On Metastability in FPU

D. Bambusi, A. Ponno

03.05.05

Abstract

We present an analytical study of the Fermi–Pasta–Ulam (FPU) α–model with periodic boundary conditions. We analyze the dynamics corresponding to initial data with some low frequency Fourier modes excited. We show that, correspondignly, a pair of KdV equations constitute the resonant normal form of the system. We also use such a normal form in order to prove the existence of a metastability phenomenon. More precisely, we show that the time average of the modal energy spectrum rapidly attains a well defined distribution corresponding to a packet of low frequencies modes. Subsequently, the distribution remains unchanged up to the time scales of validity of our approximation. The phenomenon is controlled by the specific energy.

1 Introduction

In this paper we present an analytical study of the Fermi–Pasta–Ulam (FPU) α–model with periodic boundary conditions for initial data with some low frequency Fourier modes excited. We give some rigorous results concerning the relaxation to a metastable state, in which energy sharing takes place among low frequency modes only.

The FPU model consists of a long chain of particles interacting with their nearest neighbours through nonlinear springs. It was first introduced and studied numerically by FPU [FPU65] in order to determine the time of approach to equilibrium of the system. In the FPU original experiment all the energy was initially given to a single low frequency Fourier mode and the energies of the Fourier modes were plotted vs. time. The result was surprising: energy sharing occurred only among a few low frequency modes and an almost recurrent behaviour of the solution were observed. On the contrary, a fast approach to a state characterized by equipartition of all modal energies was expected.

The chain numerically integrated by FPU was composed by a relatively small number of particles; a problem that naturally arises is that of understanding whether the unexpected lack of equipartition persists when the number of particles grows. Actually a huge number of numerical computations have been performed [BGG05, BKL03, LPRV85, PL90], but the situation is not yet clear.

From the theoretical point of view, in the FPU problem there were initially two lines of research. The first one originated from the paper [ZK65] by Zabusky and Kruskal, who numerically studied the dynamics of the Kortweg de Vries equation (KdV) which was heuristically known to describe the long wave solutions of the FPU. The authors observed a recurrent behaviour in the KdV and interpreted it as a possible explanation of the FPU recurrence. The paper by Zabusky and Kruskal constituted the starting point of the theory of Lax–integrable partial differential equations, but, as far as we know, the relevance of the KdV equation for the FPU relaxation problem was never completely clarified.

A second line of research was initiated by Izrailev and Chirikov [IC66] and is based on the Kolmogorov Arnold Moser (KAM) theorem, or more generally on the application of canonical perturbation theory to the study of FPU. The idea of Izrailev and Chirikov is that an energy threshold exists, below which KAM theory is (in principle) applicable (actually the applicability of KAM theory to FPU is a delicate question, since one has to verify the validity of the KAM nondegeneracy condition, which was accomplished only recently in the paper [Rin01]). The main point is the dependence of such a threshold on the number of degrees of freedom. The thesis by Izrailev and Chirikov is that, if the mode initially excited has high frequency, then the threshold goes to zero as the number of degrees of freedom increases, so that the region of recurrent motions becomes irrelevant. Afterwards, many heuristic arguments have been developed in order to support and refine Chirikov's thesis. In particular, Shepeliansky gave some heuristic arguments according to which Chirikov's thesis should hold also for low frequency initial excitations [She97]. Anyway, up to now no rigorous result is available.

It has to be noticed that the thesis of Izrailev–Chirikov–Shepeliansky is hardly compatible with the result of Zabusky–Kruskal: according to the former authors the FPU phenomenon disappears when the number of degrees of freedom is large, while the latter explain the FPU recurrence by making use of a PDE, which requires a large number of degrees of freedom.

Finally a new theoretical scenario, which we call the metastability scenario, was proposed for the FPU problem in the paper [FMM+82] (see also [LPRV85]). The thesis is that the FPU system approaches, in a relatively short time, a first state whose modal energy spectrum displays a plateau of equipartition among low frequency modes, followed by an exponentially decreasing tail in the region of high frequencies. Complete equipartition is eventually reached on a second very long time–scale. In [FMM+82] the presence of the exponential tail in the energy spectrum of the metastable state was explicitly referred to as "similar to Wien's law for black–body radiation". Actually such an analogy was previously pointed out by Galgani and Scotti [GS72], who fitted the FPU energy spectrum to a Planck–like distribution. A new emphasis to such a metastability scenario was given in the papers [CG99, CG01, CGG05].

2 Main ideas

In the present paper we consider low frequency initial data and, following the line sketched in [PB05b], we unify the first two approaches presented above, in the sense that we show that canonical perturbation theory leads to the Zabusky–Kruskal result. More precisely, we show that *a pair of KdV equations constitute the resonant normal form of FPU in the standard sense of canonical perturbation theory*. We also use such a normal form in order to give a first rigorous result on energy sharing among the modes. In doing this, we show that the result of Zabusky–Kruskal is controlled by specific energy, so that the stability phenomenon should persist in the thermodynamic limit, against the thesis of Izrailev–Chirikov–Shepeliansky. On the other hand, we make a bridge with the metastability scenario of [FMM$^+$82], because we point out the relevance of the time scales over which different qualitative descriptions of the dynamics holds.

More precisely, we consider a very long chain with periodic boundary conditions, and focus on initial data in which only one Fourier mode with very small index (i.e. with low frequency) and its higher harmonics are initially excited. It is useful to describe the system using an interpolating function, namely a function whose values at integers are the displacements of the particles from equilibrium. It turns out that such an interpolating function has to fulfil a differential-difference equation which is well approximated (for long wavelengths) by a partial differential equation coinciding, at first order, with the linear string equation. More precisely, the Hamiltonian of the system describing the interpolating function has the structure

$$H_0 + P + \mathcal{R}_1 \ , \tag{2.1}$$

where H_0 is the Hamiltonian of the linear wave equation, P contains the lowest order (nonlinear and dispersive) corrections, and $\mathcal{R}_1$ contains higher order corrections. In order to take into account the corrections to the dynamics due to P we use the methods of Hamiltonian perturbation theory. In particular we apply the Galerkin averaging method of [Bam05]. Thus we construct a canonical transformation conjugating the original system to a system with Hamiltonian

$$H_0 + \langle P \rangle + \mathcal{R}$$

where $\langle P \rangle$ is the average of P with respect to the flow of H_0 (which coincides with the normal form of the system), and $\mathcal{R}$ is a remainder whose size is here rigorously estimated (uniformly with respect to the length of the chain) for states with small specific energy (but large total energy).

Then, we explicitly compute the averaged Hamiltonian $H_0 + \langle P \rangle$, and show that its equations of motions consist of a pair of uncoupled KdV equations with periodic boundary conditions on a ring of length 2 (independently of the number of particles).

As a third step we use these KdV's to construct approximate solutions of the FPU chain and we estimate the error with respect to a true solution. We point out that, denoting by μ the wave number of the initially excited mode, which

we assume to have specific energy $\mathcal{E} \equiv E/N \sim \mu^4$ (where N is the number of particles and E the total energy), the dynamics of the KdV equations gives rise to finite size effects over a time–scale μ^{-3}. In order to get an estimate of the error valid over such a time–scale we use a technique by Schneider and Wayne [SW00]. It turns out that, having fixed an arbitrarily long time T_f, the KdV's describe the solutions of the FPU up to a time $T_f \mu^{-3}$.

Finally we use known results on the KdV dynamics with periodic boundary conditions in order to compute the energy per mode along an approximate solution of the FPU system. In particular, denoting by E_k the energy in the k-th mode and by $\mathcal{E}_k := E_k/N$ the corresponding specific energy, we prove that, for the considered initial data, $\mathcal{E}_k$ decreases as $\exp(-\sigma\kappa/\mu)$, with $\kappa = k/N$ and $\sigma > 0$, at least for times such that the approximation is valid. Moreover, if we consider the time average of $\mathcal{E}_k$, we prove that it quickly relaxes to a certain energy distribution, and then remains unchanged up to the times accessible within our approximation.

Notice that the time–scale $\mu^{-3} \sim \mathcal{E}^{-3/4}$ for the formation of the packet and the width $\mu \sim \mathcal{E}^{1/4}$ display the same dependence on the specific energy $\mathcal{E}$ as numerically observed in [BGG05] and [BKL03], and heuristically predicted in [PB05b, Pon03, Pon05, PB05a]. As far as we now, this is the first rigorous result on a large FPU chain with finite specific energy. Moreover, this is a first rigorous description of the fast formation of a metastable packet of modes of the type of that observed by FPU.

The main limitation of our result concerns the choice of the initial data: one would like to consider initial data involving e.g. a small packet of nearby modes as in most numerical computations. The reason of our limitation is that the manifold consisting of states with only one mode and its higher harmonics excited is invariant and we exploit such a property. The fact that the result involves specific energy and is in agreement with numerical results suggested that this limitation could be just a technical one.

From the technical point of view, the core of our paper consists in the proof that a pair of KdV's is the normal form of the FPU problem and in an estimation of the error. We point out that a previous result on the justification of KdV as a modulation equation for FPU was obtained by Schneider and Wayne in [SW00]. In their paper the attention was restricted to the case of solutions fast decreasing in space, whereas we deal here with space–periodic ones. The fact that a pair of uncoupled KdV equations describes well the FPU dynamics when the initial datum is space periodic is quite surprising. Indeed, the two waves travelling in the chain and described by the KdV equations continue to interact forever and one might expect some constructive interference to occur. This is not the case, essentially due to the structure of the FPU nonlinearity. This is in sharp contrast with the typical behaviour for short waves; see [PW95].

We also mention the papers [**?**] where a remarkable connection between the FPU and the KdV has been obtained. However, also this series of papers referrs to initial data that decay fast in space and thus is not directly connected with the problem of thermalization.

3 Main Result

Consider the Hamiltonian system

$$H(q,p) = \sum_{j=-N}^{N-1} \frac{p_j^2}{2} + U(q_{j+1} - q_j) \quad , \tag{3.1}$$

$$U(x) = \frac{x^2}{2} + \frac{x^3}{3} \quad , \tag{3.2}$$

$$q_{j+2N} = q_j \quad , \quad p_{j+2N} = p_j \quad , \tag{3.3}$$

describing a chain composed by $2N$ particles interacting through nonlinear springs. The canonical variables are $q = (q_{-N}, \ldots, q_{N-1})$, $p = (p_{-N}, \ldots, p_{N-1})$. The Hamiltonian (3.1) is known as the Fermi, Pasta and Ulam (FPU) α-model (with $\alpha = 1$). Remark that, due to the periodic boundary conditions (3.3), the total linear momentum of the system is preserved. So one can restrict oneself to the case $\sum_j p_j = \sum_j q_j = 0$.

Introduce the Fourier coefficients by

$$p_j = \frac{1}{\sqrt{2N}} \sum_{k=-N}^{N-1} \hat{p}_k e^{i\frac{jk\pi}{N}} \tag{3.4}$$

and similarly for q_j. We denote by

$$E_k := \frac{|\hat{p}_k|^2 + \omega_k^2 |\hat{q}_k|^2}{2} \quad , \quad k = -N...., N-1 \tag{3.5}$$

the energy of the k-th mode, where $\omega_k := 2|\sin\left(\frac{k\pi}{2N}\right)|$.

Remark 3.1. For real states one has $E_k = E_{-k}$ for all k, thus we will consider only positive indexes.

It is convenient to state our main result in terms of "specific quantities", thus we will label the modes with the index

$$\kappa := \frac{k}{N} \quad ;$$

correspondingly we denote by

$$\mathcal{E}_\kappa := \frac{E_k}{N} \tag{3.6}$$

the specific energy in the mode with index κ, and by

$$\langle \mathcal{E}_\kappa \rangle (t) := \frac{1}{t} \int_0^t \mathcal{E}_\kappa(s) ds \tag{3.7}$$

its time average.

In the following a small but finite index $\frac{k_0}{N} \equiv \kappa_0 \equiv \mu \ll 1$ will appear.

Theorem 3.2. *Fix a constant C_0 and a positive (large) time T_f; then there exist positive constants μ_*, C_1, C_2, dependent only on C_0 and on T_f, so that the following holds. Consider an initial datum with*

$$\mathcal{E}_\kappa(0) \equiv \mathcal{E}_\kappa(t)\big|_{t=0} = 0 \ , \quad \forall \kappa \neq n\kappa_0 \ , \quad n \in \mathbb{N} \ , \quad \mathcal{E}_{\kappa_0}(0) = C_0\mu^4 \ , \qquad (3.8)$$

and

$$\mathcal{E}_\kappa(0) \leq \mu^4 C_0 e^{-\rho\kappa/\mu} \ , \quad \forall \kappa = n\kappa_0 \ , \quad n \in \mathbb{N} \ , \quad \rho > 0 \ , \qquad (3.9)$$

where $\mu < \mu_*$. Then, there exists $\sigma = \sigma(\rho) > 0$ such that, along the corresponding solution, for all κ one has

(i)

$$\mathcal{E}_\kappa(t) \leq \mu^4 C_1 e^{-\sigma\kappa/\mu} + C_2\mu^5 \ , \quad \text{for } |t| \leq \frac{T_f}{\mu^3} \qquad (3.10)$$

(ii) There exists a sequence $\{\bar{\mathcal{F}}_n\}_{n\in\mathbb{N}}$ such that, defining the specific energy distribution

$$\bar{\mathcal{E}}_{n\kappa_0} = \mu^4 \mathcal{F}_n \ , \quad \mathcal{E}_\kappa = 0 \text{ if } \kappa \neq n\kappa_0$$

for any δ small enough, there exists a time T_0, independent of μ and of T_f, such that, for any $\kappa > 0$ one has

$$\left| \langle \mathcal{E}_\kappa \rangle (t) - \bar{\mathcal{E}}_\kappa \right| \leq C_1\mu^4\delta + C_2\mu^5 \ , \quad \frac{T_0}{\mu^3} \leq |t| \leq \frac{T_f}{\mu^3} \ . \qquad (3.11)$$

Remark 3.3. One can give heuristic arguments to show that the (rescaled) limit distribution $\mathcal{F}_n$ is the same for all initial data in a set of full measure. Moreover such a limit distribution was computed explicitly in [Pon03] obtaining a result in very good agreement with the numerical observations by [BGG05]. However, we were unable to transform the heuristic argument into a rigorous one.

Remark 3.4. There exist numerical results showing that the time T_e of approach to equipartition in FPU systems is a stretched exponential of the inverse of the specific energy $\mathcal{E}$: $T_e \sim \exp[(1/\mathcal{E})^a]$ [PL90] [BGP04]. The existence of such a time–scale *a la* Nekhoroshev was first conjectured in [FMM$^+$82] making use of probabilistic arguments. It is not yet clear whether the metastable state with energy distribution $\mathcal{E}_\kappa$ may survive on such a time–scale. The only rigorous result in this direction was obtained in [BN98] (see also [PBC01]), where the exponential stability of the fundamental mode of a nonlinear string was proved.

Remark 3.5. We expect theorem 3.2 to hold also in the β–FPU, model (the time scale should be substituted by μ^{-4}). Indeed, the theory of sections 4, 5 can be trivially generalized to the β model, the only difference being that the KdV equation has to be substituted by the modified KdV equation (mKdV). However, the study of the modified KdV, is less developed then the study of the KdV equation, so, even if the results of section 6 are expected to hold also in the case of the mKdV, there are not ready to use theorems available.

6

Remark 3.6. One could study the dynamics corresponding to states in which there is no relation between the specific energy and the index of the excited modes. As shown in [PB05b] this would involve the study of the zero dispersion limit of the KdV [Ven87]. Up to now we were not able to obtain a satisfactory (rigorous) theory for the connection between FPU and KdV in such a limit (for a partial result in this direction see [PB05b]).

Remark 3.7. With an extension of our theory we would (probably) be able to prove stability of the solutions constructed in theorem 3.2 with respect to excitation involving a small packet of modes, but only on a time–scale of order μ^{-2}. Over such a time–scale the effects of the nonlinearity are not visible, so this extension has to be considered not satisfactory.

On the time–scale μ^{-3}, at present, we are only able to prove stability of the solutions we constructed for perturbations of the initial data that decay fast in space (i.e. with vanishing specific energy). Thus the energy spectrum of the initial data that we can control has the shape of a sequence of peaks of height proportional to N, but decreasing exponentially with κ, each with superimposed a bump of modes of small height. Work is in progress in order to deal with more general initial data.

4   Normal Form

In this section we compute the normal form of the FPU and we give a rigorous estimate of the remainder.

From now on, instead of the "specific index κ" we will use integers to label the modes and the energy per mode E_k instead of the specific energy per mode.

As above, corresponding to an integer index $1 \leq k_0 \leq N$ we define the parameter

$$\mu := \frac{k_0}{N} \quad . \tag{4.1}$$

Rewrite the FPU system in terms of new rescaled variables r_j defined by

$$\mu^2 r_j := q_j - q_{j-1} \quad . \tag{4.2}$$

Imposing the constraint of zero average $\sum_j r_j = 0$, one has that the change of variables $q \to r$ is well defined and invertible. Introducing also the operator of second difference Δ_1 by

$$(\Delta_1 r)_j := r_{j+1} + r_{j-1} - 2r_j \quad , \tag{4.3}$$

the FPU equations take the form

$$\ddot{r}_j = (\Delta_1 (r + \mu^2 r^2))_j \quad . \tag{4.4}$$

Remark 4.1. Introducing also the momenta s_j defined by

$$p_j = \mu^2 (s_j - s_{j+1}) \tag{4.5}$$

and by the zero average condition, one gets that the transformation $(p, q) \to (s, r)$ is canonical. Moreover, it is easy to verify that in these variables one has

$$E_k = \mu^4 \frac{|\hat{r}_k|^2 + \omega_k^2 |\hat{s}_k|^2}{2} \tag{4.6}$$

with $\hat{r}_k$ and $\hat{s}_k$ the Fourier coefficients of r and s, respectively.

We introduce now an interpolating function $r = r(x, t)$ for the sequence r_j, namely a (smooth) function with the property that the sequence

$$r_j(t) \equiv r(j, t) \tag{4.7}$$

fulfils the FPU equations (4.4). Moreover we will assume that the function $r(x)$ is $2/\mu$ periodic and has zero average, namely that

$$r(x + 2/\mu, t) = r(x, t) \tag{4.8}$$

for any (x, t). Thus we postulate that the function r fulfils

$$\ddot{r} = \Delta_1(r + \mu^2 r^2) . \tag{4.9}$$

with an obvious extension of the definition of Δ_1 to smooth functions. It is easy to verify that this system is Hamiltonian with Hamiltonian function

$$H(r, s) := \int_{-1/\mu}^{1/\mu} \left(\frac{-s\Delta_1 s + r^2}{2} + \mu^2 \frac{r^3}{3} \right) dx \tag{4.10}$$

and with s a periodic function with zero average, playing the role of the momentum conjugated to the function $r(x)$. The momentum $s(x)$ is actually an interpolating function for the momentum introduced in remark 4.1. Actually one has $s_j(t) = s(j, t)$.

It is now convenient to rescale the length of the annulus and the size of the momentum s, by introducing as new phase variables two function (u, v) periodic of period 2, defined by

$$v(\mu x) = \mu s(x) , \quad u(\mu x) = r(x) . \tag{4.11}$$

In the following we will denote by y the rescaled space variable, namely $y = \mu x$.

The coordinate transformation (4.11) is not canonical, but it turns out that the equations for the variables (u, v) are still Hamiltonian *with the original symplectic structure*, and with Hamiltonian function

$$H(u, v) = \mu K(u, v) \tag{4.12}$$

with

$$K(u, v) = \int_{-1}^{1} \left(\frac{-v\Delta_\mu v}{2\mu^2} + \frac{u^2}{2} + \frac{\mu^2 u^3}{3} \right) dy \tag{4.13}$$

where we introduced the difference operator

$$(\Delta_\mu v)(y) := v(y + \mu) + v(y - \mu) - 2v(y) . \tag{4.14}$$

8

Remark 4.2. From now on we will study the system (4.13). This clearly amounts to introducing a new time $\tau \equiv \mu t$. More precisely, denote by $u(\tau), v(\tau)$ a solution of the equations of motion of K, namely of

$$\frac{du}{d\tau} = \frac{\delta K}{\delta v} \ , \quad \frac{dv}{d\tau} = -\frac{\delta K}{\delta u} \tag{4.15}$$

with $\frac{\delta K}{\delta v}$ denoting the L^2 gradient of K with respect to v and similarly for $\frac{\delta K}{\delta u}$. Then $u(\mu t), v(\mu t)$ is a solution of the equations of motion of H.

The formal expansion of the operator Δ_μ, defined in (4.14), gives

$$\frac{\Delta_\mu}{\mu^2} = \partial_y^2 + \frac{\mu^2 \partial_y^4}{12} + O(\mu^4) \ , \tag{4.16}$$

so that one has

$$K = H_0 + P + \mathcal{R}_1 \ , \tag{4.17}$$

with

$$H_0(u, v) := \int_{-1}^{1} \left[\frac{v(-\partial_y^2 v) + u^2}{2} \right] dy \ , \tag{4.18}$$

$$P(u, v) := \int_{-1}^{1} \left[-\mu^2 \frac{v \partial_y^4 v}{24} + \frac{\mu^2 u^3}{3} \right] dy \ , \tag{4.19}$$

$\mathcal{R}_1$ being the remainder of the expansion.

Remark 4.3. The equations of motion of the Hamiltonian H_0 are

$$u_\tau = -\partial_y^2 v \ , \quad v_\tau = -u \tag{4.20}$$

and thus they are equivalent to the linear wave equation. Its flow will be denoted $\Psi^\tau(v, u)$ and is periodic in time with period 2.

Following [Bam05] we are going to use a Galerkin averaging method in order to compute the corrections to the dynamics due to the presence of P and to estimate the effect of $\mathcal{R}_1$.

To this end we first have to introduce a topology in the phase space. This is conveniently done in terms of Fourier coefficients.

Definition 4.4. Having fixed two positive constants s, σ consider the Hilbert space $\ell_{\sigma,s}^2$ of the complex sequences $v \equiv \{v_K\}_{K \in \mathbb{Z} - \{0\}}$ such that

$$\|v\|_{\sigma,s}^2 := \sum_K |v_K|^2 |K|^{2s} e^{2\sigma|K|} < \infty \ . \tag{4.21}$$

We will identify a 2 periodic function v with its Fourier coefficients $\hat{v}_K$ defined by

$$v(y) = \frac{1}{\sqrt{2}} \sum_{K \in \mathbb{Z}} \hat{v}_K e^{i\pi K y} \ ,$$

and we will say that $v \in \ell^2_{\sigma,s}$ if its Fourier coefficients have this property. Moreover in what follows the coefficient σ will be kept fixed. We will study the system $K(u,v)$ in the phase spaces $\mathcal{P}_s$ defined by

$$\mathcal{P}_s := \ell^2_{\sigma,s+1} \times \ell^2_{\sigma,s} \ni (v,u) , \tag{4.22}$$

endowed by the norm

$$\|(v,u)\|^2_s := \|v\|^2_{\sigma,s+1} + \|u\|^2_{\sigma,s} . \tag{4.23}$$

A phase point (v,u) will also be denoted by z, and the ball of radius R centered at the origin of $\mathcal{P}_s$ will be denoted by $B_s(R)$.

It is easy to see that the flow Ψ^τ of the system H_0 is unitary in all the spaces $\mathcal{P}_s$.

Theorem 4.5. *For any $r \geq 5$ there exists a constant $\mu_* \equiv \mu_{*r}$, such that, if*

$$\mu < \mu_*$$

then there exists an analytic canonical transformation $\mathcal{T} : B_r(1) \to B_r(2)$ which averages K, namely such that

$$K \circ \mathcal{T} = H_0 + \langle P \rangle + \mathcal{R} \tag{4.24}$$

where

$$\langle P \rangle(z) := \frac{1}{2} \int_0^2 P(\Psi^\tau(z))d\tau \tag{4.25}$$

and the vector field $X_\mathcal{R}$ of the remainder is analytic in a complex ball of radius 1 and fulfils the estimate

$$\sup_{\|z\|_r \leq 1} \|X_\mathcal{R}(z)\|_0 \leq C_r \mu^{4-\frac{12}{6+r}} . \tag{4.26}$$

Moreover for any $1 \leq r_1 \leq r$ the transformation $\mathcal{T}$ maps B_{r_1} into $\mathcal{P}_{r_1}$ and fulfils

$$\sup_{\|z\|_{r_1} \leq 1} \|z - \mathcal{T}(z)\|_{r_1} \leq C \mu^{2-\frac{6}{6+r}} . \tag{4.27}$$

The proof is an application of the techniques of [Bam05] and, for the sake of completeness, it will be given in appendix A.

Remark 4.6. We recall that a heuristic discussion on the possibility of putting the FPU system in normal form corresponding to initial data with low frequency was given in [She97]. The above theorem rigorously proves such a possibility. Below we give the explicit expression of the normal form, *which is integrable!* As a consequence we think that some of the conclusions of the paper [She97], which are based on the heuristic argument that resonances enforce chaos, could be incorrect.

In the rest of this section we will perform the explicit computation of the averaged equations showing that they coincide with two uncoupled KdV equations.

To obtain the result it is useful to introduce new variables in which the unperturbed flow Ψ^τ assumes a simpler form. To this end we introduce the non canonical transformation

$$\xi := \frac{u + v_y}{\sqrt{2}} \ , \quad \eta := \frac{u - v_y}{\sqrt{2}} \ . \tag{4.28}$$

Since the transformation is not canonical one has to modify the Poisson tensor in order to deduce the equations of motion from the Hamiltonian.

Lemma 4.7. *In terms of the variables ξ, η the Poisson tensor takes the form*

$$J = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \partial_y \ , \tag{4.29}$$

i.e. the Hamilton equations associated to a Hamiltonian function H take the form

$$\frac{dz}{d\tau} = J \ \nabla H(z) \ , \quad \Longleftrightarrow \quad \left(\ \xi_\tau = -\partial_y \frac{\delta H}{\delta \xi} \ , \ \eta_\tau = \partial_y \frac{\delta H}{\delta \eta} \ \right) \ , \tag{4.30}$$

where ∇H denotes the L^2 gradient and $z = (\xi, \eta)$.

In the variables (ξ, η) the various parts of the Hamiltonian take the form

$$H_0(\xi, \eta) \ = \ \int_{-1}^{1} \frac{\xi^2 + \eta^2}{2} dy \tag{4.31}$$

$$P(\xi, \eta) \ = \ \int_{-1}^{1} \left[-\mu^2 \frac{[\partial_y(\xi - \eta)]^2}{48} + \mu^2 \frac{(\xi + \eta)^3}{6\sqrt{2}} \right] \ dy \ , \tag{4.32}$$

and in particular the equations of motion of H_0 assume the simple form

$$[\xi_\tau = -\xi_y \ , \quad \eta_\tau = \eta_y] \quad \Longleftrightarrow \quad [\xi(y, \tau) = \xi_0(y - \tau) \ , \ \eta(y, \tau) = \eta_0(y + \tau)] \ . \tag{4.33}$$

It is now easy to obtain the following

Proposition 4.8. *In the variables ξ, η the average of the perturbation is given by*

$$\langle P \rangle(\xi, \eta) = \int_{-1}^{1} \left[-\mu^2 \frac{\xi_y^2 + \eta_y^2}{48} + \mu^2 \frac{(\xi^3 + \eta^3)}{6\sqrt{2}} \right] \ dy \ , \tag{4.34}$$

and the equations of motion of $H_0 + \langle P \rangle$ are given by

$$\xi_\tau = -\xi_y - \mu^2 \frac{1}{24} \ \xi_{yyy} - \mu^2 \frac{1}{2\sqrt{2}} \ \xi\xi_y \ , \tag{4.35}$$

$$\eta_\tau = \eta_y + \mu^2 \frac{1}{24} \ \eta_{yyy} + \mu^2 \frac{1}{2\sqrt{2}} \ \eta\eta_y \ , \tag{4.36}$$

11

i.e. two uncoupled KdV equations in translating frames, which therefore constitute the resonant normal form of FPU in the region of the phase space corresponding to long wavelength excitations.

Proof. One has to compute the average of the different terms composing equation (4.32). As an example we deal explicitly with the term proportional to $\int_{-1}^{1} dy \xi_y \eta_y$. One has

$$\left\langle \int_{-1}^{1} dy \xi_y \eta_y \right\rangle = \int_{0}^{2} ds \int_{-1}^{1} dy \xi_y(y-s) \eta_y(y+s) = \frac{1}{4} \int_{-2}^{2} d\alpha \int_{0}^{4} d\beta \xi_y(\alpha) \eta_y(\beta) \tag{4.37}$$

which is zero due to the fact that ξ_y has zero average. Performing the same computation over all the terms one gets the result. $\qquad\square$

Since we are interested in the energy per mode we give now the relation of E_k with the Fourier coefficients of ξ and η, which in turn are defined by

$$\xi(y) = \frac{1}{\sqrt{2}} \sum_{K \in \mathbb{Z}} \hat{\xi}_K e^{iKy\pi} \tag{4.38}$$

and similarly for η.

Proposition 4.9. *Let $\xi(y), \eta(y)$ be a pair of functions belonging to $\mathcal{P}_0$; denote by E_k the energy in the $k-th$ mode as defined by (3.5) in terms of the original variables. Then, for μ small enough, one has*

$$\left| \frac{E_k}{N} - \mu^4 \frac{|\xi_K|^2 + |\eta_K|^2}{2} \right| \leq C\mu^{\frac{11}{2}} \|(\xi, \eta)\|_0^2 \tag{4.39}$$

for all k such that $\frac{k}{N} = \mu K$ with $|K| \leq \frac{|\ln \mu|}{2\sigma}$;

$$\frac{|E_k|}{N} \leq \mu^{\frac{11}{2}} \|(\xi, \eta)\|_0^2 \tag{4.40}$$

for all k such that $\frac{k}{N} = \mu K$ and $|K| > \frac{|\ln \mu|}{2\sigma}$, and $E_k = 0$ otherwise.

The elementary proof is based on the exponential decay of the Fourier coefficients of a function in $\ell_{\sigma,0}^2$. It is deferred to appendix B.

5 Estimate of the error

Here we use the normal form to construct approximate solutions of FPU and we estimate their difference from true solutions. First we construct explicitly the approximate solutions.

Consider the following pair of KdV equations

$$\xi_{\tau_1} = -\frac{1}{24} \xi_{yyy} - \frac{1}{2\sqrt{2}} \xi\xi_y , \tag{5.1}$$

$$\eta_{\tau_1} = \frac{1}{24} \eta_{yyy} + \frac{1}{2\sqrt{2}} \eta\eta_y , \tag{5.2}$$

obtained by rescaling time to $\tau_1 = \mu^2\tau$. Let $\xi^a(y, \tau_1), \eta^a(y, \tau_1)$ be a solution of such a pair of equations with the property that it belongs to $\mathcal{P}_r$ for all times τ_1, with a given r. Correspondingly, we define an approximate solutions $z^a \equiv (r^a, s^a)$ of the FPU by

$$r^a(x, t) := \frac{\xi^a(\mu(x-t), \mu^3 t) + \eta^a(\mu(x+t), \mu^3 t)}{\sqrt{2}} \tag{5.3}$$

$$s_x^a(x, t) := \frac{\xi^a(\mu(x-t), \mu^3 t) - \eta^a(\mu(x+t), \mu^3 t)}{\sqrt{2}} \tag{5.4}$$

The main result of this section is a theorem comparing the approximate solution with a corresponding true solution. Precisely, consider an initial datum $(r_{0,j}, s_{0,j})$ and the corresponding Fourier coefficients $(\hat{r}_{0,k}, \hat{s}_{0,k})$ as defined by eq.(3.4). We assume that they are different from zero only if $k/N = \mu K$ and that there exist two positive constants C and ρ such that

$$\frac{|\hat{r}_{0,k}|^2 + \omega_k^2 |\hat{s}_{0,k}|^2}{N} \leq C e^{-2\rho \left| \frac{k}{\mu N} \right|} .$$

Finally, we define uniquely a corresponding interpolating function for the initial datum by

$$r_0(y) := \frac{1}{\sqrt{2N}} \sum_K \hat{r}_{0,k} e^{i\pi \mu K y}$$

where the sum runs over the integers K such that $|K|\mu = |k|/N \leq 1$, and in the formula one has to intend $k = \mu K N$. We will consider a similar interpolating function for $s_{0,j}$ and corresponding initial data for the KdV equations.

Theorem 5.1. *Consider an initial datum for the FPU system with the above properties and denote by $(r_j(t), s_j(t))$ the corresponding solution. Consider the approximate solution $\xi^a(y, t), \eta^a(y, t)$ with the corresponding initial datum just constructed. Assume that for all times t the approximate solution is such that $(\xi^a, \eta^a) \in \mathcal{P}_{78}$ with some $\sigma > 0$, and fix an arbitrary $T_f > 0$. Then there exists μ_* depending on T_f and on $\left\| (\xi^a(t), \eta^a(t)) \right\|_{78}$ only, such that, if $\mu < \mu_*$ then for all times t fulfilling*

$$|t| \leq \frac{T_f}{\mu^3} \tag{5.5}$$

one has

$$\sup_j \left(|r_j(t) - r^a(j, t)| + |s_j(t) - s^a(j, t)| \right) \leq C\mu , \tag{5.6}$$

where r^a, s^a are given by (5.3), (5.4); moreover

$$\left| \frac{E_k(t)}{N} - \mu^4 \frac{|\xi_K^a(t)|^2 + |\eta_K^a(t)|^2}{2} \right| \leq C\mu^5 \tag{5.7}$$

for all k such that $\frac{k}{N} = \mu K$ with $|K| \leq \frac{|\ln \mu|}{2\sigma}$, and

$$|E_k(t)| \leq \mu^5 \tag{5.8}$$

for all k such that $\frac{k}{N} = \mu K$ with $|K| > \frac{|\ln \mu|}{2\sigma}$, whereas $E_k(t) = 0$ otherwise.

The proof of the theorem, which follows closely the strategy of [SW00], is deferred to Appendix C.

6 Dynamics of KdV and conclusion of the proof

In this section we recall some known facts on the dynamics of the KdV equation with periodic boundary conditions and we use them to prove the results of sect. 3.

Consider the KdV equation (5.1), namely

$$\xi_{\tau_1} = -\frac{1}{24}\,\xi_{yyy} - \frac{1}{2\sqrt{2}}\,\xi\xi_y$$

It is a well known consequence of the Lax pair formulation that the spectrum of the Sturm Liouville operator

$$L_\xi := -\partial_{yy} + 6\sqrt{2}\xi(y,\tau_1) \tag{6.1}$$

with periodic boundary conditions on $[0,4]$ is invariant under the KdV evolution, i.e. it is independent of τ_1.

The spectrum of L_ξ with periodic boundary conditions on $[0,4]$, will be simply called periodic spectrum of ξ.

Such a periodic spectrum is pure point and consists of a sequence of eigenvalues

$$\lambda_0 < \lambda_1 \leq \lambda_2 < \lambda_3 \leq \lambda_4 < ... \quad , \tag{6.2}$$

where equality may occur only in place of a sign '$\leq$'. The quantities

$$\gamma_n := \lambda_{2n} - \lambda_{2n-1} \tag{6.3}$$

are called *the gaps* of the spectrum. From standard asymptotic properties of the spectrum one has $\gamma_n \in \ell^2$ for any L^2 potential ξ. Moreover, it has been proved by Garnett and Trubowitz that the sequence of the γ_n entirely determine the periodic spectrum of ξ.

A further, very important, feature of the above Sturm Liouville problem is the relation between the sequence of the gaps and the regularity of the corresponding potential ξ. Indeed, up to a certain extent the correspondence between the regularity of ξ and the property of the sequence γ_n is the same existing between the regularity of a function and its Fourier coefficients (see [Mar86]). Precisely, the following theorem (from [Pös04]) holds:

Theorem 6.1. *Suppose $\xi \in L^2$ then $\xi \in \ell_{0,s}$ if and only if its gap lengths satisfy*

$$\sum_{n\geq 1} n^{2s}\,|\gamma_n|^2 < \infty \quad . \tag{6.4}$$

Moreover, if $\xi \in \ell_{\sigma,s}$ then

$$\sum_{n\geq 1} n^{2s}\mathrm{e}^{2\sigma n}\,|\gamma_n|^2 < \infty \tag{6.5}$$

conversely, if (6.5) holds, then $\xi \in \ell_{\sigma',0}$ with some $\sigma' > 0$.

From a Hamiltonian point of view the KdV is an integrable infinite dimensional system. It has been shown that it admits global action angle coordinates, the actions being exactly the quantities

$$I_n := \gamma_n^2 \ .$$

More precisely, the following result holds

Theorem 6.2. *[Kappeler-Pöschel [KP03]] There exists a diffeomorphism Ω : $L^2 \to \ell_{0,1/2}^2 \times \ell_{0,1/2}^2$ with the following properties[1]:*

i) *Ω is one-to-one, onto, bianalytic, and canonical.*

ii) *For each $s \geq 0$, the restriction of Ω to $\ell_{0,s}^2$ is a map*

$$\Omega : \ell_{0,s}^2 \to \ell_{0,s+1/2}^2 \times \ell_{0,s+1/2}^2$$

which is one-to-one, onto, and bianalytic as well.

iii) *the coordinates $(x,y) \in \ell_{0,3/2}^2 \times \ell_{0,3/2}^2$ are Birkhoff coordinates for the KdV equation. That is to say, in terms of the coordinates (x,y) the Hamiltonian H_{KdV} of the KdV depends only on $I_n := (x_n^2 + y_n^2)/2$, $n \geq 1$, with (x,y) canonically conjugated coordinates.*

In terms of the variables (x,y) the dynamics of the KdV is trivial. To describe the latter, fix an initial datum (x^0, y^0), and define

$$\nu_n(x^0, y^0) := \frac{\partial H_{KdV}}{\partial I_n}(x^0, y^0) \ ;$$

then the equations of motion take the form

$$\dot{x}_n = \nu_n y_n \ , \quad \dot{y}_n = -\nu_n x_n \ , \tag{6.6}$$

Thus, it is immediately seen that any solution is periodic, quasiperiodic or almost periodic, depending on the number of gaps (actions) initially different from zero.

With these tools at hand it is easy to obtain the

Proof of theorem 3.2. We begin by proving (i). Consider an initial datum as in the statement of the theorem. This corresponds to initial data for ξ and η analytic in a strip of width ρ. By theorem 6.1 the corresponding sequence of gaps decreases exponentially with coefficient ρ in the exponential. This property is then conserved along the corresponding solution. Going back to Fourier coefficients one immediately deduces that the corresponding solution $\xi(\tau_1)$ is analytic

---

[1]By abuse of notation, here $\ell_{0,\alpha}^2$ is the space of the sequences $\{x_n\}_{n \geq 1}$ such that $\sum_n n^{2\alpha}|x_n|^2 < \infty$.

in the y variable in a complex strip of width $\sigma(\tau_1)$. Taking the minimum of such quantities one finds the coefficient σ of theorem 3.2. This is the result for the solution of the KdV equations. Using theorem 5.1, eq. (5.7), one goes back to the quantities E_k and obtains the desired result.

In order to prove statement (ii) we use the fact that any solution is almost periodic in time. Denote the quantity

$$E_K^{(1)} := \left| \hat{\xi}_K \right|^2 \; ;$$

then, since $E_K^{(1)}(x(\tau_1), y(\tau_1))$ is an almost periodic function its time average exists (see e.g. [Fin74]). Thus we set

$$\bar{E}_K^{(1)} := \lim_{T \to \infty} \langle E_K^{(1)} \rangle(T) := \lim_{T \to \infty} \frac{1}{T} \int_0^T E_K^{(1)}(x(\tau_1), y(\tau_1))d\tau_1 \qquad (6.7)$$

Define also

$$E_K^{(2)} := |\hat{\eta}_K|^2$$

and $\bar{E}_K^{(2)}$ in analogy to $\bar{E}_K^{(1)}$, then we define $\bar{E}_K := (\bar{E}_K^{(1)} + \bar{E}_K^{(2)})/2$. Scaling back to physical variables and using again theorem 5.1, eq. (5.7), one gets statement (ii). $\qquad \square$

A   Appendix: Proof of theorem 4.5

Since the Hamiltonian (and its vector field) is analytic, it is useful to complexify the phase space. Thus, from now on we will think of the phase variable z as a complex variable. The main reason is that, through Cauchy inequality the sup norm of a function controls also the supremum of the derivatives of the function.

First we prove the following simple

Lemma A.1. *For any $s \geq 0$ one has*

$$\|X_{\mathcal{R}_1}(z)\|_s \quad \leq \quad 2\mu^4 \; , \quad \forall z \,:\, \|z\|_{s+5} \leq 2 \qquad (A.1)$$

$$\|X_P(z)\|_s \quad \leq \quad C\mu^2 \; , \quad \forall z \,:\, \|z\|_{s+3} \leq 2 \qquad (A.2)$$

Proof. The estimate of X_P is an immediate consequence of the definition of the norm and of the fact that $\ell_{\sigma,s}^2$ is an algebra for $s \geq 1$. Concerning $X_{\mathcal{R}_1}$ just remark that the K–th Fourier coefficient of its u component is given (and estimated) by

$$\left| (X_{\mathcal{R}_1}(u,v)^u)_K^{\wedge} \right| = \left| \left[\frac{4}{\mu^2} \sin^2(K\mu\pi) - \pi^2 K^2 + \frac{K^4 \pi^4 \mu^2}{24} \right] \hat{v}_K \right| \qquad (A.3)$$

$$\leq \frac{\pi^6}{6!} K^6 \mu^4 |\hat{v}_K|$$

from which the thesis follows. $\qquad \square$

16

Then we perform a Galerkin cutoff of P. Precisely, define the projector Π_n on the Fourier modes with index smaller then n, i.e.

$$\Pi_n(\hat{u}_{-\infty}...\hat{u}_{-K}...\hat{u}_K...\hat{u}_\infty) = (\hat{u}_{-n}...\hat{u}_n)$$

define also $\Pi_n(u,v) := (\Pi_n u, \Pi_n v)$, and finally define

$$P^{(n)}(z) := P(\Pi_n(z)) . \tag{A.4}$$

Following [Bam03] we have the following

Lemma A.2. *For any $s \geq 1$ there exists a constant C such that, for any $r \geq 0$, and any $n \geq 0$, one has*

$$\left\| X_{P-P^{(n)}}(z) \right\|_s \leq \frac{\mu^2 C_s}{n^r} , \quad \forall z : \|z\|_{s+r+3} \leq 3/2 \tag{A.5}$$

For the proof see [Bam03] proof of lemma 5.2.

Moreover it is easy to show that $X_{P^{(n)}}$ is analytic as a map from $\mathcal{P}_s$ to itself and that

$$\|X_{P^{(n)}}(z)\|_s \leq \mu^2 C_s n^3 \quad \forall z : \|z\|_s \leq 2 \tag{A.6}$$

We now use Lie transform to construct a canonical transformation averaging the Hamiltonian up to order μ^4 (or more precisely, slightly less).

Thus consider an auxiliary Hamiltonian function χ (of order μ^2), assume that the corresponding Hamiltonian vector field is analytic as a map from $\mathcal{P}_s$ to itself $\forall s \geq 1$, and consider the corresponding Hamilton equations

$$\dot{z} = X_\chi(z) . \tag{A.7}$$

Denote by $\mathcal{T}^\tau$ the corresponding time τ flow and by $\mathcal{T}$ the time 1 flow. We use such a $\mathcal{T}$ in order to transform our Hamiltonian system K. One has

$$K \circ \mathcal{T} = H_0 + P^{(n)} + \{\chi, H_0\} + \mathcal{R} \tag{A.8}$$

where

$$\mathcal{R} = (P - P^{(n)}) \circ \mathcal{T} + \mathcal{R}_1 \circ \mathcal{T} + \left[P^{(n)} \circ \mathcal{T} - P^{(n)} \right] + [H_0 \circ \mathcal{T} - H_0 - \{\chi, H_0\}] \tag{A.9}$$

is the sum of the higher order terms (they will be estimated in a while).

First of all we choose χ in such a way that

$$P^{(n)} + \{\chi, H_0\} = \left\langle P^{(n)} \right\rangle ,$$

according to lemma 8.4 of [Bam99] (a simple computation) this is given by

$$\chi(z) := \frac{1}{2} \int_0^2 \tau \left[P^{(n)}(\Psi^\tau(z)) - \langle P^{(n)} \rangle(\Psi^\tau(z)) \right] d\tau \tag{A.10}$$

and its vector field is analytic and estimated by

$$\|X_\chi(z)\|_s \leq \mu^2 C_s n^3 \quad \forall z : \|z\|_s \leq 2 \tag{A.11}$$

It also follows that the transformation $\mathcal{T}$ exists and fulfils the estimates (4.27). Moreover the various terms of (A.9) are estimated by

17

Lemma A.3. *The following estimates hold*

$$\left\|X_{(P-P^{(n)})\circ\mathcal{T}}\right\|_s \leq \frac{\mu^2 C_s}{n^r} \ , \quad \forall z \ : \ \|z\|_{s+r+3} \leq 1 \tag{A.12}$$

$$\|X_{\mathcal{R}_1\circ\mathcal{T}}\|_s \leq C\mu^4 \ , \quad \forall z \ : \ \|z\|_{s+5} \leq 1 \tag{A.13}$$

$$\left\|X_{P^{(n)}\circ\mathcal{T}-P^{(n)}}\right\|_s \leq C\mu^4 n^6 \ , \quad \forall z \ : \ \|z\|_s \leq 1 \tag{A.14}$$

$$\left\|X_{H_0\circ\mathcal{T}-H_0-\{\chi,H_0\}}\right\|_s \leq C\mu^4 n^6 \ , \quad \forall z \ : \ \|z\|_s \leq 1 \tag{A.15}$$

Proof. All these estimates are a direct application of some lemmas already proved in [Bam99]. In particular (A.12) and (A.13) follow from lemma 8.2 with $R = 3/2$ and $\delta = 1/2$ the first one, and $R = 2$ and $\delta = 1$ the second one. Eq. (A.14) is a consequence of lemma 8.3 with $R = 2$ and $\delta = 1$. Eq.(A.15) is a consequence of lemma 8.5 with $R = 2$ and $\delta = 1$. □

We choose now n in such a way that (A.12) and (A.14) are of the same order of magnitude. This leads to the choice $n = \mu^{-\frac{2}{r+6}}$ which gives the estimate (4.26) for the remainder. Up to now we have shown that

$$K \circ \mathcal{T} = H_0 + \left\langle P^{(n)} \right\rangle + \mathcal{R} \tag{A.16}$$

with $\mathcal{R}$ fulfilling the wonted estimate. To conclude the proof it is enough to remark that

$$\left\|X_{\langle P\rangle-\langle P^{(n)}\rangle}(z)\right\|_s \leq \frac{\mu^2 C_s}{n^r} \leq C\mu^{4-\frac{12}{6+r}} \ , \quad \forall z \ : \ \|z\|_{s+r+3} \leq 3/2 \tag{A.17}$$

and thus one can simply substitute $\langle P\rangle$ in place of $\langle P^{(n)}\rangle$ including the difference in the remainder. □

B Appendix: Proof of proposition 4.9.

Define the Fourier coefficients of the function u by

$$\hat{u}_K := \frac{1}{\sqrt{2}} \int_{-1}^{1} u(y)\mathrm{e}^{-\mathrm{i}\pi Ky}dy \tag{B.1}$$

and similarly for v, then

Lemma B.1. *For a state of the FPU corresponding to a pair of functions (u,v) one has*

$$\frac{E_k}{N} = \sum_{L\in\mathcal{L}} |\hat{u}_{K+L}|^2 + \omega_k^2 \left|\frac{\hat{v}_{K+L}}{\mu}\right|^2 \ , \quad \forall k \ : \ \mu K = \frac{k}{N} \ , \tag{B.2}$$

where

$$\mathcal{L} := \{L \in \mathbb{Z} \ : \ L\mu = 2l \ with \ l \in \mathbb{Z}\} \tag{B.3}$$

and $E_k = 0$ otherwise.

Proof. First introduce a $2N$–periodic interpolating function for r_j, namely a smooth function $r^N(x)$ such that

$$r_j = r^N(j) , \quad r^N(x + 2N) = r^N(x) . \tag{B.4}$$

Denote

$$\hat{r}_k^N := \frac{1}{\sqrt{2N}} \int_{-N}^{N} r^N(x) e^{\frac{-ik\pi x}{N}} dx , \tag{B.5}$$

then one has

$$r_j = r^N(j) = \frac{1}{\sqrt{2N}} \sum_{k \in \mathbb{Z}} \hat{r}_k^N e^{\frac{ik\pi j}{N}} = \frac{1}{\sqrt{2N}} \sum_{k=-N}^{N-1} \left[\sum_{l \in \mathbb{Z}} \hat{r}_{k+2Nl}^N \right] e^{\frac{ik\pi j}{N}}$$

which implies

$$\hat{r}_k = \sum_{l \in \mathbb{Z}} \hat{r}_{k+2Nl}^N . \tag{B.6}$$

Then the relation between $\hat{r}_k^N$ and $\hat{u}_K$ is easily obtained remarking that

$$r^N(j) = \mu^2 u(\mu j) = \frac{\mu^2}{\sqrt{2}} \sum_{K \in \mathbb{Z}} \hat{u}_K e^{iK\mu j\pi} = \frac{1}{\sqrt{2N}} \sum_{k \in \mathbb{Z}} \hat{r}_k^N e^{\frac{ik\pi j}{N}} . \tag{B.7}$$

$\square$

Proof of proposition 4.9. We start from eq.(B.2), and as a first step we remark that, for $K\mu = k/N$, one has

$$\left| \frac{\omega_k}{\mu} \right| = \frac{2}{\mu} \left| \sin\left(\frac{k\pi}{2N} \right) \right| = \frac{2}{\mu} \left| \sin\left(\frac{\mu K \pi}{2} \right) \right| \leq \pi |K| , \tag{B.8}$$

and that, for $|K| \geq 2|\ln\mu|/\sigma$ one has

$$\frac{|\hat{u}_K|^2 + \pi^2 K^2 |\hat{v}_K|^2}{2} \leq \pi^2 \mu^4 \|(u,v)\|_0^2 . \tag{B.9}$$

Using the relation between (u,v) and (ξ, η) one gets

$$\frac{\left| \hat{\xi}_K \right|^2 + |\hat{\eta}_K|^2}{2} = \frac{|\hat{u}_K|^2 + \pi^2 K^2 |\hat{v}_K|^2}{2} \tag{B.10}$$

from which, using (B.8), eq. (4.40) immediately follows. Concerning (4.39) one has, for $|K| \leq 2|\ln\mu|/\sigma$,

$$\left| \frac{E_k}{\mu^4} - \frac{\left| \hat{\xi}_K \right|^2 + |\hat{\eta}_K|^2}{2} \right| \leq \left| \frac{\omega_k^2 - (\mu K)^2}{\mu^2} \right| |\hat{v}_K|^2 + \sum_{\substack{L \neq 0 \\ L \in \mathcal{L}}} \frac{1}{2} \left[|\hat{u}_{K+L}|^2 + \frac{\omega_k^2}{\mu^2} |\hat{v}_{K+L}|^2 \right]$$

$$\leq \frac{(\mu K)^4}{\mu^2} |\hat{v}_K|^2 + \sum_{\substack{L \neq 0 \\ L \in \mathcal{L}}} \frac{1}{2} \left[|\hat{u}_{K+L}|^2 + |K+L|^2 |\hat{v}_{K+L}|^2 \right]$$

$$\leq \mu^2 (2|\ln\mu|)^2 \|v\|_{\sigma,1}^2 + \sum_{l \neq 0} \|(v,u)\|_0^2 e^{-2\sigma \frac{l}{\mu}}$$

19

The logarithm of μ can obviously be estimated by $\mu^{-1/2}$, while the sum is exponentially small with μ. Thus the thesis follows. □

C   Appendix: Proof of Theorem 5.1

It is useful to use also the variables (u,v), so define

$$u^a(y,\tau) \;:=\; \frac{\xi^a(y-\tau,\mu^2\tau) + \eta^a(y+\tau,\mu^2\tau)}{\sqrt{2}} \tag{C.1}$$

$$v^a_y(y,\tau) \;:=\; \frac{\xi^a(y-\tau,\mu^2\tau) - \eta^a(y+\tau,\mu^2\tau)}{\sqrt{2}} \tag{C.2}$$

and denote $z^a(y,\tau) = (u^a(y,\tau), v^a(y,\tau))$. Then, in order to get a better approximation we define

$$(\tilde{u},\tilde{v}) \equiv \tilde{z} = \mathcal{T}(z^a) = z^a + \psi^a(z^a) \tag{C.3}$$

where

$$\|\psi^a\|_r \leq C\mu^{2-\frac{6}{6+r}} \tag{C.4}$$

and $(\tilde{u},\tilde{v})$ fulfils the equations

$$\frac{\tilde{v}_t}{\mu} \;=\; -\tilde{u} - \mu^2\pi_0\tilde{u}^2 + \mathcal{R}^v \tag{C.5}$$

$$\tilde{u}_t \;=\; -\Delta_1\frac{\tilde{v}}{\mu} + \mu\mathcal{R}^u \tag{C.6}$$

where the operator Δ_1 acts in terms of the x variables, the remainders are functions of y,τ which fulfil

$$\|\mathcal{R}^v\|_{\sigma,1} \leq C\mu^{4-\frac{12}{6+r}} \;,\quad \|\mathcal{R}^u\|_{\sigma,0} \leq C\mu^{4-\frac{12}{6+r}} \;. \tag{C.7}$$

and π_0 is the projector on the space of the functions with zero average.

We restrict the space variable to integer values. If $\mu = l/n$ with l and n relatively prime integers then all the quantities involved in equations (C.5), (C.6) are periodic with period n. *In what follows we will restrict to the case $l = 1$; the case $l \neq 1$ can be dealt with by simple modifications.*
Keeping this in mind we will allow the space variable j to vary in $\{-n,...n-1\}$. For a (finite) sequence $r = \{r_j\}$ we define the norm

$$\|r\|^2_{\ell^2(j)} := \sum_{j=-n}^{n-1} |r_j|^2 \;. \tag{C.8}$$

For the quantities $\tilde{u},\tilde{v},\mathcal{R}^v,\mathcal{R}^u$ evaluated at the integers j we will retain the same notation as for the original quantities. Moreover it is useful to introduce the difference operator ∂ defined by

$$(\partial r)_j := r_j - r_{j-1} \tag{C.9}$$

where r is an arbitrary sequence.

We consider the FPU model (4.9). We rewrite it in the form

$$\dot{s} = -r - \mu^2 \pi_0 r^2 \qquad \text{(C.10)}$$
$$\dot{r} = -\Delta_1 s \qquad \text{(C.11)}$$

and we look for two sequences $E \equiv \{E_j\}$ and $F \equiv \{F_j\}$ such that

$$r = \tilde{u} + \mu E \ , \quad s = \frac{\tilde{v}}{\mu} + \mu F \qquad \text{(C.12)}$$

fulfil the FPU equation in the form (C.10) and (C.11). Then E and F have to fulfil

$$\dot{E} = -\Delta_1 F - \frac{\mu \mathcal{R}^u}{\mu} \qquad \text{(C.13)}$$

$$\dot{F} = -E - \mu^2 2\pi_0 \tilde{u} E - \mu^3 \pi_0 E^2 - \frac{\mathcal{R}^v}{\mu} \qquad \text{(C.14)}$$

Moreover, for (E, F) we impose initial conditions such that $(\tilde{u}, \tilde{v})$ has initial data corresponding to those of the true initial datum, namely we assume

$$\tilde{u}(\mu j, 0) + \mu E_{0,j} = r_{0,j} = u^a(\mu j, 0) \ , \quad \frac{\tilde{v}(\mu j, 0)}{\mu} + \mu F_{0,j} = s_{0,j} = \frac{v^a(\mu j, 0)}{\mu} \qquad \text{(C.15)}$$

Lemma C.1. *One has*

$$\|E_0\|_{\ell^2(j)} \le C\mu^{\frac{1}{2} - \frac{6}{6+r}} \ , \quad \|\partial F_0\|_{\ell^2(j)} \le C\mu^{\frac{1}{2} - \frac{6}{6+r}} \qquad \text{(C.16)}$$

Proof. From (C.3), (C.4) one has

$$E_0 = \frac{\tilde{u} - u^a}{\mu} = \frac{\psi^u}{\mu} \ , \quad F_0 = \frac{\tilde{v} - v^a}{\mu^2} \equiv \frac{\psi^v}{\mu^2}$$

and

$$\sup_y |\tilde{u}(y) - u^a(y)| \le C\mu^{2 - \frac{6}{6+r}} \ ,$$

from which

$$\|E_0\|_{\ell^2(j)}^2 \le \sum_{j=-n}^{n-1} \sup_j |E_j|^2 \le 2n \frac{C\mu^{4 - \frac{12}{6+r}}}{\mu^2} = 2 \frac{\mu^{4 - \frac{12}{6+r}}}{\mu^3}$$

from which the estimate of E_0 follows.

Concerning F we need an estimate of $\partial \psi^v$. Since ψ^v is a function of y, one has

$$|(\partial \psi^v)(j)| = |\psi^v(\mu j) - \psi^v(\mu j - \mu)| \le \mu \sup_y |\partial_y \psi^v(y)| \le C\mu^{3 - \frac{6}{6+r}}$$

From which

$$\|\partial \psi^v\|^2_{\ell^2(j)} \leq \sum_{j=-n}^{n-1} \left(\frac{|\partial \psi^v(j)|}{\mu^2} \right)^2 \leq \frac{n\mu^{6-\frac{12}{6+r}}}{\mu^4} \ . \tag{C.17}$$

$\square$

We use now an idea of Wayne and Schneider to obtain the

Theorem C.2. *Fix $r = 78$, and fix T_f and $C_{\mathcal{F}} > 0$, then provided μ is small enough one has that*

$$\|E\|^2_{\ell^2(j)} + \|(\partial F)\|^2_{\ell^2(j)} \leq C_{\mathcal{F}} \tag{C.18}$$

for all times t fulfilling

$$|t| \leq \frac{T_f}{\mu^3} \tag{C.19}$$

Proof. Define the function

$$\mathcal{F}(E,F) := \sum_j \left(\frac{E_j^2 + F_j(-\Delta_1 F)_j}{2} + \frac{2\mu^2 \tilde{u}_j E_j^2}{2} \right) \tag{C.20}$$

and remark that

$$\frac{1}{2}\mathcal{F}(E,F) \leq \|E\|^2_{\ell^2(j)} + \|(\partial F)\|^2_{\ell^2(j)} \leq 2\mathcal{F}(E,F)$$

Compute now the time derivative of $\mathcal{F}$; inserting the equations (C.13) and (C.14) one gets

$$\dot{\mathcal{F}} = \sum_j (\Delta_1 F)_j \mu^3 E_j^2 + (\Delta_1 F)_j \frac{\mathcal{R}_j^v}{\mu} - \frac{E_j \mu \mathcal{R}_j^u}{\mu} - \frac{2\mu^3 \tilde{u}_j \mathcal{R}_j^u E_j}{\mu} + \frac{2\mu^3 E_j^2}{2} \frac{\partial \tilde{u}_j}{\partial \tau} \tag{C.21}$$

In order to estimate the r.h.s. we need some preliminary estimates. The first one is

$$\sup_j |(\Delta_1 F)_j| = \sup_j |(\partial F)_{j+1} - (\partial F)_j| \leq 2\sup_j |(\partial F)_j| \leq 4\sqrt{\mathcal{F}}$$

Next we will need an estimate of $\|\mathcal{R}^u\|_{\ell^2(j)}$. This is given by

$$\|\mathcal{R}^u\|^2_{\ell^2(j)} \leq \sum_j |\mathcal{R}^u|^2 \leq 2n \sup_y |\mathcal{R}^u(y)|^2 \leq Cn\mu^{8-\frac{24}{6+r}} \tag{C.22}$$

which gives

$$\|\mathcal{R}^u\|_{\ell^2(j)} \leq C\mu^{4-\frac{1}{2}-\frac{12}{6+r}} \tag{C.23}$$

Concerning $\mathcal{R}^v$ we need an estimate of $\|\partial \mathcal{R}^v\|_{\ell^2(j)}$. This is given by

$$\|\partial \mathcal{R}^v\|_{\ell^2(j)} \leq C\mu^{4+\frac{1}{2}-\frac{12}{6+r}} \tag{C.24}$$

which is obtained by remarking that

$$|(\partial \mathcal{R}^v)_j| = |\mathcal{R}^v(\mu j) - \mathcal{R}^v(\mu j - \mu)| \le \mu \sup_y \left| \frac{\partial \mathcal{R}^v}{\partial y}(y) \right|$$

and proceeding as in the proof of (C.23).

Now, the first term of (C.21) is estimated by $4\mu^3 \mathcal{F}^{3/2}$. Concerning the second term, first remark that it coincides with $\sum_j (\partial F)_j (\partial \mathcal{R}^v)_j / \mu$ and therefore it is estimated by $C \mathcal{F}^{1/2} \mu^{4+\frac{1}{2}-\frac{12}{6+r}-1}$. The same estimate holds for the third term, the forth term is estimated by $C \mathcal{F}^{1/2} \mu^{6+\frac{1}{2}-\frac{12}{6+r}-1}$ and the last term is also easily estimated remarking that the derivative of $\tilde{u}$ with respect to τ is bounded and therefore such a term is bounded by $C \mu^3 \mathcal{F}$.

As far as $\mathcal{F} < 2 C_{\mathcal{F}}$ one thus has

$$\left| \dot{\mathcal{F}} \right| \le C(\mu^3 + \mu^3) \mathcal{F} + C\mu^{5+1/4-1}. \tag{C.25}$$

Such a differential inequality can be easily solved giving

$$\mathcal{F}(t) \le \mathcal{F}_0 e^{T_0 C} + e^{T_0 C} C T_0 \mu^{1+\frac{1}{2}-\frac{12}{6+r}-1} \tag{C.26}$$

which, inserting the value of r implies the thesis. Moreover the result on the Fourier modes is an immediate consequence of proposition 4.9 and of the fact that the error from a true solution is measured in the norm $\| \ \|_{\ell^2(j)}$ which controls the Fourier coefficients. $\qquad \square$

Acknowledgements

References

[Bam99]   D. Bambusi, *Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equation*, Math. Z. **130** (1999), 345–387.

[Bam03]   _____, *An averaging theorem for quasilinear Hamiltonian PDEs*, Ann. Henri Poincaré **4** (2003), 685–712.

[Bam05]   _____, *Galerkin averaging method and Poincaré normal form for some quasilinear PDEs*, Preprint (2005).

[BGG05]   L. Berchialla, L. Galgani, and A. Giorgilli, *Localization of energy in FPU chains*, DCDS-A **11** (2005), 855–866.

[BGP04]   L. Berchialla, A. Giorgilli, and S. Paleari, *Exponentially long times to equipartition in the thermodynamic limit*, Phys. Lett. A **321** (2004), 167–172.

[BKL03]   J. A. Biello, P. R. Kramer, and Y. V. Lvov, *Stages of energy transfer in the FPU model. dynamical systems and differential equations (Wilmington, NC, 2002).*, DCDS **Suppl.** (2003), 113–122.

[BN98]   D. Bambusi and N. N. Nekhoroshev, *A property of exponential stability in the nonlinear wave equation close to main linear mode*, Physica D **122** (1998), 73–104.

[CG99]   A. Carati and L. Galgani, *On the specific heat of FPU systems and their glassy behavior*, J. Stat. Phys. **94** (1999), 859–869.

[CG01]   _____, *Planck's formula and glassy behaviour in classical nonequilibrium statistical mechanics*, Physica A **280** (2001), 105–114.

[CGG05]   A. Carati, L. Galgani, and A. Giorgilli, *The Fermi–Pasta–Ulam problem as a challenge for the foundations of physics.*, Chaos (to appear) (2005).

[Fin74]   A. Fink, *Almost periodic differential equations*, Springer-Verlag, Berlin, 1974.

[FMM+82]   F. Fucito, F. Marchesoni, E. Marinari, G. Parisi, L. Peliti, S. Ruffo, and A. Vulpiani, *Approach to equilibrium in a chain of nonlinear oscillators*, J. de Physique **43** (1982), 707–713.

[FPU65]   E. Fermi, J.R. Pasta, and S.M. Ulam, *Studies of nonlinear problems*, Collected works of E. Fermi, vol.2, Chicago University Press, Chicago, 1965.

[GS72]   L. Galgani and A. Scotti, *Planck-like distribution in classical nonlinear mechanics*, Phys. Rev. Lett. **28** (1972), 1173–1176.

[IC66]   F. M. Izrailev and B. V. Chirikov, *Statistical properties of a nonlinear string*, Sov. Phys. Dokl. **11** (1966), 30–32.

[KP03]   T. Kappeler and J. Pöschel, *KAM & KdV*, Springer, 2003.

[LPRV85]   R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani, *Further results on the equipartition threshold in large nonlinear Hamiltonian systems*, Phys. Rev. A **31** (1985), 2741–2742.

[Mar86]   V. Marchenko, *Sturm-Liouville operators and applications*, Birkhäuser, Basel, 1986.

[PB05a]   A. Ponno and D. Bambusi, *Energy cascade in Fermi–Pasta–Ulam model*, Symmetry and Perturbation Theory 2004, World Scientific, 2005, pp. 263–270.

[PB05b]   ———, *KdV equation and energy sharing in FPU*, Chaos (to appear) (2005).

[PBC01]   S. Paleari, D. Bambusi, and S. Cacciatori, *Normal form and exponential stability for some nonlinear string equations*, ZAMP **52** (2001), 1033–1052.

[PL90]   M. Pettini and M. Landolfi, *Relaxation properties and ergodicity breaking in nonlinear Hamiltonian dynamics.*, Phys. Rev. A **41** (1990), 768–783.

[Pon03]   A. Ponno, *Soliton theory and the fermi-pasta-ulam problem in the thermodynamic limit*, Europhys. Lett. **64** (2003), 606–612.

[Pon05]   ———, *The Fermi–Pasta–Ulam problem in the thermodynamic limit*, Proceedings of the Cargése Summer School 2003 on Chaotic Dynamics and Transport in Classical and Quantum Systems, Kluwer Academic Publishers, 2005, pp. 431–440.

[Pös04]   J. Pöschel, *Hill's potentials in weighted Sobolev spaces and their spectral gaps*, Preprint (2004).

[PW95]   R. D. Pierce and C. E. Wayne, *On the validity of mean-field amplitude equations for counterpropagating wavetrains*, Nonlinearity **8** (1995), 769–780.

[Rin01]   B. Rink, *Symmetry and resonance in periodic FPU chains*, Commun. Math. Phys. **218** (2001), 665–685.

[She97]   D. L. Shepelyansky, *Low-energy chaos in the Fermi–Pasta–Ulam problem*, Nonlinearity **10** (1997), 1331–1338.

[SW00]   G. Schneider and C. E. Wayne, *Counter-propagating waves on fluid surfaces and the continuun limit of the Fermi Pasta Ulam model*, Proceedings of the International Conference on Differential Equations, Berlin 1999 (Rivere Edge, NJ), World Scientific, 2000.

[Ven87]   S. Venakides, *The zero dispersion limit of the Korteweg-de Vries equation with periodic initial data*, Transactions of AMS **301** (1987), 189–226.

[ZK65]   N. J. Zabusky and M. D. Kruskal, *Interaction of solitons in a collisionless plasma and the recurrence of initial states*, Phys. Rev. Lett. **15** (1965), 240–243.

Dario Bambusi, Antonio Ponno
Dipartimento di Matematica
Via Saldini 50, 20133 Milano, Italy

E-mail: bambusi@mat.unimi.it, ponno@mat.unimi.it

web-address: http://users.mat.unimi.it/users/bambusi/