J Stat Phys (2009) 135: 873–893
DOI 10.1007/s10955-008-9660-6

The Fermi-Pasta-Ulam Problem: Scaling Laws vs. Initial Conditions

G. Benettin · R. Livi · A. Ponno

Received: 23 July 2008 / Accepted: 3 December 2008 / Published online: 19 December 2008

Abstract Numerical evidence on the relevance of the initial conditions to the Fermi-Pasta-Ulam problem is reported, supported by analytic estimates. In particular, we analyze the special, crucial role played by the phases of the low frequency normal modes initially excited, their energy being the same. The results found are the following. When the phases of the initially excited modes are randomly chosen, the parameter ruling the first stage of the transfer of energy to higher frequency modes turns out to be the energy per degree of freedom (or specific energy) of the system, i.e. an intensive parameter. On the other hand, if the initial phases are "coherently" selected (e.g. they are all equal or equispaced on the unit circle), then the energy cascade is ruled by the total energy of the system, i.e. an extensive parameter. Finally, when a few modes are initially excited, in which case specifying the randomness or coherence of the phases becomes meaningless, the relevant parameter turns out to be again the specific energy (this is the case of the original Fermi-Pasta-Ulam experiment).

Keywords Fermi-Pasta-Ulam problem · Energy transfer · Scaling laws

1 Introduction

In the present paper we focus our attention on the classical Fermi-Pasta-Ulam problem [1], which consists, as is well known, in understanding and characterizing the path to energy equipartition among normal modes of a one-dimensional oscillator chain, when a small subset of long wavelength modes are initially excited. Since the original FPU experiment, the most studied model is defined by the Hamiltonian

G. Benettin · A. Ponno
Dipartmento di Matematica Pura e Applicata, Università di Padova Via Trieste 63, 35121 Padova, Italy

R. Livi (✉)
Dipartimento di Fisica–CSDC, Universitá di Firenze and Sezione INFN di Firenze, Via G. Sansone 1, 50019 Sesto Fiorentino, Italy
e-mail: roberto.livi@fi.infn.it

$$H(q,p)=\sum_{n=1}^{N}\frac{p_n^2}{2}+\sum_{n=0}^{N}V(q_{n+1}-q_n), \tag{1}$$

where

$$V(x)=\frac{x^2}{2}+\alpha\frac{x^3}{3}+\beta\frac{x^4}{4}, \tag{2}$$

α and β being the nonlinearity constants. We consider the case of fixed ends, i.e. $q_0 = q_{N+1} = 0$; N is the number of degrees of freedom of the system. For such a model the only quantities that determine the dynamics up to a rescaling are the products $\alpha^2\varepsilon$ and $\beta\varepsilon$, where $\varepsilon = E/N$ is the specific energy, with E denoting the total energy of the chain. In our numerical analysis we have chosen[1] $\alpha = -1$ and $\beta = 0.6$. The normal modes of the system are given by

$$Q_k=\sqrt{\frac{2}{N+1}}\sum_{n=1}^{N}q_n\sin\frac{\pi kn}{N+1},\qquad P_k=\sqrt{\frac{2}{N+1}}\sum_{n=1}^{N}p_n\sin\frac{\pi kn}{N+1}, \tag{3}$$

while the dispersion relation (frequency vs. mode index k) is

$$\omega_k=2\sin\frac{\pi k}{2N+2},\quad k=1,\ldots,N. \tag{4}$$

To each mode k one can associate a harmonic energy

$$E_k=\frac{1}{2}\left(P_k^2+\omega_k^2Q_k^2\right) \tag{5}$$

and a phase φ_k defined by

$$Q_k=\sqrt{2E_k/\omega_k^2}\sin\varphi_k,\qquad P_k=\sqrt{2E_k}\cos\varphi_k. \tag{6}$$

Relevant variables of the problem are the time averages of the mode energies,

$$\overline{E}_k(t)=\frac{1}{t}\int_0^t E_k(Q_k(t'),P_k(t'))\,\mathrm{d}t',\quad k=1,\ldots,N.$$

As was first shown by FPU, and then studied systematically in many papers (for a survey see [2, 3]), the first stage of the dynamics, taking place on a relatively short time-scale τ_1, consists in a partial transfer of energy from long to smaller wavelengths till a cut-off, so as to produce a long-life metastable state in which an "effective number" $M_1 < N$ of degrees of freedom are practically involved in the energy sharing. A possible definition of the effective number M of energy sharing degrees of freedom, associated to a given energy spectrum $\overline{E}_k$, is [4, 5]

$$M=\exp\eta, \tag{7}$$

[1] Let us recall that the q_n's have the meaning of displacement of the particles from the equilibrium positions, so that the argument $q_{n+1} - q_n$ of the interaction potential V represents the excess length of the nonlinear "spring" connecting the particles; the negative sign of α is then spontaneous, if one thinks that compressed springs get harder (as with molecular potentials). Note that changing the sign of α is also equivalent to advancing all phases of the normal modes by π.

where η, the so-called spectral entropy [6], is in turn defined as

$$\eta = -\sum_{k=1}^{N} F_k \log F_k, \quad \text{where } F_k = \frac{\overline{E}_k}{\sum_j \overline{E}_j}.$$

We shall often refer to the fraction

$$f = M/N$$

of the effectively energy-sharing modes, and denote in particular $f_1 = M_1/N$.

Nowadays it is generally accepted [6–10], on the basis of both numerical calculations and analytical arguments, that on a second time-scale τ_2, much larger than τ_1 for small energy, the metastable state breaks down and the system eventually relaxes to an equilibrium state compatible with the laws of statistical mechanics and consistent with the equipartition of the energy: indeed for $t > \tau_2$ it is $M(t) \simeq N$, so that all modes share on average the same amount of energy.

An example of such a behavior—often called "the FPU scenario"—is presented in Fig. 1. The left panel shows the energy spectrum, more precisely the ratio $\overline{E}_k(t)/\varepsilon$ plotted vs. k/N, at different times in geometric progression; the chain contains $N = 1024$ oscillators, the energy density is $\varepsilon = 2.5 \times 10^{-4}$, while the total energy has been initially equidistributed among the 10% of lowest frequency modes, with random initial phases φ_k, uniformly distributed in the interval $[0, 2\pi]$. Quite soon, for ($t \simeq 10^3$) a well defined profile is formed, in which only some low frequency modes effectively take part to energy-sharing, the energies of the remaining ones dacaying exponentially with k/N. The energy profile keeps its form nearly unchanged for a rather large time-scale, definitely much larger than the time needed to form it. Afterwards the dynamics slowly evolves towards energy equipartition, the high-frequency modes being progressively involved into the energy-sharing mechanism.[2] The

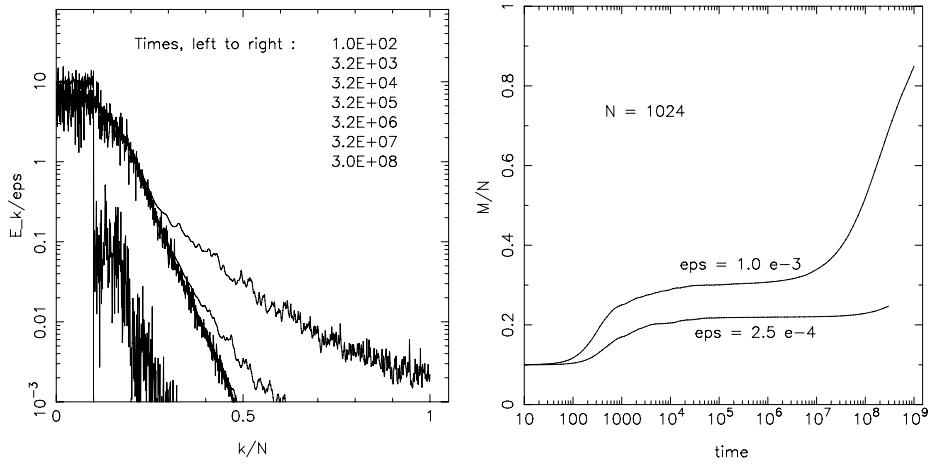

Fig. 1 *Left*: the energy distribution among normal modes at different times; data: $N = 1024$, $\varepsilon = 2.5 \times 10^{-4}$, energy initially equidistributed among 10% of modes of low frequency; times as indicated inside the figure. *Right*: the effective number of degrees of freedom M as a function of time, for the same dynamics and for $\varepsilon = 10^{-3}$

[2]The curves in the figure are a little smoothed: each point is indeed the average of a few nearby points.

right panel of the figure reports, for the same dynamics, f versus t (for comparison, we also show the same dynamics for $\varepsilon = 10^{-3}$).

The onset of the *plateau* in the figure identifies rather clearly, although qualitatively, the time-scale τ_1 which is necessary for the formation of a metastable state, which lasts until f remains close to f_1. Such a scenario is widely described and emphasized in [11] and [12] (for further comments see also [13]).[3]

In the light of this scenario, a fundamental question poses, and in fact was raised since the very beginning of the literature on FPU [14], namely *whether the scenario does persist in the so-called thermodynamic limit, or it is instead a finite-size effect which disappears if* $N \to \infty$ *at fixed specific energy* $\varepsilon = E/N$. In this perspective, it is crucial to analyze the dependence of quantities like τ_1, f_1 and τ_2 on the thermodynamic limit. Another relevant parameter to be taken into account is the number M_0 of the initially excited modes. Indeed, when $N \to \infty$ one can keep fixed either the *fraction* $f_0 = M_0/N$ of initially excited modes (as it should be done in a genuine thermodynamic perspective, since this corresponds to exciting modes up to a fixed frequency), or their *number* M_0 (in this case f_0 vanishes in the thermodynamic limit).

As a matter of fact, one finds in the literature quite different answers to the above question, with numerical results that apparently are not compatible. To be definite, we shall focus our attention on two manifestly diverging results, especially stressed in paper [15] and in the papers [10, 11] (see also, however, [12] and [9], and for comments [13, 16]). In the former study, numerical evidence is produced that the FPU scenario persists, in the thermodynamic limit, only if M_0 is kept fixed; instead for fixed f_0 the FPU scenario disappears, unless the total energy E, rather than the specific energy ε, is kept fixed when $N \to \infty$. In [10, 11] instead, an opposite numerical evidence is obtained: indeed the whole scenario, and specifically two of the above mentioned quantities, f_1 and τ_2, are shown to be independent of the thermodynamic limit, even for thermodynamic initial data with $f_0 = M_0/N$ kept fixed.

The particular models considered in these papers are not identical, nor is identical the definition of M, which in turn is at the basis of the computation of τ_1, f_1 and τ_2. On the other hand, after a careful inspection one can conclude that none of such differences seems to be really crucial. However, there remains an apparently irrelevant difference concerning *the recipe adopted for the choice of the phases of the initially excited normal modes*. Indeed, in [15] the phases are coherent (more precisely, $\varphi_k = \pi/2$, so that the initial energy is purely potential), while in papers [10, 11] the phases are incoherent, namely chosen randomly and uniformly distributed in the interval $[0, 2\pi]$. *As we are going to show, this makes a big difference: unexpectedly, initial phases turn out to play a crucial role in determining the FPU scenario*.

Illustrating the role of phases in the classical FPU problem is the principal aim of the present paper. In the first part of the paper (Sect. 2) we present a detailed numerical study on the formation and persistence of the metastable state. In particular, we show that, for given values of N, E and M_0, the quantity $f(t)$ depends in an important way on the phases: coherent phases produce definitely larger values of $f(t)$ than random phases; moreover, different choices of coherent phases produce different results, while different samples of random phases give rise to almost identical $f(t)$. Beyond this qualitative remark, we produce evidence that, when N, ε and M_0 (or f_0) are varied, the dynamics obeys some simple scaling laws. The main outcomes in this direction can be summarized as follows:

[3] Different quantitative definitions of τ_1 could be provided, and are indeed provided in the literature by the different authors, depending on their aims. The scaling analysis we shall produce in this paper, however, looks at the curve $M(t)$ as a whole, and does not need such a quantitative notion. A precise identification of τ_2 is even less necessary, since we are not investigating here the eventual relaxation of the system to equilibrium.

A. If M_0 is kept fixed and N is large, then *for any choice of the initial phases* one has $f_1 \sim \varepsilon^{1/4}$ and $\tau_1 \sim \varepsilon^{-3/8} N^{3/2}$. Such a result is in agreement with both [15] and [10, 11]. We wish to point out that in the thermodynamic limit, in order to keep ε fixed, one must store an increasing amount of energy in the excited modes. Nonetheless, this energy needs to be transferred to an increasing number of modes, and eventually τ_1 is found to diverge with N. The stability of this result for large N at fixed M_0 and ε was observed in particular in [15] (in the case of coherent phases). Let us stress again that such conditions do not correspond to the genuine thermodynamic ones, where both ε and f_0 should be fixed in the limit $N \to \infty$.

B. If instead $f_0 = M_0/N$ is kept fixed, it becomes necessary to distinguish between random and coherent initial phases:

   B1. *For random initial phases,* one has $f_1 \sim \varepsilon^{1/4}$ and $\tau_1 \sim \varepsilon^{-3/8} f_0^{-3/2}$, both quantities being independent of N. Accordingly, such conditions imply the persistence of the FPU scenario in the thermodynamic limit. This result is in agreement with [11] (where, in particular, the law $f_1 \sim \varepsilon^{1/4}$ is stressed) and [10].

   B2. *For coherent initial phases,* one has instead $f_1 \sim (f_0 E)^{1/4}$ and $\tau_1 \sim E^{-3/8} f_0^{-3/2}$. In such a case, E rather than ε is the relevant scaling variable and the persistence of the FPU scenario in the thermodynamic limit is ruled out, as observed in [15].

Let us remark that the energy cascade from low to high frequencies always implies $f_0 < f_1$. Thus, in the above mentioned case B1 one has $\tau_1 \sim \varepsilon^{-3/8} f_0^{-3/2} > \varepsilon^{-3/8} f_1^{-3/2} \sim \varepsilon^{-3/4}$, which gives a sharp lower bound to the time-scale τ_1 in agreement with the numerical estimate given in [12] (in the case B2, ε should be replaced by E). As far as we know, the scaling $f_1 \sim \varepsilon^{1/4}$ and the presence of a time-scale $\sim \varepsilon^{-3/4}$ were first analytically predicted in [17].

In the second part of the paper (Sect. 3), we then produce some simple theoretical estimates, which provide a preliminary explanation of the above mentioned numerical results. Such a theoretical investigation is based on a suitable resonant normal form of the FPU system, on its relation to the Korteweg-de Vries (KdV) equation and on the scaling properties of the latter equation. A more complete theoretical understanding of the problem, however, certainly demands a further effort; see for a sketch the last section, where perspective and conclusions are discussed.

2 Numerical Studies

We begin this section by producing qualitative evidence that initial phases are relevant for the FPU scenario. Figure 2 reports $f = M/N$ vs. t for four different samples of random phases (lower bunch of curves) and five different choices of coherent phases, namely $\varphi_k(0)$ constant for all $k < M_0$ and equal to $0, \pi/4, \pi/2, 3\pi/4, \pi$. The parameters are $N = 1024$, $\varepsilon = 2.5 \times 10^{-4}$, $M_0 = 102$. The different behavior between random and coherent phases, and among different choices of coherent phases, is rather evident. Other choices of coherent phases, for instance $\varphi_k(0) = k\psi$ with some ψ for $k \leq M_0$, produce curves similar to those obtained for coherent phases, dispersed as they are in the figure.

We then proceed by stating and testing numerically some scaling laws, which are obeyed by the dynamics and imply, and make precise statements A and B in the introduction. Scaling laws concern the behavior of f as a function of t, N, ε and either M_0 or f_0. In order to distinguish the two cases we introduce slightly different notations:

$$f(t, N, \varepsilon, M_0) \quad \text{or} \quad \hat{f}(t, N, \varepsilon, f_0).$$

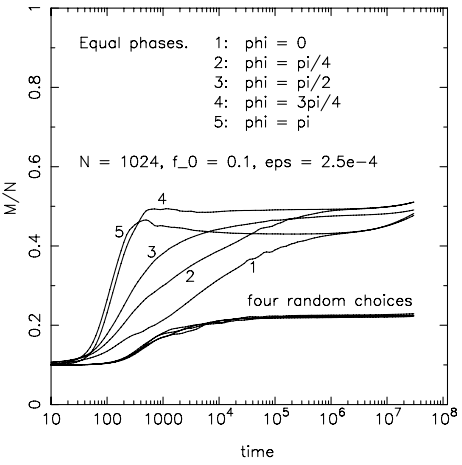

Fig. 2 M vs. t for four different samples of random phases (*lower bunch of curves*) and five different choices of coherent phases (*upper curves*), namely $\varphi_k(0) = 0, \pi/4, \pi/2, 3\pi/4$ and π for all of the excited modes. Parameters: $N = 1024$, $\varepsilon = 2.5 \times 10^{-4}$, $M_0 = 102$

A. Scaling Laws Valid for all Initial Phases Here we state two scaling laws, which concern conditions in which the number M_0 of initially excited modes is kept fixed, and are independent of the choice of the phases. The first of them, which we consider to be in a sense the fundamental one, is

$$f(\lambda^{-3/4}t, \lambda^{-1/4}N, \lambda\varepsilon, M_0) = \lambda^{1/4} f(t, N, \varepsilon, M_0). \tag{8}$$

Figure 3 shows the scaling law at work, for random phases (upper panels) and regular phases (lower panels; actually $\varphi_k(0) = \pi/2$ for all the excited modes). The left panels show f vs. t as it is, without rescaling, while in the right panels the curves are rescaled according to (8). More precisely, we took eight different values N_j and ε_j according to

$$N_j \simeq \lambda_j^{-1/4} N^*, \qquad \varepsilon_j = \lambda_j \varepsilon^*,$$

with $N^* = 4096$, ε^* equal to 2.5×10^{-4} for random phases and to 10^{-5} for coherent phases, and

$$\lambda_j = 2^{-j+2}, \quad j = 0, \ldots, 7.$$

In the right panels the curves are multiplied by $\lambda_j^{-1/4}$ and left shifted by $\frac{3}{4}\log\lambda_j$; quite clearly, both for the random and for the coherent phases, the rescaled curves beautifully superimpose. One could wonder how a cumbersome scaling relation like (8) could come to our mind. The expert reader should have realized that such a law reproduces *exactly* the scaling group of the KdV equation. There are of course deep reasons for this: they will be explained in some detail in Sect. 3 (see in particular Sect. 3.2) and more extensively analyzed in a paper in preparation [18].

The second scaling law, independent of the previous one and valid asymptotically for large N, is

$$f(\mu^{3/2}t, \mu N, \varepsilon, M_0) = f(t, N, \varepsilon, M_0). \tag{9}$$

It shows that by varying N at constant M_0 and ε, the time scale changes, but the height f_1 of the *plateau* does not. Figure 4 illustrates such a scaling law. As in Fig. 3, the upper panels refer to random initial phases, while the lower panels concern coherent phases, with $\varphi_k(0) = \pi/2$ for $k \le M_0$. The left panels show f as a function of t, while the right panels

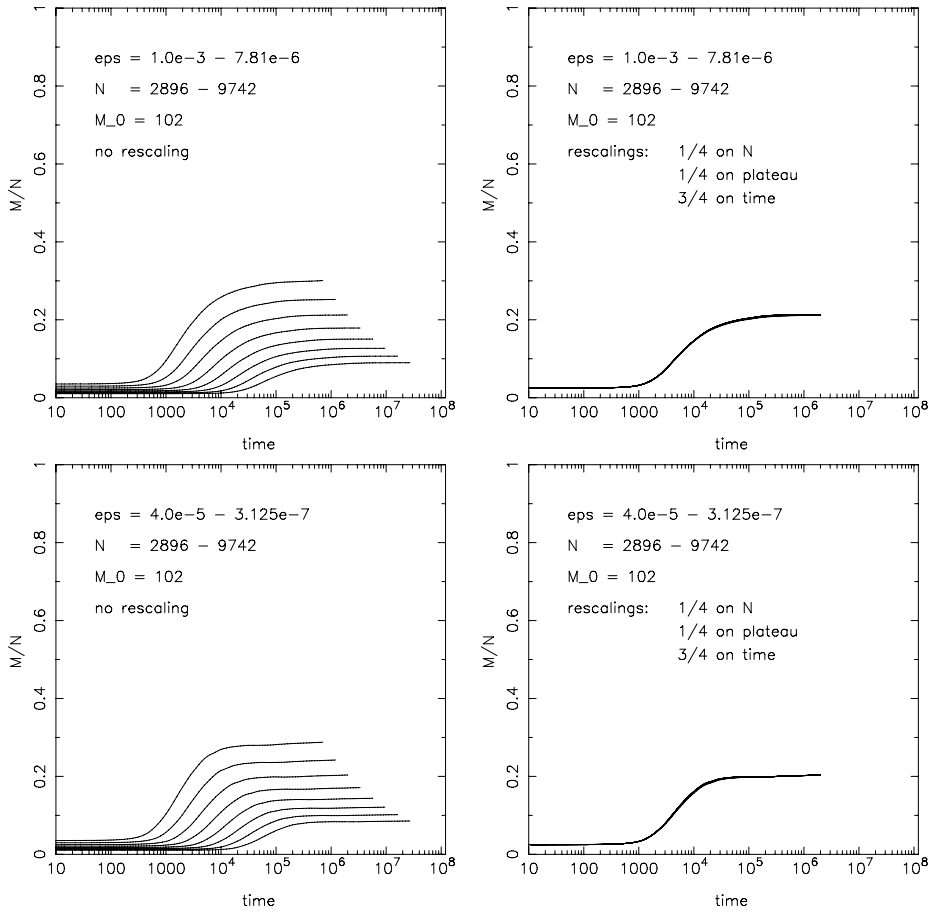

Fig. 3 A test of the basic scaling law (8). *Upper panels*: random initial phases; *lower panels*: coherent initial phases ($\varphi_k = 0$ for $k \le M_0$). See the text for the value of the parameters. The plotted quantity is f vs. t, as it is (*left*) and after rescaling (*right*)

include rescaling. Notice that in this case the *plateau* is not touched by the rescaling, which reduces here to a horizontal translation of the curves in the panel by $-\frac{3}{2}\log\mu$. The chosen parameter values are $\varepsilon = 2.5 \times 10^{-4}$, $N = 1024, \ldots, 16384$, $M_0 = 26$. The rescaled curves superimpose exactly on the *plateau*, while in the transient they do only for large N, when $f_0 = M_0/N$ is negligible (in the absence of vertical rescaling, curves with different f_0 cannot superimpose for small times).

By combining together (8) and (9) one obtains

$$f(\mu^{3/2}\lambda^{-3/4}t, \mu\lambda^{-1/4}N, \lambda\varepsilon, M_0) = \lambda^{1/4} f(t, N, \varepsilon, M_0); \tag{10}$$

in particular, for $\mu = \lambda^{1/4}$, (10) becomes

$$f(\lambda^{-3/8}t, N, \lambda\varepsilon, M_0) = \lambda^{1/4} f(t, N, \varepsilon, M_0). \tag{11}$$

As an exercise, such a scaling law can be tested directly. The result is shown in Fig. 5 for $N = 1024$, $M_0 = 3$, $\varepsilon = 2^{-j} \times 10^{-3}$ with $j = 0, \ldots, 6$, coherent phases $\varphi_k(0) = \pi/2$ for

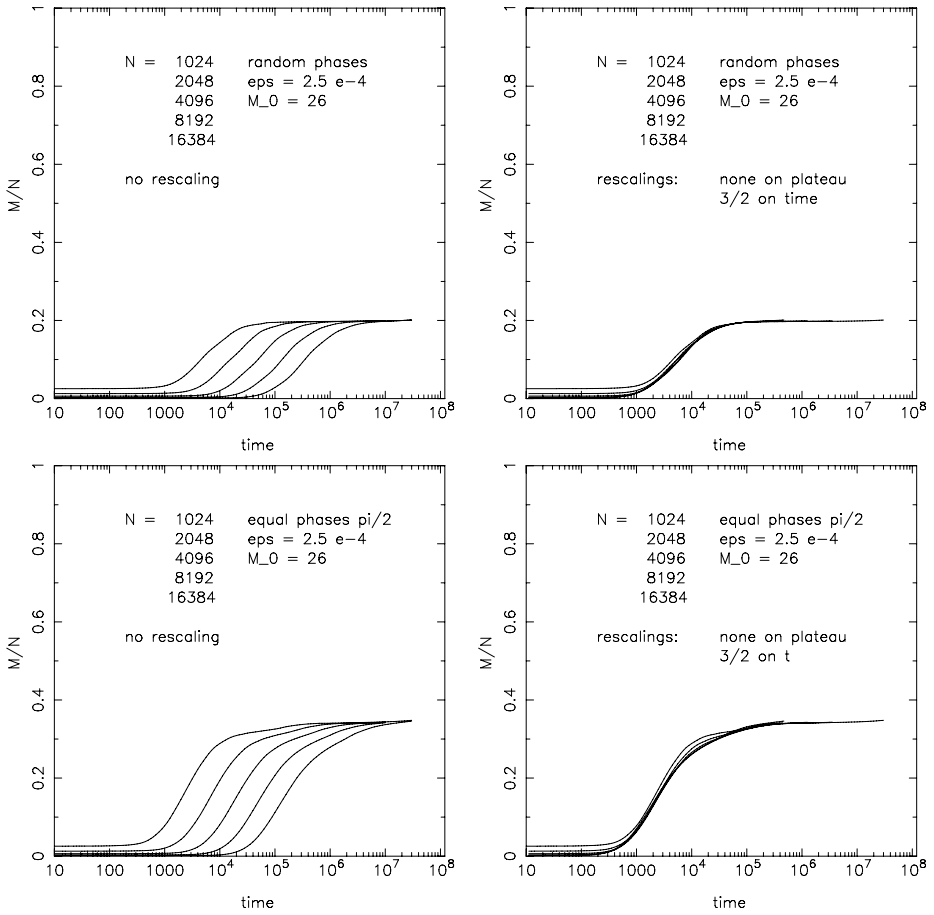

Fig. 4 A test of the scaling law (9). *Upper panels*: random initial phases; *lower panels*: coherent initial phases ($\varphi_k = \pi/2$ for $k \le M_0$). Parameters: $\varepsilon = 2.5 \times 10^{-4}$, $N = 1024, \ldots, 16384$, $M_0 = 26$. The plotted quantity is f vs. t, as it is (*left*) and after rescaling (*right*)

$k = 1, 2, 3$ (random phases would be meaningless for $M_0 = 3$). As before, the right panel includes rescaling. It is worthwhile to observe that the rescaled curves do superimpose till they reach the *plateau*, but do not anymore for larger times, when they leave the *plateau* towards equipartition. We shall come back to this point in the conclusions.

An equivalent formulation of (10) is

$$f(t, N, \varepsilon, M_0) = \varepsilon^{1/4} \mathcal{F}(\varepsilon^{3/8} N^{-3/2} t, M_0), \tag{12}$$

$\mathcal{F}$ being a suitable function of only two variables (depending of course on the choice of the initial phases). Such a relation implies statement A in the introduction.

B1. The Thermodynamic Limit for Random Initial Phases If the initial phases are chosen randomly, ε and f_0 (rather than M_0) are kept fixed, and N is sufficiently large, then f is independent of N, i.e. the model exhibits nontrivial thermodynamic limit properties. This

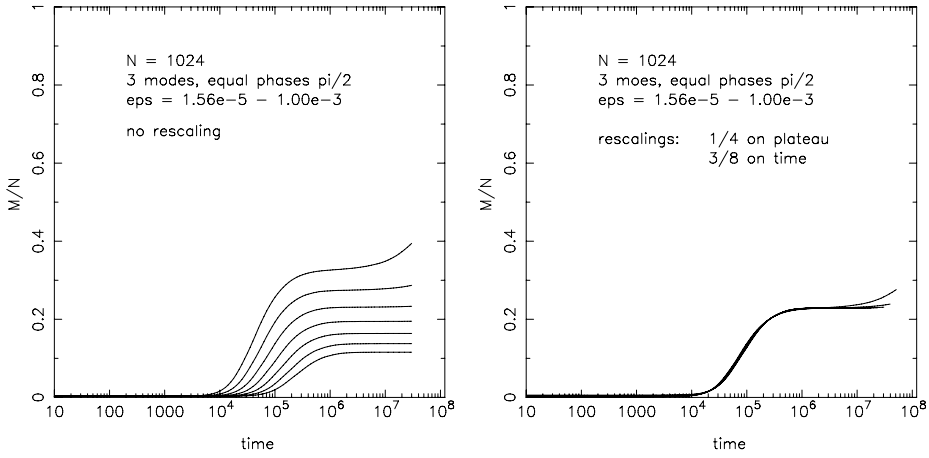

Fig. 5 A test of the scaling law (11). Parameters: $N = 1024, \ldots, 8192$, $\varepsilon = 10^{-3}, \ldots, 1.56 \times 10^{-5}$; $M_0 = 3$, $\varphi_k = \pi/2$ for $k = 1, \ldots, 3$. The plotted quantity is f vs. t, as it is (*left*) and after rescaling (*right*)

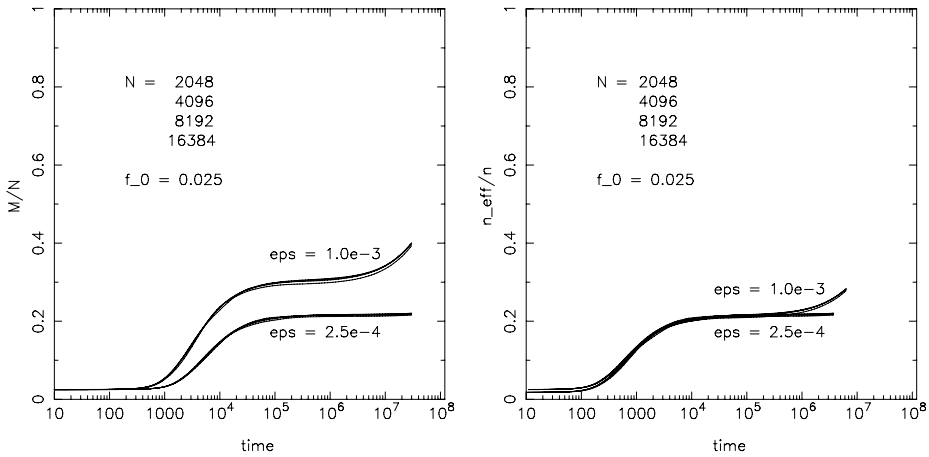

Fig. 6 A test of the scaling law (13) (*left panel*) and of the scaling law (15); the parameter values are reported inside the figures

situation can be translated in the language of scaling laws as follows:

$$f(t, \varrho N, \varepsilon, \varrho M_0) = f(t, N, \varepsilon, M_0), \tag{13}$$

or equivalently

$$\hat{f}(t, \varrho N, \varepsilon, f_0) = \hat{f}(t, N, \varepsilon, f_0). \tag{14}$$

Figure 6 illustrates this scenario. The left panel shows two bunches of curves obtained for $f_0 = 0.025$ and $N = 2048, 4096, 8182, 16384$; the lower and the upper ones correspond to $\varepsilon = 2.4 \times 10^{-4}$ and $\varepsilon = 10^{-3}$, respectively. The number N of degrees of freedom is manifestly irrelevant. According to the previous scaling laws, the two bunches must superimpose their *plateau* if they are rescaled vertically proportionally to $\varepsilon^{1/4}$; the result is shown in the

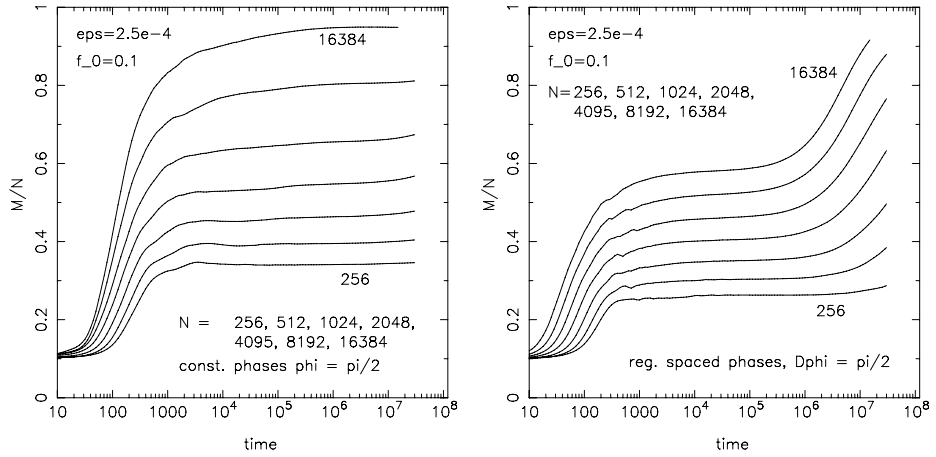

Fig. 7 The failure of the thermodynamic limit for coherent initial phases, namely $\varphi_k(0) = \pi/2$ (*left panel*) and $\varphi_k(0) = k\pi/2$ (*right panel*); the values of the parameters are reported inside the figures

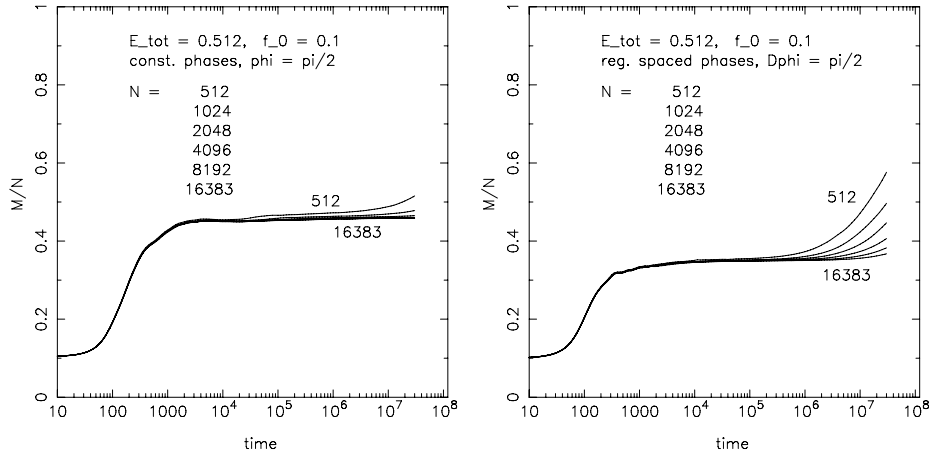

Fig. 8 The failure of the thermodynamic limit for coherent initial phases: a test of the scaling law (16) for $\varphi_k(0) = \pi/2$ (*left panel*) and $\varphi_k(0) = k\pi/2$ (*right panel*); the parameter values are reported inside the figures

left panel. An equivalent formulation of (14), keeping into account (12), is

$$\hat{f}(t, N, \varepsilon, f_0) = \varepsilon^{1/4} \mathcal{G}(\varepsilon^{3/8} f_0^{3/2} t), \tag{15}$$

$\mathcal{G}$ being a suitable function of a single variable.

B2. The Failure of the Thermodynamic Limit for Coherent Phases For coherent initial phases, thermodynamic limit properties are lost when N is increased for fixed ε and f_0. This is evident from Fig. 7, where we report f as function of t for fixed $\varepsilon = 2.5 \times 10^{-4}$ and $f_0 = 0.1$, N ranging between 256 and 16384. The left panel refers to equal initial phases, namely $\varphi_k(0) = \pi/2$ for all excited modes, while the right panel refers to $\varphi_k(0) = k\pi/2$.

As remarked in [15], a scaling law is recovered if the total energy E is kept fixed, instead of ε. More precisely, both the level of the *plateau* and the time needed to reach it are found to scale according to the relation

$$\hat{f}(t, \varrho N, \varrho^{-1}\varepsilon, f_0) = \hat{f}(t, N, \varepsilon, f_0). \tag{16}$$

Such a behavior is illustrated in Fig. 8, which refers to $E = 0.512$ and $f_0 = 0.1$, with N ranging between 512 and 16384. The left and right panels refer to the same kind of coherent initial phases as in Fig. 7. A good superposition is obtained for N large enough.

We wish to remark that the choice of coherent initial phases favours the flow of energy to higher wavelength modes, although a *plateau* is formed also in this case. This again confirms the peculiarity of the "packet-like" initial conditions, which always yield the formation of a metastable state, irrespectively of the choice of the initial phases. Nonetheless, only random initial phases guarantee the persistence of the FPU scenario in the thermodynamic limit. In the next section we are going to provide some theoretical argument to explain the origin of this unexpected situation.

3 Theory

In order to describe the FPU scenario, it looks reasonable to exploit an effective representation of the model adapted to the class of initial conditions (packets of long-wavelength modes) taken into account in the early numerical experiments as well as in the above described ones. Indeed, passing directly to a continuum-like limit representation (which seems reasonable in the case of long wavelength excitations), as Zabusky and Kruskal first did in their celebrated paper [19], easily leads to the KdV equation, which with our choice of notations and constants writes

$$U_t = \frac{1}{24} U_{xxx} + \frac{\alpha}{\sqrt{2}} U U_x, \tag{17}$$

with periodic boundary conditions:

$$U(x + L, t) = U(x, t), \quad L = 2N + 2. \tag{18}$$

Such an equation turns out to be fundamental not only in explaining the recurrence of the initial datum in the problem, but also the quasi-integrability properties of the FPU dynamics on short time-scales [20, 21], as well as the form of the energy spectrum [22]. However, such a continuum limit, even though mathematically correct, hides somehow the way one performs the large N limit, in that one approximates a finite-dimensional system with an infinite-dimensional one. Instead, a "bridge" between the continuous approximation and the standard perturbation theory for finite systems must be built up, and the discrete nature of the problem has to be kept, in order to obtain a consistent approach to the thermodynamic limit. As we are going to discuss in this section, one can obtain a suitable description of the FPU scenario by considering the dynamics of the resonant normal form of the FPU Hamiltonian, first introduced in [17]. As we shall see, a suitable rescaling transforms the normal form equations of FPU into the so-called Fourier–Galerkin truncation to N modes of the $2N + 2$-periodic KdV equation. This provides a twofold support to the numerical results shown in the previous section: (i) it explains the scaling properties of the observable f; (ii) it allows one to identify a criterion for understanding the different thermodynamic limit properties emerging from the different choices of the initial phases.

As a preliminary step, it is convenient to pass from the real coordinates (3) to the complex coordinates

$$z_k = \frac{P_k + i\omega_k Q_k}{\sqrt{2\omega_k}}, \qquad z_k^* = \frac{P_k - i\omega_k Q_k}{\sqrt{2\omega_k}}. \tag{19}$$

These are nearly canonical, the only nonvanishing Poisson brackets being $\{z_k, z_k^*\} = i$, for $k = 1, \ldots, N$. Elementary substitutions and trivial trigonometric identities transform Hamiltonian (1) into the new Hamiltonian (for which we do not change the notation)

$$H(z, z^*) = \sum_{k=1}^{N} \omega_k |z_k|^2 + \frac{i\alpha}{12\sqrt{N+1}} \times \sum_{j,k,l=1}^{N} S_3(j,k,l)\sqrt{\omega_j \omega_k \omega_l}\,(z_j - z_j^*)(z_k - z_k^*)(z_l - z_l^*) + O(\beta z^4), \tag{20}$$

where the selector S_3 of the mode-coupling is given by

$$S_3(j,k,l) = \delta_{j+k,l} + \delta_{j+l,k} + \delta_{k+l,j} - \delta_{j+k+l,2N+2}, \tag{21}$$

$\delta_{n,m}$ denoting the usual Kronecker delta. We wish to point out that in (20) we did not write explicitly the fourth-order term, since it does not influence the result of the first step of normalization that we are going to perform.[4] The Hamilton equations read now $\dot{z}_k = i\partial H/\partial z_k^*$, $\dot{z}_k^* = -i\partial H/\partial z_k$.

3.1 Resonant Normal Form

Let us shortly recall the idea underlying the construction of the resonant normal form for the FPU problem first introduced in [17]. We follow here the approach first adopted in [23], also reported in [24] (where it is also extended to the pure β-model), which consists in normalizing the Hamiltonian in Fourier space; a complementary and equivalent normalization procedure in ordinary space is instead used in papers [20, 21]. The core of the method resides in the expansion of the dispersion relation (4)

$$\omega_k = \xi_k - \frac{1}{24}\,\xi_k^3 + O(\xi_k^5), \quad \xi_k \equiv \frac{\pi k}{N+1}; \tag{22}$$

the expansion is clearly adapted to the acoustic modes (small k), and to their approximate resonance relation

$$\omega_1 \simeq \omega_2/2 \simeq \omega_3/3 \cdots,$$

with a decreasing level of approximation. As first pointed out in [25], this expansion provides a suitable description of the energy cascade mechanism in the FPU problem.

The quadratic part of Hamiltonian (20) is now split as follows:

$$\sum_{k=1}^{N} \omega_k |z_k|^2 = \sum_{k=1}^{N} \xi_k |z_k|^2 + \sum_{k=1}^{N} (-\xi_k^3/24 + \cdots)|z_k|^2,$$

[4] In absence of the cubic nonlinearity, the normalization procedure sketched in this section yields a different normal form, which corresponds to the periodic version of the modified KdV-equation (see [18]).

and for initial excitations of acoustic modes, the sum

$$J(z, z^*) \equiv \sum_{k=1}^{N} \xi_k |z_k|^2 \tag{23}$$

is regarded as the unperturbed Hamiltonian of the problem: $H = J +$perturbation. Then, one can perform a standard step of normalization (in the sense of canonical perturbation theory), removing all the non-resonant terms in the perturbation at the leading order. Neglecting the remainder, one is then left with the *first order resonant normal form Hamiltonian* of the system

$$H_{res}(z, z^*) = J(z, z^*) + \sum_{k=1}^{N} \left(-\xi_k^3/24\right) |z_k|^2 + \frac{i\alpha}{4\sqrt{N+1}} \sum_{j,k,l=1}^{N} \delta_{j+k,l} \sqrt{\xi_j \xi_k \xi_l}\ (z_j^* z_k^* z_l - z_j z_k z_l^*). \tag{24}$$

(A minor abuse of notation was introduced, since we keep on calling z the new canonical variables appearing in H_{res}.) Now, as a consequence of the normalization, one gets $\{J, H_{res}\} = 0$, i.e. the dynamics of the normal form is characterized by the presence of a second constant of motion [17]. This fact is conveniently taken into account by passing to the co-rotating coordinates ζ_k [23] defined by

$$z_k = i e^{i\xi_k t} \zeta_k, \tag{25}$$

where a $\pi/2$-phase shift is introduced for later convenience. The transformation $z \mapsto \zeta$ is canonical[5] and depends on time, in such a way to exactly erase the second integral J from the right hand side of (24). One is thus left with the *equivalent resonant normal form Hamiltonian*

$$K(\zeta, \zeta^*) = \sum_{k=1}^{N} \left(-\xi_k^3/24\right) |\zeta_k|^2 + \frac{\alpha}{4\sqrt{N+1}} \sum_{j,k,l=1}^{N} \delta_{j+k,l} \sqrt{\xi_j \xi_k \xi_l} (\zeta_j^* \zeta_k^* \zeta_l + \zeta_j \zeta_k \zeta_l^*). \tag{26}$$

Remarks

(i) The latter normal form is obtained by making use everywhere of the expansion (22) of the dispersion relation of the system, which implicitly assumes that the active modes participating to the dynamics are the low ones, only.

(ii) A more direct procedure to work out Hamiltonian (26) amounts to performing the time-dependent canonical transformation (25) on Hamiltonian (20): one is left with $H = K + R$, where R is the remainder, that displays explicit dependence on time through oscillating factors, so that its formal time average, with the ζ's kept constant, vanishes (this is the resonant averaging over the unperturbed motion *a la* Bogolyubov).

(iii) From (25) it follows that the initial phases of the ζ_k are those of the z_k (namely the initial values φ_k^0) decreased by $\pi/2$.

[5] As for the z_k, the only nonvanishing Poisson brackets are $\{\zeta_k, \zeta_k^*\} = i$, $k = 1, \ldots, N$.

The Hamilton equations associated to (26) are given by $\dot{\zeta}_k = i\,\partial K/\partial \zeta_k^*$, namely

$$\dot{\zeta}_k = -\frac{i\xi_k^3}{24}\,\zeta_k + \frac{i\sqrt{\xi_k}\alpha}{4\sqrt{N+1}}\left[\sum_{j=1}^{k-1}\sqrt{\xi_{k-j}\xi_j}\;\zeta_{k-j}\zeta_j + 2\sum_{j=1}^{N-k}\sqrt{\xi_j\xi_{k+j}}\;\zeta_j^*\zeta_{k+j}\right]. \tag{27}$$

(If expressed in terms of the action-angle variables (I,θ) defined by $\zeta_k = \sqrt{I_k}e^{i\theta_k}$, the Hamiltonian (26) coincides with the one first derived in [17]; the use of the Cartesian-like complex variables (ζ,ζ^*), however, will be convenient below.)

3.2 The Truncated KdV Equation and the Basic Scaling Law (8)

It is convenient to introduce the new *noncanonically* rescaled energy variables

$$u_k = \sqrt{\xi_k}\zeta_k, \quad k = 1,\ldots,N, \tag{28}$$

in terms of which the equations of motion (27) transform to

$$\dot{u}_k = -\frac{i\xi_k^3}{24}u_k + \frac{i\xi_k\alpha}{4\sqrt{N+1}}\left[\sum_{j=1}^{k-1}u_{k-j}u_j + 2\sum_{j=1}^{N-k}u_j^*u_{k+j}\right]. \tag{29}$$

There is a strong relation between (29) and the KdV equation (17):

Proposition *The normal form* (29) *coincide with the Fourier–Galerkin truncation to the first N modes of the KdV equation* (17), *with periodic zero average initial datum, namely* $U(x+L,0) = U(x,0)$, $\int_0^L U(x,0)\,\mathrm{d}x = 0$, *where* $L = 2N+2$.

We recall that the Fourier–Galerkin truncation to N modes of the L-periodic field variable $U(x,t)$ is defined as

$$U^N(x,t) = \frac{1}{\sqrt{L}}\sum_{k=-N}^{N}\widehat{U}_k(t)e^{2\pi ikx/L}, \tag{30}$$

where of course $\widehat{U}_k(t) = L^{-1/2}\int_0^L U(x,t)e^{2\pi ikx/L}\,\mathrm{d}x$. The zero average condition on U implies $\widehat{U}_0 = 0$. The reason of the choice $L = 2N+2$ appearing in the proposition is simply understood: the original discrete system with fixed ends at sites $n = 0$ and $n = N+1$ can be regarded as half a periodic system with sites ranging from $n = -N$ to $n = N+1$, the latter being identified with site $-N-1$.

The above proposition, whose simple proof is reported in Appendix, has deep consequences. First of all, it shows that the KdV equation is not just one among other possible continuum limits of the FPU chain, rather it plays a unique role, since it is close to the resonant normal form equations derived by the means of Hamiltonian perturbation theory. Second, it predicts that by increasing N, many features of the dynamics tend to regularize, since the normal form dynamics, close to the true dynamics, becomes closer and closer to the dynamics of the KdV equation (17). This provides, in our opinion, a strong support to the idea [19] that the integrability properties of the KdV equation might explain straight away the strong memory of the initial data displayed by the FPU dynamics.

For what concerns the thermodynamic limit, we stress that one should not be tempted to roughly perform the limit $N \to \infty$ in the KdV problem (17), (18). Indeed, as N increases,

the length of the periodicity interval also increases. One can actually rescale the space and time variables x and t setting $x = LX$, and $t = LT$, in terms of which (17) reads

$$V_T = \frac{1}{24L^2}\, V_{XXX} + \frac{\alpha}{\sqrt{2}}\, V V_X, \tag{31}$$

for the 1-periodic unknown field $V(X, T) = U(LX, LT)$. One thus sees that performing the limit $N \to \infty$ in the KdV equation (17) amounts to studying the *zero dispersion* limit $L \to \infty$ of the KdV equation (31) on a fixed interval of length 1, which is a highly nontrivial problem; see e.g. [26]. In such a respect, notice that in papers [20, 21] a particular subclass of initial condition was chosen, such that the KdV equation (31) does not display the small coefficient $1/L^2$ in front of the dispersive term. Such a nontrivial "simplification" takes place by exciting a mode (and its higher harmonics) of index k_0 such that $k_0/N \sim \varepsilon^{1/4}$, and allows to rigorously justify the normal form construction. Unfortunately, the results of [20, 21] do not apply in a simple way to the case considered here of extended, packet-like initial excitations.

We do not proceed further in such a direction, which will be reconsidered elsewhere [18], but limit ourselves to use the normal form (29) to support, by means of a simple scaling analysis, the numerical results reported in Sect. 2.

The first of them, namely the scaling law (8), is an immediate expression of the fundamental group of invariance of the KdV equation (17), which turns out to be

$$U \to \lambda^{1/2} U, \qquad t \to \lambda^{-3/4} t, \qquad x \to \lambda^{-1/4} x, \qquad L \to \lambda^{-1/4} L,$$

λ being any positive number. The connection with (8) is indeed straightforward, if we take into account that the length L of the periodicity interval of the FPU problem is (almost) proportional to N, and recall that f is, basically, a wave number, which scales as an inverse length and so as $\lambda^{-1/4}$.

Explaining the other numerically observed scaling laws requires some further work, which will occupy the next Sects. 3.3 and 3.4.

3.3 Rescaling

For later purposes, we observe that in terms of the u-variables the second integral J now reads

$$J = \sum_{k=1}^{N} |u_k|^2 \equiv N\varepsilon_a, \tag{32}$$

where we have defined the *specific acoustic energy* $\varepsilon_a = J/N$. In the perturbative regime considered here (where $H = J + \text{perturbation}$) this is close to the true specific energy $\varepsilon = H/N$. Indeed,

$$|u_k(t)|^2 = \xi_k |\zeta_k(t)|^2 = \xi_k |z_k(t)|^2 \simeq \omega_k |z_k(t)|^2 = E_k(t), \tag{33}$$

this approximation being valid as long as the dynamics involves long wavelength modes only, so that $\omega_k \simeq \xi_k$. Accordingly, the variables u_k represent the complex amplitudes of the modal energies E_k, which are used to build up the spectral entropy. Moreover, it follows from (25) that

$$u_k(0) = -i\sqrt{\xi_k}\; z_k(0) = -i\sqrt{\xi_k E_k(0)/\omega_k}\; e^{i\varphi_k(0)} \simeq \sqrt{E_k(0)}\; e^{i(\varphi_k(0) - \frac{\pi}{2})},$$

the approximation being valid as long as only low-wavenumber modes are initially excited. The initial data in terms of the new u variables read

$$u_k(0) = \begin{cases} -i\sqrt{\mathcal{E}}\, e^{i\varphi_k^0} & \text{for } k = 1, \ldots, M_0, \\ 0 & \text{for } k = M_0 + 1, \ldots, N \end{cases} \tag{34}$$

and correspondingly the conservation law (32) implies

$$M_0 \mathcal{E} = N \varepsilon_a. \tag{35}$$

This is a basic relation that can be used to eliminate the quantity $\mathcal{E}$ and rewrite the initial value problem (29), (34) as follows:

$$i\dot{u}_k = \frac{\xi_k^3}{24}\, u_k - \frac{\xi_k \alpha}{4\sqrt{N+1}} \left[\sum_{j=1}^{k-1} u_{k-j} u_j + 2 \sum_{j=1}^{N-k} u_j^* u_{k+j} \right], \tag{36}$$

$$u_k(0) = \begin{cases} \sqrt{\frac{N\varepsilon_a}{M_0}}\, e^{(i\varphi_k^0 - \pi/2)} & \text{for } k = 1, \ldots, M_0, \\ 0 & \text{for } k = M_0 + 1, \ldots, N. \end{cases} \tag{37}$$

We now perform a suitable rescaling $(u_k, N, t) \longmapsto (v_k, \mathcal{N}, T)$ of the dynamical variables, of time and of the number of degrees of freedom N, namely

$$u_k = \sqrt{N\varepsilon_a}\, v_k, \qquad N = \mathcal{N}/(\alpha^2 \varepsilon_a)^{1/4}, \qquad t = T/(\alpha^2 \varepsilon_a)^{3/4}; \tag{38}$$

the rescaling of N also implies the rescaling of the specific wavenumber $\xi_k = \pi k/(N+1)$, namely

$$\xi_k \mapsto \eta_k = \frac{\pi k}{\mathcal{N}} = (\alpha^2 \varepsilon_a)^{-1/4} \xi_k. \tag{39}$$

In terms of the new variables v_k, $\mathcal{N}$, T and η_k, (36)–(37) for large N transform to:

$$i \frac{\mathrm{d}v_k}{\mathrm{d}T} = \frac{\eta_k^3}{24}\, v_k - \frac{\eta_k}{4} \left[\sum_{j=1}^{k-1} v_{k-j} v_j + 2 \sum_{j=1}^{\mathcal{N}/(\alpha^2\varepsilon_a)^{1/4} - k} v_j^* v_{k+j} \right], \tag{40}$$

$$v_k(0) = \begin{cases} \frac{1}{\sqrt{M_0}}\, e^{i(\varphi_k^0 + \pi/2)} & \text{for } k = 1, \ldots, M_0, \\ 0 & \text{for } k = M_0 + 1, \ldots, \mathcal{N}/(\alpha^2\varepsilon_a)^{1/4}. \end{cases} \tag{41}$$

Notice that the prefactor $\sqrt{N/(N+1)}$ in front of the nonlinear term in the equations of motion (40) has been set equal to one under the hypothesis that N is very large.

3.4 Estimates

In order to find an estimate for the effective number of degrees of freedom involved in the dynamics, we propose a simple argument, along the lines already sketched in [23, 24].

As a first fundamental remark, we observe that the small parameter $\alpha^2\varepsilon_a$ enters (40)–(41) only through the rescaling (38) of N. Suppose now that only consecutive low modes take part to the dynamics. Then, their number will be of order k_c, k_c being the index of the highest significantly excited mode. *If, as a rough approximation, modes up to k_c are excited with comparable amplitude, then the height f_1 of the plateau identifies with k_c* (or is proportional

to it), so what we need is finding a good scaling law for k_c. If N is large, then the second sum in (39) practically extends to infinity; correspondingly the rescaled wavenumber η_{k_c} is not affected by the exact value of N, and this amounts to say that η_{k_c} can only depend on the number M_0 of initially excited modes, i.e.

$$\eta_{k_c} = g(M_0), \tag{42}$$

where $g(M_0)$ is some unknown function to be determined. This immediately leads to

$$f_1 \approx \frac{k_c}{N} \approx g(M_0)(\alpha^2 \varepsilon_a)^{1/4}. \tag{43}$$

The function $g(M_0)$ is expected to depend on the choice of the initial phases.

So we are left with the problem of determining $g(M_0)$. In order to do this, we observe that the right hand side of (40) consists of two terms: the first one takes into account the effect of dispersion, i.e. the detuning of the modes from the acoustic resonance, whereas the second one takes into account the nonlinear mode coupling. If the first term prevails, for a given η_k, then mode k is weakly coupled to the other ones and, if initially switched off, it is expected to remain unexcited. On the other hand, if the mode-coupling term prevails, the same mode k is expected to be involved in the dynamics and to effectively exchange energy with the other excited modes. Notice that the linear term in (40) is proportional to η_k^3, while the nonlinear one is proportional to η_k, so that modes with very low wavenumber are expected to be strongly coupled to each other, with a strength that gets weaker the higher is k. In conclusion, the energy initially stored into the first consecutive M_0 modes is expected to flow to higher modes, the cascade slowing down at that critical mode k_c, such that for $k \simeq k_c$ the above described dispersion-nonlinearity balancing holds:

$$\eta_k^3 v_k \approx \eta_k \left[\sum_{j=1}^{k-1} v_{k-j} v_j + 2 \sum_{j=1}^{\mathcal{N}/(\alpha^2\varepsilon_a)^{1/4}-k} v_j^* v_{k+j} \right]. \tag{44}$$

As mentioned above, we identify the numerically measured specific effective number of degrees of freedom f_1 with k_c/N. In order to get an estimate of k_c we rewrite condition (44) as

$$\eta_k^2 \approx \frac{1}{v_k} \left[\sum_{j=1}^{k-1} v_{k-j} v_j + 2 \sum_{j=1}^{\mathcal{N}/(\alpha^2\varepsilon_a)^{1/4}-k} v_j^* v_{k+j} \right] \tag{45}$$

and, once more as a rough first approximation, we evaluate the right hand side of (45) on the initial datum (41). To do this, we need to distinguish between the two cases of random and of coherent initial phases, and as we shall see, in both cases we shall find indications, which are fully consistent with the numerical results illustrated in Sect. 2.

3.4.1 Coherent Phases

If the phases φ_k^0 in the initial datum (41) are coherently linked to each other (e.g. they are all equal), then the square bracket on the right hand side of (45), evaluated on the initial datum (41), is of the order one: indeed the sum receives about M_0 contributions, each of order $1/M_0$. The factor in front of the square bracket is instead of order $\sqrt{M_0}$, which yields

$$\eta_{k_c} \approx M_0^{1/4}. \tag{46}$$

Taking into account the definition (39) of η_k, the latter estimate yields

$$f_1 \approx \frac{k_c}{N} \approx (M_0 \alpha^2 \varepsilon_a)^{1/4}. \tag{47}$$

Now, if one initially excites a number of modes proportional to N, i.e. if $M_0 = N f_0$ with a small but finite f_0, then the estimate (47) implies that the true control parameter of the system is the total energy $E \simeq N\varepsilon_a$, and one has $f_1 \sim E^{1/4}$. On the other hand, if one initially excites a fixed number M_0 of modes (as in the case of the original FPU experiment), the same estimate holds but tells in this case that the control parameter is the specific energy ε_a and $f_1 \sim \varepsilon_a^{1/4}$. So for coherent phases, for what concerns the behavior of f_1, the estimate (47) is in agreement with numerical results, both in the case of fixed M_0 and of fixed $f_0 = M_0/N$.

Concerning the time-scale t_c over which the energy sharing is expected to reach the mode k_c, we simply observe that from the rescaled equations (40), if $\eta_k = \eta_{k_c} \approx M_0^{1/4}$, then $T_c \sim 1/\eta_{k_c}^3 \approx M_0^{-3/4}$; thus, by the rescaling of time performed in (38) one gets

$$t_c \approx 1/f_1^3 \approx (M_0 \alpha^2 \varepsilon_a)^{-3/4}. \tag{48}$$

Such a time-scale yields only a lower bound to the numerically measured time-scale τ_1; see however the remark at the end of the introduction.

3.4.2 Random Phases

Suppose now that the phases φ_k^0 appearing in (41) are independent random variables, each being uniformly distributed over the interval $[0, 2\pi]$. Then the square bracket on the right hand side of (45) has zero expectation value and its order of magnitude is given by its root mean square, which is proportional to $1/\sqrt{M_0}$. Since the factor in front of it is $1/v_k \approx 1/\sqrt{M_0}$ as well, condition (45), in the case of random phases, implies that η_{k_c} is independent of M_0, and this in turn implies that

$$f_1 \approx \frac{k_c}{N} \approx (\alpha^2 \varepsilon_a)^{1/4}, \tag{49}$$

independent of M_0. For what concerns the time scale t_c, a reasoning identical to the one performed above leads to the estimate

$$t_c \approx f_1^{-3} \approx (\alpha^2 \varepsilon_a)^{-3/4}, \tag{50}$$

which constitutes again a lower bound to τ_1.

We stress that in the case of random phases, estimate (49) predicts that the specific number of degrees of freedom sharing the energy in the system does not depend on the number of the initially excited modes, and is a function of the specific energy only. Thus, the case of random phases is the only one where the FPU scenario is maintained in the thermodynamic limit.

4 Concluding Remarks

The FPU problem has attracted the attention of many researchers over more than half a century, thus generating a fallout of questions, results, debates that contributed significantly to the progress of the physics of nonlinear systems. The very inprint on modern science of the

FPU problem is the idea that nonlinearity should reconcile dynamics with thermodynamics, by identifying the main mechanisms leading to equilibrium in large systems. We cannot list here the huge amount of achievements induced by this brilliant hypothesis, whose great success is paradoxically a consequence of its early failure. Here, we think we made clear under which conditions the FPU scenario, i.e. the presence of a long-life metastable state, can be considered to be a genuine thermodynamic limit property. This depends in an essential way on the choice of the phases in the initial conditions: a random choice is needed to maintain the FPU scenario in the limit $N \to \infty$, while coherent phases wipe it out. We have also provided theoretical arguments and estimates, which explain the results obtained by direct numerical simulations of the model.

This notwithstanding, the question of the physical consequences of these results remains open. The main reason is that it is not clear how one could produce the ideal long-wavelength "packet-like" initial conditions with different choices of the phases in a real chain of oscillators (e.g., atoms). Actually, selective long-wavelength acoustic excitations are quite difficult to be produced in any experimental set-up, *a fortiori* if one would like to impose random initial phases. In fact, a ballistic excitation in the form of a short energy pulse would correspond to constant initial phases, while a suitable time modulation of the initial pulse should be imposed in order to reproduce the conditions for random initial phases.

We want to conclude by mentioning shortly some interesting future perspectives. In fact, we plan to extend the numerical studies and the analytic approach presented in this paper to the purely quartic FPU model and also to initial conditions made of initially excited packets of short-wavelength Fourier modes. We expect that these studies should unveil further "messages" which still lay incripted in the FPU model.

Acknowledgements We wish to thank Giacomo Gradenigo for a helpful discussion. This work has been partially supported by MiUR, PRIN program no. 2005017208.

Appendix: Proof of the Proposition

Consider the KdV equation (17). The zero average, $L = 2N + 2$-periodic, unknown field U admits the Fourier expansion

$$U(x,t) = \frac{1}{\sqrt{L}} \sum_{k \in \mathbb{Z}\setminus\{0\}} \widehat{U}_k(t) e^{2\pi i k x/L}, \tag{A.1}$$

with

$$\widehat{U}_k(t) = \frac{1}{\sqrt{L}} \int_0^L U(x,t) e^{-2\pi i k x/L}\, \mathrm{d}x, \qquad \widehat{U}_k(t) = \widehat{U}^*_{-k}(t). \tag{A.2}$$

One immediately finds

$$U_t = \frac{1}{\sqrt{L}} \sum_k \frac{\mathrm{d}\widehat{U}_k}{\mathrm{d}t}\, e^{2\pi i k x/L}, \qquad U_{xxx} = \frac{1}{\sqrt{L}} \sum_k \left[\left(\frac{2\pi i k}{L} \right)^3 \widehat{U}_k \right] e^{2\pi i k x/L}, \tag{A.3}$$

as well as

$$U U_x = \frac{\partial}{\partial x} \frac{U^2}{2} = \frac{\partial}{\partial x} \frac{1}{2L} \sum_{q,p} \widehat{U}_q \widehat{U}_p e^{2\pi i (q+p) x/L}$$

$$= \frac{1}{\sqrt{L}} \sum_k \left[\frac{2\pi i k}{L} \frac{1}{2\sqrt{L}} \sum_q \widehat{U}_{k-q} \widehat{U}_q \right] e^{2\pi i k x/L}; \tag{A.4}$$

thus, the generic Fourier coefficient $\widehat{U}_k$ evolves according to

$$\frac{\mathrm{d}\widehat{U}_k}{\mathrm{d}t} = -i \left(\frac{2\pi k}{L} \right)^3 \widehat{U}_k + \frac{2\pi k}{L} \frac{i\alpha}{2\sqrt{2L}} \sum_q \widehat{U}_{k-q} \widehat{U}_q. \tag{A.5}$$

Letting k run in $\mathbb{Z}$ yields a system of infinitely many coupled ODEs that is equivalent to the KdV equation (17).

We now work a little on the convolution sum on the right hand side of (A.5), by explicitly separating the terms $\widehat{U}_{k-q}\widehat{U}_q$ with both q and $k-q$ in the range from $-N$ to N. Using repeatedly the latter of (A.2), for any $k > 0$ one gets

$$\begin{aligned}
\sum_{q \in \mathbb{Z}\setminus\{0\}} \widehat{U}_{k-q}\widehat{U}_q &= \sum_{q=1}^{+\infty} \widehat{U}_{k-q}\widehat{U}_q + \sum_{q=-\infty}^{-1} \widehat{U}_{k-q}\widehat{U}_q \\
&= \sum_{q=1}^{k-1} \widehat{U}_{k-q}\widehat{U}_q + \sum_{q=k+1}^{+\infty} \widehat{U}_{-(q-k)}\widehat{U}_q + \sum_{q=1}^{+\infty} \widehat{U}_{k+q}\widehat{U}_{-q} \\
&= \sum_{q=1}^{k-1} \widehat{U}_{k-q}\widehat{U}_q + \sum_{q=1}^{+\infty} \widehat{U}_{k+q}\widehat{U}_q^* + \sum_{q=1}^{+\infty} \widehat{U}_{k+q}\widehat{U}_q^* \\
&= \sum_{q=1}^{k-1} \widehat{U}_{k-q}\widehat{U}_q + 2\sum_{q=1}^{N-k} \widehat{U}_q^*\widehat{U}_{k+q} + \left(2 \sum_{q=N-k+1}^{+\infty} \widehat{U}_q^*\widehat{U}_{k+q} \right).
\end{aligned} \tag{A.6}$$

Now, for any $1 \le k \le N$, in the last row above the terms of the first two sums have indices ranging from 1 to N, while the terms of the sum in brackets display at least one index ranging from $N+1$ on (precisely $k+q \ge N+1$). *The Fourier–Galerkin truncation to the first N modes consists exactly in neglecting the contribution of the latter quantity in the convolution sum above*. So, defining

$$\varrho_k^{(N)} = 2 \sum_{q=N-k+1}^{+\infty} \widehat{U}_q^*\widehat{U}_{k+q}$$

and recalling that

$$L = 2(N+1), \quad \frac{2\pi k}{L} = \frac{\pi k}{N+1} = \xi_k,$$

for any $k = 1, \ldots, N$, (A.5) can be rewritten as

$$\frac{\mathrm{d}\widehat{U}_k}{\mathrm{d}t} = -i\xi_k^3 \widehat{U}_k + \frac{i\alpha\xi_k}{4\sqrt{N+1}} \left(\sum_{j=1}^{k-1} \widehat{U}_{k-j}\widehat{U}_j + 2\sum_{j=1}^{N-k} \widehat{U}_j^*\widehat{U}_{k+j} + \varrho_k^{(N)} \right). \tag{A.7}$$

The latter system of equations, upon neglecting $\varrho_k^{(N)}$ and renaming $\widehat{U}_k(t) = u_k(t)$, exactly coincides with the system of equations (29) of the FPU resonant normal form. This ends the proof of the proposition.

References

1. Fermi, E., Pasta, J., Ulam, S.: Studies of non linear problems, Los-Alamos internal report, document LA-1940 (1955). In: Enrico Fermi Collected Papers, vol. II, pp. 977–988. University of Chicago Press/Accad. Naz. Lincei, Chicago/Roma (1965). (Also reproduced in [3])
2. Berman, G.P., Izrailev, F.M.: The "Fermi–Pasta–Ulam" problem—the first 50 years. Chaos **15**, 015104 (2005)
3. Gallavotti, G. (ed.): The Fermi-Pasta-Ulam Problem: A Status Report. Lect. Notes Phys., vol. 728. Springer, Berlin (2008)
4. Livi, R., Pettini, M., Ruffo, S., Sparpaglione, M., Vulpiani, A.: Equipartition threshold in nonlinear large Hamiltonian systems: the Fermi–Pasta–Ulam model. Phys. Rev. A **31**, 1039–1045 (1985)
5. Goedde, C.G., Lichtenberg, A.J., Lieberman, M.A.: Chaos and the approach to equilibrium in a discrete Sine-Gordon equation. Physica D **59**, 200–225 (1992)
6. De Luca, J., Lichtenberg, A.J., Ruffo, S.: Energy transition and time scales to equipartition in the Fermi–Pasta–Ulam oscillator chain. Phys. Rev. E **51**, 2877–2885 (1995)
7. Fucito, E., Marchesoni, F., Marinari, E., Parisi, G., Peliti, L., Ruffo, S., Vulpiani, A.: Approach to equilibrium in a chain of nonlinear oscillators. J. Phys. **43**, 707–713 (1982)
8. Livi, R., Pettini, M., Ruffo, S., Sparpaglione, M., Vulpiani, A.: Relaxation to different stationary states in the Fermi-Pasta-Ulam model. Phys. Rev. A **28**, 3544–3552 (1983)
9. Pettini, M., Landolfi, M.: Relaxation properties and ergodicity breaking in nonlinear Hamiltonian dynamics. Phys. Rev. A **41**, 768–783 (1990)
10. Berchialla, L., Giorgilli, A., Paleari, S.: Esponentially long times to equipartition in the thermodynamic limit. Phys. Lett. A **321**, 167–172 (2004)
11. Berchialla, L., Galgani, L., Giorgilli, A.: Localization of energy in FPU chains. Discrete Contin. Dyn. Syst. Ser. A **11**, 855–866 (2004)
12. Biello, J.A., Kramer, P.R., L'vov, Y.V.: Stages of energy transfer in the FPU model. Discrete Contin. Dyn. Syst. **2003**(Suppl.), 113–122 (2003). (Special number devoted to the Proceedings of the Fourth International Conference on Dynamical Systems and Differential Equations, 24–27 May 2002, Wilmington, NC)
13. Benettin, G., Carati, A., Galgani, L., Giorgilli, A.: The Fermi–Pasta–Ulam problem and the metastability perspective. In: Gallavotti, G. (ed.) The Fermi-Pasta-Ulam Problem: A Status Report. Lect. Notes Phys., vol. 728, pp. 151–189. Springer, Berlin (2008)
14. Izrailev, F.M., Chirikov, B.V.: Statistical properties of a nonlinear string. Sov. Phys. Dokl. **11**, 30–32 (1966)
15. Kantz, H., Livi, R., Ruffo, S.: Equipartition thresholds in chains of anharmonic oscillators. J. Stat. Phys. **76**, 627–643 (1994)
16. Lichtenberg, A.J., Livi, R., Pettini, M., Ruffo, S.: Dynamics of oscillator chains. In: Gallavotti, G. (ed.) The Fermi-Pasta-Ulam Problem: A Status Report. Lect. Notes Phys., vol. 728, pp. 21–121. Springer, Berlin (2008)
17. Shepelyansky, D.L.: Low-energy chaos in the Fermi-Pasta-Ulam problem. Nonlinearity **10**, 1331–1338 (1997)
18. Ponno, A.: A theory of the energy cascade in FPU models. Preprint (2008)
19. Zabusky, N.J., Kruskal, M.D.: Interaction of "Solitons" in a collisionless plasma and the recurrence of initial states. Phys. Rev. Lett. **15**, 240–243 (1965)
20. Bambusi, D., Ponno, A.: On metastability in FPU. Commun. Math. Phys. **264**, 539–561 (2006)
21. Bambusi, D., Ponno, A.: Resonance, metastability and blow-up in FPU. In: Gallavotti, G. (ed.) The Fermi-Pasta-Ulam Problem: A Status Report. Lect. Notes Phys., vol. 728, pp. 191–205. Springer, Berlin (2008)
22. Ponno, A.: Soliton theory and the Fermi-Pasta-Ulam problem in the thermodynamic limit. Europhys. Lett. **64**, 606–612 (2003)
23. Ponno, A.: The Fermi-Pasta-Ulam problem in the thermodynamic limit: scaling laws of the energy cascade. In: Collet, P., et al. (eds.) Proceedings of the Cargèse Summer School 2003 on Chaotic Dynamics and Transport in Classical and Quantum Systems, pp. 431–440. Kluwer Academic, Dordrecht (2005)
24. Flach, S., Ponno, A.: The Fermi-Pasta-Ulam problem: Periodic orbits, normal forms and resonance overlap criteria. Physica D **237**, 908–917 (2008)
25. Ford, J.: Equipartition of energy for nonlinear systems. J. Math. Phys. **2**, 387–393 (1961)
26. Venakides, S.: The zero dispersion limit of the Korteweg-de Vries equation with periodic initial data. Trans. Am. Math. Soc. **301**, 189–226 (1987)