Some Analytic Results on the FPU Paradox

Dario Bambusi, Andrea Carati, Alberto Maiocchi, and Alberto Maspero

Dedicato a Walter Craig il cui entusiasmo è sempre contagioso.

Abstract We present some analytic results aiming at explaining the lack of thermalization observed by Fermi Pasta and Ulam in their celebrated numerical experiment. In particular we focus on results which persist as the number N of particles tends to infinity. After recalling the FPU experiment and some classical heuristic ideas that have been used for its explanation, we concentrate on more recent rigorous results which are based on the use of (i) canonical perturbation theory and KdV equation, (ii) Toda lattice, (iii) a new approach based on the construction of functions which are adiabatic invariants with large probability in the Gibbs measure.

1 Introduction

In their celebrated paper of the year 1954 Fermi Pasta and Ulam (see [20]) studied the dynamics of a chain of nonlinear oscillators by numerical integration of the equations of motion, with the aim of testing the dynamical foundation of equilibrium statistical mechanics. They looked at the evolution of the normal mode energies and of their time averages. FPU considered initial data with all the energy in the first Fourier mode and observed that, for the initial data and the ranges of time considered, (1) the harmonic energies seem to have a recurrent behaviour, and (2) the time averages of the harmonic energies quickly relax to a distribution which is exponentially decreasing with the wave number (the so called FPU packet of modes). This was quite surprising since the solution was expected to explore a

D. Bambusi (✉) • A. Carati • A. Maspero
Dipartimento di Matematica, Università degli Studi di Milano,
Via Saldini 50, 20133 Milano, Italy
e-mail: dario.bambusi@unimi.it; andrea.carati@unimi.it; alberto.maspero@unimi.it

A. Maiocchi
Laboratoire de Mathématiques, Université de Cergy-Pontoise,
2 avenue Adolphe Chauvin, Cergy-Pontoise, France

235

whole energy surface in phase space in such a way that the normal mode energies would relax to equipartition, according to the prescription of equilibrium statistical mechanics. For this reason the FPU result is sometimes referred to as the FPU paradox.

A qualitatively new result was then obtained by Izrailev and Chirikov in the year 1966 [29] (confirmed by Bocchieri et al. [16]), who discovered that the paradox disappears if initial data with sufficiently high energy are considered. In the same period Zabuski and Kruskal [39] used the KdV equation in order to try to describe analytically the recurrent behaviour observed by FPU.

Thus two kind of analytic problems naturally arise. The first one is to describe the FPU recurrent behaviour, maybe along the lines of Zabuski and Kruskal; the second one is to establish whether the FPU paradox persists in the thermodynamic limit, i.e. the limit in which the number N of particles in the chain tends to infinity while keeping the specific energy E/N fixed, which is the relevant limit for the foundations of statistical mechanics.

The aim of the present paper is to present a short review of the status of the research, focusing only on analytic results and in particular on a couple of results recently obtained by the authors [5, 31].

The paper is organized as follows: in Sect. 2 we present some numerical computations which essentially coincide with those by FPU. We also add a further numerical computation showing the existence of an energy threshold above which the paradox disappears. In Sect. 3 we will discuss some theoretical heuristic ideas which have been used in order to try to explain and to understand the FPU paradox. In particular, in Sect. 3.1 we will discuss the relation between FPU lattice and KdV equation, while in Sect. 3.2 we discuss the use of KAM theory and canonical perturbation theory (and Nekhoroshev's theorem) in the context of FPU dynamics. In Sect. 4 we present some rigorous results that have been obtained in the last ten years on the problem and which give some explanation of the existence of the so called FPU packet of modes. The limitation of all these results is that they apply to a regime in which the specific energy goes to zero as $N \to \infty$. The section is split into three subsection, the first one deals with a result based on the KdV approximation, the second one deals with a result based on multifrequency expansion and the third one deals with a result based on the approximation by Toda lattice. The subsection on Toda lattice contains the best results now available on the FPU packet of modes. In Sect. 5 we will present an averaging theorem for the FPU chain valid in the thermodynamic limit. This last result in particular deals with a slightly different problem, namely the exchange of energy among the different degrees of freedom when one starts with an initial datum belonging to a set of large Gibbs measure. We conclude the paper with a short discussion in Sect. 6.

2   The FPU Paradox

The Hamiltonian of the FPU–chain can be written, in suitably rescaled variables, as

$$H_{FPU} = H_0 + H_1 + H_2 \tag{1}$$

where

$$H_0 \stackrel{\text{def}}{=} \sum_j \left(\frac{p_j^2}{2} + \frac{(q_{j+1} - q_j)^2}{2} \right) ,$$

$$H_1 \stackrel{\text{def}}{=} \frac{1}{3!} \sum_j (q_{j+1} - q_j)^3$$

$$H_2 \stackrel{\text{def}}{=} \frac{A}{4!} \sum_j (q_{j+1} - q_j)^4 ,$$

where (p, q) are canonically conjugated variables. We will consider either periodic boundary conditions or Dirichlet boundary conditions: the index j runs from 0 to N in the case of Dirichlet boundary conditions, namely $q_0 = q_{N+1} = 0 = p_0 = p_{N+1}$, while it runs from $-(N + 1)$ to N in the case of periodic boundary conditions, i.e. $q_{-N-1} = q_{N+1}$ and $p_{-N-1} = p_{N+1}$.

In order to introduce the Fourier basis consider the vectors

$$\hat{e}_k(j) = \begin{cases} \frac{\delta_{PD}}{\sqrt{N+1}} \sin\left(\frac{jk\pi}{N+1} \right), & k = 1, \ldots, N, \\ \frac{1}{\sqrt{N+1}} \cos\left(\frac{jk\pi}{N+1} \right), & k = -1, \ldots, -N, \\ \frac{1}{\sqrt{2N+2}}, & k = 0, \\ \frac{(-1)^j}{\sqrt{2N+2}}, & k = -N - 1. \end{cases} \tag{2}$$

Then the Fourier basis is formed by $\hat{e}_k$, $k = 1, \ldots, N$ and $\delta_{PD} = \sqrt{2}$ in the case of Dirichlet boundary conditions, and by $\hat{e}_k$, $k = -N - 1, \ldots, N$ and $\delta_{PD} = 1$ in the case of periodic boundary conditions.

Introducing the Fourier variables $(\hat{p}_k, \hat{q}_k)$ by

$$p_j = \sum_k \hat{p}_k \hat{e}_k(j) , \qquad q_j = \sum_k \hat{q}_k \hat{e}_k(j) \tag{3}$$

with

$$\omega_k = 2 \sin\left(\frac{|k|\pi}{2(N + 1)} \right) , \tag{4}$$

the system takes the form

$$H = H_0 + H_1 + H_2 \tag{5}$$

Dario.Bambusi@unimi.it

where

$$H_0(\hat{p}, \hat{q}) = \sum_k \frac{\hat{p}_k^2 + \omega_k^2 \hat{q}_k^2}{2} \,, \quad H_1 = H_1(\hat{q}) \,, \quad H_2 = H_2(\hat{q}) \,. \qquad (6)$$

We also introduce the harmonic energies

$$E_k = \frac{\hat{p}_k^2 + \omega_k^2 \hat{q}_k^2}{2},$$

and their time averages

$$\overline{E_k}(T) := \frac{1}{T} \int_0^T E_k(t) dt \,. \qquad (7)$$

We will often use also the *specific harmonic energies* defined by

$$e_k := \frac{E_k}{N} \,. \qquad (8)$$

We recall that according to the principles of classical statistical mechanics, at equilibrium at temperature T, each of the harmonic oscillators should have an energy equal to β^{-1}, where $\beta = (k_B T)^{-1}$ is the standard parameter entering in the Gibbs measure (and k_B is the Boltzmann constant). Furthermore, if the system has good statistical properties, the time averages of the different quantities should quickly relax to their equilibrium value.

Fermi Pasta and Ulam studied the time evolution of E_k and of the corresponding time averages $\overline{E_k}$ under Dirichlet boundary conditions. Figure 1 shows the results of the numerical computations by FPU[1]; the initial data are chosen with $E_1(0) \neq 0$ and $E_k(0) = 0$ for any $|k| > 1$.

From Fig. 1 one sees that the energy flows quickly to some modes of low frequency, but after a short period it returns almost completely to the first mode; in the right part of the figure the final values of $\overline{E_k}(t)$ are plotted in a linear scale. The final distribution turns out to be exponentially decreasing with k.

If one continues the integration one sees that the return phenomenon repeats itself almost identically for a very long time (see Fig. 2). The distribution of the $\overline{E_k}(t)$, too, is almost unchanged: one usually says that a packet of modes has formed.

In Fig. 3 the time averages $\overline{E_k}(t)$ are plotted versus time in a semi-log scale. One sees that the quantities $\overline{E_k}(t)$ quickly relax to well defined values which, as shown in Fig. 2 (right figure), decay exponentially with the wave number. To describe the situation with the words by Fermi Pasta and Ulam "The result shows very little, if

---

[1]Actually these figures where obtained more recently by Antonio Giorgilli by repeating the computations of FPU.
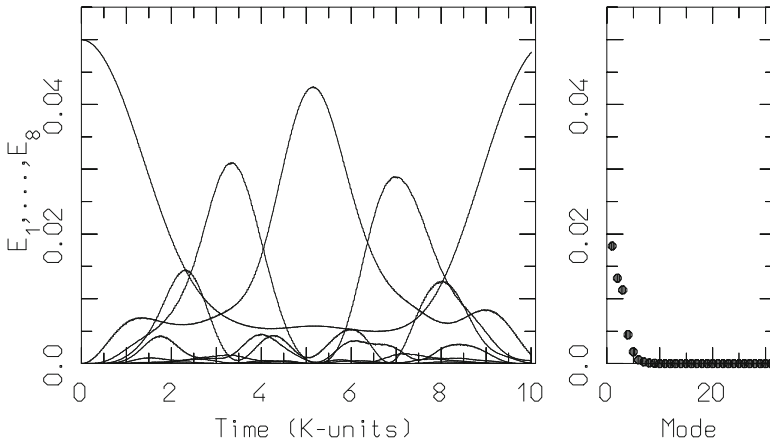
Fig. 1 Mode energies vs time (*left*) and final values of their time averages vs mode number k (*right*)
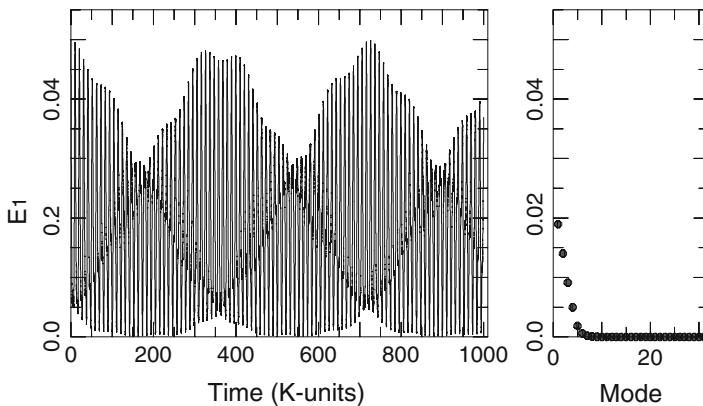

Fig. 2 Energy of the first mode and final values of $\overline{E_k}(t)$ at longer time scales

any, tendency towards equipartition of energy among the degrees of freedom." This is what is usually known as the Fermi Pasta Ulam paradox.

All the above results correspond to initial data with small energy. It was however discovered by Izrailev and Chirikov [29] that the results qualitatively change when the energy per particle is increased. This is illustrated in Fig. 4 from which one sees that the FPU paradox disappears in this regime, because equipartition is quickly attained.

The FPU numerical experiment originated a huge amount of scientific research and in particular subsequent numerical computations have established the shape of the packet of modes to which energy flows (see e.g. [13]) into FPU regime and have put into evidence that the FPU packet is only metastable [21], namely that after a quite long time, whose precise length is not yet precisely established, the system relaxes to equipartition (see e.g. [12, 14]).
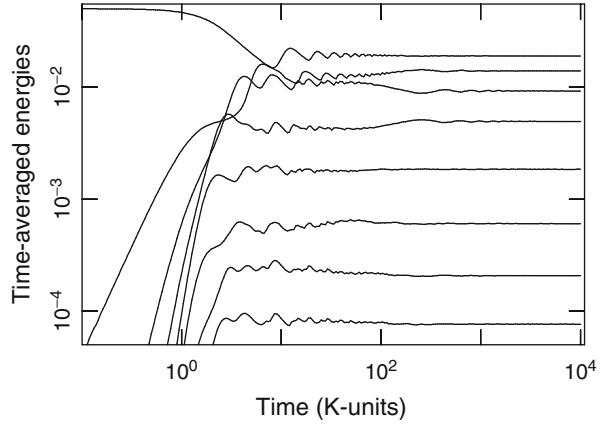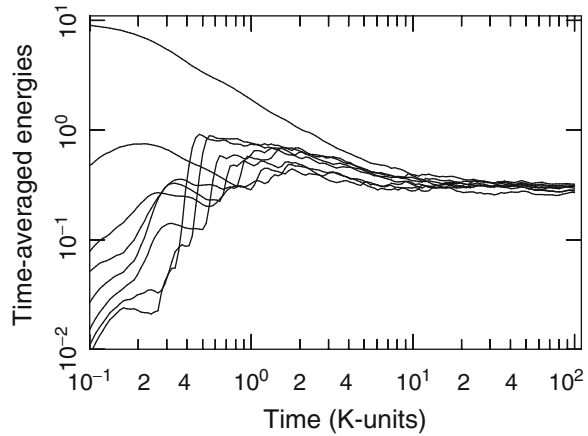
Fig. 3 $\overline{E_k}(t)$ versus time

Fig. 4 $\overline{E_k}(t)$ versus time at large energy

3 Heuristic Theoretical Analysis

We remark that the theoretical understanding of the FPU paradox would be absolutely fundamental: indeed it is clear that the phenomenon would have a strong relevance for the foundations of statistical mechanics if it were proven to persist in the thermodynamic limit, i.e., in the limit in which the number of particles $N \to \infty$ while the energy per mode, namely $\sum_k E_k/N$, is kept fixed. Of course numerics can just give some indications, while a definitive result can only come from a theoretical result, which is the only one able to attain the limit $N = \infty$.

3.1   KdV

One of the first attempts to explain the FPU paradox was based on the use of the Korteweg de Vries equation (KdV). The point is that on the one hand KdV is known to approximate the FPU and on the other one KdV is also known to be integrable, so that it displays a recurrent behaviour.

We now recall briefly the way KdV is introduced as a modulation equation for the FPU. We consider the case of periodic boundary conditions and confine the study to the subspace

$$\sum_j q_j = 0 = \sum_j p_j \tag{9}$$

which is invariant under the dynamics. The idea is to consider initial data with large wavelength and small amplitude, namely to interpolate the difference $q_j - q_{j+1}$ through a smooth small function slowly changing in space (and time). This is obtained through an Ansatz of the form

$$q_j - q_{j+1} = \epsilon u(\mu j, t), \quad \mu := \frac{1}{N}, \quad \epsilon \ll 1 \tag{10}$$

with u periodic of period 2. It turns out that in order to fulfill the FPU equations, the function u should have the form

$$u(x, t) = f(x - t, \mu^3 t) + g(x + t, \mu^3 t)$$

with $f(y, \tau)$ and $g(y, \tau)$ fulfilling the equations

$$f_\tau + \frac{\mu^2}{\epsilon} f_{yyy} + f f_y = O(\mu^2), \quad g_\tau - \frac{\mu^2}{\epsilon} g_{yyy} - g g_y = O(\mu^2), \tag{11}$$

namely, up to higher order corrections, the system is described by a couple of KdV equations with dispersion of order μ^2/ϵ. The origin of this group of ideas is the celebrated paper by Zabuski and Kruskal [39] on the dynamics of the KdV equation, which was the starting point of soliton theory and led in particular to the understanding that KdV is integrable. Thus, the enthusiasm for the discovery of such a beautiful and important phenomenology, led to the idea that also the FPU paradox might be an integrable phenomenon, or more precisely could be the shadow of the fact that KdV is an integrable system nearby FPU.

In order to transform such a heuristic idea into a theorem one should fill two gaps. The first one consists in showing that in the KdV equation a phenomenon of the kind of the formation and persistence of the packet of modes occurs; the second one consisting in showing that the solutions of the KdV equation actually describe well the dynamics of the FPU, namely that the higher order corrections neglected in (11), are actually small.

As discussed below, both problems can be solved in the case $\epsilon = \mu^2$, in which the KdV equation turns out to be the standard one.

In particular, in this case one can exploit some analytic properties of action angle variables for KdV (see [30]) in order to show that if one puts all the energy in the first Fourier mode, then the energy remains forever localized in an exponentially localized packet of Fourier modes. However, if one wants to take the limit $N \to \infty$ while keeping ϵ fixed (as needed in order to get a result valid in the thermodynamic limit), one has to study the dispersionless limit of the KdV equation and very little is known on the behaviour of action angle variables in such a limit. Thus we can say that, **in the KdV equation**, the phenomenon of formation and persistence of the packet is not explained in the limit which corresponds to the thermodynamic limit of the FPU lattice.

The second problem (justification of KdV as an approximation of FPU) is far from trivial, since FPU is a singular perturbation of KdV, namely the $O(\mu^2)$ terms in (11) contain higher order derivatives: the proof of theorems connecting the solutions of KdV and the solutions of FPU have only recently been obtained [6, 36], and only in the case $\epsilon = \mu^2$.

3.2   KAM Theory and Canonical Perturbation Theory

Izrailev and Chirikov [29] in 1966 suggested an explanation of the FPU paradox through KAM theory. We recall that KAM theory deals with perturbations of integrable systems and ensures that, provided the perturbation is small enough, most of the invariant tori in which the phase space of the unperturbed system is foliated persist in the complete system. In the case of FPU the simplest integrable system is the linearized chain for which the perturbation is provided by the nonlinearity. So the size of the nonlinearity increases with the energy of the initial datum and KAM theory should apply for energy smaller than some N-dependent threshold ϵ_N. This approach has the advantage of potentially explaining the FPU paradox *and also* to predict that it should disappear for energy larger then some threshold (as actually observed numerically). From the argument of Izrailev and Chirikov (based on Chirikov's criterion of overlapping of resonances) one can extract also an explicit estimate of ϵ_N which should go to zero like $N^{-4} \equiv \mu^4$. Such an estimate is derived by Izrailev and Chirikov by considering initial data on high frequency Fourier modes, while they do not deduce any explicit estimate for the case of initial data on low frequency modes. Their argument was extended to initial data on low frequency Fourier modes by Shepeliansky [37] leading to the claim that also in correspondence to such a kind of initial data the FPU phenomenon should disappear as $N \to \infty$. However a subsequent reanalysis of the problem led Ponno [33] to different conclusions, so, we can at least say that the situation is not yet clear.

We emphasize that the actual application of KAM theory to the FPU lattice is quite delicate since the hypotheses of KAM theory involve a Diophantine type nonresonance condition and also a nondegeneracy condition. The two conditions have been verified only much later by Rink [35] (see also [26, 32]). Then one has to estimate the dependence of the threshold ϵ_N on N and it turns out that a rough estimate gives that ϵ_N goes to zero exponentially with N (essentially due to the Diophantine type nonresonance condition). This is the main reason which led Izrailev and Chirikov to conjecture that the FPU paradox disappears in the thermodynamic limit.

In order to weaken this condition on ϵ_N, Benettin, Galgani, Giorgilli and collaborators [1, 8–11, 22] started to investigate the possibility of using averaging theory and Nekhoroshev's theorem to explain the FPU paradox. This a quite remarkable change of point of view, since at variance with KAM theory averaging theory and Nekhoroshev's theorem give results controlling the dynamics over long, but finite times, so that such a point of view leaves open the possibility that the FPU paradox disappears after a finite but long time, which is what is actually observed in numerical investigations (see also the remarkable theoretical paper [21]). Results along this line have been obtained for chains of rotators ([1, 9]) and FPU chains with alternate masses [1, 22]. An application to the true FPU model is given in the next section.

4 Some Rigorous Results at Vanishing Specific Energy

4.1 KdV and FPU

The unification of the two points of view illustrated above was obtained in the paper [6]. In that paper, first of all canonical perturbation theory is used in order to deduce a couple of KdV equations as resonant normal form for the FPU lattice and, secondly, the KdV equations are used in order to describe the phenomenon of formation and metastability of the FPU packet. We briefly recall the result of [6].

We consider here the case of periodic boundary conditions. Consider a state of the form (10) and write the equation for the evolution of the function u. Then it turns out that such an equation is a Hamiltonian perturbation of the wave equation, so one can use canonical perturbation theory for PDEs in order to simplify the equation. Passing to the variables f, g the normal form turns out to be the Hamiltonian of a couple of non interacting KdV equations. In [6] a rigorous theory estimating the error was developed, and the main results of that paper are contained in Theorem 1 and Corollary 1 below.

Consider the KdV equation

$$f_\tau + f_{yyy} + ff_y = 0 \; ;$$

it is well known [30] that if the initial datum extends to a function analytic in a complex strip of width σ, then the solution (as a function of the space variable y) is also analytic (in general in a smaller complex strip).

Consider now a couple of solutions f, g of KdV with analytic initial data and let $q_j^{KdV}(t)$ be the unique sequence such that

$$q_j^{KdV}(t) - q_{j+1}^{KdV}(t) = \mu^2 \left[f \left(\mu(j - t), \mu^3 t \right) + g \left(-\mu(j + t), \mu^3 t \right) \right],$$

$$\sum_j q_j^{KdV}(t) \equiv 0, \tag{12}$$

where, as above, $\mu := N^{-1}$. Then the result of Theorem 1 below is that q_j^{KdV} approximates well the true solution of the FPU lattice.

Let $q_j(t)$ be the solution of the FPU equations with initial data $q_j(0) = q_j^{KdV}(0)$, $\dot{q}_j(0) = \dot{q}_j^{KdV}(0)$; denote by $E_k(t)$ the energy in the kth Fourier mode of the solution of the FPU with such initial data and $e_k := E_k/N$.

The following theorem holds

Theorem 1 ([6]). *Fix an arbitrary $T_f > 0$. Then there exists μ_* such that, if $\mu < \mu_*$ then for all times t fulfilling*

$$|t| \leq \frac{T_f}{\mu^3} \tag{13}$$

one has

$$\sup_j \left| r_j(t) - r_j^{KdV}(t) \right| \leq C\mu^3, \tag{14}$$

where $r_j := q_j - q_{j+1}$ and similarly for r_j^{KdV}. Furthermore, there exists $\sigma > 0$ s.t., for the same times, one has

$$e_k(t) \leq C\mu^4 e^{-\sigma|k|} + C\mu^5. \tag{15}$$

Exploiting known results on the dynamics of KdV (and of Hill's operators [34]) one gets the following corollary which is directly relevant to the FPU paradox.

Corollary 1. *Fix a positive R and a positive T_f; then there exists a positive constant μ_*, with the following property: assume $\mu < \mu_*$ and consider the FPU system with an initial datum fulfilling*

$$e_1(0) = e_{-1}(0) = R^2\mu^4, \quad e_k(0) \equiv e_k(t)\big|_{t=0} = 0, \quad \forall |k| \neq 1. \tag{16}$$

Then, along the corresponding solution, Eq. (15) holds for the times (13).

Furthermore there exists a sequence of almost periodic functions $\{F_k(t)\}$ such that one has

$$\left| e_k(t) - \mu^4 F_k(t) \right| \leq C_2 \mu^5 , \quad |t| \leq \frac{T_f}{\mu^3} . \tag{17}$$

Remark 1. One can show that the following limit exists

$$\bar{F}_k := \lim_{T \to \infty} \frac{1}{T} \int_0^T F_k(t) dt . \tag{18}$$

It follows that up to a small error the time average of $e_k(t)$ relaxes to the limit distribution obtained by rescaling $\bar{F}_k$. Of course $\bar{F}_k$ is exponentially decreasing with k, but one can also show that actually one has $\bar{F}_k \neq 0 \; \forall k \neq 0$

The strong limitation of the above results rests in the fact that they only apply to initial data with specific energy of order μ^4, thus they do not apply to the thermodynamic limit.

4.2 Longer Time Scales at Smaller Energy

We present here a result by Hairer and Lubich [24] which is valid in a regime of specific energy smaller then that considered above, but controls the dynamics for longer time scales. The proof of the result is based on the technique of modulated Fourier expansion developed by the authors and their collaborators. In some sense such a technique can be considered as a variant of classical perturbation theory. The key tool that they use for the proof is an accurate analysis of the small denominators entering in the perturbative construction.

To be precise [24] deals with the case of periodic boundary conditions.

Theorem 2. *There exist positive constants R_*, N_*, T, with the following property: consider the FPU system with an initial datum fulfilling* (16) **with $\mathbf{R} < \mathbf{R}_*$**. *Then, along the corresponding solution, one has*

$$e_k(t) \leq R^2 \mu^4 R^{2(|k|-1)} , \quad \forall \, 1 \leq |k| \leq N , \quad \forall |t| \leq \frac{T}{\mu^2 R^5} , \tag{19}$$

where as above $e_k := E_k/N$.

It is interesting to compare the time scale covered by this theorem with the time scale of Corollary 1. It is clear that the time scale (19) is longer than (13) as far as

$$R < N^{-1/5} \tag{20}$$

(where we made the choice $T_f := T$), namely in a regime where the specific energy goes to zero faster than in Theorem 1.

One has also to remark that in Theorem 2 one gets an exponential decay of the Fourier modes valid for all k's (the term of order μ^5 present in (15) is here absent).

4.3 Toda Lattice

It is well known that close to the FPU lattice there exists a remarkable integrable system, namely the Toda lattice [25, 38] whose Hamiltonian is given by

$$H_{Toda}(p, q) = \frac{1}{2} \sum_j p_j^2 + \sum_j e^{q_j - q_{j+1}}, \tag{21}$$

so that one has

$$H_{FPU}(p, q) = H_{Toda}(p, q) + (A - 1)H_2(q) + H^{(3)}(q),$$

where

$$H_l(q) := \sum_j \frac{(q_j - q_{j+1})^{l+2}}{(l+2)!}, \quad \forall l \geq 2,$$

$$H^{(3)} := -\sum_{l \geq 3} H_l,$$

which shows the vicinity of H_{FPU} and H_{Toda}.

The idea of exploiting the Toda lattice in order to deduce information on the dynamics of the FPU chain is an old one; however in order to make it effective, one has first to deduce information on the dynamics of the Toda lattice itself, and this is far from trivial. The most obvious way to proceed consists in constructing action angle coordinates for the Toda lattice and using them to study the dynamics of the Toda lattice itself. An important result along this program was obtained by Henrici and Kappeler [26, 27] who constructed action angle coordinates and Birkhoff coordinates (a kind of cartesian action angle coordinates) showing that, for any N, such coordinates are globally analytic (see Theorem 3 below for a precise statement). However the construction by Henrici and Kappeler is not uniform in the number N of particles, thus it is not possible to exploit it directly in order to get results in the limit $N \to \infty$.

Results on the behaviour of the integrable structure of Toda for large N have been recently obtained in a series of papers [2–5, 15]. In particular in [2–4], exploiting ideas from [15], it has been shown that, as $N \to \infty$, the actions and the frequencies of the Toda lattice are well described by the actions and the frequencies of a couple of KdV equations, at least in a regime equal to that of Theorem 1, namely of specific energy of order μ^4.

Further results (exploiting some ideas from [2–4]) directly applicable to the FPU metastability problem have been obtained in [5] and now we are going to present them. In [5] the regularity properties of the Birkhoff map, namely the map introducing Birkhoff coordinates for the FPU lattice, have been studied and lower and upper bounds to the radius of the ball over which such a map is analytic have been given.

To come to a precise statement we start by recalling the result by Henrici and Kappeler.

Consider the Toda lattice in the submanifold (9) and introduce the linear Birkhoff variables

$$X_k = \frac{\hat{p}_k}{\sqrt{\omega_k}} , \quad Y_k = \sqrt{\omega_k}\hat{q}_k , \quad |k| = 1, \ldots, N ; \tag{22}$$

using such coordinates, H_0 takes the form

$$H_0 = \sum_{|k|=1}^{N} \omega_k \frac{X_k^2 + Y_k^2}{2} . \tag{23}$$

With an abuse of notations, we re-denote by H_{Toda} the Hamiltonian (21) written in the coordinates (X, Y).

Theorem 3 ([28]). *For any integer $N \geq 2$ there exists a global real analytic canonical diffeomorphism $\Phi_N : \mathbb{R}^{2N} \times \mathbb{R}^{2N} \to \mathbb{R}^{2N} \times \mathbb{R}^{2N}$, $(X, Y) = \Phi_N(x, y)$ with the following properties:*

(i) The Hamiltonian $H_{Toda} \circ \Phi_N$ is a function of the actions $I_k := \frac{x_k^2 + y_k^2}{2}$ only, i.e. (x_k, y_k) are Birkhoff variables for the Toda Lattice.
(ii) The differential at the origin is the identity: $d\Phi_N(0,0) = \mathbb{1}$.

In order to state the analyticity properties fulfilled by the map Φ_N as $N \to \infty$ we need to introduce suitable norms: for any $\sigma \geq 0$ define

$$\|(X, Y)\|_\sigma^2 := \frac{1}{N} \sum_k e^{2\sigma|k|} \omega_k \frac{|X_k|^2 + |Y_k|^2}{2}. \tag{24}$$

We denote by $B^\sigma(R)$ the ball in $\mathbb{C}^{2N} \times \mathbb{C}^{2N}$ of radius R and center 0 in the topology defined by the norm $\|.\|_\sigma$. We will also denote by $B_\mathbb{R}^\sigma := B^\sigma(R) \cap (\mathbb{R}^{2N} \times \mathbb{R}^{2N})$ the *real* ball of radius R.

Remark 2. We are particularly interested in the case $\sigma > 0$ since, in such a case, states with finite norm are exponentially decreasing in Fourier space.

The main result of [5] is the following Theorem.

Theorem 4 ([5]). *Fix $\sigma \geq 0$ then there exist $R, R' > 0$ s.t. Φ_N is analytic on $B^\sigma\left(\frac{R}{N^\alpha}\right)$ and fulfills*

$$\Phi_N\left(B^\sigma\left(\frac{R}{N^\alpha}\right)\right) \subset B^\sigma\left(\frac{R'}{N^\alpha}\right), \quad \forall N \geq 2 \tag{25}$$

if and only if $\alpha \geq 2$. The same is true for the inverse map Φ_N^{-1}.

Remark 3. A state (X, Y) is in the ball $B^\sigma(R/N^2)$ if and only if there exist interpolating periodic functions (β, α), namely functions s.t.

$$p_j = \beta\left(\frac{j}{N}\right), \quad q_j - q_{j+1} = \alpha\left(\frac{j}{N}\right), \tag{26}$$

which are analytic in a strip of width σ and have an analytic norm of size R/N^2. Thus we are in the same regime to which Theorem 1 apply.

Theorem 4 shows that the Birkhoff coordinates are analytic only in a ball of radius of order N^{-2}, which corresponds to initial data with specific energy of order N^{-4}.

We think this is a strong indication of the fact that standard integrable techniques cannot be used beyond such a regime.

As a corollary of Theorem 4, one immediately gets that in the Toda Lattice the analogous of the FPU metastable packet of modes is actually stable, namely it persists for infinite times. Precisely one has the following result.

Corollary 2. *Consider the Toda lattice (21). Fix $\sigma > 0$, then there exist constants R_0, C_1, such that the following holds true. Consider an initial datum fulfilling (16) with $\mathbf{R} < \mathbf{R_0}$. Then, along the corresponding solution, one has*

$$e_k(t) \leq R^2(1 + C_1 R)\mu^4 e^{-2\sigma|k|}, \quad \forall 1 \leq |k| \leq N, \quad \forall t \in \mathbb{R}. \tag{27}$$

We recall that this was observed numerically by Benettin and Ponno [7, 12]. One has to remark that according to the numerical computations of [12], the packet exists and is stable over infinite times also in a regime of finite specific energy (which would correspond to the case $\alpha = 0$ in Theorem 4). The understanding of this behaviour in such a regime is still a completely open problem.

Concerning the FPU chain, Theorem 4 yields the following result.

Theorem 5. *Consider the FPU system. Fix $\sigma \geq 0$; then there exist constants R_0', C_2, T, such that the following holds true. Consider a real initial datum fulfilling (16) with $\mathbf{R} < \mathbf{R_0'}$, then, along the corresponding solution, one has*

$$e_k(t) \leq 16R^2\mu^4 e^{-2\sigma|k|}, \quad \forall 1 \leq |k| \leq N, \quad |t| \leq \frac{T}{R^2\mu^4} \cdot \frac{1}{|A-1| + C_2 R\mu^2}. \tag{28}$$

Furthermore, for $1 \leq |k| \leq N$, consider the action $I_k := \frac{x_k^2 + y_k^2}{2}$ of the Toda lattice and let $I_k(t)$ be its evolution according to the FPU flow. Then one has

$$\frac{1}{N} \sum_{|k|=1}^{N} e^{2\sigma|k|} \omega_k |I_k(t) - I_k(0)| \leq C_3 R^2 \mu^5 \qquad \text{for } t \text{ fulfilling (28)}. \qquad (29)$$

So this theorem gives a result which covers times one order of magnitude longer then those covered by Theorem 1. Furthermore the small parameter controlling the time scale is the distance between the FPU and the Toda

This is particularly relevant in view of the fact that, according to Theorem 1 the time scale of formation of the packet is μ^{-3}, thus the present theorem shows that the packet persists at least over a time scale one order of magnitude longer then the time needed for its formation.

5 An Averaging Theorem in the Thermodynamic Limit

In this section we discuss a different approach to the study of the dynamics of the FPU system, which allows to give some results valid in the thermodynamic limit. Such a method is a development of the one introduced in [17] in order to deal with a chain of rotators (see also [19]), and developed in [18] in order to study a Klein Gordon chain.

We consider here the case of Dirichlet boundary conditions and endow the phase space with the Gibbs measure μ_β at inverse temperature β, namely

$$d\mu_\beta(p, q) \overset{\text{def}}{=} \frac{e^{-\beta H_{FPU}(p,q)}}{Z(\beta)} dp dq \; ; \qquad (30)$$

where as usual

$$Z(\beta) := \int e^{-\beta H_{FPU}(p,q)} dp dq$$

is the partition function (the integral is over the whole phase space). In the following we will omit the index β from μ. Given a function F on the phase space, we define

$$\langle F \rangle \overset{\text{def}}{=} \int F d\mu \; , \qquad (31)$$

$$\|F\|^2 \overset{\text{def}}{=} \langle F^2 \rangle \equiv \int |F|^2 d\mu \; , \qquad (32)$$

$$\sigma_F^2 \overset{\text{def}}{=} \|F - \langle F \rangle\|^2 \; , \qquad (33)$$

which are called respectively the average, the L^2 norm and the variance of F. The time autocorrelation function C_F of a dynamical variable F is defined by

$$C_F(t) := \langle F(t)F \rangle - \langle F^2 \rangle , \tag{34}$$

where $F(t) := F \circ G^t$ and G^t is the flow of the FPU system.

Remark that the Gibbs measure is asymptotically concentrated on the energy surface of energy N/β. Thus, when studying the system in the above setting one is typically considering data with specific energy equal to β^{-1}.

Let $g \in \mathcal{C}^2([0,1], \mathbb{R}^+)$ be a twice differentiable function; we are interested in the time evolution of quantities of the form

$$\Phi_g \stackrel{\text{def}}{=} \sum_{k=1}^{N} E_k \, g\left(\frac{k}{N+1}\right) .$$

We are thinking of a function g with a small support close to a fixed wave vector $\bar{k}/(N+1)$, so that the quantity Φ_g represents the energy of a packet of modes centered at $\bar{k}/(N+1)$.

The following theorem was proved in [31]

Theorem 6. *Let $g \in \mathcal{C}^2([0,1]; \mathbb{R}^+)$ be a function fulfilling $g'(0) = 0$. There exist constants $\beta^* > 0$, $N^* > 0$ and $C > 0$ s.t., for any $\beta > \beta^*$ and for any $N > N^*$, any $\delta_1, \delta_2 > 0$ one has*

$$\mu\left(\left|\Phi_g(t) - \Phi_g(0)\right| \geq \delta_1 \sigma_{\Phi_g}\right) \leq \delta_2 , \quad |t| \leq \frac{\delta_1 \sqrt{\delta_2}}{C} \beta \tag{35}$$

where, as above, $\Phi_g(t) = \Phi_g \circ G^t$.

This theorem shows that, with large probability, the energy of the packet of modes with profile defined by the function g remains constant over a time scale of order β^{-1}. We also emphasize that the change in the quantity Φ_g is small compared to its variance, which establishes the order of magnitude of the difference between the biggest and the smallest value of Φ_g on the energy surface.

Theorem 6 is actually a corollary of a result controlling the evolution of the time autocorrelation function of Φ_g. We point out that, in some sense the time autocorrelation function is a more important object, at least if one is interested in the problem of the dynamical foundations of thermodynamics. Indeed, it is known by Kubo linear response theory, that the specific heat of system in contact with a thermostat is the time autocorrelation function of its energy. Of course we are here dealing with an isolated system, so the previous theorem is not directly relevant to the problem of foundations of statistical mechanics.

Remark 4. Of course one can repeat the argument for different choices of the function g. For example one can partition the interval $[0, 1]$ of the variable $k/(N+1)$ in K sub-intervals and define K different functions $g^{(1)}, g^{(2)}, \dots, g^{(K)}$, with disjoint

support, each one fulfilling the assumptions of Theorem 6, so that one gets that the quantities $\Phi_{g^{(l)}} \stackrel{\text{def}}{=} \sum_k g^{(l)} \left(\frac{k}{N+1} \right) E_k$ are adiabatic invariants, i.e. the energy essentially does not move from one packet to another one.

The scheme of the proof of Theorem 6 is as follows: first, following standard ideas in perturbation theory (see [23]), one performs a formal construction of an integral of motion as a power series in the phase space variables. As usual, already at the first step one has to solve the so called homological equation in order to find the third order correction of the quadratic integral of motion. The solution of such an equation involves some small denominators which are usually the source of one of the problems arising when one wants to control the behaviour of the system in the thermodynamical limit. We show that, if one takes as the quadratic part of the integral the quantity Φ_g, then every small denominator appears with a numerator which is also small, so that the ratio is bounded. The subsequent step consists in adding rigorous estimates on the variance of the time derivative of the so constructed approximate integral of motion. This allows to conclude the proof.

We emphasize that this procedure completely avoids to impose the time invariance of the domain in which the theory is developed, which is the requirement that usually prevents the applicability of canonical perturbation theory to systems in the thermodynamic limit. Indeed in the probabilistic framework the relevant estimates are global in the phase space.

6 Conclusions

Summarizing the above results, we can say that all analytical results available nowadays can be split into two groups: the first group consisting of those which describe the formation of the packet observed by FPU and give some estimates on its time of persistence. Such results do not survive in the thermodynamic limit; indeed they are all confined to the regime in which the specific energy is order N^{-4}. We find particularly surprising the fact that very different methods lead to the same regime and of course this raises the suspect that there is some reality in this limitation. However one has to say that numerics do not provide any evidence of changes in the dynamics when energy is increased beyond this limit.

A few more comments on this point are the following ones: the limitations appearing in constructing the Birkhoff variables in the Toda lattice (which are the source of the limitations in the applicability of Theorem 5) are related to the fact that one is implicitly looking for an integrable behaviour of the system, namely a behaviour in which the system is essentially decoupled into non interacting oscillators. On the contrary the construction leading to Theorem 1 is based on a resonant perturbative construction in which the small denominators are not present. The main limitation for the applicability of Theorem 1 comes from the need of

considering the zero dispersion limit of the KdV equation. So, it is surprising that the regime at which the two results apply is the same.

So the question on whether the phenomenon of formation of a metastable packet persists in the thermodynamic limit or not is still completely open. An even more open question is that of the length of the time interval over which it persists. Up to now the best result we know is that of Theorem 5, but from the numerical experiments one would expect longer time scales (furthermore in the thermodynamic limit). How to reach them is by now not known.

At present the only known result valid in the thermodynamic limit is that of Theorem 6. However we think that this should be considered only as a preliminary one. Indeed it leaves open many important questions. The first one is the optimality of the time scale of validity: the technique used for its proof does not extend to higher order constructions. This is due to the fact that at order four new kinds of small denominators appear and up to now we were unable to control them. Furthermore there is no numerical evidence of the optimality of the time scale controlled by such a theorem.

An even more important question is the relevance of the result for the foundations of statistical mechanics. Indeed, one expects that the existence of many integrals of motion independent of the energy should have some influence on the measurement of thermodynamic quantities, for example the specific heat. In particular, since the time needed to exchange energy among different packets of modes increases as one lower the temperature, one would expect that some new behaviour appears as the temperature is lowered towards the absolute zero. However up to now we were unable to put into evidence some clear effect of the mathematical phenomenon described by Theorem 6. This is one of the main goals of our group for the next future.

References

1. Bambusi, D., Giorgilli, A.: Exponential stability of states close to resonance in infinite-dimensional Hamiltonian systems. J. Statist. Phys. **71**, 569–606 (1993)
2. Bambusi, D., Kappeler, T., Paul, T.: De Toda à KdV. C. R. Math. Acad. Sci. Paris, **347**, 1025–1030 (2009)
3. Bambusi, D., Kappeler, T., Paul, T.: Dynamics of periodic Toda chains with a large number of particles (2013) (ArXiv e-prints, arXiv:1309.5441 [math.AP])
4. Bambusi, D., Kappeler, T., Paul, T.: From Toda to KdV. ArXiv e-prints, arXiv:1309.5324 [math.AP], Sept. (2013).
5. Bambusi D, Maspero A. Birkhoff coordinates for the Toda lattice in the limit of infinitely many particles with an application to FPU. Preprint (2014)
6. Bambusi, D., Ponno A.: On metastability in FPU. Comm. Math. Phys. **264**, 539–561 (2006)

7. Benettin, G., Christodoulidi, H., Ponno, A.: The Fermi-Pasta-Ulam problem and its underlying integrable dynamics. J. Stat. Phys. **152**, 195–212 (2013)
8. Benettin, G., Galgani, L., Giorgilli, A.: Classical perturbation theory for systems of weakly coupled rotators. Nuovo Cimento B (11) **89**, 89–102 (1985)
9. Benettin, G., Galgani, L., Giorgilli, A.: Numerical investigations on a chain of weakly coupled rotators in the light of classical perturbation theory. Nuovo Cimento B (11), **89**, 103–119 (1985)
10. Benettin, G., Galgani, L., Giorgilli, A.: Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. I. Comm. Math. Phys. **113**, 87–103 (1987)
11. Benettin, G., Galgani, L., Giorgilli, A.: Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. II. Comm. Math. Phys. **121**, 557–601 (1989)
12. Benettin, G., Ponno, A.: Time-scales to equipartition in the Fermi-Pasta-Ulam problem: finite-size effects and thermodynamic limit. J. Stat. Phys. **144**, 793–812 (2011)
13. Berchialla, L., Galgani, L., Giorgilli, A.: Localization of energy in FPU chains. Discrete Contin. Dyn. Syst. **11**, 855–866 (2004)
14. Berchialla, L., Giorgilli, A., Paleari, S.: Exponentially long times to equipartition in the thermodynamic limit. Phys. Lett. A **321**, 167–172 (2004)
15. Bloch, A., Golse, F, Paul T., Uribe, A.: Dispersionless toda and toeplitz operators. Duke Math. J. **117**, 157–196 (2003)
16. Bocchieri, P., Scotti, A., Bearzi, B., Loinger, A.: Anharmonic chains with Lenard–Jones interactions. Phys. Rev. A **2**, 2013–2019 (1970)
17. Carati, A.: An averaging theorem for Hamiltonian dynamical systems in the thermodynamic limit. J. Stat. Phys. **128**, 1057–1077 (2007)
18. Carati, A., Maiocchi, A.M.: Exponentially long stability times for a nonlinear lattice in the thermodynamic limit. Comm. Math. Phys. **314**, 129–161 (2012)
19. De Roeck, W., Huveneers, F.: Asymptotic localization of energy in non-disordered oscillator chains (2013) [arXiv:1305.512]
20. Fermi, E., Pasta, J., Ulam, S.: Studies of nonlinear problems. In Collected works of E. Fermi, vol.2. Chicago University Press, Chicago (1965)
21. Fucito, F., Marchesoni, F., Marinari, E., Parisi, G., Peliti, L., Ruffo, S., Vulpiani, A.: Approach to equilibrium in a chain of nonlinear oscillators. J. de Physique **43**, 707–713 (1982)
22. Galgani, L., Giorgilli, A., Martinoli, A., Vanzini, S.: On the problem of energy equipartition for large systems of the Fermi-Pasta-Ulam type: analytical and numerical estimates. Phys. D **59**, 334–348 (1992)
23. Giorgilli, A., Galgani, L.: Formal integrals for an autonomous Hamiltonian system near an equilibrium point. Celestial Mech. **17**, 267–280 (1978)
24. Hairer, E., Lubich, C.: On the energy distribution in Fermi-Pasta-Ulam lattices. Arch. Ration. Mech. Anal. **205**, 993–1029 (2012)
25. Hénon, M.: Integrals of the Toda lattice. Phys. Rev. B (3), **9**, 1921–1923 (1974)
26. Henrici, A., Kappeler, T.: Birkhoff normal form for the periodic Toda lattice. In: Integrable systems and random matrices, Contemporary Math, vol. 458, pp. 11–29. American Mathematical Society, Providence (2008)
27. Henrici, A., Kappeler, T.: Global action-angle variables for the periodic Toda lattice. Int. Math. Res. Not. IMRN **11**, 52 (2008). (Art. ID rnn031)
28. Henrici, A., Kappeler, T.: Global Birkhoff coordinates for the periodic Toda lattice. Nonlinearity **21**, 2731–2758 (2008)
29. Izrailev, F.M., Chirikov, B.V.: Statistical properties of a nonlinear string. Sov. Phys. Dokl. **11**, 30–32 (1966)
30. Kappeler, T., Pöschel, J.: KdV & KAM. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, vol. 45 (A Series of Modern Surveys in Mathematics) [Results in Mathematics and Related Areas, 3rd Series. A Series of Modern Surveys in Mathematics]. Springer, Berlin (2003)

31. Maiocchi, A.M., Bambusi, D., Carati, A.: An averaging theorem for fpu in the thermodynamic limit. J. Stat. Phys. **155**, 300–322 (2014)
32. Nishida T. A note on an existence of conditionally periodic oscillation in a one-dimensional anharmonic lattice. Mem. Fac. Eng. Kyoto Univ. **33**, 27–34 (1971)
33. Ponno, A.: The Fermi-Pasta-Ulam problem in the thermodynamic limit. In: Chaotic dynamics and transport in classical and quantum systems, NATO Science Series II Mathematics Physics and Chemistry, vol. 182, pp. 431–440. Kluwer Acadamic, Dordrecht (2005)
34. Pöschel, J.: Hill's potentials in weighted Sobolev spaces and their spectral gaps. Math. Ann. **349**, 433–458 (2011)
35. Rink, B.: Symmetry and resonance in periodic FPU chains. Comm. Math. Phys. **218**, 665–685 (2001)
36. Schneider, G., Wayne, C.E.: Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model. In: International Conference on Differential Equations (Berlin, 1999), vols. 1, 2, pp. 390–404. World Science, River Edge (2000)
37. Shepelyansky, D.L.: Low-energy chaos in the Fermi–Pasta–Ulam problem. Nonlinearity **10**, 1331–1338 (1997)
38. Toda, T.: Vibration of a chain with nonlinear interaction. J. Phys. Soc. Jpn. **22**, 431 (1967)
39. 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)