



A Survey on Rigorous Results for the Dynamics of Periodic FPU Chains

Dario Bambusi¹ · Andrea Carati¹ · Alberto Maiocchi²

Received: 27 May 2026 / Accepted: 1 October 2026
© The Author(s) 2026

Abstract

In this paper we review some analytic results on the dynamics of the FPU system. In the first part of the paper, having in mind that the FPU Hamiltonian and the Toda Hamiltonian are close to each other, we present some results on the action angle variables of the Toda system and deduce some stability properties for the dynamics of the FPU system. We first focus on the case of finitely many particles and then we study the limit $N \rightarrow \infty$. We also present some results on the continuous limit of the Toda chain showing that it is well described by a couple of KdV equations. Then we study directly the dynamics of the function interpolating the FPU system and show that the dynamics is Hamiltonian and that the Hamiltonian is very close to a function of the first three Hamiltonians of the KdV hierarchy. In the second part of the paper we present some results valid in the thermodynamic limit, according to which the time auto correlation functions of some suitably constructed observables decay slowly implying lower bounds on the thermalization times of the system.

Keywords Metastability in FPU · Toda lattice · KdV hierarchy · Decay of correlations

1 Introduction

In this paper we review some recent results on the dynamics of the Fermi Pasta Ulam (FPU) model. There is a huge number of papers studying the dynamics of the FPU model mostly with the aim of understanding in a heuristic or rigorous way the result of the computer simulation by Fermi Pasta Ulam with the collaboration of Tsingou [18] according to which

Communicated by Antonio Ponso.

Andrea Carati and Alberto Maiocchi contributed equally to this work.

✉ Andrea Carati
andrea.carati@unimi.it

Dario Bambusi
dario.bambusi@unimi.it

Alberto Maiocchi
alberto.maiocchi@unimib.it

¹ Dipartimento di Matematica, Università di Milano, Via Saldini 50, 20133 Milan, Italy

² Dipartimento di Matematica e Applicazioni - Edificio U5, Università degli Studi di Milano-Bicocca, Via Cozzi 21, 20125 Milan, Italy

a chain of coupled particles attains a long lived meta stable state and only over a different time scale possibly attains thermalization. The computer simulation in [18] was done with 64 particles, but of course the main problem is that of the persistence of the phenomenon in the thermodynamic limit in which the number of particles N tends to infinity and at the same time the total energy of the system diverges proportionally to N .

Up to now there are no rigorous results which persist in the thermodynamic limit and explain the FPU simulations, but some phenomena have been understood mathematically and several dynamical results are known. In the present paper we will present two kinds of results: the first dealing with the dynamics of the FPU at finite small energy, and the second dealing with the problem of the fluctuations and the decay of correlations at equilibrium.

The first group of results is mainly related to the approximation of the FPU system with the Toda system and with the Korteweg de Vries equation (KdV).

We will first work at fixed number of particles N and present some results on the existence of action angle variables for the Toda system. Using such variables one can apply KAM and Nekhoroshev theory considering FPU as a perturbation of Toda. We will give a precise statement of the Nekhoroshev's theorem one deduces for the FPU system.

Then we will discuss the behavior of these results when the number of particles tends to infinity. In particular we will present a result on the behavior of the action angle coordinates for the Toda system when $N \rightarrow \infty$. It turns out that they develop a singularity at a distance of order $N^{-3/2}$ from the origin. We will also deduce some results for the FPU dynamics.

Then we will study the related problem of the continuous interpolation. We will first consider the interpolation of the Toda lattice and present some results showing that the actions and the frequencies of the Toda system converge in the continuum limit to the actions and the frequencies of a couple of KdV equations. These results hold in the regime where the total energy of the system is of order N^{-3} .

Then we will study the continuous limit directly on the FPU system and present some recent results according to which the FPU dynamics is accurately approximate up to order six in a suitable small parameter by a combination of the first three equations of the KdV hierarchy.

Finally we present some non rigorous results on the Fourier spectrum of the FPU system which have been recently obtained exploiting the fact that in some limit the FPU is described by the Burgers equation and some recent results on the deduction of the wave kinetic equation for the β FPU system.

The second group of results contains a set of results on the stability for long times of the orbits of the FPU system and related models in a statistical sense, and is presented in Sect. 6. The focus is set on time auto correlation of a dynamical variable at thermodynamic equilibrium, and the failure of the auto correlation of some physically relevant variables to decay to zero in a given macroscopic time is seen as an indicator that, up to that time, thermalization cannot be attained.

We will discuss the results obtained so far on the slow decay of correlations for the energy of suitable "packets" of modes for the FPU system and for the FPU system with alternating masses, for which the times are much longer if the ratio of the heavy and the light mass is large. Besides these results, valid for long, but finite times, we present a related analytical method which allows to control also the asymptotic decay of correlations, and has been used to study numerically such a decay in the FPU system.

2 Chains of Particles with Nearest-Neighbor Interaction

Consider the Hamiltonian of the FPU chain of N particles with periodic boundary conditions, namely

$$H(p, q) := \sum_{j=1}^N \frac{p_j^2}{2} + \sum_{j=1}^N V(q_{j+1} - q_j), \quad (2.1)$$

where $q_{j+N} = q_j$, $p_{j+N} = p_j$ and V is a potential. Then the Hamilton equations are

$$\ddot{q}_j = V'(q_{j+1} - q_j) - V'(q_j - q_{j-1}), \quad j = 1, \dots, N. \quad (2.2)$$

In the case of the FPU α - β model one has (after a simple rescaling)

$$V(\delta) = V_{FPU}(\delta) := \frac{\delta^2}{2} - \frac{\delta^3}{6} + \beta \frac{\delta^4}{4!}. \quad (2.3)$$

In this case we will write $H = H_{FPU}$. We decompose the FPU Hamiltonian as follows

$$H_{FPU} = H_0 + H_1 + H_2 \quad (2.4)$$

where

$$\begin{aligned} H_0 &:= \sum_{j=1}^N \left(\frac{p_j^2}{2} + \frac{(q_{j+1} - q_j)^2}{2} \right) \\ H_1 &:= -\frac{1}{3!} \sum_{j=1}^N (q_{j+1} - q_j)^3 \\ H_2 &:= \frac{\beta}{4!} \sum_{j=1}^N (q_{j+1} - q_j)^4. \end{aligned}$$

Introduce the discrete Fourier transform defined by

$$\hat{q}_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} q_j e^{2i\pi jk/N}, \quad k \in \mathbb{Z}, \quad (2.5)$$

and consider $\hat{p}_k$ defined analogously. One has $\hat{p}_k = \hat{p}_{k+N}$, $\hat{q}_k = \hat{q}_{k+N}$, $\forall k \in \mathbb{Z}$, so we restrict to $\{\hat{p}_k, \hat{q}_k\}_{k=0}^{N-1}$. Corresponding to real sequences (p_j, q_j) one has $\hat{q}_k = \hat{q}_{N-k}$ and $\hat{p}_k = \hat{p}_{N-k}$.

Introduce the real Fourier variables

$$X_k = \frac{\hat{p}_k + \hat{p}_{N-k} - i\omega_k(\hat{q}_k - \hat{q}_{N-k})}{\sqrt{2\omega_k}}, \quad Y_k = \frac{\hat{p}_k - \hat{p}_{N-k} + i\omega_k(\hat{q}_k + \hat{q}_{N-k})}{i\sqrt{2\omega_k}}, \quad (2.6)$$

$$k = 1, \dots, N-1,$$

where $\omega_k \equiv \omega\left(\frac{k}{N}\right) := 2 \sin(k\pi/N)$; using such coordinates, which are symplectic, the quadratic part

$$H_0 := \sum_{j=1}^N \frac{p_j^2 + (q_{j+1} - q_j)^2}{2} \quad (2.7)$$

of the Hamiltonian takes the form

$$H_0 = \sum_{k=1}^{N-1} \omega_k \frac{X_k^2 + Y_k^2}{2} + \frac{\hat{p}_0^2}{2}. \quad (2.8)$$

We remark that due to translation invariance the quantity

$$\sum_{j=1}^N p_j = P = \hat{p}_0 \sqrt{N} \quad (2.9)$$

is an integral of motion and as a consequence the “center of mass”

$$\sum_{j=1}^N q_j = Q = \hat{q}_0 \sqrt{N}, \quad (2.10)$$

moves uniformly. Without loss of generality we restrict here to the surface $P = 0$, so that also $Q = 0$ is invariant. After this restriction the Hamiltonian (2.8) takes the form

$$H_0 = \sum_{k=1}^{N-1} \omega_k \frac{X_k^2 + Y_k^2}{2} = \sum_{k=1}^{N-1} E_k, \quad (2.11)$$

where

$$E_k := \omega_k \frac{X_k^2 + Y_k^2}{2}$$

3 FPU and Toda

In the case where the potential in (2.1) is exponential, namely

$$V(\delta) = V_{Toda}(\delta) := e^{-\delta} \quad (3.1)$$

one gets the Toda lattice, which is well known to be integrable. In this case we shall denote the Hamiltonian by H_{Toda} .

The trivial remark that

$$|V_{Toda}(\delta) - (1 - \delta + V_{FPU}(\delta))| \leq |1 - \beta| \delta^4 + \mathcal{O}(\delta^5) \quad (3.2)$$

gives the connection with the FPU system and allows to use results on Toda in order to deduce results on the FPU system. We remark that the term 1 in (3.2) changes the Hamiltonian by a constant, while the term δ gives a correction to H given by $\sum_j (q_{j+1} - q_j) \equiv 0$, so they are both irrelevant. Actually in the following we always concentrate on the case where $\beta \neq 1$. When $\beta = 1$ FPU and Toda are close up to terms of order 5 and in this particular case the result can be improved in an obvious way.

3.1 Finite Number of Particles N

The first idea is to use the KAM and the Nekhoroshev theory considering FPU as a perturbation of the linear system. This has been done starting from Chirikov [31] and has been discussed in several papers mainly with the aim of understanding if such theorems can be applied to infinite chains and to states with finite (small) specific energy (see [12, 16, 19]).

Here we take a different point of view, which is more natural once one realizes that (3.2) holds, namely we consider FPU as a perturbation of Toda. To this end one has to know that the Toda Hamiltonian, when written in action angle variables, fulfills the non degeneracy properties required for the application of the KAM or the Nekhoroshev theorems. A proof of the fact that FPU fulfills the KAM non degeneracy property was done by Rink [41, 42] exploiting the symmetries of the FPU system, while a detailed study of the action angle variables for Toda, as well as a proof that they are globally defined was done by Henrici and Kappeler. We state here the result that was proved in [26].

Theorem 3.1 (Theorem 1.1 of [26]). *For any integer $N \geq 2$ there exists a global real analytic symplectic diffeomorphism $\Phi_N : \mathbb{R}^{N-1} \times \mathbb{R}^{N-1} \rightarrow \mathbb{R}^{N-1} \times \mathbb{R}^{N-1}$, $(X, Y) = \Phi_N(x, y)$ with the following properties:*

(i) *The Hamiltonian*

$$\tilde{H}_{Toda} := H_{Toda} \circ \Phi_N$$

is a function of the actions $I_k^{Toda} := \frac{x_k^2 + y_k^2}{2}$ only.

(ii) *The differential of Φ_N at the origin is the identity.*

Actually Henrici and Kappeler also computed explicitly the quadratic term of the Taylor expansion of $\tilde{H}(I^{Toda})$ at the origin. Such an expression is different in the case of an even number of particle [29] and in the case of an odd number of particles [27], but in both cases the authors succeeded to prove that such a Hamiltonian is convex as a function of the action variables [28] (see Theorem 1.2). In particular one can thus apply both KAM and Nekhoroshev theorem, still restricting to the manifold $Q = P = 0$. We state just the corresponding Nekhoroshev theorem. The result is due to Henrici Kappeler [28] and the stability exponent that we give are computed applying the version of Nekhoroshev theorem due to Lochak and Neishtadt [34] (see also [25]).

Theorem 3.2 *Consider the Cauchy problem for the FPU system (2.1) with $\beta \neq 1$, then there exist positive constants ϵ_* , C , such that for any initial datum (p^0, q^0) fulfilling $P = Q = 0$ and also*

$$\epsilon := \|(p^0, q^0)\| < \epsilon_*, \quad (3.3)$$

one has

$$\left| I_j^{Toda}(t) - I_j^{Toda}(0) \right| \leq C \epsilon^{2+b}, \quad \forall |t| \leq \exp\left(\frac{\epsilon_*}{\epsilon}\right)^a \quad (3.4)$$

with I_j^{Toda} the actions of the Toda system and

$$a = \frac{1}{2(N-1)} = b. \quad (3.5)$$

In general the constant ϵ_* depends very strongly on the dimension N . In the case of chains of rotators ϵ_* has been explicitly estimated and it tends to zero as N^{-N} (See Theorem 3.1 of [3]).

We remark that a slightly better exponent a was obtained in [47] at the expenses of worsening b . In any case one always has that, as $N \rightarrow \infty$,

$$a \sim \frac{1}{2N}. \quad (3.6)$$

Now, the main interest when studying the FPU chain pertains the limit $N \rightarrow \infty$, and the question is whether the above results, namely Theorems 3.1 and 3.2, persist when $N \rightarrow \infty$.

It is very easy to see that Theorem 3.2 becomes empty when $N \rightarrow \infty$, indeed both exponents a and b tend to zero and thus both the control of the actions (which are of order ϵ^2) and the time of control, become empty.

However the problem is even harder, indeed also Theorem 3.1 gives a result for the case of a fixed number of particles, so it is not at all clear if it remains valid in the limit $N \rightarrow \infty$.

3.2 Limit $N \rightarrow \infty$

The behavior of the transformation introducing Birkhoff coordinates for the Toda system in the limit $N \rightarrow \infty$ has been studied in [6]. It was proved that, on the one hand it develops a singularity at a distance of order N^{-2} from the origin, but on the other hand it is well defined in a ball of such an order of magnitude, even in a very strong topology and this allows to deduce a result on FPU, which, up to now is the best available concerning the times over which the dynamics is controlled. To state such results in a slightly simplified, but precise way, for any $\sigma \geq 0$, we introduce in $\mathbb{C}^{N-1} \times \mathbb{C}^{N-1}$ the discrete Sobolev-analytic norm

$$\|(\hat{p}, \hat{q})\|_{h^\sigma}^2 := \sum_{k \neq 0} [k]_N^2 e^{2\sigma[k]_N} \omega_k \frac{|X_k|^2 + |Y_k|^2}{2} \quad (3.7)$$

where

$$[k]_N := \min \{k, N - k\}.$$

Remark 1 This norm differs from the norm used in [6] by a factor $N^{-1/2}$. Indeed the present norm is of order 1 in a space containing states with finite energy and with only low modes excited, while the norm in [6] was of order 1 for states with finite specific energy and concentrated on small modes. We made this choice in order to have a norm more adapted to the developments of Sect. 4.

The space $\mathbb{C}^{N-1} \times \mathbb{C}^{N-1}$ endowed with such a norm will be denoted by h^σ . We denote by $B^\sigma(R)$ the ball of radius R and center 0 in the topology defined by the norm $\|\cdot\|_{h^\sigma}$.

From now on we will denote

$$\mu := \frac{1}{N}, \quad (3.8)$$

which will play the role of a small parameter.

Theorem 3.3 (Theorem 1.3 of [6]). *For any $\sigma \geq 0$ there exist strictly positive constants R_* , C , such that for any $N \geq 2$, the map Φ_N of Theorem 3.1 is analytic as a map from $B^\sigma(R_*\mu^{3/2})$ to h^σ and fulfills*

$$\sup_{\|(x,y)\|_{h^\sigma} \leq R\mu^{3/2}} \|\Phi_N(x, y) - (x, y)\|_{h^\sigma} \leq C R^2 \mu^{3/2}, \quad \forall R < R_*. \quad (3.9)$$

The same estimate is fulfilled by the inverse map Φ_N^{-1} .

Thus the Birkhoff coordinates of the Toda system are well defined, at least in a ball of radius of order $\mu^{3/2}$. We remark that, in such a ball, the energy is of order N^{-3} .

One would like to apply also Nekhoroshev's theorem, but here one has to face the problem that the best known threshold for the applicability of such a theorem decrease more than exponentially with the number of degrees of freedom.

What one can easily prove is a result which exploits the fact that FPU is a perturbation of Toda lattice with a perturbation of fourth order. The following theorem has been proved in [6] as a corollary of Theorem 3.3.

Theorem 3.4 (Theorem 1.16 of [6]). *Fix $\sigma \geq 0$ and consider the Cauchy problem for the FPU system with an initial datum (p^0, q^0) and let R be s.t. $\|(p^0, q^0)\|_{h^\sigma} = R\mu^{3/2}$. There exists $R_* > 0$, $T > 0$ such that, if $R < R_*$, then*

$$\|(p(t), q(t))\|_{h^\sigma} \leq 2R_*\mu^{3/2}, \quad \forall |t| \leq \frac{T}{R^2\mu^4}. \quad (3.10)$$

We remark that since the norm has an exponential weight, it follows that the energy in the k -th mode decreases exponentially with $|k|$ for the times (3.10). This is the phenomenon usually referred to as existence of a metastable packet of modes. Actually the times of order μ^{-4} obtained by this theorem represent the longest lower estimate that has been rigorously proven for the time of existence of the FPU meta stable packet.

3.3 Toda and KdV

It has long been known that the KdV equation is obtained in some formal continuous limit from FPU and also from Toda. Since both Toda and KdV are integrable, it is natural to investigate the possible convergence of the integrable structure of Toda to the integrable structure of KdV. This was done in [4, 5, 8]. In such papers it was shown that in the continuous limit both the actions and the frequencies of the Toda system converge to the actions and the frequencies of a couple of KdV equations. We are now going to state this result.

First of all we interpolate the state of the Toda chain. It is convenient to proceed as follows: given two sequences q_j, p_j , we consider two functions $u(x)$ and $w(x)$ periodic of period 1 in x with the property that

$$\delta_j := q_{j+1} - q_j = \mu^2 u(\mu j), \quad p_j = \mu^2 w(\mu j), \quad (3.11)$$

and define the functions

$$r = \frac{u + w}{\sqrt{2}}, \quad s = \frac{u - w}{\sqrt{2}}, \quad (3.12)$$

then, as $\mu \equiv N^{-1} \rightarrow 0$ it turns out that r and s converge to solutions of the KdV equations, one controlling right traveling waves, the other controlling left traveling waves.

To state the corresponding result we denote again by I_j^{Toda} the action variables of the Toda system and we remark that associated to a system of action angle coordinates one has an important dynamical object which is the set of the frequencies, which are just the derivatives of the Hamiltonian with respect to the actions and that control the evolution of the corresponding angles. We denote by $\omega_j^{Toda} = \omega_j^{Toda}(I^{Toda})$ the j th frequency of the Toda system.

One has corresponding objects in the KdV hierarchy. Namely given a periodic function u one can represent it in action angle variables, which are common to all the equations of the KdV hierarchy. Denoting by $I_j^{KdV} = I_j^{KdV}(u)$ the j th action of a function $u = u(x)$ according to the KdV hierarchy, we have the following theorem

Theorem 3.5 (Theorem 1.2 of [4]). *Assume that u and w are of class $C^2(\mathbb{T})$, then, provided N is large enough one has the following asymptotic*

$$I_j^{Toda} = \frac{1}{2N^2} \left[I_j^{KdV}(s) + \mathcal{O}\left(\frac{1}{N^{1/16}}\right) \right], \quad 0 < j < N^{1/16}, \quad (3.13)$$

$$I_{N-j}^{Toda} = \frac{1}{2N^2} \left[I_j^{KdV}(r) + \mathcal{O}\left(\frac{1}{N^{1/16}}\right) \right], \quad 0 < j < N^{1/16}, \quad (3.14)$$

$$I_j^{Toda} = \mathcal{O}\left(\frac{1}{N^{36/16}}\right), \quad N^{1/16} \leq j \leq N - N^{1/16}. \quad (3.15)$$

Such a theorem shows that the actions with small index behave in the limit as the actions of a corresponding KdV equation. The remaining actions tend to zero in the limit faster than N^{-2} .

Actually this theorem is a rescaling of Theorem 1.2 of [4]. Indeed the potentials $q_{\pm}$ of [4] are periodic of period $1/2$ while the functions r, s are periodic of period 1 . Furthermore we made here a choice of the scaling different from that of [4, 5] in order to produce a result directly comparable with the results obtained by normal form methods. The connection is $q_{-}(x) = 4\sqrt{2}s(2x)$ and similarly for the other potential. In order to produce the present version of the main result of [4] one has to study the behavior of the eigenvalues and the discriminant of the Sturm–Liouville problem under such a scaling, and then use the formula by Flaschka and McLaughlin for the actions of the KdV equation (see the formula at p. 64 of [32]). Such formulae can be scaled, too, and a long, straightforward and uninteresting computation produces the previous statement.

Concerning the dynamics, we consider the frequencies of the Toda lattice. They will be approximated by the frequencies of a linear combination of the first two equations of the KdV hierarchy, precisely by the combination

$$K^{KdV} := \frac{2}{N} K_1 - \frac{1}{12N^3} K_2, \quad (3.16)$$

where K_1 and K_2 in our scaling, are given by

$$K_1(u) = \frac{1}{2} \int_0^1 u^2 dx, \quad (3.17)$$

$$K_2(u) = \int_0^1 \left(\frac{1}{2} u_x^2 + \sqrt{2} u^3 \right) dx, \quad (3.18)$$

The frequencies corresponding to this Hamiltonian, computed for a given periodic function u will be denoted by $\omega^{KdV}(u)$ (even if they depend only on the action variables pertaining to u). Notice that, since the KdV Hamiltonian are scaled by a factor proportional to N , the frequencies are of order μ . Concerning the frequencies we have the following result

Theorem 3.6 (Theorem 1.1 of [4]). *There exists a value c_2 s.t., if u and w are $C^2(\mathbb{T})$, then the frequencies of the Toda system have the following asymptotic.*

$$\omega_j^{Toda} = \omega_j^{KdV}(s) + \mathcal{O}\left(\frac{1}{N^{193/64}}\right), \quad 0 < j < N^{1/64}, \quad (3.19)$$

$$\omega_{N-j}^{Toda} = \omega_j^{KdV}(r) + \mathcal{O}\left(\frac{1}{N^{193/64}}\right), \quad 0 < j < N^{1/64}, \quad (3.20)$$

$$\omega_j^{Toda} = \frac{2\pi j}{N} + \mathcal{O}\left(\frac{j^3}{N^3}\right), \quad N^{1/64} < j < N^{1/8}, \quad (3.21)$$

$$\omega_{N-j}^{Toda} = \frac{2\pi j}{N} + \mathcal{O}\left(\frac{j^3}{N^3}\right), \quad N^{1/64} < j < N^{1/8}, \quad (3.22)$$

$$\omega_j^{Toda} = 2 \sin\left(\frac{\pi j}{N}\right) \left(1 + \mathcal{O}\left(\frac{\log N}{N^{1/4}}\right)\right), \quad N^{1/8} < j < N - N^{1/8}. \quad (3.23)$$

We remark that the choice of the powers $1/16$ or analogous quantities has been done here in order to simplify the statements: in the original statement one has quite a lot of freedom to make different choices.

Remark 2 We remark that states of the form (3.11), (3.12) have total energy of order N^{-3} , while the first actions are of order N^{-2} . The two facts are compatible since the harmonic energy is $\sum_j \omega_j I_j$ and for small values of j the frequencies are of order j/N .

4 Continuous Limit and KdV

The results of the previous section show that in some precise sense Toda converges to KdV in the continuum limit and thus the same is true for the FPU. In this section we analyze how the equations of the KdV hierarchy can be obtained directly working on the FPU system. Quite surprisingly it turns out that the order of approximation can be pushed further.

4.1 Interpolation

First we rewrite the FPU equations in terms of the variables

$$\delta_j := q_{j+1} - q_j, \quad (4.1)$$

namely

$$\ddot{\delta} = \Delta_1 V'(\delta), \quad (4.2)$$

with Δ_1 the operator of finite difference, namely

$$(\Delta_1 \delta)_j := \delta_{j+1} + \delta_{j-1} - 2\delta_j.$$

Then we interpolate and scale amplitude and time: let $u(\tau, x)$ be a smooth periodic function of period 1 in x with the property that

$$\delta_j(t) = \epsilon u(\mu t, \mu j), \quad (4.3)$$

where $\mu := N^{-1}$, while ϵ is a “small” parameter in principle independent of μ (but see below). In order to get that δ fulfills the correct equations we assume that u fulfills the equations

$$u_{\tau\tau} = \Delta_\mu \frac{V'(\epsilon u(x))}{\epsilon}, \quad (4.4)$$

where Δ_μ is the finite discretization of the second derivative, namely

$$(\Delta_\mu u)(x) := \frac{u(x + \mu) - 2u(x) + u(x - \mu)}{\mu^2}.$$

It is easy to see that this equation is Hamiltonian with Hamiltonian function

$$H(u, v) := \int_0^1 \left\{ \frac{v(x)[- \Delta_\mu v(x)]}{2} + \frac{V(\epsilon u(x))}{\epsilon^2} \right\} dx, \quad (4.5)$$

where v is the variable conjugated to u .

Notice that if $\epsilon = \mu^2$ then one has $w = v_x + \mathcal{O}(\mu)$, where w is the function interpolating the sequence p_j . Expanding in power of ϵ and μ we get

$$H(u, v) = H_0(u, v) + \epsilon H_1(u, v) + \epsilon^2 H_2(u, v) + \text{higher order terms}, \quad (4.6)$$

with

$$H_0(u, v) = \int \frac{v_x^2 + u^2}{2} dx, \quad (4.7)$$

$$H_1(u, v) = -\frac{1}{6} \int \left(\frac{\mu^2}{4\epsilon} v_{xx}^2 + u^3 \right) dx, \quad (4.8)$$

$$H_2(u, v) = \int \left(\frac{\mu^4}{\epsilon^2} \frac{v_{xxx}^2}{720} + \frac{\beta}{24} u^4 \right) dx. \quad (4.9)$$

Following [9, 13], it is very useful to make the non canonical change of variables

$$r := \frac{u + v_x}{\sqrt{2}}, \quad s := \frac{u - v_x}{\sqrt{2}}. \quad (4.10)$$

Notice that these quantities do not coincide with the quantities (r, s) introduced in SubSect. 3.3, but differ by terms of order $\mathcal{O}(\mu)$.

In the variables (r, s) the Poisson tensor transforms in such a way that the Hamilton equations of a Hamiltonian H become

$$\frac{dr}{d\tau} = -\partial_x \nabla_r H, \quad \frac{ds}{d\tau} = \partial_x \nabla_s H, \quad (4.11)$$

where ∇_r is the L^2 gradient with respect to the variable r and similarly ∇_s . The first three terms (4.7)–(4.9) of the power expansion of the Hamiltonian, take the form

$$H_0 = \int_0^1 \frac{r^2 + s^2}{2} dx \quad (4.12)$$

$$H_1 = -\frac{1}{48} \int_0^1 \left[\frac{\mu^2}{\epsilon} r_x^2 + 2\sqrt{2}r^3 \right] dx - \frac{1}{48} \int_0^1 \left[\frac{\mu^2}{\epsilon} s_x^2 + 2\sqrt{2}s^3 \right] dx \quad (4.13)$$

$$- \int_0^1 \left[\frac{1}{4\sqrt{2}} (r^2 s + s^2 r) - \frac{1}{24} \frac{\mu^2}{\epsilon} r_x s_x \right] dx \quad (4.14)$$

$$H_2 = \int_0^1 \left[\frac{\mu^4}{\epsilon^2} \frac{(r_{xx} - s_{xx})^2}{1440} + \beta \frac{(r + s)^4}{96} \right] dx. \quad (4.15)$$

We write explicitly the equation of r (the one for s is obtained by exchanging s with r and add a minus sign):

$$r_\tau = -r_x - \epsilon \left(\frac{\mu^2}{\epsilon 24} r_{xxx} - \frac{1}{2\sqrt{2}} r r_x \right) - \epsilon^2 \left(\frac{\mu^4}{\epsilon^2} \frac{\partial_x^5 r}{720} + \frac{\beta}{24} \partial_x r^3 \right) \quad (4.16)$$

$$+ \frac{\mu^2}{24} s_{xxx} + \frac{\epsilon}{2\sqrt{2}} (r s_x + s r_x + s s_x) + \frac{\mu^4}{720} \partial_x^5 s - \frac{\epsilon^2 \beta}{24} \partial_x (3r^2 s + 3r s^2 + s^3). \quad (4.17)$$

First we remark that if we neglect the coupling (4.17), then one gets an equation for r alone, and, furthermore, such an equation is a perturbation of order ϵ^2 of a KdV equation in a moving frame. A naive idea to analyze the dynamics (but which in any case leads to interesting results) consists in assuming that s and its derivatives remain negligible for all times of interest. Then, passing to a moving frame, namely defining $y = x - \tau$, one is led to Eq.

$$\frac{6}{\epsilon} r_\tau = -\frac{\mu^2}{4\epsilon} r_{yyy} + \frac{3}{\sqrt{2}} r r_y, \quad (4.18)$$

which, after rescaling again time by $\tau_1 := \epsilon\tau$, becomes the standard KdV equation. This equation describes states with $N = \mu^{-1}$ particles, whose Fourier components are concentrated on modes with very small wave number $k \simeq \mu$ and amplitude ϵ . Thus, if one is interested in studying the limit $N \rightarrow \infty$ without sending to zero the amplitude, one has to study the KdV equation in the zero dispersion limit. This is a quite complicated problem to which we will come back in Sect. 5.1.

4.2 KdV₃: A Good Approximation

In the case $\epsilon = \mu^2$ the limiting equation is no more a singular KdV equation. Furthermore it has been shown in a series of papers that one can actually get rid of the coupling terms between r and s by a procedure of Birkhoff normal form. Such a procedure can be carried out until the order μ^4 , while in order to eliminate the coupling between r and s at order μ^6 there are some obstructions. Furthermore one can use Kodama theory [30] in order to show that, after a change of coordinates, the FPU equations are very well approximated by a function of the first three equations of the KdV hierarchy. In order to state precisely the corresponding result, we start by recalling that the first three Hamiltonians of the KdV hierarchy are (3.17), (3.18) and

$$K_3(u) = \int_0^1 \left(\frac{1}{40} u_{xx}^2 + \sqrt{2} u_x^2 u + u^4 \right) dx, \quad (4.19)$$

and the corresponding equations are obtained by using Gardner Hamiltonian structure, namely they are given by

$$u_\tau = \partial_x \nabla K_n, \quad n = 1, 2, 3. \quad (4.20)$$

One recognizes that for $n = 1$ one gets the transport equation, while for $n = 2$ one gets the standard KdV equation and for $n = 3$ one gets the second equation of the KdV hierarchy. We recall that these equations are commuting, in the sense that the function K_n is an integral of motion for each one of the equations of the hierarchy and also for the dynamics of any function of the Hamiltonians of the KdV hierarchy.

The Hamiltonian that approximates the FPU is constructed as follows: given constants c_2, c_3, c_4, c_5 define

$$K(r) := K_1(r) + \mu^2 c_2 K_2(r) + \mu^4 (c_3 K_3(r) + c_4 [K_2(r)]^2) \quad (4.21)$$

and then consider

$$K(r) + K(s) + c_5 \mu^4 [K_2(r)][K_2(s)]. \quad (4.22)$$

In order to state the main theorem we use the Sobolev norms H^ℓ which are defined in the standard way. We will denote by $B_\ell(R)$ the ball of radius R centered at the origin with respect to the H^ℓ norm. The following theorem is essentially due to Gallone, Ponso, Rink [22, 23].

Theorem 4.1 *There exists ℓ' and a map $T_\mu : B_{\ell+\ell'}(R) \times B_{\ell+\ell'}(R) \rightarrow B_\ell(R) \times B_\ell(R)$, with the following properties*

- (i) $\sup_{(r,s) \in B_{\ell+\ell'}(1)} \|T_\mu(r, s) - (r, s)\|_{H^\ell} \leq C\mu^2$,
- (ii) *There exist constants $c_2, \dots, c_5$ with the following property: let I be a possibly infinite interval of rescaled time containing the origin and $z(\cdot) = (r(\cdot), s(\cdot)) \in C^1(I; B_{\ell+\ell'}(1))$ be a solution of the Hamiltonian system (4.22). Define*

$$z_a \equiv (r_a, s_a) := T_\mu(r, s). \quad (4.23)$$

Then there exists $R \in C^1(I, H^\ell \times H^\ell)$ s.t. one has

$$\dot{z}_a(\tau) = J \nabla H_{FPU}(z_a(\tau)) + \mu^6 R(\tau), \quad \forall \tau \in I, \quad (4.24)$$

where H_{FPU} is the Hamiltonian (4.12)–(4.15) of the FPU system rewritten in the variables (r, s) .

Actually the above theorem is the result of a combination of several different arguments: the first is Proposition 10 of [22] which allows to transform (up to order μ^6) the Hamiltonian (4.6) (with $\epsilon = \mu^2$) in the Hamiltonian of two decoupled systems, the first only involving the function r and the second only involving the variable s . Then one can use Kodama theory as developed in [23] in order to transform each one of these Hamiltonians in combinations of the first three Hamiltonian of the KdV hierarchy (see Eq. (1.1) of [23] which is written in a more explicit way in Eq. (4.79) of [20]). Finally one has to use the theory developed in [7] (See Corollary 4.12, which however has to be adapted to the simpler present case of r and s which are periodic functions). Notice that a similar result for the water wave problem on $\mathbb{R}$ was proved in [7].

Then, following the proof of the main result of [9] one can probably deduce a corollary showing that the solution of the FPU system remains $\mathcal{O}(\mu^3)$ close, to the solutions constructed with the Hamiltonian (4.22), for a time of order μ^{-3} (although the approximation is of order μ^6). However, on such a time scale, the effects of the KdV_3 are at most of order μ , so they are not yet visible on a macroscopic scale. This is a limitation that the theory shares with all the known results relative to the study of the modulation equations.

A further, more serious limitation is that the result is valid only in a regime of energy E such that $EN^3 \ll 1$.

5 Turbulent Behavior and Approach to Equilibrium

5.1 Continuous Limit and Approach to Equilibrium

As shown in [21] the continuous limit can be used also to deduce some interesting results on the route to thermalization of the FPU model. To present their result consider again Eq. (4.18): by rescaling the time to $\tilde{\tau} := \epsilon \tau / 6$ and the amplitudes, one is led to the rescaled form

$$r_{\tilde{\tau}} = \frac{\mu^2}{\epsilon} r_{xxx} - r r_x. \quad (5.1)$$

The idea of [21] is that the formal zero dispersion limit $\mu = 0$, i.e. the Burgers equation, should describe well the solution of the FPU as far as the solution has a bounded third derivative. It is well known that a solution of the Burgers equation develops a singularity after a finite time and, at such time, the approximation fails for sure. However the authors take advantage of the fact that the evolution of the Fourier coefficients of the solution of the Burgers equation can be explicitly found: such evolution does not show any singularity. Thus one can analyze the behavior of the Fourier coefficients also when the solution in physical space is no more defined. In [21] the authors compute explicitly the Fourier spectrum of the solution at the time when the singularity in physical space first appear and they find the very surprising result that the energy of the k th mode decreases as $|k|^{-8/3}$, namely the spectrum has a power law decay, as predicted by Kolmogorov theory of turbulence for the Navier–Stokes equation. They also compare the result with the numerical simulations finding that they are

in a quite good agreement with the theoretical predictions. The numerical simulations also give more results, but we will not discuss them here.

We remark that it is not clear how to deduce a theorem from the above method, but similar ideas have been used by Kuksin in order to discuss the development of turbulence in NLS [33]. In particular considering Eq. (5.1) as a perturbation of the Burgers' equation one can prove the following theorem

Theorem 5.1 *Consider Eq. (5.1) with initial datum $r_0(x) = \sin x$, and assume $\mu' := \mu^2/\epsilon$ small enough, then there exists a time t_* such that the corresponding solution fulfills*

$$\max_{k \leq 5} \sup_x |\partial_x^k r(x, t_*)| \geq \frac{1}{(\mu')^{1/4}}. \quad (5.2)$$

Such a theorem shows that, although in a very weak sense, one has a transfer of energy to high frequency modes. The advantage is that the result holds in the thermodynamic limit. Application of such result to the FPU model is in progress.

5.2 Wave Kinetic Equation

A very important result has been obtained quite recently, namely the proof of the validity of the Boltzmann equation for a dilute gas of interacting particles [15]. The method of proof is very robust, and it can be used to deduce the validity of the wave kinetic equation for the dynamics in the NLS equation (see [14]) or to study the dynamics of the β -FPU model as done more recently in [45].

The idea of the deduction of the wave kinetic equation is to use a variant of the Birkhoff normal form to extract the relevant interactions from the non linearity and then to make a random phase approximation in order to extract the contributions which are relevant for almost all the values of the phases. Indeed, it is well known that corresponding to random choices of the phases most of the interactions among different normal modes are negligible. Moreover, the authors also take the large limit box, which, in the case of FPU, is the limit $N \rightarrow \infty$.

A detailed analysis of the FPU dynamics based on these ideas has been done in a series of papers by Onorato et al. [35, 40] in which the authors have been able to put into evidence the main interactions driving the thermalization process in the FPU dynamics. We refer to the original papers for more details (see in particular the review [39]).

Up to now, not much is rigorously known on the validity of the wave kinetic equation and the corresponding thermalization processes in the FPU model. In particular to deduce rigorously the wave kinetic equation is necessary to prove that the random phase approximation is preserved by the dynamics, and this is not easy. We just mention the result by Wu [45] ensuring the validity of the wave kinetic equation in the β FPU model, provided $\beta \sim N^{-\gamma}$ with some positive γ . However this result is valid only up to a fraction of the kinetic time, so that it does not guarantee that the predictions of the wave kinetic equation are rigorously justified for the dynamics of the FPU model.

We also mention the recent paper [46] in which the authors study the normal form transformation removing the non resonant interactions in the FPU system. They show that the transformation is well defined for $\beta \leq N^{-(1+\epsilon)}$.

6 Results at the Thermodynamic Limit

The results of the previous Sects. 3 and 4 imply that the dynamics of the system remains close to integrable for long times. In particular, it was shown that the solutions to the equations of motion are well approximated by the solutions of a suitable integrable system. The reported results hold in the limit in which $N \rightarrow \infty$ with a total energy $E \sim N^{-3}$. However if one thinks of the original FPU numerical experiment, it is really important to study the system in the thermodynamic limit, in which the specific energy $\epsilon = E/N$ is kept fixed as $N \rightarrow \infty$. The study of this limit is what matters for statistical mechanics, where, however, the object of study are no longer the system's trajectories but its ergodic properties.

In the ergodic context, the equilibrium state is characterized by an invariant measure μ , and thermalization refers to the process by which, starting from an initial measure μ_0 , the evolved one μ_t converges to the invariant measure μ . From this perspective, the numerical experiment by Fermi Pasta Ulam shows that starting from a measure concentrated just on the first few modes, as it evolves, it does not converge to the micro canonical measure, which, at least for small specific energy, weights all normal modes of the system approximately equally, up to an extremely long time scale. In this sense, this experiment suggests the thermalization time to be extremely long at low specific energies.

To obtain rigorous results, one exploits the property that the convergence of μ_t to μ is equivalent to the vanishing of the time auto correlation for any square-integrable function. Therefore, by showing that the auto correlation of a suitable dynamical variable remains close to its initial value for long times, we obtain an estimate from below of the thermalization time. The first subsection reports the rigorous estimates obtained using this approach for the FPU system in the thermodynamic limit.

But the study of auto correlations has also a second aspect, in which we report in the second subsection. In fact, according to linear response theory, if a system starts from an equilibrium condition, the response to an external force, or the way it adjusts to a new equilibrium after a perturbation, are determined by the time correlations of suitable dynamical variables, provided that the perturbation be sufficiently small. For example, the electrical susceptibility of a crystal is determined by the Fourier transform of the auto correlations of the optical normal modes¹ $\hat{q}_k^+$ with $k = 0$. In this way, the persistence of an integrable behavior of the correlations over long times heavily influences the physically observable quantities. In particular, the way in which the auto correlations vanish for $t \rightarrow \infty$, whether they decay exponentially, rather than as stretched exponentials, or even as a power law, is particularly relevant. Notice that the study of the correlations decay is a very active field in experimental physics, for what concerns for example, glasses, super cooled liquids and more in general the field of soft matter (see for example the review [17]).

So, in the second subsection are reported some known analytical results on the decay of time correlation, and in particular a criterion that ensures that the decay is not exponential.

6.1 Finite Time Stability of Correlations

The most straightforward example is the study of the thermalization time of the system, that is, the time in which the system reaches thermodynamic equilibrium after a displacement from it. A lower bound on this time can be obtained by providing an estimate of the time during which the time auto correlation of a dynamical variable remains far from zero. Recall

¹ Optical normal modes are the normal modes corresponding to that branch of the dispersion relation that does not vanish at $k = 0$. See below.

that, given a dynamical variable $f : \mathcal{F} \rightarrow \mathbb{R}$ and denoting by $\Phi^t : \mathcal{F} \rightarrow \mathcal{F}$, the flow induced on the phase space $\mathcal{F}$ by the solutions of the equations of motion, the auto correlation $\mathcal{C}_f(t)$ is simply defined as

$$\mathcal{C}_f(t) = \langle f(t)f(0) \rangle - \langle f \rangle^2,$$

where we have denoted $f(t) := f \circ \Phi^t$, while the average $\langle \cdot \rangle$ is taken with respect to a flow-invariant μ measure. As often in statistical mechanics, the invariant measure we will consider is the Gibbs measure. As it is well known, it is defined starting from the Hamiltonian of the system $H(q, p)$, using the formula

$$d\mu = \frac{e^{-\beta_T H}}{Z(\beta_T)} d^N q d^N p,$$

where the partition function $Z(\beta_T)$ is the normalization factor of the measure (i.e. $Z = \int \exp(-\beta_T H) d^N q d^N p$) and the parameter β_T is usually identified with the inverse temperature of the system.²

An interesting result was obtained in [36] concerning the FPU systems: the energy of “packets” of normal mode (defined shortly) remains correlated for times of the order of β_T , that is, the thermalization time increases at least with the inverse of the temperature as the temperature decreases towards zero. The idea of considering a “packet” of normal modes as a significant dynamic variable comes from the fact that in the limit $N \rightarrow \infty$ the frequencies of the normal modes become a continuum, and therefore, every measurement, due to the physical limitations, will always concern a group of modes with approximately the same frequency. Let us then define the observable

$$J_x = \sum_{k=1}^{N-1} g(k/N) E_k,$$

where $g : [0, 1] \rightarrow [0, 1]$ is a smooth function defined by the interval $[0, 1]$ in itself, peaked around a value $x \in [0, 1]$, while E_k is the energy of the k th normal mode (see (2.11)). J_x therefore represents the (mean) energy of the packet of oscillators of frequency $\omega(x)$.

For this observable, the following Theorem holds (Corollary 1 of the paper [36]):

Theorem 6.1 *There exist constants $\beta_T^* > 0$, $N^* > 0$, and $C > 0$ such that, for any $\beta_T > \beta_T^*$ and for any $N > N^*$, one has*

$$\mathcal{C}_{J_x}(t) > \frac{1}{2} \sigma_{J_x}^2 \quad \text{for all } |t| \leq C \beta_T.$$

Therefore, according to this theorem, the correlation time goes to infinity as the temperature tends to zero. This estimate is probably not optimal, because numerical work in paper [10] seems to indicate a more than linear growth of the correlation times with β_T .

A similar result has been obtained in [24] by exploiting the proximity between FPU and Toda potentials and considering the time evolution of the integrals of motion for the Toda chain under the flow induced by the closest FPU Hamiltonian. Denote by $J^{(m)}$ the trace of the m th power of the Lax matrix associated to the Toda chain, divided by m (see (2.12) in [24]). Such quantities are integral of motions for the Toda Hamiltonian flow and have a rather simple analytic expression in terms of the canonical variables (q_j, p_j) . For them, an analog of Theorem 6.1 holds, with $J^{(m)}$ replacing J_x (see Theorem 2.1 in [24]). Moreover, this

² Recall (see Sect. 2) that we are always thinking of the motion as occurring on the invariant manifold where the center of mass of the system is fixed, so that the measure is well defined for any fixed N .

provides in turn an estimate involving some special packets of normal modes, since $J^{(m)}$ is well approximated, for low specific energies, by a suitable combination of energies of normal modes. This is expressed by the following Theorem (see Theorem 2.5 in [24]):

Theorem 6.2 Fix $m \in \mathbb{N}$ and let $\mathbf{g} = (g_0, \dots, g_{N-1})$ be a vector such that $g_j = g_{N-j}$ for $0 \leq j \leq m$, $g_j = 0$ otherwise and there exist $K, K' > 0$ for which $K' < \sum_j |g_j| < K$. Consider the Hartley transform

$$\hat{g}_k := \frac{1}{\sqrt{m}} \sum_{j=-m}^m g_j \left(\cos\left(2\pi \frac{jk}{m}\right) + \sin\left(2\pi \frac{jk}{m}\right) \right)$$

of the vector $\mathbf{g}$ and define the function

$$\Phi := \sum_{k=0}^{N-1} \hat{g}_k E_k.$$

Then there exist constants $\beta_T^* > 0$, $N^* > 0$, and $C > 0$ such that, for any $\beta_T > \beta_T^*$ and for any $N > N^*$, one has

$$C_\Phi(t) > \frac{1}{2} \sigma_\Phi^2 \quad \text{for all } |t| \leq C\beta_T.$$

The estimates of the previous theorems are obtained by exploiting just the vicinity of the Toda and the FPU systems and this gives results valid up to 4-th order corrections. Better estimates would be obtained by carrying out the construction of adiabatic invariants at higher orders. As is known, this involves small denominators which up to now it is not known how to control. Nevertheless numerical simulations seem to suggest that stability holds also at slightly higher order.

One model in which the small divisors problem can be overcome is the FPU alternating-mass model, a model which goes back to Born and Von Kármán [44] (see paper [19] for a recent study) and studied in the thermodynamic limit in paper [37]. One considers a model with two kind of particles with different masses, whose Hamiltonian is the following

$$H = \sum_{j=1}^N \left(\frac{p_{j,1}^2}{2m_1} + \frac{p_{j,2}^2}{2m_2} \right) + \sum_{j=1}^N V(q_{j,2} - q_{j,1}) + V(q_{j+1,1} - q_{j,2}),$$

with periodic boundary conditions, while the potential is assumed to be a quartic one, that is, $V(r) = r^2/2 + Ar^3/3 + Br^4/4$. Normal modes can be introduced to diagonalize the quadratic part H_0 of the Hamiltonian, which takes the form

$$H_0 = \frac{1}{2} \sum_k \left(|\hat{p}_k^+|^2 + \omega_+^2(k) |\hat{q}_k^+|^2 \right) + \frac{1}{2} \sum_k \left(|\hat{p}_k^-|^2 + \omega_-^2(k) |\hat{q}_k^-|^2 \right),$$

where $\hat{p}_k^\pm$ and $\hat{q}_k^\pm$, for $k = \lfloor -N/2 \rfloor + 1, \dots, \lfloor N/2 \rfloor$, are canonically conjugate complex variables and the frequencies are given by

$$\omega_\pm(k) = \sqrt{\frac{m_1 + m_2 \pm \sqrt{m_1^2 + m_2^2 + 2m_1m_2 \cos(2\pi k/N)}}{m_1m_2}}$$

The branches $\omega_-(k)$ and $\omega_+(k)$, called acoustic and optical respectively, are separate from each other, as the former takes values between 0 and $\sqrt{2/m_1}$, while the latter takes values greater than $\sqrt{2/m_2}$. Note that the distance increases with the difference in masses, and it

can also be seen that when $m_1 \gg m_2$ the frequencies of the optical band assume a nearly constant value.

Therefore, for each integer k , there are two normal modes. Two macroscopic dynamical variables can be introduced: the harmonic energy E^+ of the “optical” normal modes, defined by $E^+ = 1/2 \sum |\hat{p}_k^+|^2 + \omega_+^2(k) |\hat{q}_k^+|^2$ (i.e., the normal modes corresponding to the frequencies of the optical band), and the harmonic energy E^- of the “acoustic” normal modes, defined by $E^- = 1/2 \sum |\hat{p}_k^-|^2 + \omega_-^2(k) |\hat{q}_k^-|^2$ (i.e., the normal modes corresponding to the frequencies of the acoustic band). If the frequencies are very different, the two subsystems are weakly coupled, and therefore their energies change slowly over time. In this way, they remain correlated for a long time. An estimate of this time is given in the following theorem (Theorem 1 of the paper [37])

Theorem 6.3 *There exist constants $\beta_T^* > 0$, $N^* > 0$, $M > 2$, and $C_1, C_2 > 0$ such that, for any $\beta_T > \beta_T^*$, any $N > N^*$ and for any $m_1/m_2 > M$ one has*

$$|C_{E^\pm}(t) - C_{E^\pm}(0)| \leq C_1 \left(\frac{1}{\sqrt{\beta_T}} + \frac{m_2}{m_1} \right) \sigma_{E^\pm}^2 \quad \text{for } |t| \leq C_2 \beta_T^{S/2},$$

where $S = \lfloor \sqrt{m_1/m_2}/2 \rfloor$.

We therefore see that for sufficiently low temperatures, for a sufficiently small mass ratio, the auto correlations of the harmonic energies $E^\pm$ are close to their initial value (and therefore different from zero), for times that increase as a power of β_T : the exponent becomes larger as the ratio between the two kind of masses increase. This is due to the fact that the perturbative construction can be carried to higher orders, unlike the case of equal masses. In fact, in the case of alternating masses, the small divisors appear at a perturbative order that is increasingly higher the greater the distance between the acoustic and the optical band. And as already mentioned, this distance depends on the mass ratio.

6.2 Asymptotic Decay of the Correlations

The results cited in the previous subsection, are based on the simple estimate introduced in the paper [11]

$$C_f(t) \geq \sigma_f^2 - \frac{t^2}{2} \|\{f; H\}\|^2, \quad (6.1)$$

where $\{\cdot; \cdot\}$ denotes the Poisson bracket, while, at the variance with the notations used in the previous Sections, now $\|\cdot\|$ denotes the norm L^2 in the phase space $\mathcal{F}$ endowed with the invariant measure μ . This estimate can be thought of as an estimate from below of the slope of the curve $C_f(t)$. Although it allows us to estimate the thermalization time, this relation does not allow us to control the asymptotic nature of the decay, that is, the actual decay over long timescales. In a certain sense, it only controls the initial stage of the process of the decaying of correlation. The study of the asymptotic nature of correlation decay, requires a generalization of this relation. This generalization was given in the paper [38], where the following theorem was proved. Defined by recurrence the k th order Lie derivative of f , for $k \geq 1$, namely $f^{(k)} := \{f^{(k-1)}; H\}$, where $f^{(0)} := f$, one has

Theorem 6.4 *Let μ be a probability measure on the phase space $\mathcal{F}$, invariant for the flow generated by H , and let $n > 0$. Then, for any dynamical variable $f \in L^2(\mu, \mathcal{F})$ such that*

$f^{(k)} \in L^2(\mu, \mathcal{F})$ for all $k \leq n$, one has

$$\begin{aligned} \mathcal{C}_f(t) &= \sigma_f^2 + \sum_{k=1}^n (-1)^k \|f^{(k)}\|^2 \frac{t^{2k}}{(2k)!} \\ &\quad + (-1)^{n+1} \int_0^t dt_1 \dots \int_0^{t_{2n-1}} dt_{2n} \|f_{t_{2n}}^{(n)} - f^{(n)}\|^2. \end{aligned} \quad (6.2)$$

We therefore obtain an expansion with an integral remainder that has a definite sign, negative for even n and positive for odd n . This allows to give bounds from above and below to the decay of the correlations, once we know how to calculate (for instance, numerically) the coefficients

$$c_0 := \sigma_f^2, \quad c_n := \|f^{(n)}\|^2 \text{ for } n > 0. \quad (6.3)$$

However, a more interesting theoretical use can be obtained from this development, to establish the asymptotic character of the decay of the correlations over long timescales, in particular, to decide whether the decay is exponential or not. To this end, it is useful to consider the Laplace transform $\hat{\mathcal{C}}_f(p)$ of the correlations

$$\hat{\mathcal{C}}_f(p) := \int_0^\infty e^{-pt} \mathcal{C}_f(t) dp,$$

since a necessary and sufficient condition for the decay to be exponential is that the transform $\hat{\mathcal{C}}_f(p)$ be analytic in a half-plane $\Re p > \gamma$ with $\gamma < 0$. The connection between the coefficients c_n and the Laplace transform is established in the following theorem (Theorem 3 of the paper [38])

Theorem 6.5 *Let $\mathcal{C}_f(t)$ be analytic in t about the origin and continuous for any $t \in \mathbb{R}$. Then the following statements hold:*

1. *in the half-plane $\Re p > 0$ the Laplace transform $\hat{\mathcal{C}}_f(p)$ is analytic and there exists a positive Borel measure $\alpha(s)$ such that*

$$\hat{\mathcal{C}}_f(p) = \frac{p}{\pi} \int_0^{+\infty} \frac{d\alpha(s)}{p^2 + s^2}, \quad (6.4)$$

2. *The positive Borel measure $\alpha(s)/\pi$ itself solves the symmetric Hamburger moment problem with moments $c'_{2n} = c_n$, $c'_{2n+1} = 0$.*

We recall that the Hamburger moment problem is the problem of reconstructing a probability measure on the real line from the sequence of its momenta. The symmetric Hamburger problem is the Hamburger problem in the particular case where the momenta of odd parity vanish.

The analytic properties of the Laplace transform can then be immediately read from the properties of the measure α , since the following corollary holds (Corollary 1 of the of the paper [38])

Corollary 6.6 *$\mathcal{C}_f(t)$ decays exponentially fast as t goes to infinity if and only if the coefficients c_n are such that $\alpha'(s)$ exists and is analytic.*

Such a result opens up the possibility to study the asymptotic decay of correlations through the methods developed for the moment problem (see [1, 43]), a well-known topic of probability theory. In particular, necessary and sufficient conditions on the sequence of coefficients c_n (see a summary in Appendix A of [38]) have been found to determine whether the measure α is regular, and therefore to determine whether the decay of the correlations are exponential

or not. Since such conditions usually require knowing the entire sequence of coefficients in sufficient detail, this method could not, however, be implemented yet to obtain rigorous results.

In any case, as already mentioned, the sequence of the first coefficients can be calculated numerically, and provide an approximation for the Laplace transform (6.4) from which some indications of the character of the decay of correlations can be extracted. In the same paper [38], an application is done to the FPU system, which suggests that at low temperatures the measure α could be singular. Another application was in the paper [2], in which a quartic FPU model was studied, with a choice of potential parameters such that the model simulates a glassy solid.

Funding Open access funding provided by Università degli Studi di Milano within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Akhiezer, N.I.: The Classical Moment Problem. Oliver and Boyd, London (1965)
2. Amati, G., Meyer, H., Schilling, T.: Memory effects in the Fermi–Pasta–Ulam model. *J. Stat. Phys.* **174**, 219–257 (2019)
3. Bambusi, D., Giorgilli, A.: Exponential stability of states close to resonance in infinite-dimensional Hamiltonian systems. *J. Stat. Phys.* **71**, 569–606 (1993)
4. Bambusi, D., Kappeler, T., Paul, T.: Dynamics of periodic Toda chains with a large number of particles. *J. Differ. Equ.* **258**(12), 4209–4274 (2015)
5. Bambusi, D., Kappeler, T., Paul, T.: From Toda to KdV. *Nonlinearity* **28**(7), 2461–2496 (2015)
6. Bambusi, D., Maspero, A.: Birkhoff coordinates for the Toda lattice in the limit of infinitely many particles with an application to FPU. *J. Funct. Anal.* **270**(5), 1818–1887 (2016)
7. Bambusi, D.: Hamiltonian studies on counter-propagating water waves. *Water Waves* **3**(1), 49–83 (2021)
8. Bambusi, D., Kappeler, T., Paul, T.: De Toda à KdV. *C. R. Math. Acad. Sci. Paris* **347**(17–18), 1025–1030 (2009)
9. Bambusi, D., Ponno, A.: On metastability in FPU. *Commun. Math. Phys.* **264**(2), 539–561 (2006)
10. Benettin, G., Christodoulidi, H., Ponno, A.: The Fermi–Pasta–Ulam problem and its underlying integrable dynamics. *J. Stat. Phys.* **152**, 195–212 (2013)
11. Carati, A.: An averaging theorem for hamiltonian dynamical systems in the thermodynamic limit. *J. Stat. Phys.* **128**, 1057–1077 (2007)
12. Casartelli, M., Diana, E., Galgani, L., Scotti, A.: Numerical computations on a stochastic parameter related to the Kolmogorov entropy. *Phys. Rev. A* **13**, 1921–1925 (1976)
13. Craig, W., Groves, M.D.: Hamiltonian long-wave approximations to the water-wave problem. *Wave Motion* **19**(4), 367–389 (1994)
14. Deng, Yu., Hani, Z.: Full derivation of the wave kinetic equation. *Invent. Math.* **233**(2), 543–724 (2023)
15. Deng, Y., Hani, Z., Ma, X.: Long time derivation of the Boltzmann equation from hard sphere dynamics. *arXiv Preprint* (2024). [arXiv:2408.07818](https://arxiv.org/abs/2408.07818)
16. Diana, E., Galgani, L., Casartelli, G., Casati, G., Scotti, A.: Stochastic transition in a classical nonlinear dynamical system: a Lennard–Jones chain. *Theor. Math. Phys.* **29**, 1022–1027 (1976)
17. Douglas, J.F., Yuan, Q.-L., Zhang, J., Zhang, H., Wen-Sheng, X.: A dynamical system approach to relaxation in glass-forming liquids. *Soft Matter* **20**, 9140–9160 (2024)
18. Fermi, E., Pasta, J.R., Ulam, S.M.: Studies of nonlinear problems. In: *Collected works of E. Fermi*, vol. 2. Chicago University Press, Chicago (1965)

19. 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. *Physica D* **59**(4), 334–348 (1992)
20. Gallone, M.: Hydrodynamics of the FPU problem and its integrable aspects. Master's thesis, Università degli Studi di Padova (2016)
21. Gallone, M., Marian, M., Ponno, A., Ruffo, S.: Burgers turbulence in the Fermi–Pasta–Ulam–Tsingou chain. *Phys. Rev. Lett.* **129**(112), 114101 (2022)
22. Gallone, M., Ponno, A.: Hamiltonian field theory close to the wave equation: from Fermi–Pasta–Ulam to water waves. In: Georgiev, V., Michelangeli, A., Scandone, R. (eds.) *Qualitative Properties of Dispersive PDEs*, pp. 205–244. Springer, Singapore (2022)
23. Gallone, M., Ponno, A., Rink, B.: Korteweg–de Vries and Fermi–Pasta–Ulam–Tsingou: asymptotic integrability of quasi unidirectional waves. *J. Phys. A* **54**(30), 305701 (2021)
24. Grava, T., Maspero, A., Mazzuca, G., Ponno, A.: Adiabatic invariants for the FPUT and Toda chain in the thermodynamic limit. *Commun. Math. Phys.* **380**, 811–851 (2020)
25. Guzzo, M., Chierchia, L., Benettin, G.: The steep Nekhoroshev's theorem. *Commun. Math. Phys.* **342**(2), 569–601 (2016)
26. Henrici, A., Kappeler, T.: Global Birkhoff coordinates for the periodic Toda lattice. *Nonlinearity* **21**(12), 2731–2758 (2008)
27. Henrici, A., Kappeler, T.: Results on normal forms for FPU chains. *Commun. Math. Phys.* **278**(1), 145–177 (2008)
28. Henrici, A., Kappeler, T.: Nekhoroshev theorem for the periodic Toda lattice. *Chaos* **19**(3), 033120 (2009)
29. Henrici, A., Kappeler, T.: Resonant normal form for even periodic FPU chains. *J. Eur. Math. Soc. (JEMS)* **11**(5), 1025–1056 (2009)
30. Hiraoka, Y., Kodama, Y.: Normal form and solitons. In: *Integrability*, vol. 767. Lecture Notes in Physics. Springer, Berlin, pp. 175–214 (2009)
31. Izrailev, F.M., Chirikov, B.V.: Statistical properties of a nonlinear string. *Dokl. Akad. Nauk SSSR* **166**, 57–59 (1966)
32. Kappeler, T., Pöschel, J.: *KdV & KAM. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge*, vol. 45. Springer, Berlin (2003)
33. Kuksin, S.B.: Growth and oscillations of solutions of nonlinear Schrödinger equation. *Commun. Math. Phys.* **178**(2), 265–280 (1996)
34. Lochak, P., Neishtadt, A.I.: Estimates of stability time for nearly integrable systems with a quasiconvex Hamiltonian. *Chaos* **2**(4), 495–499 (1992)
35. Lvov, Y.V., Onorato, M.: Double scaling in the relaxation time in the β -Fermi–Pasta–Ulam–Tsingou model. *Phys. Rev. Lett.* **120**(14), 144301 (2018)
36. Maiocchi, A., Bambusi, D., Carati, A.: An averaging theorem for FPU in the thermodynamic limit. *J. Stat. Phys.* **155**, 300–322 (2014)
37. Maiocchi, A.M.: Freezing of the optical-branch energy in a diatomic FPU chain. *Commun. Math. Phys.* **372**, 91–117 (2019)
38. Maiocchi, A.M., Carati, A., Giorgilli, A.: A series expansion for the time autocorrelation of dynamical variables. *J. Stat. Phys.* **148**, 1054–1071 (2012)
39. Onorato, M., Lvov, Y.V., Dematteis, G., Chibbaro, S.: Wave turbulence and thermalization in one-dimensional chains. *Phys. Rep.* **1040**, 1–36 (2023)
40. Onorato, M., Vozella, L., Proment, D., Lvov, Y.V.: Route to thermalization in the α -Fermi–Pasta–Ulam system. *Proc. Natl. Acad. Sci.* **112**(14), 4208–4213 (March 2015)
41. Rink, B.: Symmetry and resonance in periodic FPU chains. *Commun. Math. Phys.* **218**(3), 665–685 (2001)
42. Rink, B.: Symmetric invariant manifolds in the Fermi–Pasta–Ulam lattice. *Physica D* **175**(1–2), 31–42 (2003)
43. Schmüdgen, K.: *The Moment Problem. Graduate Texts in Mathematics*, vol. 277. Springer, Cham (2017)
44. von Kármán, T., Born, M.: Über Schwingungen in Raumgittern. *Phys. Z.* **13**, 297–309 (1912)
45. Wu, B.: Rigorous derivation of the Wave Kinetic Equation for β -FPUT system. *arXiv Preprint* (2025). [arXiv:2506.02948](https://arxiv.org/abs/2506.02948)
46. Wu, B., Onorato, M., Hani, Z., Pan, Y.: Validity condition of normal form transformation for the β -FPUT system. *arXiv Preprint* (2025). [arxiv:2510.04831](https://arxiv.org/abs/2510.04831)
47. Zhang, J., Zhang, K.: Improved stability for analytic quasi-convex nearly integrable systems and optimal speed of Arnold diffusion. *Nonlinearity* **30**(7), 2918–2929 (2017)