations

$$p = \frac{\partial S}{\partial q}(P,q), \qquad\qquad Q = \frac{\partial S}{\partial P}(P,Q) \tag{2.20}$$

defines implicitly a canonical coordinate transformation.

Another way to perform canonical transformations is via Lie series as the following proposition states.

Proposition 2.1.17. *We denote with Φ_G^s the flow at time s along the vector field X_G. Given a Hamiltonian system with Hamiltonian $H \in \mathscr{A}(\Gamma)$, for every Hamiltonian $G \in \mathscr{A}(\Gamma)$ the coordinate transformation*

$$(Q,P) = \Phi_G^s(q,p) \tag{2.21}$$

is canonical.

In this context we define also the Lie derivative along the Hamiltonian vector field X_H as

$$\mathcal{L}_H = \{\,\cdot\,, H\}. \tag{2.22}$$

This leads to the following expression for the canonical change of coordinates

$$(Q,P) = e^{s\mathcal{L}_G}(q,p). \tag{2.23}$$

We recall that, if X, Y are vector fields on a n-dimensional manifold Γ, in differential geometry we usually define the Lie derivative of X along Y as

$$[X,Y] = L_Y X \tag{2.24}$$

where $[X,Y]$ is the Lie bracket of the above two vector fields

$$[X,Y]^i = \sum_{j=1}^{n}\left(\frac{\partial X^i}{\partial x^j}Y^j - \frac{\partial Y^i}{\partial x^j}X^j\right). \tag{2.25}$$

These two apparently different Lie derivatives are connected by the following proposition which justifies the same name for them.

Proposition 2.1.18. *Let X, Y, Z vector fields defined on a Poisson manifold Γ and let us suppose there exist three functions $A, B, C : \mathscr{A}(\Gamma) \to \mathbb{R}$ such that*

$$X = J\nabla A \qquad\qquad Y = J\nabla B \qquad\qquad Z = J\nabla C \tag{2.26}$$

then

$$\{A,B\} = C \qquad \Leftrightarrow \qquad [X,Y] = Z \tag{2.27}$$

Proof. For the sake of simplicity we will do this proof for the finite dimensional case and using the Einstein sum convention. It is a straightforward computation. We start computing the vector field associated to C:

$$C = \partial_k A J^{kl} \partial_l B$$

So that

$$X_C^i = J^{ij}\partial_j(\partial_k A J^{kl}\partial_l B) = J^{ij}(\partial^2_{jk} A J^{kl}\partial_l B + \partial_k A \partial_j J^{kl}\partial_l B + \partial_k A J^{kl}\partial^2_{lj} B).$$

We calculate Z then:

$$\begin{aligned} Z^i &= \partial_j X^i Y^j - \partial_j Y^i X^j = \partial_j(J^{ik}\partial_k A)J^{jl}\partial_l B - \partial_j(J^{il}\partial_l B)J^{jk}\partial_k A = \\ &= J^{jl}\partial_j J^{ik}\partial_k A\partial_l B + J^{ik}J^{jl}\partial^2_{jk}A\partial_l B - J^{jk}\partial_j J^{il}\partial_l B\partial_k A - J^{il}J^{jk}\partial^2_{lj}B\partial_k A = \\ &= (J^{lj}\partial_j J^{ki} + J^{kj}\partial_j J^{il})\partial_l B\partial_k A + J^{ik}(\partial^2_{jk}AJ^{jl}\partial_l B + \partial_l A J^{lj}\partial^2_{kj}B) \end{aligned}$$

which is equal to X_C^i if we use (2.4). In the last step we renamed several indices and we used skew-symmetric property of J. $\square$

We can use Lie derivative also to express the derivative with respect to time of a function $f : \Gamma \times \mathbb{R} \to \mathbb{R}$ along the flow of a Hamiltonian system wit Hamiltonian H. Let $\gamma : \mathbb{R} \to \Gamma$ be a trajectory of the Hamiltonian system, from `Proposition 2.1.8` follows directly

$$\frac{df(t,\gamma(t))}{dt} = \frac{\partial f}{\partial t}(t,\gamma(t)) + \mathcal{L}_H f(t,\gamma(t)). \tag{2.28}$$

We conclude this subsection with a remark on the Lie method. It is proved (for example in [10]) that on symplectic manifolds there is a one-to-one corrispondence between one-parameter group of canonical transformations and Hamiltonians in the following sense. We have that for every one parameter group of canonical transformation there exists a Hamiltonian G such that

$$\mathcal{G}(\lambda; x) = e^{\lambda \mathcal{L}_G} x \tag{2.29}$$

where $\mathcal{G}(\lambda; x)$ denotes the action of the element of the group $\mathcal{G}$ with parameter λ.

This is, in general, not true for Poisson manifolds. Therefore if the Poisson tensor is degenerate we can find a one-parameter group of canonical transformation for which there not exists a correspondent Hamiltonian. We will show this fact with an example.

Let $\Gamma = \mathbb{R}^3$ with Poisson tensor

$$J = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \tag{2.30}$$

and with coordinates (x, y, z). It is easy to see that the x coordinate is a Casimir function[7]. Let us consider the following one parameter group of canonical transformations

$$\begin{aligned} \tilde{x}(x,\lambda) &= x + \lambda \\ \tilde{y}(y) &= y \\ \tilde{z}(z) &= z \end{aligned} \tag{2.31}$$

[7] More precisely, the projection on the x-axis is a Casimir Function.

Let us now suppose that there exists a Hamiltonian function G such that this canonical transformation can be expressed as Lie series. We then have

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = e^{\lambda\{\cdot,G\}} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{2.32}$$

Expanding in series of λ we get

$$\begin{aligned} x + \lambda &= x + \lambda\{x, G\} + \dots \\ y &= y + \lambda\{y, G\} + \dots \\ z &= z + \lambda\{z, G\} + \dots \end{aligned} \tag{2.33}$$

This means $G = G(x)$ and then

$$x + \lambda = x + \lambda \sum_{ij} \partial_i x J^{ij} \partial_j G = x + \lambda J^{11} \partial_x G = x \tag{2.34}$$

which implies $\lambda = 0$. We than deduce the non-existence of such a G.

2.1.2 Integrability: Liouville

When one writes down a physical equation one is interested in finding exact solutions of such an equation. Tipically the first solutions that one can find are the so called *free solutions* which are the solutions of the equation without a potential energy term. So are, for example, the trajectories for free particles in classical mechanics or the electromagnetic waves propagation in electrodynamics.

These are not, of course, the only easy solutions to find. There usually exist some potentials for which it is possible to find exact solutions of the physical equations. Example of this are the Kepler problem or the harmonic oscillator in classical mechanics and their quantum correspondents (harmonic oscillators and hydrogen atom) for quantum mechanics.

The request of finding systems for which were possible to write an exact solution is precisely the request of finding the so called *integrable systems*. These are systems for which the solutions are written up to quadrature. This means that one can compute the exact solutions solving a finite number of algebraic equations, integrations and inversion of functions.

Historically the first answer to the question "which systems can be solved by quadrature?" was given by Liouville and this section is devoted in recalling what is nowadays called a *Liouville integrable system*.

To state Liouville's definition of integrable systems we start recalling that a function on the phase space which doesn't change along the flow of a dynamical system is called *first integral* and, by using `Proposition 2.1.8`, we can state the following definition.

Definition 2.1.19. A function I is said to be a *first integral* of a Hamiltonian system with Hamiltonian $H \in \mathscr{A}(\Gamma)$ if

$$\frac{\partial I}{\partial t} + \{I, H\} = 0. \tag{2.35}$$

Definition 2.1.20. Given two first integrals I_1 and I_2 of the same Hamiltonian system with Hamiltonian H, we say that they are *in involution* if

$$\{I_1, I_2\} = 0. \tag{2.36}$$

Definition 2.1.21. Given two first integrals I_1 and I_2 we say that these are *dependent* if $\nabla I_1 = c\nabla I_2$ for some $c \in \mathbb{R}$. Otherwise we say that they are *independent*.

Definition 2.1.22. Given a $2n$-dimensional Hamiltonian system. We say that it is *integrable in the sense of Liouville* if it admits n independent first integrals in involution with each other.

The importance of the Liouville integrable systems is due to the fact that Liouville proved that if in a Hamiltonian system we know n independent first integrals in involution, then the system is integrable by quadrature. For the exact formulation of the theorem see [1].

Thanks to the theorem it is showed that, if the $2n$-dimensional Hamiltonian system is integrable and $I_1, \dots, I_n$ are n first integrals of the system, it is possible to perform a canonical transformation $f : (q, p) \to (\vartheta, I)$ such that the transformed equations of motion are

$$\begin{aligned} \dot{I} &= 0, \\ \dot{\theta} &= \omega(I) \end{aligned} \tag{2.37}$$

and their solution, very trivially, is

$$I(t) = I_0 \qquad\qquad \vartheta(t) = \omega(I_0)t + \vartheta_0. \tag{2.38}$$

Variables I are called *action variables* while ϑ are called *angle variables*. Explicit solution shows that the motion takes place on the n-dimensional submanifold defined by $I = I_0$ and, more, it is trivial in these coordinates. To get the motion in the original coordinate system one has to invert f and this is usually very hard to do (it is the quadrature).

Example. Harmonic oscillators[(8)]. As a first example of integrable system we consider a chain of n unitary masses interacting with their first neighbours with an armonic force and periodic boundary conditions. The Hamiltonian of the system is

$$H(q, p) = \sum_{k \in \mathbb{Z}_N} \left(\frac{p_k^2}{2} + \frac{1}{2}(q_{k+1} - q_k)^2 \right). \tag{2.39}$$

It is convenient now to write the Hamiltonian in a matricial expression as

$$H(q, p) = \frac{1}{2}p^T p - \frac{1}{2}q^T S q \tag{2.40}$$

where T denotes transposition and S is the matrix

$$S = \begin{pmatrix} -2 & 1 & 0 & & 1 \\ 1 & -2 & 1 & \dots & \\ 0 & 1 & -2 & & \\ & \vdots & & \ddots & \\ 1 & & & 1 & -2 \end{pmatrix}. \tag{2.41}$$

(8) Taken from [11].

Since S is symmetric we can diagonalize it using an orthogonal matrix O:

$$O^T S O = \operatorname{diag}\left(-\omega_1^2, \dots, -\omega_n^2\right) \tag{2.42}$$

and

$$\omega_j = 2\sin(k_j/2), \qquad k_j = \frac{2\pi j}{n}. \tag{2.43}$$

We can then perform the canonical transformation

$$\tilde{q} = O\tilde{q}, \qquad \tilde{p} = Op. \tag{2.44}$$

It is easy to see that this transformation is canonical since it preserves the Liouville 1-form (`Proposition 2.1.13`):

$$\tilde{p} \cdot d\tilde{q} = Op \cdot Odq = p \cdot O^T O dq = p \cdot dq. \tag{2.45}$$

In the new coordinates the Hamiltonian takes the form

$$\tilde{H}(\tilde{q}, \tilde{p}) = \sum_{j \in \mathbb{Z}_n} \left(\frac{p_j^2}{2} + \omega_j^2 q_j^2 \right). \tag{2.46}$$

The equation of motion associated to this Hamiltonian are

$$\dot{\tilde{q}}_j = \tilde{p}_j \qquad \dot{\tilde{p}}_j = -\omega_j^2 \tilde{q}_j \tag{2.47}$$

which are easily solved

$$\tilde{q}_j = c_j \sin(\omega_j t + \theta_j^0) \qquad \tilde{q}_n = c_n t + \theta_n^0. \tag{2.48}$$

In this example we see that the $\tilde{q}_j$ are angles while the c_j are actions and they don't evolve in time.

2.2 Isospectral method and Lax pairs

Isospectral method is a very general procedure to get a series of first integrals of motion of a given Hamiltonian system if the equations of motion can be written in a very particular form, involving a pair of operators called *Lax pair*.

We supposte to have a Hamiltonian system[9] and to have a pair of operators L, M such that the equations of motion can be written in the form

$$\dot{L} = [L, M] \tag{2.49}$$

where the square bracket denotes the usual commutator $[A, B] = AB - BA$. More we suppose that M is skew-symmetric. We can then state

Theorem 2.2.1. *If the equations of motions are in the form (2.49) then eigenvalues of L are first integrals of the system.*

(9) At this level it is not required the system to be Hamiltonian.

Proof. The proof consists in showing that there exist an invertible operator U such that

$$\frac{d}{dt}\left(U(t)L(t)U^{-1}(t)\right) = 0.$$

In this way $L(t)$ and $L(0)$ have the same spectrum.**(10)** Since U is invertible we have

$$U(t)U^{-1}(t) = \mathbf{1}$$

where with **1** we denote the identity operator. Deriving both sides with respect to t we get

$$\frac{dU(t)}{dt}U^{-1}(t) + U(t)\frac{dU^{-1}(t)}{dt} = 0$$

which gives us an expression for the time derivative of U^{-1}:

$$\frac{dU^{-1}(t)}{dt} = -U^{-1}(t)\frac{dU(t)}{dt}U^{-1}(t).$$

Setting now

$$\frac{dU(t)}{dt}U^{-1}(t) = M(t) \qquad U(0) = \mathbf{1}$$

we have a Cauchy problem for U which, under suitable hypothesis on M, guarantees the existence of U. Last we verify that the spectum of L is conserved by this transformation:

$$\begin{aligned}\frac{d}{dt}\left(ULU^{-1}\right) &= U^{-1}\left(-\frac{dU}{dt}U^{-1}L + \frac{dL}{dt} + L\frac{dU}{dt}U^{-1}\right)U = \\ &= U^{-1}\left(-ML + \frac{dL}{dt} + LM\right)U = \\ &= U^{-1}\left(\frac{dL}{dt} - [L, M]\right)U = 0.\end{aligned}$$

□

We then proved that a system with a Lax pair has n first integrals.

As a first remark we see that the Lax pair is not unique. In fact if S is an invertible matrix we see that

$$L \mapsto SLS^{-1} \qquad\qquad M \mapsto SMS^{-1} + \frac{dS}{dt}S^{-1}$$

is also a Lax pair.

As a second remark it is possible to see that every Liouville-integrable system admits a formulation with Lax pairs. We don't enter in this subject but we refer to [2].

We recall now that a system is integrable in the sense of Liouville if it has n independent integrals which are in involution between them. The spectrum of L matrix simply gives n first integrals but it is not proved that these are in involution between them. We don't enter in such a general treatment.

The following adaptation of [13] shows that Toda system admits a Lax pair.**(11)**

(10) This property is often called *isospectral deformation.*

(11) In the same paper it is shown also that Toda system is integrable.

Example. Toda lattice. In this example we will show, thanks to a Lax pair, that the N-dimensional Toda lattice admits a N integrals of motion. Toda lattice is a system with N particles with Hamiltonian

$$H(q,p) = \sum_{k\in\mathbb{Z}_N} \left(\frac{p_k^2}{2} + V_0 e^{\alpha(q_{k+1}-q_k)} \right). \tag{2.50}$$

And its equation of motion are

$$\begin{aligned} \dot{q}_m &= p_m, \\ \dot{p}_m &= \alpha V_0 \left(e^{\alpha(q_{m+1}-q_m)} - e^{\alpha(q_m - q_{m-1})} \right). \end{aligned} \tag{2.51}$$

According to [11] we introduce the quantities

$$a_k = \frac{\sqrt{V_0\alpha}}{2} e^{\alpha(q_{k+1}-q_k)/2} \qquad b_k = \frac{1}{2} p_k. \tag{2.52}$$

And we see that Lax pair for Toda system is given by

$$L = \begin{pmatrix} b_1 & a_1 & & & & a_n \\ a_1 & b_2 & a_2 & & \dots & \\ & a_2 & b_3 & a_3 & & \\ & & \vdots & & \ddots & \\ a_n & & & & a_{n-1} & b_n \end{pmatrix} \tag{2.53}$$

$$M = \begin{pmatrix} 0 & a_1 & & & & -a_n \\ -a_1 & 0 & a_2 & & \dots & \\ & -a_2 & 0 & a_3 & & \\ & & \vdots & & \ddots & \\ a_n & & & & -a_{n-1} & 0 \end{pmatrix}. \tag{2.54}$$

A proof of the fact that the equation of motion for Toda system can be written as

$$\dot{L} = [M, L] \tag{2.55}$$

is a straightforward calculation. We can begin from the one involving $\dot{p}_k$:

$$\begin{aligned} \dot{L}_{kk} = [M,L]_{kk} = \sum_{j=1}^{n} (M_{kj}L_{jk} - L_{kj}M_{jk}) = \\ = M_{k,k+1}L_{k+1,k} + M_{k,k-1}L_{k-1,k} - L_{k,k+1}M_{k+1,k} - L_{k,k-1}M_{k-1,k} \end{aligned} \tag{2.56}$$

where we understood that if $k=1$, $k-1=n$. Moreover from the expressions above we get

$$M_{k,k+1} = -M_{k+1,k} = a_k \qquad M_{k,k-1} = -M_{k-1,k} = -a_{k-1} \tag{2.57}$$

$$L_{k,k+1} = L_{k+1,k} = a_k \qquad L_{k,k-1} = L_{k-1,k} = a_{k-1} \qquad L_{kk} = b_k \tag{2.58}$$

and then

$$\dot{b}_k = 2(a_k^2 - a_{k-1}^2) \tag{2.59}$$

with the help of the definitions above this reads

$$\dot{p}_k = V_0 \alpha \left(e^{\alpha(q_{k+1}-q_k)} - e^{\alpha(q_k - q_{k-1})} \right). \tag{2.60}$$

An analogous calculation gives the equation involving $\dot{q}_m$.

Thanks to Lax pair we can find n first integral of Toda lattice. These are the eigenvalues of the matrix L. Since eigenvalues of L are conserved quantities so is also their sum. We obtain in this way the (obvious)

$$I_1 = \operatorname{Tr} L = \sum_{j=1}^{n} b_j = \frac{1}{2} \sum_{j=1}^{n} p_k = \frac{1}{2} P_{tot}. \tag{2.61}$$

This states nothing more than the conservation of total momentum.

Another example of system which admits a Lax Pair is the KdV (see section 2.4).

2.3 Some tools from functional analysis

In this section we recall some notions of functional analysis that will be useful in the following. The first two subsections are entirely inspired by [17].

2.3.1 Normed spaces

The notion of linear space (also called vector space) is one of the basic notions in mathematics. It is also well known that a vector space can be endowed with different additional structures as scalar products or norms. A normed space is a vector space on which it is defined the length of a vector. Here we recall precise definitions of that.

Definition 2.3.1 (Linear space)**.** A structure $(X, +, *, K)$ where X is a set, $+ : X \times X \to X$ is a function, $* : K \times X \to X$ is a function and K is a field, is called vector space if the following properties hold:

1. commutativity of the sum: $x_1 + x_2 = x_2 + x_1 \ \forall x_1, x_2 \in X$;
2. associativity of the sum: $x_1 + (x_2 + x_3) = (x_1 + x_2) + x_3 \ \forall x_1, x_2, x_3 \in X$;
3. existence of the neutral element: $\exists e \in X$ such that $x_1 + e = 0 \ \forall x_1 \in X$;
4. $a(bx) = (ab)x \ \forall a, b \in K$ and $\forall x \in X$;
5. $1x = x \ \forall x \in X$ where 1 is the neutral element in K;
6. $(a + b)x = ax + bx \ \forall a, b \in K$ and $\forall x \in X$;
7. $a(x_1 + x_2) = ax_1 + ax_2 \ \forall a \in K$ and $\forall x_1, x_2 \in X$.

Definition 2.3.2 (Norm)**.** Given a vector space X, a norm on X is a function $\| \cdot \| : X \to \mathbb{R}$ such that

1. $\|x\| \geq 0 \ \forall x \in X$;
2. $\|x\| = 0$ if and only if $x = 0$;
3. $\|x_1 + x_2| \leq \|x_1\| + \|x_2\|$;
4. $\|ax\| = |a|\|x\|$.

A normed space is a vector space endowed with a norm.

Example. The space of square integrable functions, $L_2(\mathbb{T})$ with the L_2-norm defined as follows

$$\|f\|_{L_2} = \sqrt{\langle f, f\rangle_{L_2}} = \left(\int_{\mathbb{T}} f^2(x)\, dx\right)^{\frac{1}{2}} \tag{2.62}$$

is a normed space.

2.3.2 Differentiation on linear spaces

On linear spaces it is possible to define two kind of derivatives. This is the mirror of what happens in the real analysis when one consider the directional derivative and the differential of a function: one is stronger and one is weaker. Imaginatively these two kind of differentiations are referred as strong and weak differentiations.

Let X, Y be two normed spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ and let $f : X \to Y$ be a function.

We say that f is *differentiable* in a given point $x \in X$ if there exists a bounded linear operator $L_x : X \to Y$ such that for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|h\|_X < \delta$ implies

$$\|F(x+h) - F(x) - L_x(h)\|_Y \leq \varepsilon \|h\|_X \tag{2.63}$$

Expression $L_x(h)$ represents an element in Y which is often called *differential* of f or *Fréchet differential* of f in x. The operator L_x is called *strong derivative* of the function f at x. We will often write $f'(x)$ instead of L_x.

Remark 2.3.3. *Uniqueness of the strong derivative is obvious for linear operators. In fact, let us suppose there exist two different derivatives called $L_x^{(1)}$ and $L_x^{(2)}$. Since these operators are linear and bounded we have $\|L_x^{(1)}h - L_x^{(2)}h\| = o(h)$ implies $L_x^{(1)} = L_x^{(2)}$.*

Beside very trivial (but important) results like the sum of two differentiable function is still differentiable, multiplication for a scalar keeps differentiability, one of the most important results concerning strong derivatives is the so-called "chain rule":

Proposition 2.3.4 (Chain rule)**.** *Let X, Y, Z be three normed spaces, $U(x_0)$ an open neighbourhood of $x_0 \in X$, F a continue application of this neighbourhood in Y, $y_0 = F(x_0)$, $V(y_0)$ an open neighbourhood of $y_0 \in Y$ and G a continue application of this neighbourhood in Z. If F is differentiable in x_0, G is differentiable in y_0 then $H = GF$ defined in a neighbourhood of x_0 will be differentiable in x_0 and*

$$H'(x_0) = G'(y_0)F'(x_0) \tag{2.64}$$

Proof. From the hypothesis above we have

$$\begin{aligned} F(x_0+\xi) &= F(x_0) + F'(x_0)\xi + o(\xi), \\ G(y_0+\eta) &= G(y_0) + G'(y_0)\eta + o(\eta). \end{aligned} \tag{2.65}$$

Both F' and G' are bounded linear operators, then

$$\begin{aligned} H(x_0+\xi) = G(y_0 + F'(x_0)\xi + o(\xi)) = \\ &= G(y_0) + G'(y_0)(F'(x_0)\xi + o(\xi)) + o(F'(x_0)\xi + o(\xi)) = \\ &= G(y_0) + G'(y_0)F'(x_0)\xi + o(\xi) \end{aligned} \tag{2.66}$$

which is precisely our thesis. ∎

Let us consider again an application $f : X \to Y$. The *weak differential* (also called *Gateaux differential*) of f in $x \in X$ with increment h is the limit

$$Df(x,h) = \frac{d}{d\epsilon} f(x+\epsilon h)|_{\epsilon=0} = \lim_{\epsilon\to 0} \frac{f(x+\epsilon h) - f(x)}{\epsilon} \tag{2.67}$$

where the convergence is meant with the norm convergence in Y. If the Gateaux derivative exists for every $h \in X$ then f is said to be Gateaux-derivable in x.

Weak differential can be non linear in h. If this happens the linear operator L_x such that

$$Df(x.h) = L_x(h) \tag{2.68}$$

is called *Gateaux derivative* or *weak derivative* of f in x and we will denote it with D/Dx. It is interesting to note that for weak derivatives the theorem for the derivative of composed functions does not hold.

In the previous sections we often referred to the L_2 gradient. We define it starting from the weak derivative: if we take a real-valued functional $F : X \to \mathbb{R}$ we define its L_2 gradient as

$$DF(x,h) = \langle \nabla_{L_2} F(x), h\rangle_{L_2} \tag{2.69}$$

where $DF(x,h)$ is defined as (2.67).

2.3.3 Anti-derivative operators

Since we will focus our attention on the Fermi-Pasta-Ulam problem on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ we will often deal with periodic functions. We remind that (nearly) every function on the torus can be expanded in Fourier series as follows

$$f(x) = \sum_{k\in\mathbb{Z}} \hat{f}_k e^{2\pi i k x}. \tag{2.70}$$

We recall also how a derivative operator acts on a function represented in Fourier series

$$\partial_x f(x) = \partial_x \sum_{k\in\mathbb{Z}} \hat{f}_k e^{2\pi i k x} = 2\pi i \sum_{k\in\mathbb{Z}} k \hat{f}_k e^{2\pi i k x}. \tag{2.71}$$

It is possible to define an *anti-derivative operator*, denoted by ∂_x^{-1}, which acts on f as follows

$$\partial_x^{-1} f(x) = \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{2\pi i k} \hat{f}_k e^{2\pi i k x}. \tag{2.72}$$

From the definition one gets

$$\partial_x \partial_x^{-1} f = f - \hat{f}_0 \tag{2.73}$$

where, obviously, $\hat{f}_0 = \langle f\rangle_{\mathbb{T}} = \int_{\mathbb{T}} f(x)\,dx$. This operator has a property which is analogous to the integration by parts of the derivative, as the proposition below shows.

Proposition 2.3.5 (Integration by parts)**.** *The anti-derivative operator satisfies an analogous of the integration by parts formula which is*

$$\int_{\mathbb{T}^1} f(x)\partial_x^{-1} g(x)\,dx = -\int_{\mathbb{T}^1} g(x)\partial_x^{-1} f(x)\,dx. \tag{2.74}$$

Proof. The proof consists on a straightforward calculation:

$$\begin{aligned}
\int_{\mathbb{T}^1} f(x)\partial_x^{-1} g(x)\, dx &= \int_{\mathbb{T}^1} dx \sum_{k\in\mathbb{Z}} \sum_{k'\in\mathbb{Z}\setminus\{0\}} \hat{f}_k \frac{1}{2\pi i k'} \hat{g}_{k'} e^{2\pi i (k'+k)x} = \\
&= \sum_{k\in\mathbb{Z}} \sum_{k'\in\mathbb{Z}\setminus\{0\}} \hat{f}_k \frac{1}{2\pi i k'} \hat{g}_{k'} \delta_{k+k',0} = \\
&= -\sum_{k\in\mathbb{Z}} \sum_{k'\in\mathbb{Z}\setminus\{0\}} \hat{f}_k \frac{1}{2\pi i k} \hat{g}_{k'} \delta_{k+k',0} = \\
&= -\int_{\mathbb{T}^1} dx \sum_{k\in\mathbb{Z}} \sum_{k'\in\mathbb{Z}\setminus\{0\}} \hat{f}_k \frac{1}{2\pi i k} \hat{g}_{k'} e^{2\pi i (k'+k)x} = \\
&= -\int_{\mathbb{T}^1} g(x)\partial_x^{-1} f(x)\, dx
\end{aligned}$$

where we noted that the term with $k = 0$ vanishes because of the Kronecker delta. □

2.4 The Korteweg-de Vries equation

The Korteweg-de Vries equation (KdV) is a nonlinear partial differential equation which arises in the study of many physical systems. It was introduced in the second half of the eighteenth century by Boussinesq ([8] footnote on pag. 360[12]) and then re-discovered by Korteweg and de Vries around fifteen years later to describe the evolution of long water waves down a canal of rectangular cross section.

Among the many (equivalent) forms on which one can find it in the literature, we choose to present it in the following one

$$u_t = buu_x + au_{xxx}. \tag{2.75}$$

To connect this form with the others, one can eliminate the parameter a simply rescaling time by the same factor. One then obtains the following equation depending only on a parameter $\gamma = \frac{b}{a}$:

$$u_t = \gamma uu_x + u_{xxx}. \tag{2.76}$$

One can then choose the value of γ simply rescaling u.[13]

It will be important for us the connection anticipated in chapter 1 between KdV and FPU that will be treted in chapter 4.

In order to prove the existence of infinitely-many constants of motion we introduce some "immediate" conservation laws for the KdV. Since it depends on x and t only through differentiations, it is unchanged for every translation $t' = t + \tau$ and $x' = x + \xi$ with τ and ξ constants. More, KdV equation is unchanged if one applies the Galileian transformation

$$t' = t, \qquad x' = x - ct, \qquad u'(x', t') = u(x, t) + \frac{c}{b}, \tag{2.77}$$

[12] An historical reconstruction of the discovery of this equation can be found in [9].

[13] It is also possible, with opportune scaling, to bring KdV equation in the form $u_t = uu_x + \delta^2 u_{xxx}$ which is very useful if one studies the small dispersion limit.

where c is a constant. Physically this corresponds to going on a steady moving reference frame with velocity c.

To prove this fact we simply transform quantities in (2.75) with (2.77):

$$\begin{aligned} u_t = \frac{\partial u'}{\partial t} = \frac{\partial t'}{\partial t}\frac{\partial u'}{\partial t'} + \frac{\partial x'}{\partial t}\frac{\partial u'}{\partial x'} = u'_{t'} - cu'_{x'} \\ buu_x + au_{xxx} = b\left(u' - \frac{c}{b}\right)u'_{x'} + au'_{x'x'x'} = bu'u'_{x'} + au'_{x'x'x'} - cu'_{x'}. \end{aligned} \tag{2.78}$$

When we put (2.78) in (2.75) we get

$$u'_{t'} = bu'u'_{x'} + au'_{x'x'x'} \tag{2.79}$$

which is precisely the KdV equation (2.75) for the transformed quantities.

2.4.1 Conservation laws and integrals of the motion

Usually we refer to conservation laws as equations like

$$\rho_t + X_x = 0 \tag{2.80}$$

where ρ is a conserved density and $-X$ is its flow. Equations of this form are very common in physics (conservation of charge in electromagnetism, conservation of mass for fluids, ...). Dealing with KdV, ρ and X are functions of x, t, u and the derivatives of u, $\partial_x^n u$.

One of the interesting properties of the KdV equation is that it has an infinite number of conserved density.

To see how special is this property, we have to consider one of the possible classes of generalization of KdV equation given by the following

$$w_t = bw^p w_x + aw_{xxx} \tag{2.81}$$

where we recognize that for $p = 1$ we get exactly the KdV. The first equation in this class, the one with $p = 2$, is called *modified KdV (mKdV)* and it reads

$$v_t = bv^2 v_x + av_{xxx}. \tag{2.82}$$

This equation arises in the study of nonlinear discrete mass string with cubic force law between masses (like β-FPU, see chapter 5).

This equation is as special as KdV from the point of view of conservation laws. A study of the generalized (2.81) with $p \geq 3$ leads to the discovery of only three polynomial conservation laws. In other words the only "very special" equations in this class are the KdV and the mKdV. Assuming (for the moment) the special role of KdV, the special role of the modified KdV can be understood in terms of the following theorem by Miura [19]:

Theorem 2.4.1 (Miura). *If v is a solution of*

$$v_t = bv^2 v_x + av_{xxx} \tag{2.83}$$

then, if i is the imaginary unit,

$$u = v^2 + i\sqrt{\frac{6a}{b}}v_x \tag{2.84}$$

is a solution of

$$u_t = buu_x + au_{xxx}. \tag{2.85}$$

Proof. For the sake of simplicity we call $\sigma = i\sqrt{\frac{6a}{b}}$. Substituting (2.84) in (2.85) we get

$$\begin{aligned} u_t - buu_x - au_{xxx} &= (v^2+\sigma v_x)_t - b(v^2+\sigma v_x)(v^2+\sigma v_x)_x - a(v^2+\sigma v_x)_{xxx} = \\ &= \left(2v + \sigma\frac{\partial}{\partial x}\right)(v_t - bv^2 v_x - av_{xxx}). \end{aligned}$$

Where the last step is valid if and only if

$$\sigma^2 = -\frac{6a}{b}$$

which is precisely the statement of the theorem. □

Remark 2.4.2. *We will denote with $\mathscr{M}(v)$ the operator $\mathscr{M}(v) = \left(2v + \sigma\frac{\partial}{\partial x}\right)$.*

Remark 2.4.3. *This theorem connects KdV and mKdV. Thanks to the transformation (2.84), the infinitely-many constants of motion of KdV are mapped in infinitely many constants of motion for mKdV (see subsection 2.4.5).*

We note that the theorem works only in one sense. Anyway this is the starting point for the Gardner proof of the existence of infinitely many integrals of KdV equation. We chose to adapt this proof in [19] to KdV equation in the form (2.75) because it shows directly the existence of quantities satisfying the conservation equation (2.80) without involving Hamiltonian structures.

The idea of the proof is to generalize the transformation (2.84) and to proceed in the converse direction of `Theorem 2.4.1`. Recalling that KdV equation is Galileian invariant (2.77) whereas mKdV is not, we introduce the following transformation of variables:

$$\begin{aligned} t' &= t \\ x' &= x - \frac{3a}{2\varepsilon^2}t \\ u'(x',t') &= u(x,t) + \frac{3a}{2b\varepsilon^2} \\ v(x,t) &= \varepsilon w(x',t') + \frac{1}{2\varepsilon}\sqrt{-\frac{6a}{b}} \end{aligned} \tag{2.86}$$

where the specific dependence on the formal parameter ε has been chosen in order to get the desired results below.

The above transformation implies the following transformation rule for derivatives:

$$\begin{aligned} \frac{\partial}{\partial t} &= \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} + \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'} = \frac{\partial}{\partial t'} - \frac{3a}{2\varepsilon^2}\frac{\partial}{\partial x'}, \\ \frac{\partial}{\partial x} &= \frac{\partial x'}{\partial x}\frac{\partial}{\partial x'} + \frac{\partial t'}{\partial x}\frac{\partial}{\partial t'} = \frac{\partial}{\partial x'}. \end{aligned} \tag{2.87}$$

Hence, calling

$$P(u) = u_t - buu_x - au_{xxx}, \qquad Q(v) = v_t - bv^2 v_x - av_{xxx}, \tag{2.88}$$

we get, because of the Galileian invariance,

$$P'(u') = P(u) \tag{2.89}$$

and

$$\begin{aligned} Q'(v') = \left(\partial_{t'} - \frac{3a}{2\varepsilon^2}\partial_{x'}\right)\left(\varepsilon w + \frac{\sigma}{2\varepsilon}\right) - b\left(\varepsilon w + \frac{\sigma}{2\varepsilon}\right)^2 \varepsilon w_{x'} - aw_{x'x'x'} = \\ = \varepsilon w_{t'} - b\varepsilon(\varepsilon^2 w^2 w_{x'} + \sigma w w_{x'}) - a\varepsilon w_{x'x'x'}. \end{aligned} \tag{2.90}$$

Transforming now the operator connecting mKdV and KdV we get

$$\mathscr{M}'(v') = \left(2v + \sigma\frac{\partial}{\partial x}\right)' = \left(2\varepsilon w + \frac{\sigma}{\varepsilon} + \sigma\partial_{x'}\right). \tag{2.91}$$

If we recall now that $P(u) = 0$ if u is a solution of KdV equation, $P'(u') = P(u)$ and $P'(u') = \mathscr{M}'(v')Q'(v')$ we get $0 = P(u) = P'(u') = \mathscr{M}'(v')Q'(v')$. Dropping the primes one is thus left with

$$0 = P(u) = (\varepsilon^2 w + \sigma + \sigma\varepsilon\partial_x)(w_t - b(\varepsilon^2 w^2 w_x + \sigma w w_x) - aw_{xxx}). \tag{2.92}$$

Transforming both sides of (2.84) with (2.86) we get a relation between u and w:

$$u = \varepsilon^2 w^2 + \varepsilon\sigma w_x + \sigma w. \tag{2.93}$$

One can now suppose to expand w in a formal series of ε:

$$w(x, t; \varepsilon) = w_0(x, t) + \varepsilon w_1(x, t) + \varepsilon^2 w_2(x, t) + \dots \tag{2.94}$$

which, substituted into (2.93), provides power by power in ε some expressions connecting w_i, $i = 0, 1, \dots$, with u and its derivatives:

$$w_0 = \frac{u}{\sigma}, \qquad w_1 = -\frac{u_x}{\sigma}, \qquad w_2 = -\frac{u^2}{\sigma^2} + u_{xx}, \qquad \dots \tag{2.95}$$

Because of the structure of (2.92) one gets a conservation law for every power of ε. Since these are infinitely-many, one has infinitely many conservation laws for KdV. Here we conclude this proof finding, as an example, the first one (i.e. the one related to ε^0). Substituting (2.95) in (2.92) and taking only terms with order 0 in ε one gets:

$$u_t - \left(\frac{b}{2}u^2 + au_{xx}\right)_x = 0 \tag{2.96}$$

which is the conserved form of KdV equation.

Another way to show that KdV equation has infintely many constant of motion is noting that it admits a Lax pair

$$L = \frac{6b}{a}\partial_x^2 + u, \qquad M = 4b\partial_x^3 + \frac{a}{2}(u\partial_x + \partial_x u) \tag{2.97}$$

and then all the eigenvalues of L are constant of motion (as `Theorem 2.2.1` shows). Last, as far as I know, one can prove the existence of infinitely many integrals of motion involviong the twofold Hamiltonian structure underlying the KdV equation. In advantage this last proof shows that integrals of motion are in involution and then, consequently, KdV equation is an integrable system.[14]

2.4.2 Bi-Hamiltonian structure of KdV

We saw in the previous subsections that a Hamiltonian system lives naturally on a Poisson manifold which is a differentiable manifold with a Poisson bracket defined on the algebra of the real-valued functions over the manifold. There are particular systems for which it is possible to define two Poisson structures on the same dynamical system. Namely it is possible to define two different Poisson tensors that, applied to the gradient of two different Hamiltonians, lead to the same equations of motion. In other words one has

$$X_h = \{H_1, x\}_{J_1} = \{H_2, x\}_{J_2} \tag{2.98}$$

where the pedices denote the Poisson bracket associated respectively to the Poisson tensor J_1 and J_2.

Since we are not interested in a general treatment, we focus on the Korteweg-de Vries equation (2.75) which admits the following bi-Hamiltonian decomposition:

$$H_1 = \int \left(\frac{b}{6} u^3 + \frac{a}{2} {u_x}^2 \right) dx, \qquad J_1 = \partial_x, \tag{2.99}$$

$$H_2 = \int \frac{u^2}{2}\, dx, \qquad J_2 = a\partial_{xxx} + \frac{2b}{3} u\partial_x + \frac{b}{3} u_x. \tag{2.100}$$

We can note that $\{H_1, H_2\}_{J_1} = \{H_1, H_2\}_{J_2} = 0$ and then both Hamiltonians are conserved functionals of KdV. Noting then that, posing $F_0 = C \int u\, dx$, $F_1 = H_2$ and $F_3 = H_1$ (the lower number is the degree of the integrand of these functionals)

$$J_1 \frac{\delta F_1}{\delta u} = J_2 \frac{\delta F_0}{\delta u} \tag{2.101}$$

and

$$J_1 \frac{\delta F_2}{\delta u} = J_2 \frac{\delta F_1}{\delta u} \tag{2.102}$$

one can guess the existence of a recursion formula for finding the infinitely many first integrals on the form

$$J_1 \frac{\delta F_{n+1}}{\delta u} = J_2 \frac{\delta F_n}{\delta u}. \tag{2.103}$$

Before to go ahead we prove that this recursion formula is true. Following [18], we start proving the following proposition.

[14] Here, with integrable, we mean simply the existence of infinitely many integrals of motion in involution.

Lemma 2.4.4. *There exists a sequence G_n of polynomials in u and its derivatives up to order $2n-2$ satisfying an analogous of (2.103) in the form*[15]

$$J_1 G_{n+1} = J_2 G_n. \tag{2.104}$$

More, G_n is uniquely determined if we set the constant terms equal to zero during integrations.

Proof. We use the following calculus lemma: suppose that Q is a polynomial in u and its derivatives up to order k, such that for every periodic function u of period p

$$\int_0^p Q(u)\,dx = 0.$$

Then there exists a polynomial G in derivatives of u up to order $k-1$ such that

$$Q = J_1 G.$$

We assume that G_k has been constructed for all $k \leq n$; to construct G_{k+1} we have to solve (2.104). According to the above calculus lemma, we have to show that for all u,

$$\int J_2 G_k\,dx = 0.$$

Since we can pose $G_0 = \text{const}$ we have that this equation is equivalent to

$$\langle J_2 G_k, G_0 \rangle_{L_2} = 0$$

Using now repeatedly the skew-symmetric property of J_2 and the recursion formula (2.104) we have

$$\langle J_2 G_k, G_0 \rangle_{L_2} = -\langle G_k, J_2 G_0 \rangle_{L_2} = -\langle G_k, J_1 G_1 \rangle_{L_2} = \langle J_1 G_k, G_1 \rangle_{L_2} = \langle J_2 G_{k-1}, G_1 \rangle_{L_2}. \tag{2.105}$$

One can now repeat the procedure until one gets

$$\langle J_2 G_{n/2}, G_{n/2} \rangle_{L_2} \qquad \text{or} \qquad \langle J_1 G_{(n+1)/2}, G_{(n+1)/2} \rangle_{L_2} \tag{2.106}$$

depending on whether n is even or odd. Because of the skew-symmetry of J_i $i = 1, 2$ both these expressions vanish and then we can apply calculus lemma. □

It is possible to show that G_k are gradients, i.e. that for every k there exists a functional F_k such that

$$G_k = \frac{\delta F_k}{\delta u}. \tag{2.107}$$

For a proof see [18]. Accepting the existence of these functionals[16], we show that they are in involution between them.

(15) We simply identified $G_n = \nabla_{L_2} F_n$.

(16) The very technical part of this proof is the existence of these functionals.

Proposition 2.4.5. *Given a Poisson tensor $J = J_1, J_2$ and a set of functionals $F_0, \dots, F_n$ conserved for the Hamiltonian system under examination satisfying (2.103), we have that functionals $F_0, \dots, F_n$ are in involution with respect to $\{\cdot,\cdot\}_J$. In other words we have*

$$\{F_n, F_m\}_J = 0 \qquad \forall m, n. \tag{2.108}$$

Proof. The case $m = n$ is straightforward since Poisson brackets are skew-symmetric. We can then assume $m > n$. Writing explicitly Poisson brackets one has, assuming $J = J_1$,

$$\{F_n, F_m\}_{J_1} = \langle \frac{\delta F_n}{\delta u}, J_1 \frac{\delta F_m}{\delta u} \rangle_{L_2}.$$

Thus one is left in proving that last quantity is vanishing. To do so we use (2.103) in a iterative manner:

$$\begin{aligned}\langle \frac{\delta F_n}{\delta u}, J_1 \frac{\delta F_m}{\delta u} \rangle_{L_2} &= \langle \frac{\delta F_n}{\delta u}, J_2 \frac{\delta F_{m-1}}{\delta u} \rangle_{L_2} = -\langle J_2 \frac{\delta F_n}{\delta u}, \frac{\delta F_{m-1}}{\delta u} \rangle_{L_2} = \\ &= -\langle J_1 \frac{\delta F_{n+1}}{\delta u}, \frac{\delta F_{m-1}}{\delta u} \rangle_{L_2} = \langle \frac{\delta F_{n+1}}{\delta u}, J_1 \frac{\delta F_{m-1}}{\delta u} \rangle_{L_2} = \\ &= \cdots = \{F_\mu, F_\mu\}_{J_1} = 0.\end{aligned} \tag{2.109}$$

Where μ can be $(m+n)/2$ or $(m+n+1)/2$ depending on $m+n$ odd or even. □

Remark 2.4.6. *The very technical part of the proof is showing that G_n are precisely gradients of F_n. We don't think this is very enlightening so we simply refer to the original paper by Lax [18].*

2.4.3 KdV hierarchy

In this section we will derive explicit expressions for the conserved functionals of KdV. These, increasing with the grade of the polynomial in u, are usually referred as *KdV hierarchy*. This subsection is then a straightforward application of recursion formula (2.103). To apply such a formula we have to choose two things: first one is which Poisson tensor we want to invert; second one we have to choose a conserved functional from which we start the computation. We choose, to invert $J_1 = \partial_x$ since it is easier to invert and to start from

$$F_0 = C \int_{\mathbb{T}} u \, dx \tag{2.110}$$

which is a Casimir for J_1 and hence it is also a conserved functional because of `Proposition 2.1.9`. We now can calculate its L_2-gradient

$$G_0 = C. \tag{2.111}$$

Applying (2.104) we get the gradient of the second order first integral

$$\partial_x G_1 = \left(a\partial_{xxx} + \frac{2}{3} bu\partial_x + \frac{b}{3} u_x \right) C. \tag{2.112}$$

From this it follows that $G_1 = \frac{bC}{3} u$ and then, inverting the L_2 gradient one obtains

$$F_1 = \frac{bC}{6} \int_{\mathbb{T}} u^2 \, dx. \tag{2.113}$$

We can iterate this procedure to get the third conserved functional

$$\partial_x G_2 = \left(a\partial_{xxx} + \frac{2}{3} bu\partial_x + \frac{b}{3} u_x \right) \frac{bC}{3} u \tag{2.114}$$

which implies, after a straightforward calculation,

$$G_2 = \frac{abC}{3} u_{xx} + \frac{b^2 C}{6} u^2. \tag{2.115}$$

Which is the L_2 gradient of

$$F_2 = \int_{\mathbb{T}} \left(\frac{abC}{6} uu_{xx} + \frac{b^2 C}{18} u^3 \right) \, dx. \tag{2.116}$$

As last step we calculate F_3 by applying once again the recursion formula:

$$\partial_x G_3 = \left(a\partial_{xxx} + \frac{2}{3} bu\partial_x + \frac{b}{3} u_x \right) \left(\frac{abC}{3} u_{xxx} + \frac{b^2 C}{3} uu_x \right). \tag{2.117}$$

This time the calculation is a bit tedious and the result is

$$\partial_x G_3 = \frac{a^2 bC}{3} u_{xxxxx} + \frac{ab^2 C}{18} \left(4uu_{xxx} + 2u_x u_{xx} + 3(u^2)_{xxx} \right) + \frac{b^3 C}{18} \left(2u(u^2)_x + u_x u^2 \right). \tag{2.118}$$

We can now integrate on the torus both sides of the equation to get

$$G_3 = \frac{a^2 bC}{3} u_{xxxx} + \frac{5}{9} ab^2 C \int_{\mathbb{T}} (uu_{xxx} + 2u_x u_{xx}) \, dx + \frac{5}{18} b^3 C \int_{\mathbb{T}} u^2 u_x \, dx \tag{2.119}$$

and, after some integration by parts, we get the following expression

$$G_3 = \frac{a^2 bC}{3} u_{xxxx} + \frac{5}{9} ab^2 C uu_{xx} + \frac{5}{18} ab^2 C (u_x)^2 + \frac{5}{54} b^3 C u^3. \tag{2.120}$$

We can see that this is the L_2-gradient of the functional

$$F_3 = \int_{\mathbb{T}} \left(\frac{a^2 bC}{6} uu_{xxxx} - \frac{5}{18} ab^2 C u (u_x)^2 + \frac{5}{216} b^3 C u^4 \right) \, dx. \tag{2.121}$$

Integrating by parts we can find the following alternative expression for F_3:

$$F_3 = \int_{\mathbb{T}} \left(\frac{a^2 bC}{6} (u_{xx})^2 + \frac{5}{36} ab^2 C u^2 u_{xx} + \frac{5}{216} b^3 C u^4 \right) \, dx. \tag{2.122}$$

With some algebra we get the following expressions for the conserved functionals

$$\begin{aligned} F_0 &= C \int_{\mathbb{T}} u \, dx \\ F_1 &= \frac{bC}{6} \int_{\mathbb{T}} u^2 \, dx \\ F_2 &= \frac{abC}{6} \int_{\mathbb{T}} \left(uu_{xx} + \frac{1}{3} \gamma u^3 \right) \, dx \\ F_3 &= \frac{a^2 bC}{6} \int_{\mathbb{T}} \left(\frac{5}{36} \gamma^2 u^4 + \frac{5}{6} \gamma u^2 u_{xx} + (u_{xx})^2 \right) \, dx \end{aligned} \tag{2.123}$$

where we called $\gamma = \frac{b}{a}$.

2.4.4 Generalized KdV equations

In the above subsection we found explicit expressions for the infinitely many conserved functionals of KdV. Each of them can be taken to be a Hamiltonian for a system which will be integrable (since is has enough conserved functionals in involution between them). The physical system related to these Hamiltonians are called *generalzied KdVs* and the corresponding equation of motion *generalized KdV equations.* Since they will be useful later we spend some time to derive the generalized KdV equations till order 2.

The generalized KdV equation of order j is characterized by

$$u_t = \partial_x \frac{\delta F_j}{\delta u}. \tag{2.124}$$

The equation involving F_0 is trivial since its gradient is a constant we have

$$u_t = C \qquad\qquad C = \text{const} \tag{2.125}$$

The one involving F_1 is the so-called *transport equation* which is, roughly speaking, half of the wave equation

$$u_t = Bu_x. \tag{2.126}$$

From F_2 one gets the KdV equation

$$u_t = \gamma u u_x + u_{xxx} \tag{2.127}$$

where we put equal to 1 the multiplicative constant in front of the equation and γ is the parameter appearing in the hierarchy. For the third functional we get, after a longer calculation than above, the following vector field up to multiplicative constants

$$u_t = \frac{5}{3}\gamma^2 u^2 u_x + \frac{20}{3}\gamma u_x u_{xx} + \frac{10}{3}\gamma u u_{xxx} + 2u_{xxxxx}. \tag{2.128}$$

2.4.5 mKdV hierarchy

Previously we showed the relation between KdV and mKdV and we deduced that if KdV is integrable also is mKdV. Now, thanks to Miura transformation, it is possible to write down explicitly the conserved functionals for mKdV. Recalling (2.84):

$$u = v^2 + i\sqrt{\frac{6}{\gamma}} v_x \tag{2.129}$$

Starting from $F_0 = \int u\,dx$ we get

$$F_0^{(m)} = F_0(v^2 + i\sqrt{\frac{6}{\gamma}} v_x) = \int v^2\,dx. \tag{2.130}$$

If we keep on going

$$F_1^{(m)} = \int \left(v^2 + i\sqrt{\frac{6}{\gamma}} v_x\right)^2 dx = \int \left(v^4 - \frac{6}{\gamma} v_x^2\right) dx. \tag{2.131}$$

Last one we are interested in is

$$
\begin{aligned}
F_2^{(m)} &= \int \left(u u_{xx} + \frac{\gamma}{3} u^3 \right) dx \bigg|_{u=v^2+i\sqrt{6\gamma^{-1}}v_x} = \\
&= \int \left(\frac{\gamma}{3} v^6 + \frac{6}{\gamma} {v_{xx}}^2 - 10 v^2 v_x^2 \right) dx.
\end{aligned}
\tag{2.132}
$$

which can be written, equivalently as

$$
F_2^{(m)} = \int \left(\gamma^2 v^6 + 10\gamma v^3 v_{xx} + 18 {v_{xx}}^2 \right) \, dx. \tag{2.133}
$$

Equations (2.131) and (2.133) are the first of a infinite series of conserved functionals for mKdV equation and they constitute the mKdV hierarchy.

CHAPTER 3

Perturbation theory

In the physical world integrable systems are very rare and hard to find while there exists a lot of system "close" to an integrable one. Examples of this are very common in nature (in celestial mechanics the Solar system, in systems of masses and springs the harmonic chain and the Toda one...). It is known that also the Fermi-Pasta-Ulam system with low-frequency modes initially excited is close to another integrable system: the KdV one. Since it is already proved that at first order FPU can be written, by a canonical transformation, as the sum of the first two term of KdV Hierarchy and numerical experiments show that the FPU actions evolve on a time-scale of the order of $\varepsilon^{-9/4}$ (where ε is the specific energy) we try to see if FPU is in KdV Hierarchy at second order. In this section we will briefly recall the main ideas of perturbation theory and we develop the tools that will be used for such a proof in the next two chapters.

3.1 Classical theory of perturbations

To introduce the classical theory of perturbations we consider a Hamiltonian system whose Hamiltonian is the sum of a first integrable piece (unperturbed part) and a second piece, small with respect to the first one, which is usually called *perturbation*. In the action-angle variables of the unperturbed system we have

$$H(I,\varphi) = h(I) + \lambda Q_1(\varphi, I) + \lambda^2 Q_2(\varphi, I) + \dots \tag{3.1}$$

where h is the integrable Hamiltonian, Q_1 and Q_2 are the perturbations and λ is a small parameter. Here we implicitly assumed that Q_1 and Q_2 have the same order in λ of h.

The aim of this theory is to answer the question whether and for how long the actions of the unperturbed system remain λ-close to the actions of the perturbed system. The classical way to answer such a question is to look for a λ-close to the identity canonical transformation[1]

$$\begin{array}{rcccl} f_\lambda & : & \mathbb{T}^n \times B & \to & \mathbb{T}^n \times B' \\ & & (\varphi, I) & \mapsto & (\vartheta, J) \end{array} \tag{3.2}$$

such that in the transformed Hamiltonian, the first term of perturbation, if not removed, is simplified as much as possible. The best one can expect is that such a new perturbation term is independent of the angles. If one is able to find such a transformation and this machinery works for the first k steps, one gets a Hamiltonian in the form

$$H(f_\lambda^{-1}(\vartheta, J)) = h(J) + \lambda Z_1(J) + \cdots + \lambda^k Z_k(J) + R_k(\vartheta, J; \lambda). \tag{3.3}$$

If one supposes or proves that the remainder at the $k-$th step satisfies $R_k = O(\lambda^{k+1})$ and $\frac{\partial R_k}{\partial \vartheta^j} = O(\lambda^{k+1})$ then the time evolution of actions is governed by

$$\dot{J}_j = \{J_j, H \circ f_\lambda^{-1}\} = \{J_j, R_k\}. \tag{3.4}$$

From this we obtain the following inequality

$$|J(t) - J(0)| \leq \int_0^t \left| \frac{\partial R_k}{\partial \vartheta^j}(\vartheta(s), J(s); \lambda) \right| ds \leq c_j \lambda^{k+1} t = (c_j \lambda) \lambda^k t \tag{3.5}$$

which directly yields $|J_j(t) - J_j(0)| = O(\lambda)$ over times $t \leq \lambda^{-k}$. On the other hand, since the canonical transformation is λ-close to the identity we get

$$\begin{aligned} |I_j(t) - I_j(0)| &= |I_j(t) - J_j(t) + J_j(t) - J_j(0) + J_j(0) - I_j(0)| \leq \\ &\leq |I_j(t) - J_j(t)| + |J_j(t) - J_j(0)| + |J_j(0) - I_j(0)| = \\ &= O(\lambda) + O(\lambda) + O(\lambda) = O(\lambda) \end{aligned} \tag{3.6}$$

again over times $t \leq \lambda^{-k}$. Thus the original actions variables undergo a slow change, so that, on a long time interval $0 \leq t < \lambda^{-k}$, the dynamics of the perturbed system resembles that of the unperturbed, integrable one.

The λ-closeness to the identity of the canonical transformation f_λ prevents one from improving the estimate on the variation of the original action variables. For example, if $|J_j(t) - J_j(0)| = O(\lambda^2)$ equation (3.6) tells us that $|I_j(t) - I_j(0)| = O(\lambda)$. This shows that, in going back to the original actions such a sharper control is lost and the best one can do is decided by the canonical transformation.

3.1.1 A first perturbative step: Lie method

A way to perform the near-to-identity canonical transformation described above is the so-called Lie method (see section 2.1.1). This method consists in finding a canonical transformation as flow of another Hamiltonian system. More precisely we set

$$f_\lambda^{-1} = \Phi_{G_1}^\lambda \tag{3.7}$$

where G_1 is the Hamiltonian generating the canonical transformation at first order. By a series expansion one gets

$$H \circ f_\lambda^{-1} = (1 + \lambda \mathcal{L}_1 + \lambda^2 \ldots)(h + \lambda P_1 + \lambda^2 \ldots) = h + \lambda(\mathcal{L}_1 h + P_1) + \lambda^2 \ldots . \tag{3.8}$$

[1] Which is a family of canonical transformations depending on a parameter λ such that for $\lambda = 0$ one obtains the identity map.

Requiring the above equation to be equal to $h + \lambda Z_1 + \lambda^2 \dots$ one has the so-called *homological equation*

$$\mathcal{L}_1 h + P_1 = Z_1. \tag{3.9}$$

If one requires the latter quantity to be independent of angles and that $h = h(I)$ one gets

$$\omega(J) \cdot \frac{\partial G_1(\vartheta, J)}{\partial \vartheta} = P_1(\vartheta, J) - Z_1(J) \tag{3.10}$$

where

$$\omega(J) = \frac{\partial h(J)}{\partial J}. \tag{3.11}$$

If we average the homological equation (3.9) on the torus we get the following explicit expression for Z_1:

$$Z_1 = \int_{\mathbb{T}^n} P_1(\vartheta, J)\, d^n\vartheta. \tag{3.12}$$

Looking at (3.12) one could think that we completed our project of finding a Z_1 independent of angles. This is not completely true since till now we proceeded supposing that there exists a Hamiltonian G_1 which is $O(1)$ and which generates the canonical transformation. Equation (3.12) represent a small perturbations only if there exists an Hamiltonian with these characteristics.

We can try to find such a Hamiltonian expanding in Fourier series the homological equation. Defining

$$\begin{aligned} \hat{G}_{1,k}(J) &= \int_{\mathbb{T}^n} G_1(J, \vartheta) e^{-2\pi i k\cdot\vartheta}\, d^n\vartheta, \\ \hat{P}_{1,k}(J) &= \int_{\mathbb{T}^n} P_1(J, \vartheta) e^{-2\pi i k\cdot\vartheta}\, d^n\vartheta, \end{aligned} \tag{3.13}$$

where $\mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n$ we can set

$$G_1(J, \vartheta) = \sum_{k\in\mathbb{Z}^n} \hat{G}_{1,k}(J) e^{2\pi i k\cdot\vartheta}, \qquad P_1(J, \vartheta) = \sum_{k\in\mathbb{Z}^n} \hat{P}_{1,k}(J) e^{2\pi i k\cdot\vartheta}. \tag{3.14}$$

Inserting these expressions in the homological equation (3.9) we get

$$ik \cdot \omega(J) \hat{G}_{1,k}(J) = \hat{P}_{1,k}(J) - \delta_{k,0} P_{1,0}(J) \tag{3.15}$$

and if $\omega \cdot k \neq 0$ its solution is

$$\hat{G}_{1,k}(J) = \frac{\hat{P}_{1,k}(J)}{ik \cdot \omega(J)}. \tag{3.16}$$

There are two important things to note here. First, we have no informations about the mean of G_1 and this is a sort of "gauge freedom" of our problem. Second by, we cannot find $\hat{G}_{1,\tilde{k}}$ of order $O(1)$ if $\tilde{k}\cdot\omega(J) \sim \lambda$. In fact, if this happens, we have that $\hat{G}_{1,\tilde{k}} \sim \lambda^{-1}$.[(2)] In the literature this is referred as the *small divisors* problem and one refers to the

[(2)] At a deeper level, considering the vector fields of the transformations, one finds out that, in general, dangerous small denominators are those satisfying $\tilde{k} \cdot \omega(J) \sim \sqrt{\lambda}$.

relations $\tilde{k} \cdot \omega(J) \sim \lambda$ as to the *resonances* of the system. If we call $K^*(J)$ the set of all $\tilde{k}$ such that the divisor is very dangerous, i.e. $K^*(J) = \{k \in \mathbb{Z}^n : \omega(J) \cdot k = 0\}$,(3) we get

$$
\begin{aligned}
G_1(J, \vartheta) &= \sum_{k \in \mathbb{Z}^n \setminus K^*(J)} \frac{\hat{P}_{1,k}(J)}{ik \cdot \omega(J)} e^{2\pi i k \cdot \vartheta}, \\
Z_1(J, \vartheta) &= \int_{\mathbb{T}} P_1(\vartheta, J)\, d\vartheta + \sum_{k \in K^*(J)} \hat{P}_{1,k}(J) e^{2\pi i k \cdot \vartheta}
\end{aligned}
\tag{3.17}
$$

and we see that in general it is not possible to have a normal form at first order independent of the angles. We see also that in general, since $K^*(J)$ depends on J, the normal form can vary from a point to another one in the phase space.

To better understand the construction in the next section we have to write the expressions for G_1 and Z_1 in (3.17) in terms of time average.

We start from the following definition

Definition 3.1.1 (Quasi-periodic motion(4))**.** Let $\mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n$ be the n-dimensional torus and $\varphi = (\varphi_1, \dots, \varphi_n)$ angular coordinates.(5) Then by a *quasi periodic motion* we mean the flow $\mathbb{T}^n \times \mathbb{R} \to \mathbb{T}^n$ of the following differential equations

$$
\dot{\varphi} = \omega \qquad\qquad \omega = (\omega_1, \dots, \omega_n) = \text{const.} \tag{3.18}
$$

These differential equations are easily integrated and this integration gives

$$
\varphi(t) = \varphi(0) + \omega t. \tag{3.19}
$$

The quantities $\omega_1, \dots, \omega_n$ are called frequencies of the conditionally periodic motion. Frequencies are called (rationally) *independent* if they are linearly independent over the field of rational numbers, i.e. if $k \in \mathbb{Z}^n$ and $k \cdot \omega = 0$ implies $k = 0$.

Let now $f(\varphi)$ be an integrable function on the torus $\mathbb{T}^n$. We recall the following

Definition 3.1.2 (Space average)**.** The space average of a function f on the torus $\mathbb{T}^n$ is the number

$$
\langle f \rangle_{\mathbb{T}^n} = \int_{\mathbb{T}^n} f(\varphi)\, d^n\varphi \tag{3.20}
$$

If we now consider the value of the function $f(\varphi)$ on the trajectory of $\varphi(t) = \varphi(0) + \omega t$ we get a function of time. We can consider its time average

Definition 3.1.3 (Time average)**.** The time average of the function f on the torus $\mathbb{T}^n$ is the function

$$
f(\varphi_0) = \lim_{t \to \infty} \frac{1}{t} \int_0^t f(\varphi_0 + \omega s)\, ds \tag{3.21}
$$

defined where the limit exists.

(3)It is obvius that K^* depends on J if $\omega(J)$ depends on J.

(4)Sometimes one refers to these as *conditionally-periodic motions*, see for example [1].

(5)Our angular coordinates are $\varphi \in [0, 1)^n$.

Theorem 3.1.4. *The time average exists everywhere, and coincides with the space average if f is continuous (or merely Riemann-integrable*(6)*) and the frequencies ω_i are rationally independent.*

Proof. The proof can be found in [1] paragraph 51. □

This result is interesting because it means that we can write Z_1 exactly as a time average. Since the motion of the integrable system is conditionally periodic we can state

$$Z_1(J) = \lim_{t\to\infty} \frac{1}{t} \int_0^t e^{s\mathcal{L}_h} P_1(J,\vartheta)\, ds \tag{3.22}$$

if the frequencies are independent. If the frequencies are dependent we have $\omega(J)\cdot k = 0$ for some $k \neq 0$ and this is precisely the definition of resonant frequence. One finds

$$Z_1(J,\vartheta) = \lim_{t\to\infty} \frac{1}{t} \int_0^t e^{s\mathcal{L}_h} P_1(J,\vartheta)\, ds = \int_{\mathbb{T}} P_1(\theta, J)\, d\theta + \sum_{k\in K^*(J)} \hat{P}_{1,k}(J) e^{2\pi i k\cdot\vartheta}. \tag{3.23}$$

One can repeat the same argument to find an expression for G_1. Anyway this is not very important for the theory of perturbations above and, since our plan is to connect the classical theory of perturbations and the one in the next section, equation (3.23) is sufficient for our scope. To be more concrete we perform the calculations described in this section in an example.

Example. Let us calculate the first order normal form for a Hamiltonian system in $\mathbb{T}^2 \times \mathbb{R}^2$ with Hamiltonian

$$H = I_1 + I_2 + \varepsilon(\cos(\varphi_1 + \varphi_2) + \cos(\varphi_1 - \varphi_2)). \tag{3.24}$$

It is straightforward to calculate the normal form using (3.17) once one notes that the only resonant frequency is $\omega = (1,1)$. One gets

$$\tilde{H} = I_1 + I_2 + \varepsilon\cos(\varphi_1 - \varphi_2). \tag{3.25}$$

To use the time-average machinery one has to write the unperturbed equations of motion which are

$$\dot{\varphi}_1 = 1, \qquad \dot{\varphi}_2 = 1, \qquad \dot{I}_1 = 0, \qquad \dot{I}_2 = 0. \tag{3.26}$$

One then calculate the time average of the perturbation $P_1 = \cos(\varphi_1+\varphi_2)+\cos(\varphi_1-\varphi_2)$ using the unperturbed flux which is $\phi(t) = \phi_0 + t$:

$$P_1 = \lim_{t\to\infty} \frac{1}{t} \int_0^t (\cos(\varphi_1 + \varphi_2 + 2t) + \cos(\varphi_1 - \varphi_2))\, dx. \tag{3.27}$$

Due to the linearity of the integral one can split it. The first term vanishes while the second one remains the same. One thus get precisely (3.23)

$$\tilde{H} = I_1 + I_2 + \varepsilon\cos(\varphi_1 - \varphi_2) \tag{3.28}$$

which is the same result obtained before in the "standard way".

(6) It is sufficient Lebesgue-integrability.

3.2 Infinite-dimensional Hamiltonian perturbation theory

The perturbation theory exposed above is deeply linked to the action-angle variables. In this chapter we develop the canonical standard formalism with the aim of finding expressions independent of the choice of coordinates. This is very useful in order to treat perturbations for which expressions for action-angle variables are not friendly to handle.

3.2.1 Hamiltonian normal form

We start from a perturbed Hamiltonian system whose Hamiltonian in in the form

$$H(x) = h(x) + \lambda P_1(x) + \lambda^2 P_2(x) + \dots \tag{3.29}$$

where x is a generical variable. Our goal is to simplify as much as possible the perturbative series bringing it in the so-called *normal form*. This means that we will look for a near-to-identity canonical transformation which maps (3.29) into

$$H \circ f^{-1} = h + \lambda S_1 + \lambda^2 S_2 + \dots \tag{3.30}$$

where $S_1, S_2, \dots$ "commute" with h. More precisely we say that H is in *normal form to λ-th order* if $\{S_j, h\} = 0$ for every $j = 1, \dots, \lambda$.

To outline the procedure in a concrete way we perform a canonical transformation setting H in normal form to second order. If one is interested in a j-th order normal form one can generalize the procedure outlined here.

To perform the near-to-identity canonical transformation we use Lie method and we call K the transformed Hamiltonian. We thus get

$$K = e^{\lambda^2 \mathcal{L}_2} e^{\lambda \mathcal{L}_1} H \tag{3.31}$$

where, as in the previous section, we denoted with G_j the Hamiltonian generating the canonical trasformation at the j-th order and with $\mathcal{L}_j(\cdot)$ the Lie derivative associated to it, $\mathcal{L}_j(\cdot) = \{\cdot, G_j\}$. K will be in the form

$$K = h + \lambda S_1 + \lambda^2 S_2 + O(\lambda^3). \tag{3.32}$$

Let us perform the calculations in detail to find explicit expressions for S_1 and S_2. Expanding in series the exponentials in (3.31) till the second order we get

$$\begin{aligned} K = e^{\lambda^2 \mathcal{L}_2} e^{\lambda \mathcal{L}_1} H &= \left(1 + \lambda^2 \mathcal{L}_2\right) \left(1 + \lambda \mathcal{L}_1 + \frac{\lambda^2}{2} \mathcal{L}_1^2\right) \left(h + \lambda P_1 + \lambda^2 P_2\right) + \dots = \\ &= \left(1 + \lambda^2 \mathcal{L}_2\right) \left(h + \lambda (P_1 + \mathcal{L}_1 h) + \lambda^2 \left(P_2 + \frac{1}{2} \mathcal{L}_1^2 h + \mathcal{L}_1 P_1\right)\right) + \dots = \\ &= h + \lambda \left(P_1 + \mathcal{L}_1 h\right) + \lambda^2 \left(\mathcal{L}_2 h + P_2 + \frac{1}{2} \mathcal{L}_1^2 h + \mathcal{L}_1 P_1\right) + \dots \end{aligned} \tag{3.33}$$

This defines in a natural way the perturbative series for K with

$$\begin{aligned} S_1 &= \mathcal{L}_1 h + P_1 \\ S_2 &= \frac{1}{2}\mathcal{L}_1^2 h + \mathcal{L}_1 P_1 + P_2 + \mathcal{L}_2 h. \end{aligned} \tag{3.34}$$

From above we notice the following general feature which is supposed to be true at every order in λ: we can always write the j-th part of the perurbation as

$$S_j = \mathcal{L}_j h + P_j^\natural \qquad P_j^\natural = P_j + \mathscr{F}(P_{j-1}, \ldots, P_1). \tag{3.35}$$

Now we apply $e^{s\mathcal{L}_h}$ at both sides of the above equation. If we require S_j to be in normal form with respect to h we have $\{S_j, h\} = 0$ which implies that S_j doesn't change under the flow of h: $e^{s\mathcal{L}_h} S_j = S_j$. Using this property we can find G_j and S_j:

$$S_j = e^{s\mathcal{L}_h} S_j = e^{s\mathcal{L}_h}\left(\mathcal{L}_j h + e^{s\mathcal{L}_h} P_j^\natural\right). \tag{3.36}$$

The skew-symmetric property of the Poisson bracket tells us that $\mathcal{L}_j h = \{h, G_j\} = -\{G_j, h\} = -\mathcal{L}_h G_j$ and the above expression yields

$$S_j = e^{s\mathcal{L}_h} S_j = -e^{s\mathcal{L}_h}\mathcal{L}_h G_j + e^{s\mathcal{L}_h} P_j^\natural = -\frac{d}{ds}\left(e^{s\mathcal{L}_h} G_j\right) + e^{s\mathcal{L}_h} P_j^\natural. \tag{3.37}$$

We integrate now both members between 0 and t, then we divide by t to obtain an equation for S_j:

$$S_j = -\frac{1}{t}\left(e^{t\mathcal{L}_h} G_j - G_j\right) + \frac{1}{t}\int_0^t e^{s\mathcal{L}_h} P_j^\natural \, ds. \tag{3.38}$$

A smart way to choose t in the integration extrema is choosing, if it exists, a t^* such that the first term vanishes. In the following two cases this is for sure possible:

- *periodic flow:* if the flow of h is periodic with period t^* then $e^{t^*\mathcal{L}_h} = 1$. Under this hypotesis (3.38) becomes
$$S_j = \frac{1}{t^*}\int_0^{t^*} e^{s\mathcal{L}_h} P_j^\natural \, ds. \tag{3.39}$$
- *bounded flow:* if the flow of h is bounded for every t the first term vanishes. It follows that
$$S_j = \lim_{t\to\infty}\frac{1}{t}\int_0^t e^{s\mathcal{L}_h} P_j^\natural \, ds. \tag{3.40}$$

Comparing now (3.40) with (3.23) one can recognize some similarities. Anyway we get something more here since we didn't require the starting system to be in action-angle variables.

Assuming that our system satisfies one of the previous condition, from (3.40) or (3.39) we have now the perturbation in normal form, that is

$$S_j = \langle P_j^\natural \rangle_h \tag{3.41}$$

where we denoted with $\langle\cdot\rangle_h = \lim_{t\to\infty}\frac{1}{t}\int_0^t(\cdot)ds$.

As we see from (3.34) to calculate the j-th order we need all Hamiltonians generating the canonical transformations till $(j-1)$-th order. Looking for such a Hamiltonian we start from (3.35):

$$S_j = -\mathcal{L}_h G_j + P_j^\natural \tag{3.42}$$

that, combined together to the expression above for S_j, reads

$$\mathcal{L}_h G_j = P_j^\natural - \langle P_j^\natural \rangle_h. \tag{3.43}$$

For the simplicty of notations we define $\tilde{P}_j^\natural = P_j^\natural - \langle P_j^\natural \rangle_h$. Our problem consists in the inversion of (3.43).

Let us multiply both sides of (3.43) by $(s-t^*)$ and integrate them between 0 and t^*. We get

$$\int_0^{t^*} (s-t^*) e^{s\mathcal{L}_h} \mathcal{L}_h G_j \, ds = \int_0^{t^*} (s-t^*) e^{s\mathcal{L}_h} \tilde{P}_j^\natural \, ds \tag{3.44}$$

We recognize on the left side that $\frac{d}{ds}\left(e^{s\mathcal{L}_h} G_j\right) = e^{s\mathcal{L}_h} \mathcal{L}_h G_j$ and then,

$$\begin{aligned} \text{LHS} &= \int_0^{t^*} (s-t^*) e^{s\mathcal{L}_h} \mathcal{L}_h G_j \, ds = \int_0^{t^*} (s-t^*) \frac{d}{ds}\left(e^{s\mathcal{L}_h} G_j\right) ds = \\ &= (s-t^*) e^{s\mathcal{L}_h} G_j \Big|_{s=0}^{s=t^*} - \int_0^{t^*} e^{s\mathcal{L}_h} G_j \, ds = t^* G_j - \int_0^{t^*} e^{s\mathcal{L}_h} G_j \, ds \end{aligned} \tag{3.45}$$

After equating the left-hand side to the right-hand side, with some elementary algebra we get

$$G_j = \frac{1}{t^*} \int_0^{t^*} e^{s\mathcal{L}_h} G_j \, ds + \frac{1}{t^*} \int_0^{t^*} (s-t^*) e^{s\mathcal{L}_h} \tilde{P}_j^\natural \, ds. \tag{3.46}$$

Using the above notations, this last expressions reads

$$G_j = \langle G_j \rangle_h + \frac{1}{t^*} \int_0^{t^*} (s-t^*) e^{s\mathcal{L}_h} \tilde{P}_j^\natural \, ds. \tag{3.47}$$

As in the classical case we see that we have no informations about $\langle G_j \rangle_h$ (recall equation (3.16)).

3.2.2 Some semplifications for the second order

Since the aim of the following chapters is to look for a canonical transformation to bring the FPU Hamiltonian in normal form to the second order, we dedicate this subsection to simplify as much as possible the expression for S_2.

Our starting point is given by equations (3.34) and (3.41). Putting together these equations we get

$$\begin{aligned} S_1 &= \langle P_1 \rangle_h \\ S_2 &= \langle P_2 + \mathcal{L}_1 P_1 + \frac{1}{2} \mathcal{L}_1^2 h \rangle_h. \end{aligned} \tag{3.48}$$

While the first order expression cannot be simplified more, there is a smart way to write the expression for S_2.

The way to do it is very simple: it consists in splitting every functional U in an averaged part $\langle U\rangle_h$ and in a zero-mean part $\tilde{U}$:

$$U = \langle U\rangle_h + \tilde{U}. \tag{3.49}$$

Doing these splittings on the second term of the RHS of (3.48) we have

$$\begin{aligned}\mathcal{L}_1 P_1 = \{P_1, G_1\} &= \{\langle P_1\rangle_h + \tilde{P}_1, \langle G_1\rangle_h + \tilde{G}_1\} = \\ &= \{\langle P_1\rangle_h, \langle G_1\rangle_h\} + \{\tilde{P}_1, \langle G_1\rangle_h\} + \{\langle P_1\rangle_h, \tilde{G}_1\} + \{\tilde{P}_1, \tilde{G}_1\}\end{aligned} \tag{3.50}$$

Since they have to be averaged along the flow of h, we get that all the terms with only one fluctuating (i.e. "tilded") function will be zero after the average. Thus

$$\langle \mathcal{L}_1 P_1\rangle_h = \{\langle P_1\rangle_h, \langle G_1\rangle_h\} + \langle\{\tilde{P}_1, \tilde{G}_1\}\rangle_h. \tag{3.51}$$

Doing the same procedure on the third term we get

$$\mathcal{L}_1^2 h = \mathcal{L}_1\mathcal{L}_1 h = -\mathcal{L}_1\tilde{P}_1 = -\{\tilde{P}_1, G_1\} = -\{\tilde{P}_1, \tilde{G}_1\} - \{\tilde{P}_1, \langle G_1\rangle_h\}. \tag{3.52}$$

Averaging both sides of the preceding equation we get

$$\langle \mathcal{L}_1^2 h\rangle_h = -\langle\{\tilde{P}_1, \tilde{G}_1\}\rangle_h. \tag{3.53}$$

Inserting both semplifications in (3.48) we obtain

$$S_2 = \langle P_2 + \frac{1}{2}\{\tilde{P}_1, \tilde{G}_1\}\rangle_h + \{\langle P_1\rangle_h, \langle G_1\rangle_h\}. \tag{3.54}$$

This is the final expression for S_2.

3.3 Normal form of vector fields

In this section we present the perturbative approach to general (i.e. not necessarily Hamiltonian) differential equations.(7) Because of `Proposition 2.1.10`, if the starting system is Hamiltonian, also the transformed system is Hamiltonian but with a different Poisson tensor, in general.

Let us start from the autonomous system of differential equations on a manifold M

$$\dot{x} = X(x;\lambda) \tag{3.55}$$

where x is the coordinate on M, $X \in TM$ is a vector field and λ is a small parameter in the system we are studying. We can expand X in series of λ:

$$X(x) = X_0(x) + \lambda X_1(x) + \lambda^2 X_2(x) + \dots \tag{3.56}$$

We are thus left with a natural perturbative series expansion of the vector field X. Roughly speaking we now ask if it is possible to perform a coordinate transformation such that the form of the perturbative series is "as simple as possible" for the transformed

(7) We will restrict ourselves to autonomous systems.

vector field Y. Pratically we require that the transformed field Y is in *normal form* with respect to X_0 to a certain order m:

$$[Y_k, X_0] = 0 \qquad\qquad k < m \tag{3.57}$$

where the square brackets denote the Lie bracket on the tangent bundle. In a coordinate system and using Einstein sum convention this reads

$$[X, Y]^i = \frac{\partial X^i}{\partial x^j} Y^j - \frac{\partial Y^i}{\partial x^j} X^j. \tag{3.58}$$

The relation between the definition of normal form with Lie brackets and with Poisson brackets is discussed below in subsection 3.3.3.

We are then looking for a "near to the identity" diffeomorphism[(8)] in the form

$$x(y) = y + \lambda f_1(y) + \lambda^2 f_2(y) + \dots \tag{3.59}$$

such that the transformed differential equations

$$\dot{y} = Y_0(y) + \lambda Y_1(y) + \lambda^2 Y_2(y) + \dots \tag{3.60}$$

are in normal form to every fixed order.

3.3.1 Transformation rule for vector fields

In order to find explicit expressions for Y_k we need to find the transformation rule for vector fields under the diffeomorphism (3.59). For the sake of simplicity we will use Einstein sum convention in this section.

At first we substitute (3.59) on the left hand side of the differential equation (3.55). We get

$$\dot{x}^i = \frac{d}{dt}\left(y^i + \lambda f_1^i(y) + \lambda^2 f_2^i(y) + \dots\right) = \left(\delta^i_j + \lambda \frac{\partial f_1^i}{\partial y^j}(y) + \lambda^2 \frac{\partial f_2^i}{\partial y^j}(y) + \dots\right) \dot{y}^j \tag{3.61}$$

where we used the chain rule. Substituting on the right-hand side we get

$$X_0(x(y)) + \lambda X_1(x(y)) + \lambda^2 X_2(x(y)) + \dots. \tag{3.62}$$

Expanding each term we obtain for the first one

$$\begin{aligned} X_0^i(x(y)) &= X_0^i(y + \lambda f_1(y) + \lambda^2 f_2(y) + \dots) = \\ &= X_0^i(y) + \lambda \frac{\partial X_0^i}{\partial y^j}(y) f_1^j(y) + \lambda^2 \left(\frac{\partial X_0^i}{\partial y^j} f_2^j(y) + \frac{1}{2}\frac{\partial^2 X_0^i}{\partial y^k \partial y^j} f_1^k f_1^j\right) + \dots \end{aligned} \tag{3.63}$$

for the second one

$$\lambda X_1^i(y + \lambda f_1(y) + \dots) = \lambda X_1^i(y) + \lambda^2 \frac{\partial X_1^i}{\partial y^j} f_1^j(y) + \dots \tag{3.64}$$

[(8)] Which is a one-parameter family of diffeomorphism depending on a parameter λ and such that it reduces to identity map if $\lambda = 0$.

and the last one remain unchanged to order λ^2. Grouping terms with the same order in λ we get

$$X^i(x(y);\lambda) = X_0^i(y) + \lambda\left(\frac{\partial X_0^i}{\partial y^j}(y) f_1^j(y) + X_1^i(y)\right) + \\ + \lambda^2\left(\frac{\partial X_0^i}{\partial y^j} f_2^j(y) + \frac{1}{2}\frac{\partial^2 X_0^i}{\partial y^k \partial y^j} f_1^k f_1^j + X_2^i(y) + \frac{\partial X_1^i}{\partial y^j} f_1^j(y)\right). \tag{3.65}$$

We still don't have the transformation law for vector fields since the LHS is not composed only by $\dot{y}$. We have then to invert

$$Df = \delta_j^i + \lambda\frac{\partial f_1^i}{\partial y^j}(y) + \lambda^2\frac{\partial f_2^i}{\partial y^j}(y) + \dots \tag{3.66}$$

Let us now suppose that

$$Dg = \delta_j^i + \lambda {g_1}_j^i + \lambda^2 {g_2}_j^i + \dots \tag{3.67}$$

is the inverse of Df, i.e. $Dg\,Df = \mathbf{1}$ where $\mathbf{1}$ is the identity. We can use this relation to find an expression for g_1 and g_2 as function of f_1 and f_2. This is

$$\left(\delta_j^i + \lambda\frac{\partial f_1^i}{\partial y^j}(y) + \lambda^2\frac{\partial f_2^i}{\partial y^j}(y) + \dots\right)\left(\delta_k^j + \lambda {g_1}_k^j + \lambda^2 {g_2}_k^j + \dots\right) = \delta_k^i. \tag{3.68}$$

Computing the LHS of this expression we get

$$\delta_k^i + \lambda\left(\frac{\partial f_1^i}{\partial y^k} + {g_1}_k^i\right) + \lambda^2\left(\frac{\partial f_2^i}{\partial y^k} + \frac{\partial f_1^i}{\partial y^j}{g_1}_k^j + {g_2}_k^i\right) + \dots = \delta_k^i. \tag{3.69}$$

So we have to solve equations

$$\frac{\partial f_1^i}{\partial y^k} + {g_1}_k^i = 0, \qquad \frac{\partial f_2^i}{\partial y^k} + \frac{\partial f_1^i}{\partial y^j}{g_1}_k^j + {g_2}_k^i = 0. \tag{3.70}$$

which give

$${g_1}_j^i = -\frac{\partial f_1^i}{\partial y^j}, \qquad {g_2}_j^i = \frac{\partial f_1^i}{\partial y^k}\frac{\partial f_1^k}{\partial y^j} - \frac{\partial f_2^i}{\partial y^j}. \tag{3.71}$$

Thus we get:

$$Dg_j^i = \delta_j^i - \lambda\frac{\partial f_1^i}{\partial y^j} + \lambda^2\left(\frac{\partial f_1^i}{\partial y^k}\frac{\partial f_1^k}{\partial y^j} - \frac{\partial f_2^i}{\partial y^j}\right) + \dots \tag{3.72}$$

From (3.61), (3.65) and (3.72) we have

$$\dot{y}^i = Dg_j^i\left(X_0^j + \lambda\left(\frac{\partial X_0^j}{\partial y^k} f_1^k + X_1^j\right) + \lambda^2\left(\frac{\partial X_0^j}{\partial y^k} f_2^k + \frac{1}{2}\frac{\partial^2 X_0^j}{\partial y^k \partial y^l} f_1^k f_1^l + X_2^j + \frac{\partial X^j}{\partial y^k} f_1^k\right)\right). \tag{3.73}$$

Order by order in λ we obtain

$$Y_0^i(y) = X_0^i(y) \tag{3.74}$$

and then

$$Y_1^i = \frac{\partial X_0^i}{\partial y^j} f_1^j - \frac{\partial f_1^i}{\partial y^j} X_0^j + X_1^i = [X_0, f_1]^i + X_1^i \tag{3.75}$$

For the second order we get

$$
\begin{aligned}
Y_2^i = {} & \frac{\partial X_0^i}{\partial y^j} f_2^j + \frac{1}{2}\frac{\partial^2 X_0^i}{\partial y^j \partial y^k} f_1^j f_1^k + X_2^i - \frac{\partial f_1^i}{\partial y^j}\left(\frac{\partial X_0^j}{\partial y^k} f_1^k + X_1^j\right) + \frac{\partial X^i}{\partial y^k} f_1^k + \\
& + \left(\frac{\partial f_1^i}{\partial y^j}\frac{\partial f_1^j}{\partial y^k} - \frac{\partial f_2^i}{\partial y^k}\right) X_0^k = \\
& = \frac{\partial X_0^i}{\partial y^j} f_2^j + \frac{1}{2}\frac{\partial^2 X_0^i}{\partial y^j \partial y^k} f_1^j f_1^k + X_2^i - \frac{\partial f_1^i}{\partial y^j}\left(\frac{\partial X_0^j}{\partial y^k} f_1^k - \frac{\partial f_1^j}{\partial y^k} X_0^k + X_1^j\right) - \\
& - \frac{\partial f_2^i}{\partial y^k} X_0^k + \frac{\partial X^i}{\partial y^k} f_1^k = \\
& = [X_0, f_2]^i - \frac{\partial f_1^i}{\partial y^j} Y_1^j + \frac{\partial X_1^i}{\partial y^j} f_1^j + \frac{1}{2}\frac{\partial^2 X_0^i}{\partial y^j \partial y^k} f_1^j f_1^k + X_2^i
\end{aligned}
\tag{3.76}
$$

At the end we have

$$\dot{y} = Y_0 + \lambda Y_1 + \lambda^2 Y_2 + \dots \tag{3.77}$$

With

$$
\begin{gathered}
Y_0^i = X_0^i \qquad Y_1^i = [X_0, f_1]^i + X_1^i \\
Y_2^i = [X_0, f_2]^i - \frac{\partial f_1^i}{\partial y^j} Y_1^j + \frac{\partial X_1^i}{\partial y^j} f_1^j + \frac{1}{2}\frac{\partial^2 X_0^i}{\partial y^j \partial y^k} f_1^j f_1^k + X_2^i.
\end{gathered}
\tag{3.78}
$$

These are the expressions for the transformed vector fields under the diffeomorphism (3.59).

We see clearly that if we choose $\tilde{f}_1 = f_1 + g$ with g such that $[X_0, g] = 0$ obtain a diffeomorphism which leaves unchanged the transformed vector field to first order but changes the second order. We will refer to this freedom as *gauge freedom*.

3.3.2 Continuous case

We can transport the above construction to the infinite dimensional case making suitable changes. That is replacing partial derivatives respect to y with weak differentiations with respect to the variables we are considering. Since in the FPU system we will deal with two variables, V^+ and V^- (see chapter 4), we focus our treatment on this case.

As an example we see that after these substitutions, the Lie bracket definition reads $(i = +, -)$:

$$[X, Y]^i = \frac{DX^i}{DV^j} Y^j - \frac{DY^i}{DV^j} X^j. \tag{3.79}$$

If we begin with an infinite dimensional dynamical system with equations of motion

$$V_t^i = X_0^i(V) + \lambda X_1^i(V) + \lambda^2 X_2^i(V) + \dots \tag{3.80}$$

we can repeat the above construction introducing the diffeomorphism

$$V(U) = U + \lambda f_1(U) + \lambda^2 f_2(U) + \dots . \tag{3.81}$$

We then get, for the transformed equations of motion:

$$U_t^i = Y_0^i(U) + \lambda Y_1^i(U) + \lambda^2 Y_2^i(U) + \dots. \tag{3.82}$$

We save paper in writing directly that

$$\begin{gathered} Y_0^i = X_0^i \qquad Y_1^i = [X_0, f_1]^i + X_1^i \\ Y_2^i = [X_0, f_2]^i - \frac{Df_1^i}{DV^j} Y_1^j + \frac{DX_1^i}{DV^j} f_1^j + \frac{1}{2} \frac{D^2 X_0^i}{DV^j DV^k} f_1^j f_1^k + X_2^i. \end{gathered} \tag{3.83}$$

We will be interested, later, in the case of $X_1 = Y_1$ and $D^2 X_0^i = 0$. In this case the expression for Y_2 simplifies and it reads

$$Y_2^i = [X_0, f_2]^i + [X_1, f_1]^i + X_2^i. \tag{3.84}$$

Also in these expressions we see that it is possible to introduce a function g which, added to f_1, does not modify the equations of motion to first order if $[X_0, g] = 0$. This is again the same gauge freedom we found in the previous section, but for the infinite dimensional case.

3.3.3 Relation between normal form of vector field and Hamiltonian normal form

In the section above we said that two functionals are in normal form when the Poisson-commute while two vector fields are in normal form one with respect to the other if they Lie-commute. One is justified in calling both them normal forms because of `Proposition 2.1.18`. In fact a straightforward corollary of that proposition is the following

Proposition 3.3.1. *Let X, Y vector fields defined on a Poisson manifold Γ and let us suppose there exist a couple of functionals $A, B : \mathcal{A}(\Gamma) \to \mathbb{R}$ such that*

$$X = J\nabla A, \qquad Y = J\nabla B \tag{3.85}$$

then

$$\{A, B\} = 0 \qquad \Leftrightarrow \qquad [X, Y] = 0. \tag{3.86}$$

This is why one is justified in calling both them normal forms.

By the way, such a treatment, is valid if and only if one has Hamiltonian vector fields. Once one leaves this constraint one can have non-Hamiltonian vector fields in normal form with a given Hamiltonian vector field. More precisely we have two vector fields which Lie-commute but we don't have a correspondent relation between functionals.

CHAPTER 4

Normal form of the $\alpha + \beta$-FPU system

In this chapter we will use all the techniques presented above to build the normal form Hamiltonian of the Fermi-Pasta-Ulam system to second order. Our aim is to prove that such a normal form consists of the first few terms of the KdV hierarchy. In particular, the FPU normal form to the second order is integrable.

4.1 Extension to continuum

The aim of this section is to include the Hamiltonian for the Fermi-Pasta-Ulam system and its equations in a continuous model.

We start from (1.6):

$$H_{FPU}(q_n, p_n) = \sum_{n \in \mathbb{Z}_N} \left[\frac{p_n^2}{2} + \phi(q_{n+1} - q_n) \right]. \tag{4.1}$$

Due to translational invariance, the flow of the system preserves the total momentum. As a consequence it is a constant of motion which is zero in the mass center system:

$$\sum_{n \in \mathbb{Z}_N} p_n = 0. \tag{4.2}$$

Since Hamiltonian (4.1) depends on q coordinates only through their difference we can consider new coordinates r_n such that

$$r_n := q_{n+1} - q_n. \tag{4.3}$$

If we want to preserve the Hamiltonian structure use these new coordinates we have to perform a canonical transformation $(q, p) \mapsto (s, r)$. Here s_n is the conjugate variable to r_n. Such a transformation can be found defining the generating function[(1)]

$$F(q, s) = \sum_{n \in \mathbb{Z}_N} s_n(q_n - q_{n+1}). \tag{4.4}$$

[(1)]See section 2.1.1.

Its differential is given by $dF = \sum_{n\in\mathbb{Z}_N}(p_n dq_n - r_n ds_n)$ and the transformation is found by solving

$$r_n = -\frac{\partial F}{\partial s_n} \qquad p_n = \frac{\partial F}{\partial q_n}. \tag{4.5}$$

This is simply due to the form of F and one can find s_n by inverting

$$p_n = s_n - s_{n-1}. \tag{4.6}$$

We have to notice that, among the new variables, the r_n play the role of momenta while the s_n play the role of coordinates.[2] New momenta are periodic by construction and thus, thanks to (1.20), one gets

$$\sum_{n\in\mathbb{Z}_N} r_n = 0. \tag{4.7}$$

After this change of variables, the transformed Hamiltonian (which we denote by H) takes the form:

$$H(s_n, r_n) = \sum_{n\in\mathbb{Z}_N}\left[\phi(r_n) + \frac{(s_{n+1}-s_n)^2}{2}\right]. \tag{4.8}$$

Its equations of motion are

$$\begin{cases} \dot{s}_n = \dfrac{\partial H}{\partial r_n} = \phi'(r_n) \\ \dot{r}_n = -\dfrac{\partial H}{\partial s_n} = s_{n+1} + s_{n-1} - 2s_n. \end{cases} \tag{4.9}$$

Next we have to find the physical dimension of our variables. With this purpose we can start from (4.1) and, denoting in this section with square brackets the physical dimension, calculate $[p_n]$ and $[q_n]$. From (4.8)

$$[H] = \left[\sum_{n\in\mathbb{Z}_N}\left(\frac{p_n^2}{2} + \frac{(q_{n+1}-q_n)^2}{2}\right)\right]. \tag{4.10}$$

Denoting with $N = h^{-1}$ the total number of masses in the chain, with E the total energy in the chain and with $\varepsilon = E/N$ the energy per degree of freedom, we have $[\sum_{n\in\mathbb{Z}_N}] = N$, $[H] = E$ and $[q_{n+1} - q_n] \sim [\frac{\partial q_n}{\partial x} h] = [q][h]$. We thus obtain

$$[p_n] = \sqrt{\varepsilon} \qquad [q_n] = \frac{\sqrt{\varepsilon}}{h}. \tag{4.11}$$

From (4.6) we get

$$[p_n] = \sqrt{\varepsilon} = [s_{n+1} - s_n] \approx \left[\frac{\partial s_n}{\partial x} h\right] \qquad [s_n] = \frac{\sqrt{\varepsilon}}{h} \tag{4.12}$$

Thus one can obtain the physical dimension of r_n calculating the physical dimension of $[F]$ or from the transformed Hamiltonian. The second one, by an analogy of the previous case, yields immediately to

$$[r_n] = [p_n] = \sqrt{\varepsilon}. \tag{4.13}$$

[2] This is obvious if one looks at the form of dF and recalls `Proposition 2.1.13`.

We assume now there exist a couple of dimensionless analytical functions $R, S : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ such that

$$\begin{cases} r_n(t) = \sqrt{\varepsilon} R(x,\tau)|_{x=hn,\tau=\tau(t)} \\ s_n(t) = \frac{\sqrt{\varepsilon}}{h} S(x,\tau)|_{x=hn,\tau=\tau(t)}. \end{cases} \tag{4.14}$$

As one can see from equations above, the last quantity we have to investigate is time: we are left in finding a dimensionless time τ. For this purpose we have to use the equations of motion of the system (4.9). From the first one we get

$$\frac{ds_n(t)}{dt} = \frac{d\phi(r_n)}{dr_n} \tag{4.15}$$

where the physical dimensions of ϕ are easily found from (4.1) to be $[\phi] = \varepsilon$. As a consequence we obtain for the ratio $[\phi/r_n] = \sqrt{\varepsilon}$. On the other side we have, from a dimensional point of view,

$$\frac{1}{[t]} \frac{\sqrt{\varepsilon}}{h} = \sqrt{\varepsilon}. \tag{4.16}$$

This implies that $[t] = h^{-1}$ and we can define a dimensionless time as

$$\tau = ht. \tag{4.17}$$

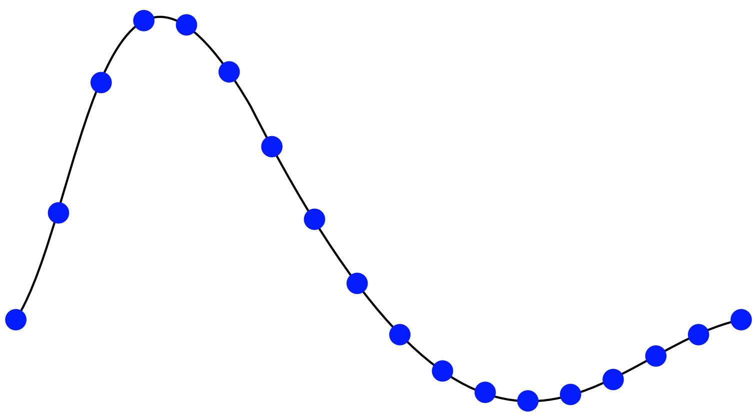

Figure 4.1: *The idea below the study of the interpolating system.*

We can now substitute (4.14) into (4.9) and transform all the time derivatives in derivatives with respect to τ. From the first substitution one gets

$$\begin{cases} \frac{\sqrt{\varepsilon}}{h} \frac{\partial S(x,\tau)}{\partial t} = \phi'(\sqrt{\varepsilon} R(x,\tau)) \\ \sqrt{\varepsilon} \frac{R(x,\tau)}{\partial t} = \frac{\sqrt{\varepsilon}}{h} (S(x+h,\tau) + S(x-h,\tau) - 2S(x,\tau)) \end{cases} \tag{4.18}$$

To be precise these equations are not equivalent to (4.9). This equivalence is true only if $x = nh$ and $\tau = ht$.

Transforming time derivatives one gets

$$\begin{cases} \dfrac{\partial S(x,\tau)}{\partial \tau} = \dfrac{1}{\sqrt{\varepsilon}}\phi'(\sqrt{\varepsilon}R(x,\tau)) \\ \dfrac{\partial R(x,\tau)}{\partial \tau} = \dfrac{1}{h^2}(S(x+h,\tau)+S(x-h,\tau)-2S(x,\tau)) \end{cases} \tag{4.19}$$

We can rewrite second equation in a clearer form recalling that if S is analytical in $\mathbb{T}$ we have

$$S(x+h,\tau) = \sum_{n=0} \frac{h^n}{n!}\partial_x^n S(x,\tau) = e^{h\partial_x}S(x,\tau) \tag{4.20}$$

Hence

$$S(x+h,\tau)+S(x-h,\tau)-2S(x,\tau) = (e^{h\partial_x}+e^{-h\partial_x}-2)S(x,\tau). \tag{4.21}$$

Defining the discrete Laplacian as

$$\Delta_h = \frac{e^{h\partial_x}+e^{-h\partial_x}-2}{h^2} = \sum_{n=1}^{\infty} \frac{2}{(2n)!}h^{2n-2}\partial_x^{2n} \tag{4.22}$$

the equations of motion read

$$\begin{cases} \dfrac{\partial S(x,\tau)}{\partial \tau} = \dfrac{1}{\sqrt{\varepsilon}}\phi'(\sqrt{\varepsilon}R(x,\tau)) \\ \dfrac{\partial R(x,\tau)}{\partial \tau} = \Delta_h S(x,\tau) \end{cases} . \tag{4.23}$$

At this point we can drop the constraints $x = nh$, $t = h\tau$ and we can assume these equations to be valid for every $x \in \mathbb{T} = \mathbb{R}/\mathbb{Z}$ and $\tau \in \mathbb{R}$. We see immediately that these equations of motion are still Hamiltonian with

$$H[S,R] = \int_{\mathbb{T}} \left(\frac{1}{\varepsilon}\phi(\sqrt{\varepsilon}R(x,\tau)) - \frac{1}{2}S(x,\tau)\Delta_h S(x,\tau) \right) dx, \tag{4.24}$$

and with Poisson tensor the standard symplectic matrix

$$\mathbb{E} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{4.25}$$

To show that (4.23) are the equations of motion related to Hamiltonian (4.24) with Poisson tensor $\mathbb{E}$ it is sufficient to calculate the L_2-gradient of H and then to verify that (4.23) are precisely $\mathbb{E}\nabla_{L_2}H$.

To find the L_2-gradient of H we can calculate

$$\begin{aligned} \frac{d}{d\epsilon}H(S+\epsilon h, R+\epsilon k)\bigg|_{\epsilon=0} &= \frac{d}{d\epsilon}\int_{\mathbb{T}} \left(\frac{1}{\varepsilon}\phi(\sqrt{\varepsilon}(R+\epsilon k)) - \frac{1}{2}(S+\epsilon h)\Delta_h(S+\epsilon h) \right) dx \\ &= \int_{\mathbb{T}} \left(\frac{1}{\sqrt{\varepsilon}}\phi'(\sqrt{\varepsilon}R)k - h\Delta_h S \right) dx \end{aligned} \tag{4.26}$$

where we used the symmetry of Δ_h. We thus get, for the L_2-gradient, the expression

$$\nabla_{L_2} H(S,R) = \begin{pmatrix} -\Delta_h S \\ \dfrac{1}{\sqrt{\varepsilon}}\phi'(\sqrt{\varepsilon}R) \end{pmatrix}. \tag{4.27}$$

Finally, we see that equations of motion are

$$\begin{pmatrix} S \\ R \end{pmatrix}_\tau = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} -\Delta_h S \\ \frac{1}{\sqrt{\varepsilon}}\phi'(\sqrt{\varepsilon}R) \end{pmatrix} \tag{4.28}$$

which are precisely (4.23).

4.2 Hamiltonian normal form of $\alpha+\beta$-FPU

In this section we apply the perturbation theory presented in the previous chapter to the Hamiltonian normal form of the continuous Fermi-Pasta-Ulam Hamiltonian system. We start from (4.24)

$$H[S,R] = \int_{\mathbb{T}} \left(\frac{1}{\varepsilon}\phi(\sqrt{\varepsilon}R) - \frac{1}{2}S\Delta_h S \right) dx \tag{4.29}$$

where we recall that Δ_h is defined as (4.22) and $\phi(\xi)$ is definded as

$$\phi(\xi) = \frac{1}{2}\xi^2 + \frac{\alpha}{3}\xi^3 + \frac{\beta}{4}\xi^4 + \dots \tag{4.30}$$

In this section our aim is to expand (4.29) in series of $\sqrt{\varepsilon}$ and h^2 and to build up a normal form of the perturbation series to second order. Our perturbative parameters are both h and ε.[(3)]

We start expanding the Hamiltonian at first order:

$$\begin{aligned} H[S,R] &= \int_{\mathbb{T}} \left(\frac{R^2}{2} + \frac{\alpha\sqrt{\varepsilon}}{3}R^3 - \frac{1}{2}S\partial_x^2 S - \frac{h^2}{4!}S\partial_x^4 S \right) dx = \\ &= \underbrace{\int_{\mathbb{T}} \left(\frac{R^2}{2} - \frac{1}{2}S\partial_x^2 S \right) dx}_{=h[S,R]} + \underbrace{\int_{\mathbb{T}} \left(\frac{\alpha\sqrt{\varepsilon}}{3}R^3 - \frac{h^2}{4!}S\partial_x^4 S \right) dx}_{=P[S,R]} \end{aligned} \tag{4.31}$$

which in the last row is written as sum of an unperturbed Hamiltonian $h[S,R]$ and a perturbation $P[S,R]$. Using (4.23) we get

$$S_\tau = \frac{\delta H}{\delta R} = R + \alpha\sqrt{\varepsilon}R^2, \tag{4.32}$$

$$R_\tau = -\frac{\delta H}{\delta S} = \partial_x^2 S + \frac{h^2}{12}\partial_x^4 S. \tag{4.33}$$

(3) The ordering will be explained in subsection 4.2.2.

Performing the derivative with respect to τ on the second one, using the first one and ignoring terms of order $\sqrt{\varepsilon}h^2$ we get, as equation of motion

$$R_{\tau\tau} = R_{xx} + \frac{h^2}{12}R_{xxxx} + 2\alpha\sqrt{\varepsilon}RR_x. \tag{4.34}$$

This is an integrable Boussinesq equation.

Using the fact that ∂_x is a skew-symmetric operator in $L_2(\mathbb{T})$ we can rewrite (4.31) integrating by parts the term with S. We get

$$\begin{aligned} H[S,R] &= \int_{\mathbb{T}} \left(\frac{R^2}{2} - \frac{1}{2}S\partial_x^2 S\right) dx + \int_{\mathbb{T}} \left(\frac{\alpha\sqrt{\varepsilon}}{3}R^3 - \frac{h^2}{4!}S\partial_x^4 S\right) dx = \\ &= \int_{\mathbb{T}} \left(\frac{R^2}{2} + \frac{(S_x)^2}{2}\right) dx + \int_{\mathbb{T}} \left(\frac{\alpha\sqrt{\varepsilon}}{3}R^3 + \frac{h^2}{4!}S_x\partial_x^2 S_x\right) dx \end{aligned} \tag{4.35}$$

Setting $\tilde{S} = S_x$ we obtain a new Hamiltonian (which, by abuse of notation, we denote again by H)

$$H[\tilde{S},R] = \int_{\mathbb{T}} \left(\frac{R^2}{2} + \frac{\tilde{S}^2}{2}\right) dx + \int_{\mathbb{T}} \left(\frac{\alpha\sqrt{\varepsilon}}{3}R^3 + \frac{h^2}{4!}\tilde{S}\partial_x^2\tilde{S}\right) dx \tag{4.36}$$

Taking the derivative of (4.32) with respect to x and substituting $S_x = \tilde{S}$ we obtain the new equations of motion

$$\tilde{S}_\tau = \partial_x\left(R + \alpha\sqrt{\varepsilon}R^2\right) = \partial_x\frac{\delta H}{\delta R}, \tag{4.37}$$

$$R_\tau = \partial_x\left(\tilde{S} + \frac{h^2}{12}\partial_x^2\tilde{S}\right) = \partial_x\frac{\delta H}{\delta\tilde{S}}. \tag{4.38}$$

In a matricial form, after the transformation, Hamilton equations are

$$\begin{pmatrix} \tilde{S} \\ R \end{pmatrix}_\tau = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \partial_x \begin{pmatrix} \frac{\delta H}{\delta\tilde{S}} \\ \frac{\delta H}{\delta R} \end{pmatrix} \tag{4.39}$$

We are now ready to perform the third change of variables. We are going to split the right-going solutions and the left-going ones in a sense that will be clear below, when we will explicitly solve the wave equation. To do so we set

$$V^\pm = \frac{R \pm \tilde{S}}{\sqrt{2}}. \tag{4.40}$$

The Hamiltonian in this new set of variables takes the form

$$\begin{aligned} H[V^+,V^-] &= \int_{\mathbb{T}} \left(\frac{1}{4}(V^+ + V^-)^2 + \frac{1}{4}(V^+ - V^-)^2\right) dx + \\ &+ \int_{\mathbb{T}} \left(\frac{\alpha\sqrt{\varepsilon}}{6\sqrt{2}}(V^+ + V^-)^3 - \frac{h^2}{4!2}\left(\partial_x(V^+ - V^-)\right)^2\right) dx = \\ &= \int_{\mathbb{T}} \frac{1}{2}\left({V^+}^2 + {V^-}^2\right) dx + \\ &+ \int_{\mathbb{T}} \frac{\alpha\sqrt{\varepsilon}}{6\sqrt{2}}\left({V^+}^3 + 3{V^+}^2V^- + 3V^+{V^-}^2 + {V^-}^3\right) dx - \\ &- \int_{\mathbb{T}} \frac{h^2}{4!2}\left({V_x^+}^2 - 2V_x^+V_x^- + {V_x^-}^2\right) dx \end{aligned} \tag{4.41}$$

To find out the form of the Poisson tensor in these new coordinates we transform the equations of motion. After calculations analogous to (4.37) and (4.38) we get

$$\begin{pmatrix} V^+ \\ V^- \end{pmatrix}_\tau = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \partial_x \begin{pmatrix} \frac{\delta H}{\delta V^+} \\ \frac{\delta H}{\delta V^-} \end{pmatrix}. \tag{4.42}$$

Before starting calculations we determine the Casimir invariants for the Poisson tensor J. This is done allows to choose the functions V^+ and V^- with vanishing spatial mean. This choice, which is the most natural one if one recalls relations (4.7) and (1.20), simplifies a lot the calculations below.

Let's start recalling the definition of Casimir invariant given in Section 2.1.

Definition 4.2.1 (Casimir invariant)**.** Given a Poisson tensor J, we say that a function $C \in \mathscr{A}(\Gamma)$ is a *Casimir invariant* associated to J if

$$\{C, H\}_J = 0 \qquad \forall H \in \mathscr{A}(\Gamma). \tag{4.43}$$

From the definition it follows that, to find the Casimirs, we have to solve the equation

$$J\nabla C = 0 \tag{4.44}$$

that takes the form

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \partial_x \begin{pmatrix} \frac{\delta C}{\delta V^+} \\ \frac{\delta C}{\delta V^-} \end{pmatrix} = 0 \tag{4.45}$$

or, more explicitly,

$$\begin{cases} \partial_x \dfrac{\delta C}{\delta V^+} = 0 \\ \partial_x \dfrac{\delta C}{\delta V^-} = 0 \end{cases}. \tag{4.46}$$

This means, by integration,

$$\frac{\delta C}{\delta V^+} = A, \qquad \frac{\delta C}{\delta V^-} = B, \tag{4.47}$$

which are satisfied for functionals of the form

$$C(V^+, V^-) = \tilde{F} + A \int V^+(t,x)\, dx + B \int V^-(t,x)\, dx \tag{4.48}$$

where F, A and B are arbitrary constants.

Both $\int_{\mathbb{T}} V^+\, dx$ and $\int_{\mathbb{T}} V^-\, dx$ are Casimirs and thus constants of motion. So they are not influent in the equations of motion. We can therefore suppose $\int_{\mathbb{T}} V^+\, dx = \int_{\mathbb{T}} V^-\, dx = 0$ from now on. Notice that from definition (4.40) one gets[(4)]

$$\int V^\pm\, dx = \frac{1}{\sqrt{2}} \int R\, dx \tag{4.49}$$

which is constant. From (4.7) and recalling the relation between R and r one sees that the natural choice is $\int R\, dx = 0$. Once this choice is taken this remains for every time. In fact the value of a Casimir does not vary along the flow of the Hamiltonian system. In fact

$$C(t) = e^{t\mathcal{L}_H} C = C + t\{C, H\} + \cdots = C. \tag{4.50}$$

Thus the value of C is preserved both from the flux and from canonical transformations.

[(4)]One has to recall that the average of $\tilde{S}$ is vanishing because $\tilde{S} = S_x$.

4.2.1 First order normal form

According to subsection 3.2.1 we have first to calculate the flow associated to the unperturbed Hamiltonian system $\Phi_h^s(V^+,V^-)$. We start writing the equations of motion of h. We get

$$\begin{cases} V_\tau^+(x,\tau) = V_x^+(x,\tau) \\ V_\tau^+(x,\tau) = -V_x^-(x,\tau) \end{cases} \tag{4.51}$$

To solve the first PDE[5] we see that if we consider a curve in the (x,t) plane like $r \mapsto (x_0+r,-r) = \gamma(r)$ we find that this PDE evaluated on the curve is reduced to the ODE

$$\frac{d}{dr}(V^+ \circ \gamma(r)) = 0 \tag{4.52}$$

with the obvious solution

$$0 = \int_0^R \frac{d(V^+ \circ \gamma(r))}{dr} dr = V^+(\gamma(r)) - V^+(\gamma(0)) = V^+(x_0+r,-r) - V^+(x_0,0). \tag{4.53}$$

At this point we define $V^+(x,\tau) = V^+(x_0+r,-r)$ and we obtain, as natural definition $x = x_0 + r$ and $\tau = -r$. We then define $V^+(x_0,0) = \tilde{V}^+(x_0)$ and so we get for free

$$\tilde{V}^+(x_0) = V^+(x,\tau) \qquad \Rightarrow \qquad V^+(x,\tau) = \tilde{V}^+(x+\tau) \tag{4.54}$$

since $x_0 = x - r = x + \tau$. With an easy adaptation of this procedure to the other case we get, the following expression for the flow

$$\begin{aligned} V^+(x,s) &= \tilde{V}^+(x+s) \\ V^-(x,s) &= \tilde{V}^-(x-s). \end{aligned} \tag{4.55}$$

From the spatial periodicity of $V^\pm$ we deduce the time periodicity of the unperturbed flow. This is important in order to apply equation (3.39) and to compute $S_1[V^+,V^-]$. To do so we state the following propositions:

Proposition 4.2.2. *Given two real valued function on the torus $\mathbb{T}$, we have*

$$\int_0^1 dx \int_0^1 ds\, u(x\pm s)v(x\mp s) = \int_{\mathbb{T}} u(x)\,dx \int_{\mathbb{T}} v(y)\,dy \tag{4.56}$$

Proof. This proof is a straightforward calculation

$$\begin{aligned} \int_0^1 dx \int_0^1 ds\, u(x\pm s)v(x\mp s) &= \int_0^1 dx \int_0^1 ds \sum_{k,k'\in\mathbb{Z}} \hat{u}_k \hat{v}_{k'} e^{2\pi i k(x\pm s)} e^{2\pi i k'(x\mp s)} = \\ &= \sum_{k,k'\in\mathbb{Z}} \hat{u}_k \hat{v}_{k'} \delta_{k+k',0}\delta_{k-k',0} = \hat{u}_0\hat{v}_0 \end{aligned}$$

where we expanded u and v in Fourier series and integrated the exponential functions over the torus and $\hat{u}_0 = \int_{\mathbb{T}} u(x)\,dx$, $\hat{v}_0 = \int_{\mathbb{T}} v(x)\,dx$. □

[5] Here we use the method of characteristics.

Proposition 4.2.3. *Given a real valued function on the torus $\mathbb{T}$, we have that*

$$\int_0^1 dx \int_0^1 ds\, u(x \pm s) = \int_0^1 u(x)\, dx \tag{4.57}$$

Proof. This proof, like the ones above, it is a straightforward calculation

$$\int_0^1 dx \int_0^1 ds\, u(x \pm s) \sum_{k\in\mathbb{Z}} \int_0^1 dx \int_0^1 ds\, \hat{u}_k e^{2\pi i k(x\pm s)} =$$
$$= \sum_{k\in\mathbb{Z}} \int_0^1 dx \hat{u}_k \delta_{k,0} = \int_0^1 dx \hat{u}_0 = \hat{u}_0 = \int_0^1 u(x)\, dx.$$

□

Using the two propositions above we can calculate $S_1[V^+, V^-]$. Starting from (4.41) and exploiting the zero-mean conditions $\int V^\pm\, dx = 0$, we obtain:

$$S_1[V^+, V^-] = \int_{\mathbb{T}} \left(\frac{\alpha\sqrt{\varepsilon}}{6\sqrt{2}} \left((V^+)^3 + (V^-)^3\right) - \frac{h^2}{4!2} \left((V_x^+)^2 + (V_x^-)^2\right) \right) dx. \tag{4.58}$$

The equations of motion associated to $h + S_1$ are two uncoupled KdVs: from $V_\tau = J\nabla_{L_2} H$ we get

$$\begin{cases} V_\tau^+ = V_x^+ + \dfrac{\alpha\sqrt{\varepsilon}}{\sqrt{2}} V^+ V_x^+ + \dfrac{h^2}{4!} V_{xxx}^+ \\ V_\tau^- = -V_x^- - \dfrac{\alpha\sqrt{\varepsilon}}{\sqrt{2}} V^- V_x^- - \dfrac{h^2}{4!} V_{xxx}^- \end{cases} \tag{4.59}$$

As showed in [15, 18, 19] KdV equations are integrable. One has then the result that normal form of the FPU system to first order consists in two KdV equations and thus is integrable. integrable and it is related to KdV. This should not make surprise since at first order FPU system is tangent to the Toda one which is the discrete analogous of KdV.

4.2.2 Analyticity of the interpolating function

The aim of this section is to discuss the perturbative parameters in our problem. We left this discussion up to this point because it is easier to discuss it using the KdV equations above. Anyway one can repeat the same argument for the equation of motion associated to the full Hamiltonian. At the end we see that the correct parameter in this analysis is $\lambda \sim \alpha\sqrt{\varepsilon} \sim h^2$.

We begin this section recalling that our working hypothesis is that the initial datum is a long wavelength oscillation. Such an initial datum is analytical on the torus.[(6)] We write equations (4.59) setting $u = V^\pm$:

$$u_t = u_x + \frac{\alpha\sqrt{\varepsilon}}{\sqrt{2}} u u_x + \frac{h^2}{4!} u_{xxx} \tag{4.60}$$

[(6)] I.e. it is an analytical function with periodic real part.

and, since u is analytical, we use Cauchy estimates to weight each term.

We recall that[7] thanks to a theorem by Cauchy, we have that

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(\xi)}{\xi - z}d\xi \tag{4.61}$$

where γ is a circle inside the analyticity region of f. If M is the maximum of f along γ we can estimate the value of f as follows

$$|f(z)| \leq \left|\frac{1}{2\pi i}\int_\gamma \frac{f(\xi)}{\xi - z}d\xi\right| \leq M\left|\frac{1}{2\pi i}\int_\gamma \frac{1}{\xi - z}d\xi\right| = M. \tag{4.62}$$

We thus get an estimate of $f(z)$. To get an estimate of its derivatives one can derive both sides of (4.61) with respect to z. One thus gets an expression for the derivative

$$\frac{d^n f(z)}{dz^n} = \frac{(-1)^{n+1}n!}{2\pi i}\int_\gamma \frac{f(\xi)}{(\xi - z)^{n+1}}d\xi \tag{4.63}$$

and, consequently, an extime

$$\left|\frac{d^n f(z)}{dz^n}\right| = n!\left|\frac{1}{2\pi i}\int_\gamma \frac{f(\xi)}{(\xi - z)^{n+1}}d\xi\right| \leq n!\frac{M}{\sigma^n} \tag{4.64}$$

where we called σ the minimum width of the strip of analyticity which is the maximum radius one can take to perform the integration if one wants to integrate the function along a circular path.

We now go back to (4.60) and we weight every term on the right hand side. Forgetting the numerical constants we concentrate on the dependence of the parameters

$$u_x \leq \frac{M}{\sigma}, \qquad \alpha\sqrt{\varepsilon}uu_x \leq \alpha\sqrt{\varepsilon}\frac{M^2}{\sigma}, \qquad h^2 u_{xxx} \leq h^2\frac{M}{\sigma^3}. \tag{4.65}$$

We recall now how the dynamics associated to the KdV works: at the beginning the dispersive term (the one involves three derivatives) is irrelevant and the dynamics is governed by the so-called Burgers equation

$$u_t = \alpha\sqrt{\varepsilon}uu_x \tag{4.66}$$

but after a finite time the solution starts to steepen a lot and then the term with three derivatives balances the formation of multivalued-solutions splitting the initial datum in solitons. At this time one has

$$\alpha\sqrt{\varepsilon}uu_x \geq h^2 u_{xxx} \tag{4.67}$$

using the estimates above one gets the following relation

$$M\sigma^2 \geq \frac{h^2}{\alpha\sqrt{\varepsilon}}. \tag{4.68}$$

(7) For an introduction see [22].

At this point if we can treat M as a number (i.e. it is independent of the parameters[(8)]). We get

$$\sigma \leq \left(\frac{h^2}{\alpha\sqrt{\varepsilon}}\right)^{\frac{1}{2}}. \tag{4.69}$$

Inserting this expressions of σ in (4.65) one sees that the nonlinear term and the dispersive one have the same weight and this is why we formally treated $h^2 \sim \sqrt{\varepsilon}$.

4.2.3 Hamiltonian generating canonical transformation to first order

If we want to go beyond the first order, from section 3.2.1 and according to (3.54), we have to find the Hamiltonian generating the canonical transformation to the first order.

We start from (3.47) which, in the present case, reads

$$G_1 = \langle G_1\rangle_h + \int_0^1 (s-1)e^{s\mathcal{L}_h}\tilde{P}_1^\natural\, ds. \tag{4.70}$$

where, we recall $\tilde{P}_1^\natural = \tilde{P}_1 = P_1 - \langle P_1\rangle_h = P_1 - S_1$ as given by (4.41) and (4.58). Making these substitutions we obtain for $\tilde{P}_1$ the expression

$$\tilde{P}_1[V^+, V^-] = \int_0^1 \left[\frac{\alpha\sqrt{\varepsilon}}{2\sqrt{2}}\left(V^+{V^-}^2 + {V^+}^2V^-\right) + \frac{h^2}{4!}V_x^+V_x^-\right] dx. \tag{4.71}$$

The integral we have to evaluate to find the Hamiltonian is then[(9)]

$$G_1 = \int_0^1\int_0^1 s\left[\frac{\alpha\sqrt{\varepsilon}}{2\sqrt{2}}\left(V^+{V^-}^2 + {V^+}^2V^-\right) + \frac{h^2}{4!}V_x^+V_x^-\right] dxds. \tag{4.72}$$

The above calculation cannot be performed using the propositions of the previous subsection because of the s factor multiplying the functions. We can, anyway, state the following proposition

Proposition 4.2.4. *Given two real-valued function on the torus $\mathbb{T}$ we have*

$$\int_0^1 ds\int_0^1 dx\, su(x\pm s)v(x\mp s) = \frac{1}{2}\left(\int_0^1 u(x)\,dx\right)\left(\int_0^1 v(x)\,dx\right) \pm \frac{1}{2}\int_0^1 v(x)\partial_x^{-1}u(x)\,dx \tag{4.73}$$

where ∂_x^{-1} is the operator defined in section 2.3.3.

---

(8) There are several reasons to trust this if the initial datum is chosen with initial "random phases" between the excited Fourier modes. Otherwise it is proportional to $\sqrt{N}$, where N is the number of particles in the discrete system.

(9) Where $s-1$ is replaced by s just because the periodicity of the flow of h guarantees that $\int e^{s\mathcal{L}_h}\tilde{P}\,dx = 0$.

Proof. We make the computation

$$\begin{aligned}\int_0^1 ds\, s \int_0^1 dx\, u(x\pm s)v(x\mp s) &= \sum_{k,k'\in\mathbb{Z}} \int_0^1 ds\, s \int_0^1 dx\, \hat{u}_k \hat{v}_{k'} e^{2\pi i k(x\pm s)} e^{2\pi i k'(x\mp s)} = \\ &= \sum_{k,k'\in\mathbb{Z}} \int_0^1 ds\, s\hat{u}_k\hat{v}_{k'} e^{\pm 2\pi i(k-k')s} \int_0^1 dx\, e^{2\pi i(k+k')x}\, dx = \\ &= \sum_{k,k'\in\mathbb{Z}} \int_0^1 ds\, s\hat{u}_k\hat{v}_{k'} e^{\pm 2\pi i(k-k')s} \delta_{k+k',0} = \\ &= \sum_{k\in\mathbb{Z}} \hat{u}_k\hat{v}_{-k} \int_0^1 s e^{\pm 4\pi i k s}\, ds.\end{aligned}$$

We have to compute carefully last integration splitting the case $k=0$ and $k\neq 0$.

$$\int_0^1 s e^{\pm 4\pi i k s}\, ds = \delta_{k,0} \int_0^1 s\, ds \pm (1-\delta_{k,0}) \int_0^1 s e^{\pm 4\pi i k s}\, ds.$$

The first integration is straightforward and gives $\frac{1}{2}$; the second one can be done by parts[10] and, recalling that $\int_0^1 e^{\pm 4\pi i k s}\, ds = 0$ we obtain

$$\int_0^1 s e^{\pm 4\pi i k s}\, ds = \pm\frac{1}{4\pi i k}.$$

Using this result we can continue the calculation above which yields

$$\begin{aligned}\sum_{k\in\mathbb{Z}} \hat{u}_k\hat{v}_{-k} \int_0^1 s e^{\pm 4\pi i k s}\, ds &= \sum_{k\in\mathbb{Z}} \hat{u}_k\hat{v}_{-k} \left[\frac{1}{2}\delta_{k,0} \pm \frac{1}{4\pi i k}(1-\delta_{k,0})\right] = \\ &= \frac{1}{2}\hat{u}_0\hat{v}_0 \pm \frac{1}{2} \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{2\pi i k}\hat{u}_k\hat{v}_{-k}.\end{aligned}$$

We recognize in this last expression the definition of the anti-derivative operator acting on u (see equation (2.72)) and since $\delta_{k+k',0} = \int_0^1 e^{2\pi i(k+k')x}\, dx$ we obtain the thesis. □

The calculation of $G_1[V^+,V^-]$ is just a straightforward application of the above proposition and it leads to the following expression

$$G_1[V^+,V^-] = \int_0^1 \left[\frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}\left(V^{-2}\partial_x^{-1}V^+ - V^{+2}\partial_x^{-1}V^-\right) + \frac{1}{2}\frac{h^2}{4!}V^+\partial_x V^-\right] dx. \tag{4.74}$$

With this explicit expression of G_1 one may prove whether all the calculations performed till now are correct or not. It is sufficient to prove that

$$P_1 + \mathcal{L}_1 h = S_1. \tag{4.75}$$

We exhibit this calculation in appendix B.1. We have now all the elements to build up the normal form to the second order.

[10] In this second computation $k\neq 0$.

4.2.4 Second order

In this subsection we will carry on the calculation of (4.31) till the second order and we will outline how to build the normal form. Since the calculations are tedious we will not repeat all the steps. In this order of ideas we can start by expanding Δ_h until the fourth order and ϕ till the second one to get

$$P_2[S,R]=\int_{\mathbb{T}}\left(\frac{\beta\varepsilon}{4}R^4+\frac{h^4}{6!}(S_{xxx})^3\right)\,dx. \tag{4.76}$$

After performing the same two change of variables as in the beginning of section 4.2, we obtain by substitution the following expression for $P_2[V^+,V^-]$:

$$\begin{aligned}P_2[V^+,V^-]=&\int_0^1\left[\frac{\beta\varepsilon}{16}\left(V^{+4}+V^{-4}+4V^{+3}V^-+4V^{-3}V^++6V^{+2}V^{-2}\right)\right]\,dx+\\&+\int_0^1\left[\frac{1}{2}\frac{h^2}{6!}\left(V_{xx}^{+2}-2V_{xx}^+V_{xx}^-+V_{xx}^{-2}\right)\right]\,dx.\end{aligned} \tag{4.77}$$

We recall then (3.54) which will be used to calculate the second order normal form of the perturbation under the hypothesis of vanishing mean of the Hamiltonian generatrix function

$$S_2=\langle P_2+\frac{1}{2}\{\tilde{P}_1,\tilde{G}_1\}\rangle_h. \tag{4.78}$$

We do not exhibit calculations here but we write directly the result[(11)]

$$\begin{aligned}S_2[V^+,V^-]=&\int_0^1\left[\left(\frac{\beta\varepsilon}{16}-\frac{\alpha^2\varepsilon}{32}\right)\left(V^{+4}+V^{-4}\right)+\frac{h^2\alpha\sqrt{\varepsilon}}{4\sqrt{24!}}\left(V^{-2}V_{xx}^-+V^{+2}V_{xx}^+\right)+\right.\\&\left.+\frac{3}{20}\frac{h^4}{(4!)^2}\left(V_{xx}^{+2}+V_{xx}^{-2}\right)\right]\,dx+\left(\frac{3\beta\varepsilon}{8}-\frac{\alpha^2\varepsilon}{4}\right)\langle V^{+2}\rangle_{\mathbb{T}}\langle V^{-2}\rangle_{\mathbb{T}}+\\&+\frac{\alpha^2\varepsilon}{32}\left(\langle V^{-2}\rangle_{\mathbb{T}}^2+\langle V^{-2}\rangle_{\mathbb{T}}^2\right).\end{aligned} \tag{4.79}$$

We don't report the equations of motion associated to this Hamiltonian since they are not very enlightening.

4.3 Integrability of the FPU hierarchy

In this section we show that the problem of mapping the normal forms we got to a system of generalized KdV equations is not trivial.

4.3.1 First order

We can compare the first order normal form of the perturbation S_1 as (4.58) and the conserved functional of KdV F_2 as (2.123). On a first sight we can recognize some analogies between them. This link is precisely the one stated at the end of subsection

(11) For the full calculation see Appendix B.2.

4.2.1. In fact with an opportune choice of the parameter γ in the expression of F_2 one obtains S_1. We begin recalling their expressions:

$$\begin{aligned} S_1[V^+,V^-] &= \int_{\mathbb{T}} \left(\frac{\alpha\sqrt{\varepsilon}}{6\sqrt{2}} \left((V^+)^3 + (V^-)^3 \right) - \frac{h^2}{4!2} \left((V_x^+)^2 + (V_x^-)^2 \right) \right) dx \\ F_2[u] &= C \int_{\mathbb{T}} \left(u u_{xx} + \frac{1}{3}\gamma u^3 \right) dx \end{aligned} \tag{4.80}$$

We see that $F_2[V^+] + F_2[V^-]$ is very similar to the above expression of the first order perturbation in normal form. This Hamiltonian can be written in a smarter form:

$$S_1 = \frac{h^2}{4!2} \int_{\mathbb{T}} \left(\frac{4\alpha\sqrt{2\varepsilon}}{h^2} (V^{+3} + V^{-3}) + V^+ V_{xx}^+ + V^- V_{xx}^- \right) dx. \tag{4.81}$$

With a straightforward comparison we obtain

$$C = \frac{h^2}{4!2} \qquad \gamma = \frac{12\alpha\sqrt{2\varepsilon}}{h^2}. \tag{4.82}$$

This means that the first order perturbation is "on KdV hierarchy" with the choice of parameters as (4.82) or, more explicitly:

$$S_1[V^+,V^-] = F_2[V^+] + F_2[V^-]. \tag{4.83}$$

This is a remarkable result already known in literature, see for example [3, 4].

4.3.2 Second order

Since the previous result was already known and we have an expression for a normalized second order we can check if we are in hierarchy or not at this order.

This time the value of γ parameter is fixed from the previous subsection. Substituting in F_3 we obtain the form of the conserved functional for the KdV equation:

$$F_3 = C \int_{\mathbb{T}} \left(\frac{40\alpha^2\varepsilon}{h^4} u^4 + \frac{10\alpha\sqrt{2\varepsilon}}{h^2} u^2 u_{xx} + (u_{xx})^2 \right) dx \tag{4.84}$$

There are some analogies between this form and $\hat{S}_2$ defined as follow:

$$\begin{aligned} \hat{S}_2 = \int_0^1 \Bigg[& \left(\frac{\beta\varepsilon}{16} - \frac{\alpha^2\varepsilon}{32} \right) \left(V^{+4} + V^{-4} \right) + \frac{h^2\alpha\sqrt{\varepsilon}}{4\sqrt{24!}} \left(V^{-2} V_{xx}^- + V^{+2} V_{xx}^+ \right) + \\ & + \frac{3}{20} \frac{h^4}{(4!)^2} \left(V_{xx}^{+2} + V_{xx}^{-2} \right) \Bigg] dx. \end{aligned} \tag{4.85}$$

We therefore see that S_2, in this form, cannot be reduced to (4.84) because of the presence of β on the first term. A deeper analysis shows that the real obstruction in this reduction is due to a factor 2 on the second term.[12]

[12] See next section.

One can now ask if it is possible to go through this difficulty inserting $\langle G_1\rangle_h$. The following sections show that this is not enough.[13]

To prove this fact and to check if it is possible to transform the FPU system in a generalized KdV one we will use the technology of general diffeomorphisms described in section 3.3.

4.4 Normalization of vector fields

Here we try to use the perturbation theory developed for general diffeomorphisms to find the normal form of FPU. As a first step we will look for the transformation at first order to find an explicit hint for the "gauge function".[14] This hint will be very helpful in finding a first order transformation which will not change the KdV at first order but which will "tune" the second order Hamiltonian to the hierarchy one.

4.4.1 First perturbative step

In this framework we deal with the equations of motion of the form

$$\begin{cases} V^+_\tau = V^+_x + \mathcal{F}^+_1 + \mathcal{F}^+_2 + \dots \\ V^-_\tau = -V^-_x - \mathcal{F}^-_1 - \mathcal{F}^-_2 - \dots \end{cases} \tag{4.86}$$

where

$$\mathcal{F}^\pm_j = \partial_x \frac{\delta P_j}{\delta V^\pm} \tag{4.87}$$

and with a diffeomorphism in the form

$$V^\pm = \mathrm{id}^\pm + \mathcal{G}^\pm_1 + \mathcal{G}^\pm_2 + \dots \tag{4.88}$$

If we require that the normal form is the KdV hierarchy vector field which we indicate with $\mathcal{F}_{KdV_3} + \mathcal{F}_{KdV_5}$ we will have

$$\begin{cases} V^+_t = V^+_x + \mathcal{F}_{KdV_3}(V^+) + \mathcal{F}_{KdV_5}(V^+) + \dots \\ V^-_t = -V^-_x - \mathcal{F}_{KdV_3}(V^-) - \mathcal{F}_{KdV_5}(V^-) - \dots \end{cases} \tag{4.89}$$

or, in a more compact form,

$$V_\tau = V_x + \mathcal{F}_{KdV_3}(V) + \mathcal{F}_{KdV_5}(V) + \dots \tag{4.90}$$

From equations (3.78) one obtains, at first order

$$[V_x, \mathcal{G}_1(V)] + \mathcal{F}_1(V) = \mathcal{F}_{KdV_3}(V) \tag{4.91}$$

which, component by component, it reads

$$+: \qquad \frac{DV^+_x}{DV^+}\mathcal{G}^+_1 - \frac{D\mathcal{G}^+_1}{DV^+}V^+_x + \frac{D\mathcal{G}^+_1}{DV^-}V^-_x + \mathcal{F}^+_3 = \mathcal{F}_{KdV_3} \tag{4.92}$$

[13] In these terms the sentence is false since if $\beta = \frac{7}{9}\alpha^2$ such a transformation is possible and it involves an Hamiltonian in the form $\langle G_1\rangle_h = G(V^+) - G(V^-)$ with $G(u) = A\int_{\mathbb{T}} u\partial_x^{-1}u^2\,dx$ for a fixed parameter A. Exhibiting these calculations here is not very interesting from a physical point of view. With some economy one can get the same solution looking at the next section.

[14] Where, with *gauge function*, we mean the g appearing in 3.3.2.

$$-:\qquad -\left(\frac{DV_x^-}{DV^-}\mathcal{G}_1^- - \frac{D\mathcal{G}_1^-}{DV^-}V_x^- + \frac{D\mathcal{G}_1^-}{DV^+}V_x^+ + \mathcal{F}_3^+\right) = -\mathcal{F}_{KdV_3}. \tag{4.93}$$

If one recalls now that $DV_x^\pm/DV^\pm = \partial_x$ and, by chain rule $\partial_x\mathcal{G}_1^\pm = \mathcal{G}_{1,+}^\pm V_x^+ + \mathcal{G}_{1,-}^\pm V_x^-$ where the low $\pm$ indicates a derivative with respect $V^\pm$, one is left with partial cancellations and thus the equations to solve are

$$\begin{cases} 2\mathcal{G}_{1,-}^+V_x^- + \mathcal{F}_1^+ = \mathrm{KdV}_3^+ \\ 2\mathcal{G}_{1,+}^-V_x^+ + \mathcal{F}_1^- = \mathrm{KdV}_3^-. \end{cases} \tag{4.94}$$

We try to solve the first equation in the unknown $\mathcal{G}_1^+$. After, using time reversal symmetry, we can deduce the form of $\mathcal{G}_1^-$. At this level the two unknowns are $\mathcal{G}_1^\pm$. Substituting in the first one the expressions (4.59) and (2.123) we get

$$2\mathcal{G}_{1,-}^+V_x^- + \frac{h^2}{4!}\left(\gamma(V^+ + V^-)(V_x^+ + V_x^-) + (V_{xxx}^+ - V_{xxx}^-)\right) = 2A(\gamma V^+V_x^+ + V_{xxx}^+) \tag{4.95}$$

Here it is natural to pose $A = \frac{h^2}{2\cdot 4!}$ to obtain

$$\mathcal{G}_{1,-}^+V_x^- = -\frac{h^2}{2\cdot 4!}\left(\gamma(V^+V_x^- + V^-V_x^+ + V^-V_x^-) - V_{xxx}^-\right) \tag{4.96}$$

Using anti-derivative operators one can rewrite RHS as

$$\mathcal{G}_{1,-}^+V_x^- = -\frac{h^2}{2\cdot 4!}(\gamma(V^+ + V_x^+\partial_x^{-1} + V^-) - \partial_x^2)V_x^-. \tag{4.97}$$

Thus one immediately has

$$\mathcal{G}_{1,-}^+ = \gamma(V^+ + V_x^+\partial_x^{-1} + V^-) - \partial_x^2). \tag{4.98}$$

One can now see that this is the derivative with respect to V^- of the following function

$$\mathcal{G}_1^+ = -\frac{h^2}{2\cdot 4!}\left(\gamma(V^+V^- + V_x^+\partial_x^{-1}V^- + \frac{1}{2}V^{-2}) - V_{xx}^-\right) + g(V^+) \tag{4.99}$$

where $g(V^+)$ is a generic function of V^+ and its derivatives. With a simple argument of time-reversal symmetry ($t \to -t \Rightarrow V^+ \leftrightarrow V^-$) we get $\mathcal{G}_1^-$.

$$\mathcal{G}_1^- = -\frac{h^2}{2\cdot 4!}\left(\gamma(V^+V^- + V_x^-\partial_x^{-1}V^+ + \frac{1}{2}V^{+2}) - V_{xx}^+\right) + g(V^-). \tag{4.100}$$

We thus find a precise expression for the gauge function g. It is, for example, possible to choose $g(f)$ in a way such that the above expressions look more simple. If one imposes that he obtains

$$g(f) = -\frac{h^2\gamma}{2\cdot 4!}\frac{1}{2}f^2 \tag{4.101}$$

And one gets, for example

$$\mathcal{G}_1^+ = -\frac{h^2}{2\cdot 4!}\left(\gamma\left(\frac{1}{2}\left(V^+ + V^-\right)^2 + V_x^+\partial_x^{-1}V^-\right) - V_{xx}^-\right) + \tilde{g}(V^+) \tag{4.102}$$

where $\tilde{g}(V^+)$ is the freedom left in the choice of the transformation which leaves invariant the first order vector field.

We will, anyway, deal with $\mathcal{G}_1$ in the form (4.99) and (4.100).

4.4.2 Relation between vector fields and Hamiltonian transformations

Here we write both the transformations at first order. First the Hamiltonian functional one and then the vector field one. We have

$$G_1 = \frac{h^2}{2\cdot 4!}\int_{\mathbb{T}} \left(\frac{\gamma}{2}\left(V^{-2}\partial_x^{-1}V^+ - V^{+2}\partial_x^{-1}V^-\right) + V^+\partial_x V^-\right)\,dx \tag{4.103}$$

$$\mathcal{G}_1^+ = -\frac{h^2}{2\cdot 4!}\left(\gamma(V^+V^- + V_x^+\partial_x^{-1}V^- + \frac{1}{2}V^{-2}) - V_{xx}^-\right) + g(V^+) \tag{4.104}$$

where we set

$$\gamma = \frac{12\alpha\sqrt{2\varepsilon}}{h^2}. \tag{4.105}$$

From relations between Poisson brackets and Lie brackets we want now to prove that

$$J\nabla G_1 = \begin{pmatrix} \mathcal{G}_1^+ \\ -\mathcal{G}_1^- \end{pmatrix}. \tag{4.106}$$

With this aim it is sufficient to calculate $\partial_x(\delta G_1/\delta V^+)$ and $\partial_x(\delta G_1/\delta V^-)$. We have thus

$$\frac{\delta G_1}{\delta V^+} = \frac{h^2}{2\cdot 4!}\left(-\frac{\gamma}{2}\partial_x^{-1}V^{-2} - \gamma V^+\partial_x^{-1}V^- + V_x^-\right). \tag{4.107}$$

If we derive now this expression with respect to x we obtain

$$\partial_x\frac{\delta G_1}{\delta V^+} = -\frac{h^2}{2\cdot 4!}\left(\gamma\left(\frac{1}{2}V^{-2} + V^+V^- + V_x^+\partial_x^{-1}V^-\right) - V_{xx}^-\right) \tag{4.108}$$

which is precisely $\mathcal{G}_1^+$ as stated in (4.99). We can perform the same calculation for $\mathcal{G}_1^-$ to obtain the agreement between the two transformations.

4.5 Second order homological equation

If one tries to write down the equations for the normal form at second order with the vector fields formalism one gets immediately sick because of the quantity of commutators to evaluate. Luckily we can start from the Hamiltonian canonical transformation and then perform a more general transformation which leaves invariant first order normal form. The advantage of this procedure is that most of the work is done by the Hamiltonian transformation. One is thus left with (3.84) where $[f_2, X_0] = 0$ since f_2 is already in normal form with respect to X_0 and $D^2X_0 = 0$ because of the form of our vector field. We have then to solve the following homological equation

$$[\mathcal{F}_1, g] = \mathcal{P} \tag{4.109}$$

where $\mathcal{P}$ is the vector field associated to a functional of the form

$$\mathscr{F} = F\int \left(a_1k^2u^4 + a_2ku^2u_{xx} + a_3(u_{xx})^2\right)\,dx \tag{4.110}$$

where F is a constant proportional to h^4; a_1, a_2, a_3 are numbers, $k = \frac{\alpha\sqrt{\varepsilon}}{h^2}$ and $\mathcal{F}_1$ and g are respectively the first order KdV vector field and a vector field generating the canonical transformation to be determined:

$$\mathcal{F}_1 = G(bkuu_x + u_{xxx}), \qquad g = \frac{F}{G}\left(kg_1 + g_0\right). \tag{4.111}$$

We recall that the Lie bracket of $\mathcal{F}_1$ and g is $[\mathcal{F}_1, g] = \mathcal{F}_1' g - g'\mathcal{F}_1$. Direct computations show that

$$\mathcal{F}_1' = G[bk(u\partial_x + u_x) + \partial_x^3] \tag{4.112}$$

and

$$\begin{aligned}\mathcal{F}_1' g &= F[bk(u\partial_x + u_x) + \partial_x^3](kg_1 + g_0) = \\ &= F\{k^2 b(u\partial_x g_1 + u_x g_1) + k[b(u\partial_x g_0 + u_x g_0) + \partial_x^3 g_1] + \partial_x^3 g_0\}\end{aligned} \tag{4.113}$$

The second term is

$$g'\mathcal{F}_1 = F(kg_1' + g_0')(kbuu_x + u_{xxx}) = F[k^2 b g_1' uu_x + k(bg_0' uu_x + g_1' u_{xxx}) + g_0' u_{xxx}]. \tag{4.114}$$

Subtracting the second equation from the first one we get the Left Hand side of equation (4.109)

$$\begin{aligned}\text{L.H.S.} &= F\{k^2 b(u\partial_x g_1 + u_x g_1 - g_1' uu_x) + k[b(u\partial_x g_0 + u_x g_0 - g_0' uu_x) + \\ &+ \partial_x^3 g_1 - g_1' u_{xxx}] + \partial_x^3 g_0 - g_0' u_{xxx}\}.\end{aligned} \tag{4.115}$$

When we pose LHS=RHS we get three equations. One for every power of k. We are then left to the following system

$$\begin{cases} b(u\partial_x g_1 + u_x g_1 - g_1' uu_x) = 12a_1 u^2 u_x \\ b(u\partial_x g_0 + u_x g_0 - g_0' uu_x) + \partial_x^3 g_1 - g_1' u_{xxx} = 4a_2(2u_x u_{xx} + uu_{xxx}) \\ \partial_x^3 g_0 - g_0' u_{xxx} = 2a_3 u_{xxxxx} \end{cases} \tag{4.116}$$

We have now to solve these two equations. To get a hint on the form of g we do a dimensional analysis of the first and third equations. Denoting with square brackets the physical dimensions

$$[g_1] = [u]^2 \qquad [g_0] = [\partial_x]^2[u]. \tag{4.117}$$

Since the Poisson tensor consists in a derivative operator we can use at most an anti-derivative operator. From this and the dimensional considerations above we can then assume that

$$g_1 = Au^2 + cu_x\partial_x^{-1}u \qquad g_0 = du_{xx}. \tag{4.118}$$

Substituting these expression in the system above we obtain from the third equation $a_3 = 0$.

To get informations from the second equation we shall calculate

$$\begin{aligned}\partial_x^3 g_1 &= A\partial_x^3 u^2 + c\partial_x^3(u_x\partial_x^{-1}u) = \\ &= A(6u_x u_{xx} + 2uu_{xxx}) + c(u_{xxxx}\partial_x^{-1}u + 3uu_{xxx} + 4u_x u_{xx})\end{aligned} \tag{4.119}$$

and

$$\begin{aligned}g_1' u_{xxx} &= (2Au + cu_x\partial_x^{-1} + c\partial_x^{-1}u\partial_x)u_{xxx} = \\ &= 2Auu_{xxx} + cu_x u_{xx} + cu_{xxxx}\partial_x^{-1}u\end{aligned} \tag{4.120}$$

substituting on the second equation we obtain

$$(6A - 2bd + 3c)u_x u_{xx} + 3cuu_{xxx} = 8a_2 u_x u_{xx} + 4a_2 uu_{xxx} \tag{4.121}$$

From this one, we get a couple of algebraic equations for A, c and d:

$$\begin{cases} 3c = 4a_2 \\ 6A - 2bd + 3c = 8a_2 \end{cases} \tag{4.122}$$

so we get

$$3c = 4a_2 \qquad\qquad d = \frac{3A - 2a_2}{b} \tag{4.123}$$

From the first one we get $A = 12\frac{a_1}{b} - \frac{2}{3}a_2$. So that we obtain for g the following expression

$$g = \frac{F}{G}\left\{k\left[\left(12\frac{a_1}{b} - \frac{2}{3}a_2\right)u^2 + \frac{4}{3}a_2 u_x \partial_x^{-1} u\right] + \left(36\frac{a_1}{b^2} - 4\frac{a_2}{b}\right)u_{xx}\right\}. \tag{4.124}$$

This treatment is just formal. We have now to find explicit expressions for the various coefficients.

We recall that $\mathcal{P}$ is the vector field related to $F_{KdV_5} - S_2$. Since the following equalities hold

$$\begin{aligned} F_{KdV_5} &= \frac{3}{20}\frac{h^4}{4!^2}\int_{\mathbb{T}}\left(40\frac{\alpha^2\varepsilon}{h^4}u^4 + \frac{10\alpha\sqrt{2\varepsilon}}{h^2}u^2 u_{xx} + {u_{xx}}^2\right)dx \\ \hat{S}_2 &= \frac{3}{20}\frac{h^4}{4!^2}\int_{\mathbb{T}}\left(\left(\frac{\beta}{\alpha^2} - \frac{1}{2}\right)\frac{240\alpha^2\varepsilon}{h^4}u^4 + \frac{20\alpha\sqrt{2\varepsilon}}{h^2}u^2 u_{xx} + {u_{xx}}^2\right)dx \\ \mathcal{F}_{KdV_3} &= \frac{h^2}{4!}\left(\frac{12\alpha\sqrt{2\varepsilon}}{h^2}uu_x + u_{xxx}\right) \end{aligned} \tag{4.125}$$

by an immediate comparison with the above definitions one gets immediately

$$F = \frac{3}{20}\frac{h^4}{4!^2} \qquad G = \frac{h^2}{4!} \qquad a_1 = 80\left(2 - 3\frac{\beta}{\alpha^2}\right) \qquad a_2 = -10\sqrt{2} \qquad b = 12\sqrt{2} \tag{4.126}$$

which shows that, if $\beta = \frac{2}{3}\alpha^2$ (Toda), $a_1 = 0$ and one sees that the real obstruction is a_2.

Remark 4.5.1. *we notice that, if we were interested only in simplifying as much as possible the expression for the vector field equation we would have chosen a_1 and a_2 in a way such that the transformed system would have looked Hamiltonian with Hamiltonian*

$$K = F\int_{\mathbb{T}}(u_{xx})^2\,dx. \tag{4.127}$$

Anyway this functional is not an integral of motion for $H_0 + \lambda S_1$ and then it is not very useful for our purposes.

Remark 4.5.2. *If we try to choose parameters α and β in order to make the coefficient in front of u_{xx} in (4.124) equal to zero we obtain $\beta = \frac{7}{9}\alpha^2$. With this choice of coefficients g has the form*

$$g = -\frac{\alpha\sqrt{\varepsilon}}{12\sqrt{2}}(u^2 + 2u_x\partial_x^{-1}) \tag{4.128}$$

which is an Hamiltonian vector field with

$$\bar{G} = \frac{\alpha\sqrt{\varepsilon}}{12\sqrt{2}}\int_{\mathbb{T}} u\partial_x^{-1}u^2\,dx. \tag{4.129}$$

In the next section it will be clear that this is the only case for which this transformation can be done in a Hamiltonian way.

4.6 Canonicity of last transformation

At a first sight it seems that we performed a canonical transformation since we mapped a Hamiltonian system with Poisson tensor J as given in equation (4.42) into another Hamiltonian system with the same Poisson tensor J. It is then natural to ask if this transformation is canonical or not. This section is devoted to answer such a question.

We first recall that a canonical transformation is a diffeomorphism which preserves in form the Poisson tensor. Here we are sure that the transformation performed via Lie method is canonical (see 2.1.1). The only doubt can arise from the last step, the one that we call gauge fixing.

The transformation is

$$\begin{aligned} V^+ &\mapsto V^+ + \lambda(\mathcal{G}_1^+ + g(V^+)) + \lambda^2\mathcal{G}_2^+ + O(\lambda^3) \\ V^- &\mapsto V^- + \lambda(\mathcal{G}_1^- + g(V^-)) + \lambda^2\mathcal{G}_2^- + O(\lambda^3) \end{aligned} \tag{4.130}$$

We recall that a transformation is canonical at order λ if it preserves the Poisson tensor till order λ. It means

$$J = (1 + \lambda D\mathcal{F}_1 + \lambda^2 Df_2)J(1 + \lambda D\mathcal{F}_1 + \lambda^2 Df_2)^T. \tag{4.131}$$

For a proof of canonicity of $\mathcal{G}_1$[15] we invite the reader to look at the appendix B.3. Thus we have to prove just that

$$DgJ + JDg^T = 0 \tag{4.132}$$

Since $g(V^+, V^-) = g(V^+) + g(V^-)$ depends only on V^+ or V^- the equation above reduces to prove

$$Dg\partial_x + \partial_x Dg^T = 0. \tag{4.133}$$

Writing g as

$$g(u) = Au^2 + Bu_x\partial_x^{-1}u + Cu_{xx} \tag{4.134}$$

we get

$$Dg(u) = 2Au + B\partial_x^{-1}u\partial_x + Bu_x\partial_x^{-1} + C\partial_x^2 \tag{4.135}$$

[15] It is obvious since the transformation $\mathcal{G}_1$.

and $Dg^T(u)$

$$Dg^T(u) = 2Au - Bu - B\partial_x^{-1}u\partial_x - B\vec{\partial}_x^{-1}u_x + C\partial_x^2. \tag{4.136}$$

Here with the arrow above the differential operator we mean the action on the operator on all what follows by multiplication. If we try to check if the above equation is satisfied we get

$$Dg(u)\partial_x = 2Au\partial_x + B\partial_x^{-1}u\partial_x^2 + Bu_x + C\partial_x^3 \tag{4.137}$$

$$\partial_x Dg^T(u) = 2Au_x + 2Au\partial_x - 2Bu_x - 2Bu\partial_x - B\partial_x^{-1}u\partial_x^2 + C\partial_x^3 \tag{4.138}$$

One thus obtains

$$Dg\partial_x + \partial_x Dg^T = 4Au\partial_x + 2Au_x - Bu_x - 2Bu\partial_x + 2C\partial_x^3 \tag{4.139}$$

and then one conclude that the Poisson tensor is not conserved in form by the transformation.

Here we summarize what we got till now. We performed a Canonical transformation via Lie method to get a Hamiltonian in "normal form" with respect to the first term. Then we performed a general transformation which mapped our system in another dynamical system which is still Hamiltonian with the same Poisson tensor of the starting system but here we found that such a transformation is not canonical so it should not preserve the Poisson tensor.

To enlighten why the Poisson tensor seemed to be unchanged during these processes we start recalling that if we have a Hamiltonian system we can write its equation of motion as

$$\dot{y} = J\nabla_y H(y) \tag{4.140}$$

or, after a diffeomorphism $y = f(x)$

$$\dot{x} = Df^{-1}J(f(x))Df^{-T}\nabla_x H(f(x)). \tag{4.141}$$

If such a transformation is a near-to identity transformation in the form $f(x) = x + \lambda f_1 + \dots$ we get

$$\dot{x} = (\mathbf{1} + \lambda Df_1)^{-1} J(x + \lambda f_1)(\mathbf{1} + \lambda Df_1)^{-T} \nabla_x H(x + \lambda f_1). \tag{4.142}$$

After some steps of algebra and assuming that J does not depend on x, one is left with

$$\dot{x} = \left(J + \lambda(-Df_1 J - JDf_1^T)\right) \nabla_x(H_0 + \lambda(H_0' f_1 + H_1)). \tag{4.143}$$

Defining $J_1 = (-Df_1 J - JDf_1^T)$, equation (4.140) can be written as

$$\begin{aligned} \dot{x} &= (J_0 + \lambda J_1)\nabla_x(H_0 + \lambda(H_0' f_1 + H_1)) = \\ &= J_0\nabla_x H_0 + \lambda\left(J_1\nabla_x H_0 + J_0\nabla_x(H_0' f_1 + H_1)\right). \end{aligned} \tag{4.144}$$

It will be possible to write RHS as $J_0\nabla_x H_1$ (which is precisely what happens in the FPU case with the first order) iff RHS is independent of f_1, i.e. if

$$J_1\nabla_x H_0 + J_0\nabla_x H_0' f_1 = 0. \tag{4.145}$$

This test is easy in our case since $f_1 = g$ and $J_1 = -DgJ - JDg^T$. So it is diagonal and it is sufficient to prove that

$$-(Dg\partial_x + \partial_x Dg^T)\nabla_u H_0 + \partial_x \nabla_u H_0' g \tag{4.146}$$

Since $\nabla_u H_0 = u$ and

$$H_0' g = \int ug\,dx = \int (Au^3 + Buu_x\partial_x^{-1} + Cuu_{xx})\,dx \tag{4.147}$$

its L_2 gradient is easy to compute

$$\begin{aligned}\nabla_u(H_0'g) &= 3Au^2 + Bu_x\partial_x^{-1}u - B\partial_x(u\partial_x^{-1}u) - B\frac{u^2}{2} + 2Cu_{xx} = \\ &= 3Au^2 - \frac{3}{2}Bu^2 + 2Cu_{xx}\end{aligned} \tag{4.148}$$

Applying J one obtains

$$J\nabla_u(H_0'g) = 6Auu_x - 3Buu_x + 2Cu_{xxx}. \tag{4.149}$$

On the other side we have

$$\begin{aligned}(DgJ + JDg^T)u = (4Au\partial_x + 2Au_x - Bu_x - 2Bu\partial_x + 2C\partial_x^3)u = \\ 6Auu_x - 3Buu_x + 2Cu_{xxx}\end{aligned} \tag{4.150}$$

and the condition is satisfied since (4.149) − (4.150) = 0.

Here we see that if $B = 2A$ and $C = 0$ both (4.149) and (4.150) vanishes. This happens only when $\beta = \frac{7}{9}\alpha^2$. Under these conditions, in fact, as showed in remark 4.5.2, the transformation is canonical. Last we see that in (4.128) one has precisely $B = 2A$.

CHAPTER 5

Normal form of the β-FPU system

The phenomenology related to the β model is quite different from the one related to the $\alpha+\beta$ one. This is probably due to the non-existence of an integrable system playing the same role of the Toda for the $\alpha+\beta$ one. One thus sees the formation of a metastable state with the modes of the mKdV equation and then one expects the breaking of integrability to happen at second order.

The study of how the breaking of integrability occurs in the beta model could help to understand how it will happen in the $\alpha+\beta$ one.

5.1 Hamiltonian first order normal form

We don't repeat the general treatment for the interpolation procedure and the choice of variables of the previous chapter in detail. We start again from (4.29):

$$H[S,R] = \int_{\mathbb{T}} \left(\frac{1}{\varepsilon}\phi(\sqrt{\varepsilon}R) - \frac{1}{2}S\Delta_h S \right) dx \tag{5.1}$$

where, instead of (4.30), we get for the potential ϕ

$$\phi(\xi) = \frac{1}{2}\xi^2 + \frac{\beta}{4}\xi^4. \tag{5.2}$$

One can, in analogy of section 4.2, define the right-travelling and left-travelling waves as

$$V^{\pm} = \frac{R \pm S_x}{\sqrt{2}} \tag{5.3}$$

and one obtains, in analogy with (4.41):

$$
\begin{aligned}
H[V^+, V^-] = & \int_{\mathbb{T}} \frac{1}{2}\left(V^{+2} + V^{-2}\right) dx + \\
& + \int_{\mathbb{T}} \left[\frac{\beta\varepsilon}{16}\left(V^{+4} + V^{-4} + 4V^{+3}V^- + 4V^+V^{-3} + 6V^{+2}V^{-2}\right) + \right. \\
& + \left. \frac{h^2}{2\cdot 4!}\left(V_x^{+2} - 2V_x^+V_x^- + V_x^{-2}\right)\right] dx + \\
& + \int_{\mathbb{T}} \frac{h^4}{2\cdot 6!}\left(V_{xx}^{+2} - 2V_{xx}^+V_{xx}^- + V_{xx}^{-2}\right) dx.
\end{aligned}
\tag{5.4}
$$

When one calculates the first order normal form averaging the first order perturbation (3.48) one gets

$$
S_1 = \int_{\mathbb{T}} \left[\frac{\beta\varepsilon}{16}\left(V^{+4} + V^{-4}\right) - \frac{h^2}{2\cdot 4!}\left(V_x^{+2} + V_x^{-2}\right)\right] dx + \frac{3\beta\varepsilon}{8}\langle V^{+2}\rangle_{\mathbb{T}}\langle V^{-2}\rangle_{\mathbb{T}}. \tag{5.5}
$$

From a comparison with mKdV hierarchy (subsection 2.4.5) we see that this Hamiltonian can be mapped into a couple of mKdV Hamiltonians as (2.131) choosing

$$
\gamma = \frac{18\beta\varepsilon}{h^2}. \tag{5.6}
$$

When we calculate the vector field related to $h + S_1$ we get

$$
\begin{cases}
V_\tau^+ = \left(1 + \frac{3\beta\varepsilon}{4}\langle V^{-2}\rangle_{\mathbb{T}}\right)V_x^+ + \frac{3\beta\varepsilon}{4}V^{+2}V_x^+ + \frac{h^2}{4!}V_{xxx}^+ \\
V_\tau^- = -\left(1 + \frac{3\beta\varepsilon}{4}\langle V^{+2}\rangle_{\mathbb{T}}\right)V_x^- - \frac{3\beta\varepsilon}{4}V^{-2}V_x^- - \frac{h^2}{4!}V_{xxx}^-
\end{cases}
\tag{5.7}
$$

To get the perturbative ordering we have to weight each term as we did in section 4.2.2. The procedure is the same and then denoting with M the maximum of V on the integration path we have

$$
V \le M \tag{5.8}
$$

and, calling σ the width of analyticity strip, we get

$$
\frac{h^2}{4!}V_{xxx} \le \frac{6h^2M}{4!\sigma^3}, \qquad \frac{3\beta\varepsilon}{4}V^2V_x \le \frac{3\beta\varepsilon M^3}{4\sigma}. \tag{5.9}
$$

In analogy of section 4.2.2 we have to balance the nonlinear term and the dispersive one to get the following relation between M and σ:

$$
M^2\sigma^2 \le \frac{h^2}{3\beta\varepsilon}. \tag{5.10}
$$

If M is independent of the parameters and, forgetting the numerical coefficients, one gets

$$
\sigma^2 \sim \frac{h^2}{\beta\varepsilon}. \tag{5.11}
$$

This last relation, once introduced in (5.7), shows that the perturbative parameter is formally $\lambda \sim \beta\varepsilon \sim h^2$. This justifies the perturbative approach above.

5.2 Hamiltonian second order normal form

As we did in the previous chapter one can now try to answer the question whether the second order is or is not in mKdV hierarchy. To do so we have to calculate the Hamiltonian generating the canonical transformation at first order. For this computation we need to use equation (3.54), where $\tilde{P}_1$ is

$$\tilde{P}_1 = \int \left[\frac{\beta\varepsilon}{16} \left(4V^{+4}V^- + 4V^{-3}V^+ + 6V^{+2}V^{-2} \right) + \frac{h^2}{4!} V_x^+ V_x^- \right] dx - \frac{3\beta\varepsilon}{8} \langle V^{+2} \rangle_{\mathbb{T}} \langle V^{-2} \rangle_{\mathbb{T}}. \tag{5.12}$$

Thus a straightforward application of (3.47) yields

$$\begin{aligned} G_1 = \int_{\mathbb{T}} \left[\frac{\beta\varepsilon}{16} \left(2V^{-3}\partial_x^{-1}V^+ - 2V^{+3}\partial_x^{-1}V^- + 3V^{-2}\partial_x^{-1}V^{+2} \right) + \frac{h^2}{2\cdot 4!} V_x^- V^+ \right] dx - \\ - \frac{3\beta\varepsilon}{16} \langle V^{+2} \rangle_{\mathbb{T}} \langle V^{-2} \rangle_{\mathbb{T}}. \end{aligned} \tag{5.13}$$

To compute the second order normal form one has to use (3.54) and thus, after a long calculation, one gets

$$\begin{aligned} S_2 = \int \left[-\frac{\beta^2\varepsilon^2}{64} \left(V^{+6} + V^{-6} \right) + \frac{\beta\varepsilon h^2}{8\cdot 4!} \left(V^{-3}V_{xx}^- + V^{+3}V_{xx}^+ \right) + \frac{3h^4}{20(4!)^2} \left(V_{xx}^{+2} + V_{xx}^{-2} \right) \right] dx \\ + \frac{\beta^2\varepsilon^2}{64} \left[\langle V^{-3} \rangle^2 + \langle V^{+3} \rangle^2 + 24 \left(\langle V^{+4} \rangle \langle V^{-2} \rangle + \langle V^{-4} \rangle \langle V^{+2} \rangle \right) + 36 \langle V^{+3} \rangle \langle V^{-3} \rangle \right] - \\ - \frac{9\beta^2\varepsilon^2}{64} \langle V^{+2} \rangle \langle V^{-2} \rangle \left(\langle V^{+2} \rangle + \langle V^{-2} \rangle \right) + \frac{3h^2\beta\varepsilon}{8\cdot 4!} \left(\langle V^{+2} \rangle \langle V_{xx}^- V^- \rangle + \langle V^+ V_{xx}^+ \rangle \langle V^{-2} \rangle \right). \end{aligned} \tag{5.14}$$

With this expression we have all the ingredients to answer the question on integrability.

5.3 Integrability at second order

Due to the quantity of terms appearing in (5.14) we try to eliminate all the terms out of hierarchy piecemeal. Starting from the ones involving only V^+ one can try to bring in hierarchy the following Hamiltonian

$$\hat{S}_2 + \tilde{S}_2 = \frac{3h^4}{20\cdot 4!^2} \left[\int \left(-\frac{5}{27}\gamma^2 V^{+6} + \frac{10}{9}\gamma V^{+3}V_{xx}^+ + V_{xx}^{+2} \right) dx + \frac{5}{27}\gamma^2 \langle V^{+3} \rangle^2 \right]. \tag{5.15}$$

Then it is convenient to write $F_2^{(m)}$, as defined in (2.133), in the following form

$$F_2^{(m)} = M \int \left(\frac{\gamma^2}{18} v^6 + \frac{5\gamma}{9} v^3 v_{xx} + {v_{xx}}^2 \right) dx \tag{5.16}$$

then we will need the first order normal form for the β model (5.5) that, after some maniplulations, reads

$$S_1 = \frac{h^2}{2\cdot 4!} \left\{ \int \left[\frac{1}{6} \frac{18\beta\varepsilon}{h^2} (V^{+4} + V^{-4}) - (V_x^{+2} + V_x^{-2}) \right] dx + \frac{18\beta\varepsilon}{h^2} \langle V^{+2} \rangle_{\mathbb{T}} \langle V^{-2} \rangle_{\mathbb{T}} \right\}. \tag{5.17}$$

Setting $\gamma = \frac{18\beta\varepsilon}{h^2}$ one gets

$$S_1 = \frac{h^2}{2\cdot 4!}\left\{\int\left[\frac{\gamma}{6}(V^{+4}+V^{-4})-(V_x^{+2}+V_x^{-2})\right]dx+\gamma\langle V^{+2}\rangle_{\mathbb{T}}\langle V^{-2}\rangle_{\mathbb{T}}\right\}. \tag{5.18}$$

From now on we call

$$A=\frac{h^2}{2\cdot 4!},\qquad B=\frac{3h^4}{20\cdot 4!^2},\qquad M=B \tag{5.19}$$

where the last condition is posed in order to get a solution for the homological equation.

We will not consider the interaction part in S_1 (i.e. the one involving $\langle V^{+2}\rangle\langle V^{-2}\rangle$ since it is already in hierarchy due to the fact that $\langle V^{\pm 2}\rangle$ is a constant of motion (see appendix A). From now on we denote with $\hat{S}_1$ the first order Hamiltonian without the interaction term. We then get for the mKdV the expression

$$v_t = 2A(\gamma v^2 v_x + v_{xxx}) \tag{5.20}$$

where we wrote v instead of $V^{\pm}$. To get equations of motion related to $\hat{S}_1$ one has to calculate its L_2 gradient and then apply the Gardner tensor. The functional derivative of $\hat{S}_1$ is

$$\hat{S}_1' = B\left(-\frac{10}{9}\gamma^2 v^5+\frac{20}{3}\gamma(v^2 v_{xx}+v{v_x}^2)+2v_{xxxx}\right). \tag{5.21}$$

Applying Gardner tensor one gets

$$v_t = B\left(-\frac{50}{9}\gamma^2 v^4 v_x+\frac{80}{3}\gamma v v_x v_{xx}+\frac{20}{3}\gamma v^2 v_{xxx}+\frac{20}{3}\gamma {v_x}^3+2v_{xxxxx}\right)=\mathcal{S}_2. \tag{5.22}$$

Doing the same procedure for $F_2^{(m)}$ one gets, for its functional derivative,

$$F_2^{(m)\prime} = B\left(\frac{\gamma^2}{3}v^5+\frac{10\gamma}{3}v^2 v_{xx}+\frac{10\gamma}{3}v{v_x}^2+2v_{xxxx}\right). \tag{5.23}$$

Applying Gardner tensor to (5.23) one gets the following equation of motion

$$v_t = B\left(\frac{5\gamma^2}{3}v^4 v_x+\frac{40\gamma}{3}v v_x v_{xx}+\frac{10\gamma}{3}v^2 v_{xxx}+\frac{10\gamma}{3}{v_x}^3+2v_{xxxxx}\right)=\mathcal{F}_2. \tag{5.24}$$

The equation we have to solve in order to bring the β-FPU system in mKdV hierarchy is

$$[\mathcal{S}_1, g] = \mathcal{S}_2-\mathcal{F}_2. \tag{5.25}$$

in the unknown g. If we are able to find such a g then our problem is solved and the β-FPU model can be mapped in mKdV hierarchy at second order. If we were able to show the non existence of such a g then the system will not be integrable at this order. We start computing RHS of (5.25)

$$\mathcal{S}_2-\mathcal{F}_2 = B\left(-\frac{65}{9}\gamma^2 v^4 v_x+\frac{10\gamma}{3}\left(4\gamma v v_x v_{xx}+v^2 v_{xx}+{v_x}^3\right)+\frac{20}{9}\langle v^3\rangle v v_x\right). \tag{5.26}$$

Recalling the definition of Lie bracket we have

$$[\mathcal{S}_1, g] = \mathcal{S}_1' g - g' \mathcal{S}_1 \tag{5.27}$$

We thus get for $\mathcal{S}_1'$ the following expression

$$\mathcal{S}_1' = 2A(2\gamma v v_x + \gamma v^2 \partial_x + \partial_x^3) \tag{5.28}$$

The LHS of (5.25) becomes

$$2AG\left[(2\gamma v v_x + \gamma v^2 \partial_x + \partial_x^3)\tilde{g} - \tilde{g}'(\gamma v^2 v_x + v_{xxx})\right] \tag{5.29}$$

where $\tilde{g} = g/G$ and $G \sim h^2$. We can choose G to get $2AG = B$. We are thus left with the following equation

$$(2\gamma v v_x + \gamma v^2 \partial_x + \partial_x^3)\tilde{g} - \tilde{g}'(\gamma v^2 v_x + v_{xxx}) = -\frac{65}{9}\gamma^2 v^4 v_x + \frac{10\gamma}{3}\left(4\gamma v v_x v_{xx} + v^2 v_{xx} + {v_x}^3\right). \tag{5.30}$$

Setting

$$\tilde{g} = \gamma g_1 + g_0 \tag{5.31}$$

and substituting in (5.30) one obtains

$$\begin{aligned}(2\gamma v v_x + \gamma v^2 \partial_x + \partial_x^3)(\gamma g_1 + g_0) - (\gamma g_1 + g_0)'(\gamma v^2 v_x + v_{xxx}) = \\ = -\frac{65}{9}\gamma^2 v^4 v_x + \frac{10\gamma}{3}\left(4 v v_x v_{xx} + v^2 v_{xx} + {v_x}^3\right).\end{aligned} \tag{5.32}$$

Writing LHS as a polynomial in γ one gets

$$\begin{aligned}&\gamma^2[(2 v v_x + v^2 \partial_x) g_1 - g_1'(v^2 v_x)] + \gamma[(2 v v_x + v^2\partial_x) g_0 + \partial_x^3 g_1 - g_1' v_{xxx} - g_0'(v^2 v_x)] + \\ &+ \partial_x^3 g_0 - g_x'(v_{xxx}) = -\frac{65}{9}\gamma^2 v^4 v_x + \frac{10\gamma}{3}\left(4 v v_x v_{xx} + v^2 v_{xx} + {v_x}^3\right).\end{aligned} \tag{5.33}$$

Proceeding as we did in section 4.5 we compare terms with the same order in γ. This gives us the system of equations

$$\begin{cases} (2 v v_x + v^2 \partial_x) g_1 - g_1'(v^2 v_x) = -\frac{65}{9} v^4 v_x + \frac{20}{9}\langle v^3\rangle v v_x + c_1 \langle v^4 \rangle v_x + c_2 \langle v^2 \rangle v^2 v_x \\ (2 v v_x + v^2 \partial_x) g_0 + \partial_x^3 g_1 - g_1' v_{xxx} - g_0'(v^2 v_x) = \\ \qquad = \frac{10}{3}\left(4 v v_x v_{xx} + v^2 v_{xx} + {v_x}^3\right) + c_3 \langle v^2 \rangle v_{xxx} + c_4 \langle v v_{xx}\rangle v_x \\ \partial_x^3 g_0 - g_x'(v_{xxx}) = 0 \end{cases} \tag{5.34}$$

From a dimensional analysis we get

$$[g_1] \sim v^3 \qquad [g_0] \sim \partial_x^2 v. \tag{5.35}$$

and then without losing generality[(1)]

$$g_1 = a_1 v^3 + a_2 v v_x \partial_x^{-1} v + a_3 v_x \partial_x^{-1}(v^2) + a_4 \langle v^3 \rangle + a_5 \langle v^2 \rangle v, \tag{5.36}$$

[(1)] In our opinion the following are all the possible terms which can appear in the transformation.

$$g_0 = bv_{xx}. \tag{5.37}$$

Our aim is now to insert (5.36) and (5.37) in (5.34) to get a relation between coefficients. Deriving (5.36) and (5.37) one gets

$$\begin{aligned} g_1' = {} & 3a_1v^2 + a_2(v_x\partial_x^{-1}v + v\partial_x^{-1}v\partial_x + vv_x\partial_x^{-1}) + a_3(\partial_x^{-1}v^2\partial_x + 2v_x\vec{\partial}_x^{-1}v) + \\ & + 3a_4\langle v^2(\cdot)\rangle + a_5(2v\langle v(\cdot)\rangle + \langle v^2\rangle), \end{aligned} \tag{5.38}$$

$$g_0' = b\partial_x^2. \tag{5.39}$$

Starting from the easier equation to solve we see that the third one of (5.34) has as solution

$$g_0 = bv_{xx} \tag{5.40}$$

while substituting (5.36) and (5.37) in the second one of (5.34) one gets, on the left hand side

$$\left(2a_1 + \frac{2}{3}a_2 + \frac{a_3}{2}\right)v^4v_x + \left(2a_4 + \frac{a_2}{3}\right)\langle v^3\rangle vv_x + (2a_5 - a_3)\langle v^2\rangle v^2v_x + \frac{a_3}{2}\langle v^4\rangle v_x \tag{5.41}$$

which has to be solved by looking for a solution of the associated homogeneus equation and a particular one. The solution of the associated homogeneus equation is

$$a_2 = -3a_1 \qquad a_4 = \frac{a_1}{2} \tag{5.42}$$

Which yields directly

$$g_1^{(homog)} = a(2v^3 - 6vv_x\partial_x^{-1}v + \langle v^3\rangle). \tag{5.43}$$

To find a particular solution one is left with the problem of finding a solution of

$$\begin{cases} 2a_1 + \frac{2}{3}a_2 + \frac{a_3}{2} = -\frac{65}{9} \\ 2a_4 + \frac{a_2}{3} = \frac{20}{9} \\ 2a_5 - a_3 = c_2 \\ \frac{a_3}{2} = c_1. \end{cases} \tag{5.44}$$

One can see immediately that

$$a_3 = 2c_1, \qquad a_5 = \frac{c_2}{2} + c_1, \qquad a_1 = -\frac{65}{18} - \frac{c_1}{2}, \qquad a_4 = \frac{10}{9}. \tag{5.45}$$

is a particular solution of (5.44). One thus obtains for $g_1^{(part)}$:

$$g_1^{(part)} = \left(-\frac{65}{18} - \frac{c_1}{2}\right)v^3 + \frac{10}{9}\langle v^3\rangle + 2c_1v_x\partial_x^{-1}v^2 + \left(\frac{c_2}{2} + c_1\right)\langle v^2\rangle v. \tag{5.46}$$

Summing (5.43) and (5.46) one gets for g_1

$$g_1 = -\left(\frac{f_1}{2} + \frac{c_1}{2} + 2a\right)v^3 + \left(\frac{f_2}{2} - a\right)\langle v^3\rangle + 6avv_x\partial_x^{-1}v + 2c_1v_x\partial_x^{-1}v^2 + \left(\frac{c_2}{2} + c_1\right)\langle v^2\rangle v. \tag{5.47}$$

The term involving v^3 when inserted in left hand side of the second equation of (5.34) gives

$$(3f_1 - 3c_1 - 12a)(v_x^3 + 3vv_xv_{xx}) \tag{5.48}$$

The term involving $\langle v^3 \rangle$, after the substitution on the same equation gives

$$-3\left(\frac{f_2}{2} - a\right)\langle v^2 v_{xxx}\rangle \tag{5.49}$$

which is a forcing term in the vector field. The term involving $vv_x\partial_x^{-1}v$ yields to

$$18a(v_{xx}^2\partial_x^{-1}v + v_xv_{xxx}\partial_x^{-1}v + 4vv_xv_{xx} + v_x^3 + v^2v_{xxx}), \tag{5.50}$$

the term involving $v_x\partial_x^{-1}v^2$ gives

$$6c_1(v^2v_{xxx} + 2vv_xv_{xx} + v_x^3 - \langle v^2\rangle v_{xxx} - 3v_x\langle v_x^2\rangle) \tag{5.51}$$

and the term involving $\langle v^2\rangle v$ vanishes. The terms involving g_0, inserted in the same equation, gives

$$-bv_x^3 - 4bvv_xv_{xx}. \tag{5.52}$$

Summing (5.48)-(5.52) one obtains an explicit equation to solve with respect to the parameters. The LHS of this equation is

$$\begin{aligned}&(3f_1 - 3c_1 - 12a)(v_x^3 + 3vv_xv_{xx}) - 3\left(\frac{f_2}{2} - a\right)\langle v^2v_{xxx}\rangle+\\&+ 18a({v_{xx}}^2\partial_x^{-1}v + v_xv_{xxx}\partial_x^{-1}v + 4vv_xv_{xx} + v_x^3 + v^2v_{xxx})+\\&+ 6c_1(v^2v_{xxx} + 2vv_xv_{xx} + v_x^3 - \langle v^2\rangle v_{xxx} - 3v_x\langle v_x^2\rangle)-\\&- bv_x^3 - 4bvv_xv_{xx}\end{aligned} \tag{5.53}$$

and it has to put equal to RHS of the second equation of (5.34). So far, comparing with the RHS, one sees that all the terms involving anti-differentiations must disappear which implies $a = 0$. Also the forcing term $\langle v^2v_{xxx}\rangle$ must disappear which means that $f_2 = 0$. But f_2 is fixed from the equation above and then it doesn't exist a transformation in the form $g = \gamma g_1 + g_0$ which maps the second order of β-FPU in mKdV hierarchy.

Up to this point we have that we were not able to map the β model normal form into the mKdV hierarchy. Of course we cannot conclude that it is not possible to do so; but we proceeded in the most general path, in our eyes.

APPENDIX A

Symmetries of the continuous FPU

In the discrete FPU system the translational invariance is a discrete group of transformations which leaves the Hamiltonian invariant. Once we study the extension to continuum this group becomes a $U(1)$ and then there is a constant of motion associated to it thanks to Nöther theorem.

The main consequence of this new constant of motion is that the L_2 norm of V^+ and V^- are conserved until the motion of the $\alpha+\beta$-FPU is due to the KdV hierarchy.

A.1 Nöther theorem for Hamiltonian systems

In this first section we state Nöther theorem for Hamiltonian system as it is stated in [10].

Theorem A.1.1 (Nöther for Hamiltonian systems)**.** *If a system with Hamiltonian $H(x)$ has a one-parameter group of symmetries $G(x,\alpha)$, the Hamiltonian $K(x)$ of which the group is the flow is a first integral for the flow associated with H.*

Proof. The invariance of K can be interpreted as its being constant along the flow generated by K. Therefore $\mathcal{L}_K H = \{H, K\} = 0$.

Conversely this implies that K is a first integral for the flow generated by H. □

A.2 Translational invariance

We take the dynamical system with Hamiltonian (4.24) and Poisson tensor (4.25):

$$H[S,R] = \int_{\mathbb{T}} \left(\frac{1}{\varepsilon}\phi(\sqrt{\varepsilon}R(x,\tau)) - \frac{1}{2}S(x,\tau)\Delta_h S(x,\tau) \right) dx, \tag{A.1}$$

$$\mathbb{E} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{A.2}$$

We recall that S is the coordinate while R is the momentum.

The one-parameter group of diffeomorphism which gives the translation $x \mapsto x+a$ is

$$S(x,\tau) \mapsto S(x+a,\tau), \qquad R(x,\tau) \mapsto R(x+a,\tau) \tag{A.3}$$

This is the flow at time a associated to the Hamiltonian

$$K = \int_{\mathbb{T}} R S_x \, dx. \tag{A.4}$$

In fact, once we write down the equations of motion we get

$$S_a = \frac{\delta K}{\delta R} = S_x, \qquad R_a = -\frac{\delta K}{\delta S} = R_x. \tag{A.5}$$

These can be solved by means of the characteristics and give

$$\Phi_K^a(S(x), R(x)) = (S(x+a), R(x+a)). \tag{A.6}$$

Introducing the right-travelling and left-travelling waves

$$V^{\pm} = \frac{R \pm S_x}{\sqrt{2}}. \tag{A.7}$$

One gets, for the conserved functional, the expression

$$\tilde{K} = \frac{1}{2} \int_{\mathbb{T}} \left(V^{+2} - V^{-2} \right) \, dx. \tag{A.8}$$

Among the conserved functionals of the FPU system, until the order on which the normal form holds, it is conserved also

$$h = \frac{1}{2} \int_{\mathbb{T}} \left(V^{+2} + V^{-2} \right) \, dx. \tag{A.9}$$

It is a straightforward consequence of the definition of normal form:

$$\{h, H\} = \{h, h + \lambda P\} = \{h, h\} + \lambda \{h, P\} \tag{A.10}$$

which vanishes if P is in normal form with respect to h.

At this point a linear combination of conserved functionals is still a conserved functional. These are $K+h$ and $h-K$ which are precisely

$$K + h = \int_{\mathbb{T}} V^{+2} \, dx \tag{A.11}$$

$$K - h = \int_{\mathbb{T}} V^{-2} \, dx \tag{A.12}$$

and this states the conservation of the L_2 norm of V^+ and V^- up to the time scale on which the dynamics is the one of the normal form.

APPENDIX B

Perturbative computations for the $\alpha + \beta$-model

B.1 Exactness of the generating Hamiltonian at first order

We recall here the result (3.42) of section 3.2.1 which states the relation between the perturbative Hamiltonian (more precisely, between its Lie derivative) and the first order normal form S_1.

$$S_1 = P_1 + \mathcal{L}_1 h \tag{B.1}$$

where the Lie derivative can be expressed by definition via Poisson brackets as (2.22) states:

$$\mathcal{L}_1 h = \{h, G_1\}. \tag{B.2}$$

It is also useful for the present computation recall that

$$\{h, G_1\} = \langle \nabla_{L_2} h | J \nabla_{L_2} G_1 \rangle_{L_2}, \tag{B.3}$$

and the coordinate expressions for L_2-gradient and Poisson tensor J are

$$\nabla_{L_2} = \begin{pmatrix} \dfrac{\delta}{\delta V^+} \\ \dfrac{\delta}{\delta V^-} \end{pmatrix} \qquad\qquad J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \partial_x \tag{B.4}$$

To show that the expression found for G_1 in (4.74) satisfies (B.1) it is sufficient to put together that expression for G_1, the expression for S_1 in (4.58) and the expressions for h and P_1 in (4.41) and verify if (B.1) is satisfied. We write down here these expressions for completeness:

$$h[V^+, V^-] = \int_0^1 \frac{1}{2} \left(V^{+2} + V^{-2} \right) dx, \tag{B.5}$$

$$\begin{aligned} P_1[V^+, V^-] = \int_0^1 \Bigg[& \frac{\alpha\sqrt{\varepsilon}}{6\sqrt{2}} \left(V^{+3} + 3V^{+2}V^- + 3V^+V^{-2} + V^{-3} - \right) \\ & - \frac{h^2}{4!2} \left(V^{+2} - 2V^+V^- + V^{-2} \right) \Bigg] dx \end{aligned} \tag{B.6}$$

$$S_1[V^+,V^-] = \int_0^1 \left[\frac{\alpha\sqrt{\varepsilon}}{6\sqrt{2}}\left(V^{+3}+V^{-3}\right) - \frac{h^2}{4!2}\left(V_x^{+2}+V_x^{-2}\right)\right] dx, \tag{B.7}$$

$$G_1[V^+,V^-] = \int_0^1 \left[\frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}\left(\partial_x^{-1}V^+V^{-2} - \partial_x^{-1}V^-V^{+2}\right) + \frac{1}{2}\frac{h^2}{4!}V^+\partial_x V^-\right] dx. \tag{B.8}$$

We start calculating the L_2-gradient of G_1 and, using the integration by parts property of anti-derivative operators stated in section 2.3.3 we get

$$\nabla_{L_2} G_1(V^+,V^-) = \begin{pmatrix} \frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}(-2V^+\partial_x^{-1}V^- - \partial_x^{-1}(V^-)^2) + \frac{h^2}{2\cdot4!}V_x^- \\ \frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}(\partial_x^{-1}(V^+)^2 + 2V^-\partial_x^{-1}V^+) - \frac{h^2}{2\cdot4!}V_x^+ \end{pmatrix}. \tag{B.9}$$

We now apply the Poisson tensor (B.4) to get[1]

$$J\nabla_{L_2} G_1 = \begin{pmatrix} -\frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}\left(2V^-V^+ + 2V_x^+\partial_x^{-1}V^- + V^{-2} - \langle V^{-2}\rangle_{\mathbb{T}}\right) + \frac{h^2}{2\cdot4!}V_{xx}^- \\ -\frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}\left(2V^+V^- + 2V_x^-\partial_x^{-1}V^+ + V^{+2} - \langle V^{+2}\rangle_{\mathbb{T}}\right) + \frac{h^2}{2\cdot4!}V_{xx}^+ \end{pmatrix}. \tag{B.10}$$

For sake of completeness we write down also the L_2 gradient of h which is

$$\nabla_{L_2} h(V^+,V^-) = \begin{pmatrix} V^+ \\ V^- \end{pmatrix}. \tag{B.11}$$

We recall then that the L_2 scalar product works as

$$\langle v|w\rangle_{L_2} = \int_{\mathbb{T}} v\cdot w\, dx \tag{B.12}$$

where with $\cdot$ we denote the standard Euclidean scalar product. We obtain then

$$\begin{aligned}\{h,G_1\} &= \langle\nabla_{L_2}h, J\nabla_{L_2}G_1\rangle_{L_2} = \\ &= \int_0^1 \Bigg[-\frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}\Big(2V^-V^{+2} + 2\partial_x^{-1}V^-V_x^+V^+ + V^+V^{-2} + V^-V^{+2} \\ &\quad +2V^-V^{-2} + 2V^+\partial_x^{-1}V^+V^-\Big) + \frac{h^2}{2\cdot4!}\left(V^+V_{xx}^- + V^-V_{xx}^+\right)\Bigg] dx = \\ &= \int_0^1 \Bigg[-\frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}\Big(3V^{+2}V^- + 3V^+V^{-2} + 2\partial_x^{-1}V^-V_x^+V^+ + \\ &\quad +2V^-\partial_x^{-1}V^+V_x^-\Big) + \frac{h^2}{2\cdot4!}\left(V^+V_{xx}^- + V^-V_{xx}^+\right)\Bigg] dx.\end{aligned} \tag{B.13}$$

We now use the identity $2V^+V_x^+ = \partial_x(V^+)^2$ and we integrate by parts to bring the derivative operator in front of the antiderivative operator. We obtain then

$$\begin{aligned}\{h,G_1\} &= \int_0^1 \Bigg[-\frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}\left(3V^{+2}V^- + 3V^+V^{-2} - V^{+2}V^- - V^{-2}V^+\right) - \\ &\quad -\frac{h^2}{4!}\left(V_x^+V_x^-\right)\Bigg] dx = \\ &= \int_0^1 \left[-\frac{\alpha\sqrt{\varepsilon}}{4\sqrt{2}}\left(2V^{+2}V^- + 2V^+V^{-2}\right) - \frac{h^2}{4!}\left(V_x^+V_x^-\right)\right] dx\end{aligned} \tag{B.14}$$

If we now sum (B.6) and (B.14) we obtain (B.7), which is what we wanted to prove.

[1] Note that it is not possible to apply here the integration by parts property of anti-derivative operators or for derivative ones since this expression is not integrated alone.

B.2 Normal form of second order perturbation

We start this section recalling that the second order normal form can be obtained from the perturbative series with (4.78):

$$S_2 = \langle P_2 + \frac{1}{2}\{\tilde{P}_1, \tilde{G}_1\}\rangle_h. \tag{B.15}$$

where from (4.77) we have

$$\begin{aligned} P_2[V^+, V^-] = &\int_0^1 \left[\frac{\beta\varepsilon}{16}\left(V^{+4} + V^{-4} + 4V^{+3}V^- + 4V^{-3}V^+ + 6V^{+2}V^{-2}\right)\right] dx + \\ &+ \int_0^1 \left[\frac{1}{2}\frac{h^2}{6!}\left(V_{xx}^{+2} - 2V_{xx}^+V_{xx}^- + V_{xx}^{-2}\right)\right] dx. \end{aligned} \tag{B.16}$$

and we recall (4.71):

$$\tilde{P}_1 = \int_0^1 \left[\frac{\alpha\sqrt{\varepsilon}}{2\sqrt{2}}\left(\partial_x^{-1}V^+V^{-2} + V^{+2}\partial_x^{-1}V^-\right) + \frac{h^2}{4!}V_x^+V_x^-\right] dx.. \tag{B.17}$$

As a first step we evaluate $\langle P_2\rangle_h$. Starting from (B.16), using `Proposition 4.2.2` and `Proposition 4.2.3` we obtain immediately

$$\langle P_2\rangle_h = \int_0^1 \left[\frac{\beta\varepsilon}{16}\left(V^{+4} + V^{-4}\right) + \frac{h^2}{2\cdot 6!}\left(V_{xx}^{+2} + V_{xx}^{-2}\right)\right] dx + \frac{3\beta}{8}\varepsilon\langle V^{+2}\rangle_{\mathbb{T}}\langle V^{-2}\rangle_{\mathbb{T}} \tag{B.18}$$

We then calculate the Poisson bracket $\{\tilde{P}_1, \tilde{G}_1\}$ where $\tilde{G}_1 = G_1 - \langle G_1\rangle_h$ and, under the assumptions of the previous section, we have $\tilde{G}_1 = G_1$. In the previous section we calculated already $J\nabla_{L_2}G_1$ in (B.10). We calculate then $\nabla_{L_2}\tilde{P}_1$:

$$\nabla_{L_2}\tilde{P}_1(V^+, V^-) = \begin{pmatrix} \frac{\alpha\sqrt{\varepsilon}}{2\sqrt{2}}\left(V^{-2} + 2V^+V^-\right) - \frac{h^2}{4!}V_{xx}^- \\ \frac{\alpha\sqrt{\varepsilon}}{2\sqrt{2}}\left(V^{+2} + 2V^+V^-\right) - \frac{h^2}{4!}V_{xx}^+ \end{pmatrix}. \tag{B.19}$$

With this result we have to calculate the L_2-scalar product. With this aim we start computing the scalar product $\nabla_{L_2}\tilde{P}\cdot J\nabla_{L_2}G_1$ component by component:

$$\begin{aligned} &J_{11}\frac{\delta G_1}{\delta V^+}\frac{\delta\tilde{P}_1}{\delta V^+} = \\ &= -\frac{\alpha^2\varepsilon}{16}\Big(V^{-4} + 4V^{-3}V^+ + 4V^{+2}V^{-2} + 2\partial_x^{-1}V^-V_x^+V^{-2} + 4V^+V^-\partial_x^{-1}V^-V_x^+ - \\ &-2\langle V^{-2}\rangle_{\mathbb{T}}V^+V^- - \langle V^{-2}\rangle_{\mathbb{T}}V^{-2}\Big) + \frac{h^2\alpha\sqrt{\varepsilon}}{4\sqrt{2}4!}V_{xx}^-\Big(2V^{-2} + 4V^-V^+ + 2\partial_x^{-1}V^-V_x^+ - \\ &-\langle V^{-2}\rangle_{\mathbb{T}}\Big) - \frac{1}{2}\frac{h^4}{(4!)^2}V_{xx}^{-2} \end{aligned} \tag{B.20}$$

$$J_{22}\frac{\delta G_1}{\delta V^-}\frac{\delta \tilde{P}_1}{\delta V^-}=$$
$$=-\frac{\alpha^2\varepsilon}{16}\Big(4V^+V^-V_x^-\partial_x^{-1}V^+ + 4V^{+2}V^{-2}+4V^{+3}V^- + 2V^{+2}V_x^-\partial_x^{-1}V^+ + V^{+4}-$$
$$-2\langle V^{+2}\rangle_{\mathbb{T}}V^-V^+ - \langle V^{+2}\rangle_{\mathbb{T}}V^{+2}\Big)+\frac{h^2\alpha\sqrt{\varepsilon}}{4\sqrt{24!}}V_{xx}^+\Big(2V_x^-\partial_x^{-1}V^+ + 4V^+V^- + 2V^{+2}-$$
$$-\langle V^{+2}\rangle_{\mathbb{T}}\Big)-\frac{1}{2}\frac{h^4}{(4!)^2}\left(V_{xx}^+\right)^2. \tag{B.21}$$

We have to sum the two expressions above, to integrate the sum on the torus and then we have to perform a time average. We don't write the sum of these two terms since it is just a waste of paper. After writing the integral, we use Leibnitz rule to simplify the expression obtained and then we perform a time average[(2)] to obtain

$$\begin{aligned}\langle\{\tilde{P}_1,G_1\}\rangle_h = \int_0^1 \Big[-\frac{\alpha^2\varepsilon}{16}\left(V^{+4}+V^{-4}\right)-\frac{h^4}{2(4!)^2}\left(V_{xx}^{+2}+V_{xx}^{-2}\right)+ \\ +\frac{h^2\alpha\sqrt{\varepsilon}}{2\sqrt{24!}}\left(V^{-2}V_{xx}^- + V^{+2}V_{xx}^+\right)\Big]\,dx+\frac{\alpha^2\varepsilon}{2}\langle V^{+2}\rangle_{\mathbb{T}}^2\langle V^{-2}\rangle_{\mathbb{T}}^2\end{aligned} \tag{B.22}$$

During the calculations we used the following statement

Proposition B.2.1. *Let f and g be two real valued functions on the torus such that, under the flow of h,*

$$f(x)\mapsto f(x+s) \qquad\qquad g(x)\mapsto g(x-s) \tag{B.23}$$

we have

$$\int_0^1 ds\int_0^1 e^{s\mathcal{L}_h}\left(4f^2g^2-2f^2g_x\partial_x^{-1}g-2g^2f_x\partial_x^{-1}f\right)\,dx = 8\langle f^2\rangle_{\mathbb{T}}\langle g^2\rangle_{\mathbb{T}}. \tag{B.24}$$

Proof. We start calculating $\langle g_x\partial_x^{-1}g\rangle_{\mathbb{T}}$:

$$\langle g_x\partial_x^{-1}g\rangle_{\mathbb{T}}=\int_{\mathbb{T}}g_x\partial_x^{-1}g\,dx \overset{\text{Leibnitz}}{=} -\int_{\mathbb{T}}g^2\,dx = -\langle g^2\rangle_{\mathbb{T}}.$$

The same holds for f, so we have

$$\langle f_x\partial_x^{-1}f\rangle_{\mathbb{T}}=-\langle f^2\rangle_{\mathbb{T}}.$$

Using now `Proposition 4.2.2` we can perform the integration (B.24) to get

$$\int_0^1 ds\int_0^1 e^{s\mathcal{L}_h}\left(4f^2g^2-2f^2g_x\partial_x^{-1}g-2g^2f_x\partial_x^{-1}f\right)\,dx =$$
$$=4\langle f^2\rangle_{\mathbb{T}}\langle g^2\rangle_{\mathbb{T}}-2\langle f^2\rangle_{\mathbb{T}}\langle g_x\partial_x^{-1}g\rangle_{\mathbb{T}}-2\langle g^2\rangle_{\mathbb{T}}\langle f_x\partial_x^{-1}f\rangle_{\mathbb{T}}$$

which can be simplified using the results above as

$$\int_0^1 ds\int_0^1 e^{s\mathcal{L}_h}\left(4f^2g^2-2f^2g_x\partial_x^{-1}g-2g^2f_x\partial_x^{-1}f\right)\,dx = 8\langle f^2\rangle_{\mathbb{T}}\langle g^2\rangle_{\mathbb{T}}$$

which is precisely our thesis. ∎

(2) We recall that $V^{\pm}$ has vanishing mean and so are its derivatives.

Summing now (B.18) and (B.22) divided by two, taking into account that $V^{\pm}_{xx}$ have vanishing mean, we get as Second order normal form

$$S_2[V^+,V^-] = \int_0^1 \left[\left(\frac{\beta\varepsilon}{16} - \frac{\alpha^2\varepsilon}{32} \right) \left(V^{+4} + V^{-4} \right) + \frac{h^2\alpha\sqrt{\varepsilon}}{4\sqrt{24!}} \left(V^{-2}V^-_{xx} + V^{+2}V^+_{xx} \right) + \right.$$
$$\left. + \frac{3}{20}\frac{h^4}{(4!)^2} \left(V^{+2}_{xx} + V^{-2}_{xx} \right) \right] dx + \left(\frac{3\beta\varepsilon}{8} - \frac{\alpha^2\varepsilon}{4} \right) \langle V^{+2} \rangle_{\mathbb{T}} \langle V^{-2} \rangle_{\mathbb{T}} +$$
$$+ \frac{\alpha^2\varepsilon}{32} \left(\langle V^{-2} \rangle_{\mathbb{T}} + \langle V^{-2} \rangle_{\mathbb{T}} \right) \tag{B.25}$$

B.3 Canonicity at first order: a direct calculation

Here we want to verify that the transformation generated by $\mathcal{G}_1$ is canonical at first order using the definition. If we recall results of section 2.1.1 we have that a transformation $\mathcal{G}_1$ is canonical at first order if and only if

$$D\mathcal{G}_1 J + J(D\mathcal{G}_1)^T = 0 \tag{B.26}$$

Where J is given by (B.4) and $\mathcal{G}_1$ is given by

$$\mathcal{G}_1 = \begin{pmatrix} \mathcal{G}_1^+ \\ \mathcal{G}_1^- \end{pmatrix} = \begin{pmatrix} -AV^{-2} - 2AV^+_x\partial_x^{-1}V^- - 2AV^+V^- + BV^-_{xx} \\ -AV^{+2} - 2AV^-_x\partial_x^{-1}V^+ - 2AV^+V^- + BV^+_{xx} \end{pmatrix} \tag{B.27}$$

We have to calculate its Jacobian and we obtain

$$D\mathcal{G}_1 = \begin{pmatrix} -2A\partial_x^{-1}V^-\partial_x - 2AV^- & -2AV^- - 2AV^+_x\partial_x^{-1} - 2AV^+ + B\partial_x^2 \\ -2AV^+ - 2AV^-_x\partial_x^{-1} - 2AV^- + B\partial_x^2 & -2A\partial_x^{-1}V^+\partial_x - 2AV^+ \end{pmatrix} \tag{B.28}$$

Recalling that given an operator O we define its transposed, denoted as O^T, as the operator satisfying the following relation

$$\langle f(u) | O^T g(u) \rangle_{L_2} = \langle O f(u) | g(u) \rangle_{L_2} \tag{B.29}$$

where f and g are vector fields on L_2. A direct calculation shows that

$$(D\mathcal{G}_1)^T = \begin{pmatrix} 2A\partial_x^{-1}V^-\partial_x & -2AV^+ + 2A\vec{\partial}_x^{-1}V^+_x - 2AV^+ + B\partial_x^2 \\ -2AV^- + 2A\vec{\partial}_x^{-1}V^+_x - 2AV^+ + B\partial_x^2 & 2A\partial_x^{-1}V^+\partial_x \end{pmatrix}. \tag{B.30}$$

We have just to prove that these two operators satisfies the canonicity relation above (B.26). Thus we calculate

$$D\mathcal{G}_1 J = \begin{pmatrix} -2A\partial_x^{-1}V^-\partial_x^2 - 2AV^-\partial_x & 2AV^-\partial_x + 2AV^+_x + 2AV^+\partial_x - B\partial_x^3 \\ -2AV^+\partial_x - 2AV^-_x - 2AV^-\partial_x + B\partial_x^3 & 2A\partial_x^{-1}V^+\partial_x^2 + 2AV^+\partial_x \end{pmatrix} \tag{B.31}$$

Similarly

$$J(D\mathcal{G})^T = \begin{pmatrix} 2A\partial_x^{-1}V^-\partial_x^2 + 2AV^-\partial_x & -2AV^-\partial_x - 2AV^+_x - 2AV^+\partial_x + B\partial_x^3 \\ 2AV^+\partial_x + 2AV^-_x + 2AV^-\partial_x - B\partial_x^3 & -2A\partial_x^{-1}V^-\partial_x^2 + 2AV^+\partial_x \end{pmatrix} \tag{B.32}$$

And we see that the condition of canonicity is completely satisfied.

Bibliography

[1] Arnol'd, V. I. (1978). *Mathematical methods of classical mechanics*. Graduate texts in mathematicals. Springer, New York, NY. Transl. from the Russian, publ. by Nauka, Moscow, 1974.

[2] Babelon, O., Bernard, D., and Talon, M. (2003). *Introduction to classical integrable systems*. Cambridge Univ. Press, Cambridge.

[3] Bambusi, D. and Ponno, A. (2006). On metastability in FPU. *Commun. Math. Phys.*, 264(2):539–561.

[4] Bambusi, D. and Ponno, A. (2008). Resonance, metastability and blow up in FPU. *Lecture notes in physics*, (728):191–205.

[5] Benettin, G., Christodoulidi, H., and Ponno, A. (2013). The Fermi-Pasta-Ulam problem and its underlying integrable dynamics. *Journal of Statistical Physics*, 152(2):195–212.

[6] Benettin, G., Livi, R., and Ponno, A. (2008). The Fermi-Pasta-Ulam problem: Scaling laws vs. initial conditions. *Journal of Statistical Physics*, 135(5-6):873–893.

[7] Benettin, G. and Ponno, A. (2011). Time-scales to equipartition in the Fermi–Pasta–Ulam problem: Finite-size effects and thermodynamic limit. *Journal of Statistical Physics*, 144(4):793–812.

[8] Boussinesq, J. (1877). Essai sur la théorie des eaux courantes. *Mémoires présentés par divers savants à l'Acad. des Sci. Inst. Nat. France*, (XXIII).

[9] de Jager, E. M. (2011). On the origin of the Korteweg-de Vries equation. *Forum der Berliner Mathematischen Gesellschaft*, 19:171–195.

[10] Fasano, A. and Marmi, S. (2006). *Analytical Mechanics*. Oxford University Press Inc., New York, 1st edition.

[11] Ferguson, W. E., Flaschka, H., and McLaughlin, W. (1982). Nonlinear normal modes for the Toda chain. *Journal of Computational Physics*, 45:157–209.

[12] Fermi, E., Pasta, J., and Ulam, S. (1955). Studies of non linear problems. *Los Alamos report LA-1940,* published later in *Collected Papers of Enrico Fermi,* Univ. Chicago Press (1965), pages 978 – 988.

[13] Flaschka, H. (1975). Discrete and periodic illustrations of some aspects of the inverse method. *Lecture Notes in Physics*, 38:441–466.

[14] Fucito, F., Marchesoni, F., Marinari, E., Parisi, G., Peliti, L., Ruffo, S., and Vulpiani, A. (1982). Approach to equilibrium in a chain of nonlinear oscillators. *J. Phys. France*, 43(5):707–713.

[15] Gardner, C. S., Greene, J. M., Kruskal, M. D., and Miura, R. M. (1967). Method for solving the Korteweg-de Vries equation. *Physical review letters*, 19(19):1095–1097.

[16] Kittel, C. (1986). *Introduction to Solid State Physics.* John Wiley & Sons, Inc., New York, 6th edition.

[17] Kolmogorov, A. N. and Fomin, S. V. (1980). *Elementi di analisi funzionale.* Edizioni MIR, Moscow, 1st edition.

[18] Lax, P. D. (1976). Almost periodic solutions of the KdV equation. *SIAM review*, 18(3):351–357.

[19] Miura, R. M. (1976). The Korteweg-de Vries equation: A survey of results. *SIAM Review*, 18(3):412–459.

[20] Ponno, A. (2003). Soliton theory and the Fermi-Pasta-Ulam problem in the thermodynamic limit. *Europhysics letters*, 64(5):606–612.

[21] Ponno, A. and Bambusi, D. (2005). Korteweg–de Vries equation and energy sharing in Fermi–Pasta–Ulam. *Chaos*, 15(1):015107.

[22] Rudin, W. (1987). *Real and complex analysis.* McGraw-Hill Book Co., New York, 3rd edition.

[23] Zabusky, N. and Kruskal, M. (1965). Interaction of solitons in a collisionless plasma and the recurrence of initial states. *Physical review letters*, 15(6):240–242.