FPU model and Toda model: a survey, a view

Giancarlo Benettin and Antonio Ponno

Abstract A huge amount of papers investigated, over more than 65 years, the Fermi-Pasta-Ulam problem. One of the leading ideas, present already in the early literature, is that the unespected regular behavior observed by the authors, quite different from the expected ergodicity, could be explained by the presence of a close hidden nonlinear integrable dynamics. This was initially searched among nonlinear wave equations, but rather soon, after the discovery of the integrability of the Toda model, it was progressively understood that Toda provides the natural integrable approximation to FPU. The aim of this paper is to provide a short updated review of the relation between the FPU dynamics and the Toda dynamics. Updated means it also includes new results, see Section 3. The paper is ideally addressed to the wide—very wide!—community of people who feel *The legacy of Carlo Cercignani*, so it includes a few introductory comments.

Foreword *It is a great honor for me to contribute to this volume in honor and memory of Carlo Cercignani. Carlo has been an important person in my life. In my scientific life of course, where Carlo, as for all of us, has been a light and a model. But beyond science, I wish to say I always felt from him warm friendship. In conferences, in particular in Ravello, I remember his lectures, but even more, I remember staying together: with him, his wife Silvana, Luigi, a few friends. Conversating, about so many different topics; walking together, when it was still possible. It was a great opportunity for me to meet Carlo, and I feel deep sincere gratitude.* [G.B.]

Giancarlo Benettin
Università di Padova, Dipartimento di Matematica “Tullio Levi-Civita”, Via Trieste 63, 35121 Padova (Italy), e-mail: benettin@math.unipd.it

Antonio Ponno,
Università di Padova, Dipartimento di Matematica “Tullio Levi-Civita”, Via Trieste 63, 35121 Padova (Italy), e-mail: ponno@math.unipd.it

1 Introduction

In 1955 Fermi, Pasta and Ulam wrote a paper [1] which was destined to have a deep influence in different branches of research.

– It started Molecular Dynamics (more generally, numerical experiments on dynamical systems), namely investigating the statistical properties of a system by numerically solving its microscopic equations of motion.
– It raised "elementary" questions in the dynamical fundations of Statistical Mechanics, which still are not clearly answered.
– It motivated a relevant branch of the theory of nonlinear oscillations, namely modern theory of nonlinear wave equations (Boussinesq, KdV...), and more generally the study of nonlinear integrability for systems with many degrees of freedom.

Hundreds of papers have been devoted to the subject, with a great variety of theoretical and numerical approaches, still far from merging in a unitary clear view.[1]

The aim of this paper is to focus on one of the main ideas, namely that the reference integrable dynamics for FPU is Toda dynamics [4], nonlinear and highly nontrivial. This view was suggested already in 1974 in [5], one of the three simultaneous papers where integrability of the Toda model was proved [6, 7, 5]. The perspective was reconsidered and widely developed in 1982 [8], but nevertheless not much exploited in the later literature, and emphasized again only recently, in a few papers; among them [9, 10, 11, 12, 13, 14].

We aim to show that viewing FPU as a perturbed Toda model, provides a unitary perspective, which can possibly give order to the complex phenomenology of FPU. We refer here to the standard, generic case of the so-called α (also called $\alpha + \beta$) FPU model. We shall not discuss instead the β-model, which is not at all close to Toda, nor we shall enter extensions to dimension two and three, although physically very important (see in particular the papers by Carati and Galgani and by Gangemi in this volume).

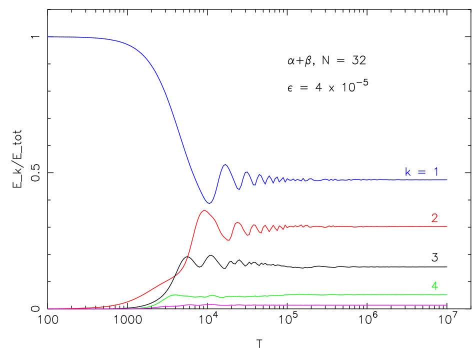
The paper is addressed, ideally, to the wide community of researchers feeling *The legacy of Carlo Cercignani*, joined together in his memory and contributing to this Conference in his honor. Carlo was indeed very interested in FPU (see [15, 16] and the comments in the paper by Carati and Galgani in this volume). Likely however, some of us are not very familiar with FPU, and so, in the remaining part of this Introduction, a very short tentative introduction to FPU is provided.

After that, Sections 2 and 3 are devoted, respectively, to the role of Toda dynamics in the long-time approach to statistical equilibrium, and to the mechanism of formation of the so-called "FPU-state", that is the first, crucial part of the FPU dynamics, where the underlying integrable Toda dynamics is particularly transparent. Section 3 also includes new results.

---

[1] See for example the collections of papers [2, 3], appeared in occasion of the 50$^{\text{th}}$ anniversary of FPU.

Fig. 1 The time average $\overline{E}_k(T)$ as a function of T, for the first few modes. FPU model with $N = 32$, $\alpha = 1$, $\beta = 2$; specific energy $\varepsilon = 4 \times 10^{-5}$, initial excitation of mode $k = 1$.



1.1 FPU in a nutshell

The specific problem Fermi, Pasta and Ulam confronted with, is the problem of energy sharing in weakly nonlinear chains of oscillators. The Hamiltonian has the form

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{N} p_i^2 + \sum_{i=0}^{N} V(q_{i+1} - q_i)\,, \tag{1}$$

V being some nearest-neighbours potential with a minimum in zero,

$$V(r) = \frac{r^2}{2} + \alpha\frac{r^3}{3} + \beta\frac{r^4}{4} + \cdots\,, \qquad \beta > 0\,. \tag{2}$$

In (1) the boundary conditions are still not specified; they are generally either fixed ends, i.e. $q_0 = q_{N+1} = 0$, like in the original FPU paper, or periodic, $q_N = q_0$, with not much difference.

The value of α, if different from zero, is irrelevant, since a trivial rescaling reports it to any prefixed value; the effective parameters determining the dynamics are indeed $|\alpha|\sqrt{\varepsilon}$, $\varepsilon = E/N$ being the specific energy, and then $\beta/\alpha^2, \ldots$ The choice of α actually fixes the energy scale, as well as the scale of β and of possibly further coefficients in (2). Througout the paper we shall use $\alpha = 1$.

Fermi, Pasta and Ulam aimed to investigate how the system reaches the statistical equilibrium, identified with the equipartition of energy among normal modes, if started very far from equilibrium, the whole energy being given to only one or two long-wavelength normal modes. This is indeed part of the general problem of the energy flowing from macroscopic to microscopic scale, so crucial in quite different fields of physics. With great surprise they found—within the time scale accessible to their computer—no equilibrium at all: the system apparently reached a stationary state, different both from the initial state and from equipartition, in which only a few normal modes significantly share energy, and dynamics looks quasi-periodic with long time almost exact recurrencies.

Consider, to be definite, the case of fixed ends. Normal modes are then

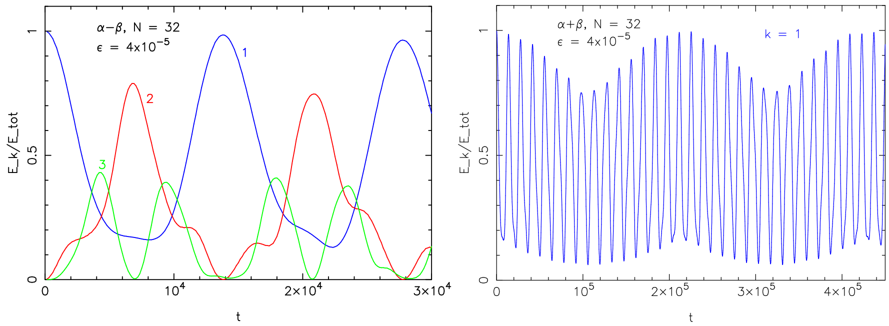


Fig. 2 The instantaneous values $E_k(t)$ as functions of t, in the same conditions as figure 1. Left: modes $k = 1, 2, 3$; right: only mode $k = 1$, for a longer time scale.

$$Q_k = \sqrt{\tfrac{2}{N+1}} \sum_{i=1}^{N} q_i \, \sin \frac{\pi k i}{N+1} \,, \qquad P_k = \sqrt{\tfrac{2}{N+1}} \sum_{i=1}^{N} p_i \, \sin \frac{\pi k i}{N+1} \;;$$

their energies E_k and frequencies ω_k are

$$E_k(P_k, Q_k) = \tfrac{1}{2}(P_k^2 + \omega_k^2 Q_k^2) \,, \qquad \omega_k = 2 \, \sin \frac{\pi k}{2(N+1)} \,,$$

and the Hamiltonian in such coordinates assumes the form

$$\tilde{H}(P, Q) = \sum_{k=1}^{N} E_k(P_k, Q_k) + \alpha U_3(Q) + \beta U_4(Q) + \cdots \,,$$

U_j being a homogeneous polynomial of degree j. For given initial data, let

$$\overline{E}_k(T) = \frac{1}{T} \int_0^T E_k(P(t), Q(t)) \, \mathrm{d}t \;;$$

statistical mechanics is based on the *ergodic hypothesis*, which implies

$$\overline{E}_k(T) \overset{T\to\infty}{\longrightarrow} \langle E_k \rangle \quad \simeq \varepsilon \,,$$

$\langle E_k \rangle$ denoting the microcanonical phase average. Figures 1 and 2 summarize the heart of the FPU results. They both refer to a model with $N = 32$ and an initial datum in which only mode $k = 1$ is excited, at small energy $\varepsilon = 4 \times 10^{-5}$. Figure 1 shows $\overline{E}_k(T)$ as function of T, for the first few modes. Quite evidently, there is no indication at all of any tendency to energy equipartition: on the contrary, an asymptotic state is apparently reached, in which only a few modes, and not at the same extent, are involved in energy sharing. Figure 2 reports, for the same dynamics, the instantaneous values $E_k(t)$ as functions of t: panel a (left) shows, on a short time scale, the behavior of modes $k = 1, 2, 3$, while panel b (right) reports only

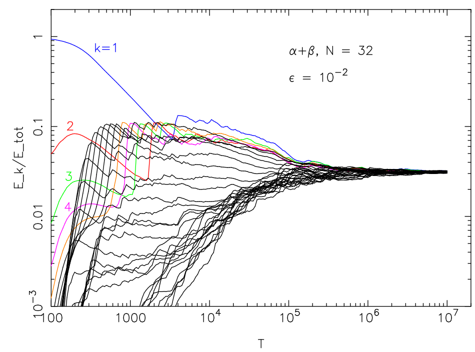


Fig. 3 Same as figure 1, at larger $\varepsilon = 10^{-2}$; all modes (logarithmic vertical axis, too).

mode $k = 1$, on a longer time scale. The presence of long time recurrencies is quite impressive, the dynamics appearing quasi periodic and thus possibly integrable. Such longer time recurrencies have been observed a few years after FPU, in [17].

1.2 The search for an underlying integrable dynamics

The suspect the dynamics is close to integrable, in the conditions studied by the authors—that is when only long-wave modes are excited, and so the discrete chain appears almost continuous—prompted the idea to approximate the FPU model with a convenient nonlinear wave equation.

The first attempt in this direction, going back to 1965, is [18], a fundamental paper which is at the basis of modern theory of nonlinear integrable wave equations. FPU appears there as a main motivation to study again, after years, the KdV equation

$$u_t = \frac{\alpha}{\sqrt{2}} u u_x + \frac{1}{24} u_{xxx} \ .$$

The progress in the field is then rapid: a couple of years later the method of inverse scattering [19] and Lax pairs [20] are introduced, and the presence of infinitely many constants of motion is established [21]. Finally, in 1971 [22], KdV is shown to be an infinite dimensional completely integrable Hamiltonian system.[2]

In parallel with the research on the integrability, numerical work clearly established that FPU is not integrable: it is enough to raise the energy, to recover the expected normal statistical behavior [24]. This is clear for example in figure 3, which differs from figure 1 only for the larger energy $\varepsilon = 10^{-2}$. It is not easy to reconcile

[2] In fact, the nonlinear wave equation which is more immediately related to FPU, if one searches for a continuum limit in which the first nonlinear and the first dispersive terms beyond the wave equation are kept, is the Boussinesq equation; a possible form is $u_{tt} = u_{xx} + 2\alpha u_x u_{xx} + \frac{1}{12} u_{xxxx}$. From Boussinesq it is possible to deduce, in a convenient limit, KdV, but the Boussinesq equation itself was soon proved, in 1973, to be integrable [23]. In [23] the connection between FPU and Boussinesq is particularly emphasized.

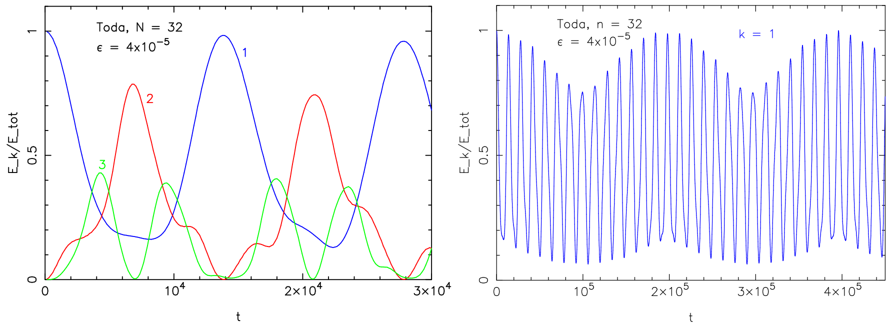


Fig. 4 Same as figure 2, for Toda rather than for FPU.

the two views; the suggestion in [23], spontaneous in that moment, was that the lack of integrability of FPU is possibly due to the discretization.

In the same years, statistical physicists become interested in the Toda model [4]. As is well known, this is a Hamiltonian system with the same form as (1), V being the Toda exponential potential

$$V_T(r) = \frac{1}{\lambda^2}(e^{\lambda r} - 1 - \lambda r) \ .$$

In 1974 the Toda model was proved to be completely integrable, remarkably in three independent papers [6, 7, 5]. Reference [5] is particularly important for FPU, because the connection with the FPU problem is there stressed. Indeed, for $\lambda = 2\alpha$ it is

$$V_T(r) = \tfrac{1}{2}r^2 + \tfrac{1}{3}\alpha r^3 + \tfrac{1}{4}\beta_T r^4 + \tfrac{1}{5}\gamma_T r^5 + \cdots \ , \qquad \beta_T = \tfrac{2}{3}\alpha^2 \ , \ \gamma_T = \tfrac{1}{3}\alpha^3 \ , \ \ldots$$

so the model has a third order contact with FPU and provides an integrable approximation better than the harmonic chain. In [5] the slow stochastization of FPU is not anymore attributed to the discretization with respect to an integrable wave equation, but to the small difference, with dominating term $\frac{1}{4}(\beta - \beta_T)r^4$, between FPU and Toda. *The reference to Toda as the best integrable approximation for FPU, is a considerable change of paradigm in the FPU problem*. For example, the distinction between long and short-wave initial excitation ceases to be important, although of course wave equations, definitely easier than a discrete model, remain useful in situations where only long waves are present.

The connection between FPU and Toda was proposed again, and emphasized, in 1982 [8]. In this paper, on the one hand, a stricking evidence is provided that the dynamics of FPU, at low energy, is hardly distinguishable from the Toda dynamics; on the other hand, the integrability of Toda is studied very constructively, and an algorithm is proposed to compute numerically the Toda actions. We shall come back to this point in Section 3. The stricking similarity of the FPU and Toda dynamics is evident in figure 4, produced under the same condition as figure 2 but for the Toda model.

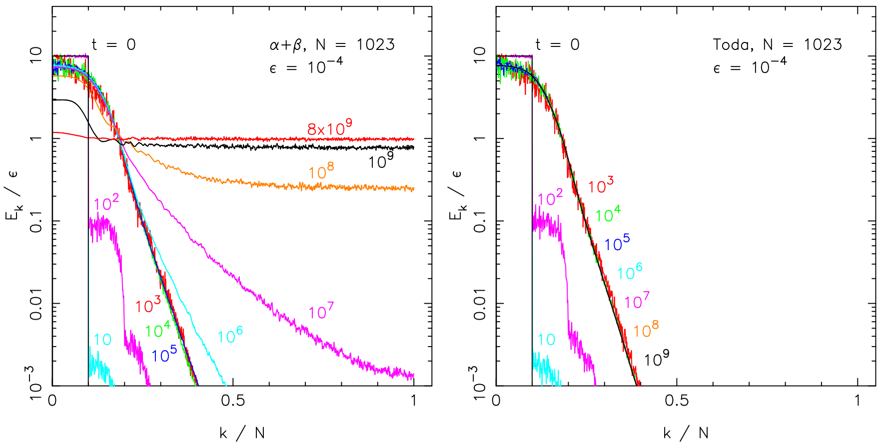


Fig. 5 The shape of the energy spectrum $\overline{E}_k(T)$ plotted vs. k/N, at selected times T (marked in the figure) in geometric progression. Left: FPU, $N = 1023$, $\alpha = 1$, $\beta = 2$, $\varepsilon = 10^{-4}$; right: the corresponding Toda model. Energy initially equidistributed among modes $0 < k/N < 0.1$, see the rectangle marked $t = 0$. Each point is the average over 24 random extractions of the initial phases.

1.3 Different phenomena at different time scales

Simultaneously with the understanding of the strong connection between FPU and Toda in [8], and completely independently, a new idea entered the literature [25, 26], sometimes referred to as the “two time-scales scenario”. The suggestion is that at least for large N the formation of the FPU state, in which energy is shared only by a small fraction of modes as in figures 1 and 2, is not the end of the story, and eventually, in a possibly *much* larger time scale, statistical equilibrium is always reached: more or less, as it happens in metastable phenomena.

To illustrate such a scenario, a good way is to look at the energy profile, i.e. the distribution of energy among normal modes, at different times; raising N is also convenient. Figure 5a shows the result for $N = 1023$ and $\varepsilon = 10^{-4}$; energy was initially equidistributed among the first 10% of normal modes, with random phases. The figure shows[3] $\overline{E}_k(T)$ vs. k/N, at selected times, marked in the figure, in geometric progression. The initial profile is the black rectangle. It can be seen that already at $T \simeq 10^3$, after a transient in which the initial discontinuity is still present, a well defined regular profile is formed, in which only some low frequency modes effectively take part to the energy sharing, the energies of the remaining ones decaying exponentially with k/N. The energy profile keeps its form nearly unchanged for a rather large time scale, $T \simeq 10^5$ or 10^6, definitely larger than the time needed to form it. Afterwords, on a much larger time scale, the dynamics slowly evolves towards energy equipartition, the high-frequency modes being progressively

[3] To be precise: time averages are computed here in a running window of width proportional to T, namely $\frac{2}{3}T \le t \le T$ (averages from the beginning are a little lazy), and moreover, to clean the curves, an average over 24 different choices of the phases is introduced.

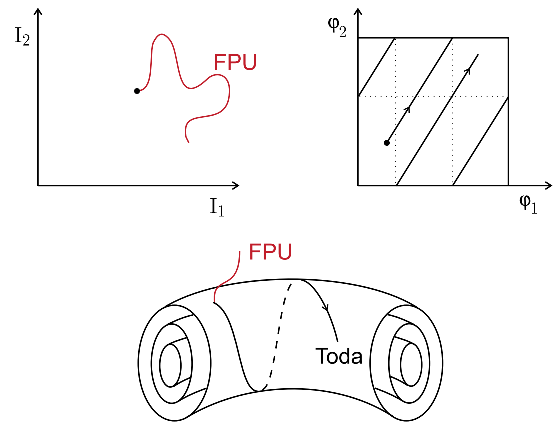


Fig. 6 Symbolically illustrating FPU as a perturbed Toda system.

involved into the energy-sharing game; in the above conditions, equipartition requires $T \simeq 10^{10}$.

The natural conjecture, at this point (natural, but explicit in the literature only after [9], in 2011), is that the first time scale is the one in which the system behaves similarly to Toda, while on larger times the difference between the two dynamics becomes evident. To confirm such an interpretation, we can repeat the above computation for the Toda model. The result is in figure 5b. Quite clearly, exactly the same profile is formed, but there is no further evolution to anything different: in Toda, only the first time scale does exist and is perpetual.

We can rephrase such a view in a better language. Toda is integrable and so, for any initial datum, the motion is confined to a torus of dimension N: actions stay constant, while angles advance linearly, and generically fill the torus. Time averages on such a motion are very partial averages in the phase space, namely averages only on the angles and not the actions. For Toda, this is all. For FPU, the lack of integrability results in an additional slow drift transversal to tori, which asymptotically (according to figure 5) results in a diffusion throughout the phase space, and makes possible statistical equilibrium. Figure 6 shows very symbolically the situation.

Both phenomena, that is the filling of the Toda torus in the dynamics common to FPU and Toda, and the diffusion across tori possibly leading to normal statistical equilibrium, are worth to be investigated. We shall start with the latter, devoting to it the next Section 2, discussing instead the former in Section 3.

2 The long-time motion across Toda tori

In the previous section we focused the attention on very special initial data, in which only a few normal modes share energy. The corresponding region of the phase space is extremely small and atypical. Measuring the rate of approach to equilibrium in such an exceptional situation can be done, see for example [9], but it is certainly more interesting to consider generic initial data, in which the energy is distributed randomly among normal modes with microcanonical distribution, and to study the drift of FPU trajectories across Toda tori in such a generic situation.

Such a study has been performed in [10], looking at the correlation time of the Toda constants of motion in the FPU dynamics. The Toda constants of motion, for a system with N degrees of freedom and fixed ends as we are dealing with, can be explicitly written by making reference to a larger system with $2(N+1)$ degrees of freedom and periodic boundary conditions, restricting the attention to skew-symmetric states such that

$$q_{N+1+i} = -q_{N+1-i} \,, \qquad p_{N+1+i} = -p_{N+1-i} \,, \qquad i = 0, \ldots, N \,, \tag{3}$$

which are easily seen to form an invariant submanifold. In turn, the constants of motion of the periodic chain are the eigenvalues of the Lax matrix $L(p,q)$ associated to the systems, or equivalently (much easier), the traces of the powers of L; see Section 3.2 for the expression of L. In the submanifold (3), precisely N constants of motion

$$F_s(p,q) = \mathrm{Tr}\, L^{2s}(p,q) \,, \qquad s = 1, \ldots, N \,,$$

are independent and nontrivial, the odd powers of L having vanishing trace. F_1 turns out to be the total energy of the system.

For given initial data, denote shortly $F_s(t)$ for $F_s(p(t),q(t))$, and let $\langle\, . \,\rangle$ denote microcanonical averaging on the initial data. The correlation function $\mathcal{G}_s$ of F_s is defined as

$$\mathcal{G}_s(t) = \frac{\langle F_s(t)\, F_s(0)\rangle - \langle\, F_s\, \rangle^2}{\langle F_s(0)^2\rangle - \langle\, F_s(0)\, \rangle^2} \;; \tag{4}$$

the decay time of such functions provides the desired time-scale of the motion transversal to tori, in a generic situation. It is worthwhile to remark that looking at the decay of correlation functions means looking at mixing, and this is fully in the spirit of the original FPU paper.[4] Practically, the microcanonical distribution in (4) is approximated by a Gaussian distribution of the normal modes coordinates P_k, Q_k, rescaled so as to fit the desired energy.

Figure 7a shows the decay of $\mathcal{G}_s(t)$, $s = 2, \ldots, 12$, for FPU with $N = 1023$, $\beta = 2$, $\varepsilon = 8 \times 10^{-4}$. Quite clearly, curves accumulate on a line $\mathcal{G}^*(t)$, which in the semi-log scale of the figure corresponds to an exponential

$$\mathcal{G}^*(t) = e^{-t/t^*} \;;$$

the inverse slope t^*, to be thought of as depending, in principle, on N, β and ε, provides the time scale of the motion transversal to tori we are looking at, ideally the mixing time that Fermi and coworkers aimed to observe. A similar accumulation is

[4] From [1]: “Instead of a gradual increase of all the higher modes, the energy is exchanged, essentially, among only a certain few. It is, therefore, very hard to observe the rate of ‘thermalization’ or mixing in our problem, and this was the initial purpose of the calculation.” For a previous study of the decay properties of the correlation functions of normal modes energies, done in the same spirit and actually inspiring [10], see [27].

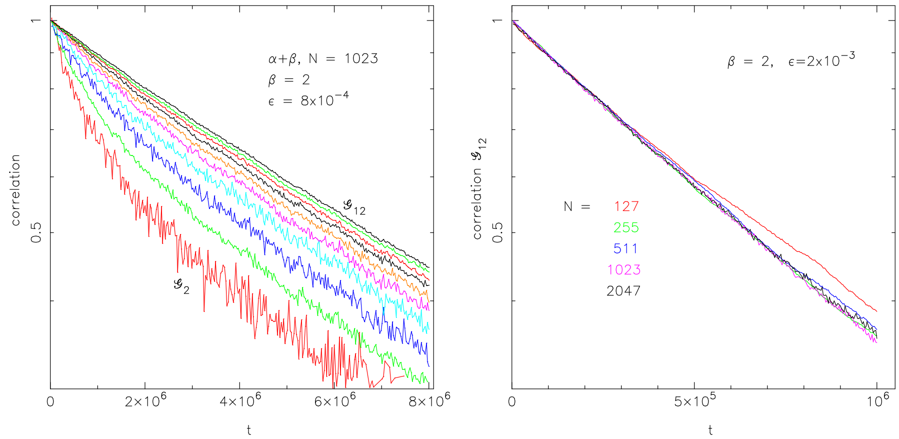


Fig. 7 Left: the time correlations $\mathcal{G}_s(t)$, $s = 2, \ldots, 12$, for FPU with $N = 1023$, $\beta = 2$, $\varepsilon = 8 \times 10^{-4}$; semi-log scale. Gaussian random extraction of 20, 000 initial data. Right: the time correlation $\mathcal{G}_{12}(t)$ of F_{12}, for $\beta = 2$, $\varepsilon = 2 \times 10^{-3}$, $N = 127, \ldots, 2047$.

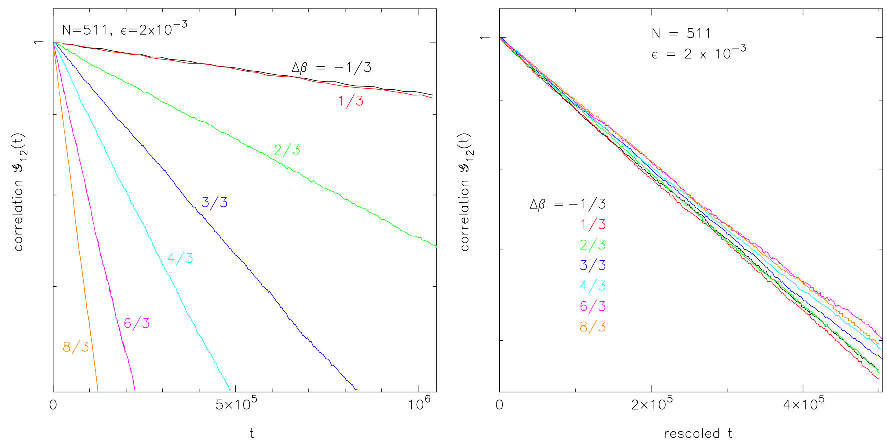


Fig. 8 The time correlations $\mathcal{G}_{12}(t)$ of F_{12}, for $N = 511$, $\varepsilon = 2 \times 10^{-3}$, $\Delta\beta = \beta - \beta_T$ as marked in the figure. Left: no rescaling. Right: time axis rescaled by a factor $\frac{9}{16}\Delta\beta^2$.

observed at different N and ε, $\mathcal{G}_{12}$ always appearing as a reasonable approximation of the limit curve $\mathcal{G}^*$. In the following, $\mathcal{G}^*$ will be identified with $\mathcal{G}_{12}$.

Figure 7b reports $\mathcal{G}^*(t)$ as function of t at fixed $\beta = 2$ and $\varepsilon = 2 \times 10^3$, for different N between 127 and 2047. The stability in N is quite evident. To investigate the dependence on β, computations have been repeated for fixed $N = 511$ and $\varepsilon = 2 \times 10^{-3}$, at different values of $\Delta\beta = \beta - \beta_T$ in the range $-1/3 \le \Delta\beta \le 8/3$. The result is in figure 8: indeed figure 8a reports the lines as they are, showing a rather dramatic dependence of the slope of $\mathcal{G}^*(t)$ on $\Delta\beta$ (although, remarkably, lines are identical for $\Delta\beta = \pm 1/3$); figure 8b shows the same curves, reported however as functions of the rescaled time

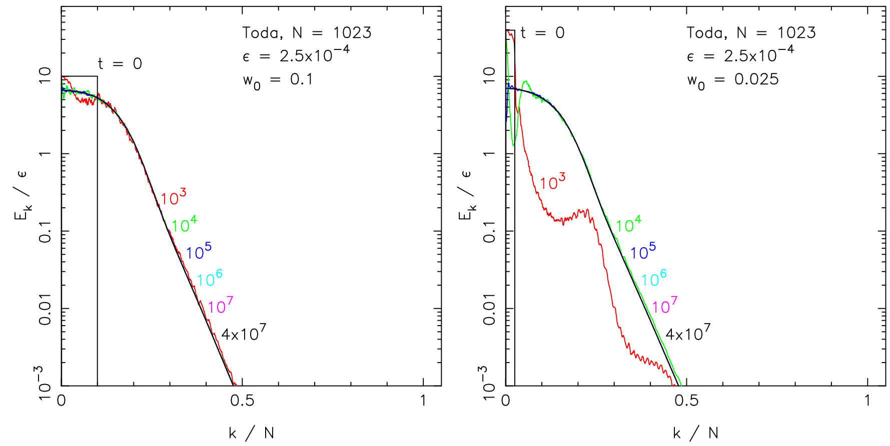


Fig. 9 Similar to figure 5b, higher $\varepsilon = 2.5 \times 10^{-4}$. Left: initial excitation of lower 10% of modes; right: lower 2.5%.

$$t' = \tfrac{16}{9} \Delta\beta^{-2} t$$

(no rescaling for $\beta = 2$, $\Delta\beta = 4/3$). Curves, although roughly, collapse into one, thus indicating a rough growth of t^* as $\Delta\beta^{-2}$. In a very similar way, one observes an ε-dependence of t^* which approximately follows the power law $\varepsilon^{-5/2}$, see [10] for details. So, the overall behavior of the correlation time looks

$$t^* \sim (\beta - \beta_T)^{-2} \varepsilon^{-5/2} .$$

Unfortunately, there is no theory at all in support of such a law. From the point of view we are exploiting in this paper, the most important point is the dependence of t^* on the difference $\beta - \beta_T$, and thus, so to speak, on the distance between FPU and Toda.[5] It is worthwhile to mention that the analysis made in [9] of the time scale of the phenomenon illustrated in figure 5a, gives exactly the same dependence on $\Delta\beta$, although with a slightly different dependence on ε, namely with exponent $9/4$ in place of $5/2$ (a slightly faster phenomenon, for small ε).

3 Investigating the FPU state

The formation of the FPU state, in the first part of the FPU dynamics common to FPU and Toda, is the process of filling a Toda torus. Should it be possible to observe it in the Toda action-angle variables $(\mathcal{I}, \varphi)$, it would appear completely trivial, i.e.

[5] The divergence of t^* for $\Delta\beta \to 0$ should not be taken literally: for vanishing β, FPU has a higher order contact with Toda, but the difference between the two Hamiltonians does not vanish, and a crossover to a different dependence of t^* on ε, faster for small ε, is expected; see [9] for a very similar situation.

$$(\mathcal{I}^0, \varphi^0) \mapsto (\mathcal{I}^0, \varphi^0 + \omega(\mathcal{I}^0)t)\,, \tag{5}$$

as for any integrable system. Observed instead in the normal modes coordinates, it appears as a progressive partial sharing of energy among some of the modes, as illustrated in figures 1–5. Understanding the formation and the properties of the FPU state means, ultimately, understanding the relation between the Toda action-angle coordinates, with their simple behavior (5), and the normal modes coordinates (P, Q), or equivalently, the harmonic action-angle variables (I, θ), related to (P, Q) by

$$P_k = \sqrt{2\omega_k I_k}\,\cos\theta_k\,, \qquad Q_k = \sqrt{2I_k/\omega_k}\,\sin\theta_k\,, \qquad E_k = \omega_k I_k\,.$$

Before entering such a delicate question, let us examine, in the next subsection, some important scaling properties of the FPU state, as described in the literature.

3.1 Scalings laws from the dynamics

Figure 9 refers to Toda and shows a process similar to the one in figure 5b. Figure 9a differs from figure 5b only for the higher energy $\varepsilon = 2.5 \times 10^{-4}$; the profiles are similar, but for higher energy the width of the spectrum gets larger, i.e. the FPU state includes a larger number of modes. In figure 9b, in addition, the initial state (the black rectangle) is narrower; quite clearly, the asymptotic situation is identical to the previous one, but the process of formation of the FPU state gets slower.

To be quantitative, we need to assign to any profile of the spectrum an "effective number" M of excited modes. This can be done in a rather standard way: if $\overline{E}_k$ is the energy spectrum at a certain time, let

$$h = -\sum_{k=1}^{N} p_k \log p_k\,, \qquad p_k = \frac{\overline{E}_k}{\sum_j \overline{E}_j}\,, \tag{6}$$

denote the so-called "spectral entropy", $0 \le h \le \log N$; then

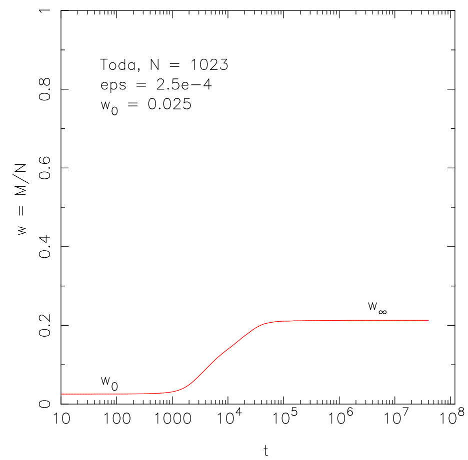
$$M = e^h\,, \qquad 1 \le M \le N\,. \tag{7}$$

In support to the definition, it is worthwhile to observe that in a situation in which exactly M modes equally share energy, while the others are at rest, the definition gives precisely M. We shall call *width* of a state the ratio $w = M/N$. Figure 10 shows, in the situation of figure 9b, the growth of w in time, from the initial value $w_0 = 0.025$ to the asymptotic value $w_\infty \simeq 0.21$.

For FPU-like initial states, i.e. with the energy shared initially by low frequency modes, the width w of the spectrum is a function, in principle, of t, N, ε and the initial width w_0. Figure 10 shows a process of the form

$$w(t, N, \varepsilon, w_0) \xrightarrow{t\to\infty} w_\infty(N, \varepsilon, w_0)\,.$$

Fig. 10 The growth of $w(T)$, same conditions as figure 9b.



In several papers, $w(t, N, \varepsilon, w_0)$ has been observed to follow some elementary scaling laws, or homogeneity relations, which are well established numerically and also partially understood theoretically, although at a very heuristic level.

The first and better established scaling law [28, 29, 30] concerns the asymptotic width w_∞, and states that if w_0 is sufficiently small, and N sufficiently large, then w_∞ is independent of both N and w_0, and it is

$$w_\infty \sim \varepsilon^{1/4} \ . \tag{8}$$

A more systematic investigation of the scaling laws satisfied by $w(t, N, \varepsilon, w_0)$ can be found in [31] for FPU, and in [14] for Toda, with identical results. The width w is there shown to satisfy three homogeneity relations, which reduce the variables from four to only one. The resulting scaling law depends on whether the phases θ_k^0 of the initially excited modes are chosen randomly or are coherent.

– For random initial phases, the law is

$$w(t, N, \varepsilon, w_0) = \varepsilon^{1/4}\, \mathcal{G}(\varepsilon^{3/8} w_0^{3/2}\, t) \,, \tag{9}$$

$\mathcal{G}$ being a suitable function of a single variable, with a sigmoid profile as in figure 10.[6] This holds also for w_0 very small, including the case of a fixed small number of excited modes; in such a case the assumption of random initial phases is obviously meaningless and in fact unnecessary.

– Instead for coherent initial phases, for example equal to each other or following some easy pattern (see [31, 14] for details), then the total energy $E = N\varepsilon$ rather than the specific energy is relevant, and (9) is replaced by

$$w(t, N, \varepsilon, w_0) = E^{1/4} w_0^{1/4} \mathcal{G}'(E^{3/8} w_0^{15/8} t) \ . \tag{10}$$

[6] The time scale for the formation of the state has been first studied in [29], in the particular (but important) case $w_0 \sim \varepsilon^{1/4}$. The result $t \sim \varepsilon^{-3/4}$ there reported is coherent with (9).

3.2 Scaling laws from Toda actions

In principle, should one know the transformation from the Toda action angle variables $(\mathcal{I}, \varphi)$ to the harmonic variables (I, θ), and conversely, one could understand everything according to the scheme

$$(I^0, \theta^0) \mapsto (\mathcal{I}^0, \varphi^0) \mapsto (\mathcal{I}^0, \varphi(t)) \mapsto (I(t), \theta(t)) \,. \tag{11}$$

Practically, in spite of the quite considerable theoretical progress [32, 33, 34, 35, 36], the relation between the Toda and the harmonic variables is not really understood, other than in the regime, very far from statistical mechanics,

$$\varepsilon \ll N^{-4} \,.$$

Numerically the situation is hard as well, but something can be done. Indeed, as already mentioned, there exists an algorithm to compute the Toda actions $\mathcal{I}_k$ in any configuration of the chain [8]. This is a little part of (11), but sufficient to support an "elementary" conjecture concerning the FPU state, namely:

In the FPU state, the number $\mathcal{M}$ of Toda actions which are substantially different from zero, that is the effective dimensionality of the Toda torus, scales as $N\varepsilon^{1/4}$.

$\mathcal{M}$ can be defined similarly to M, namely via (6) and (7), using however the Toda equivalent energy $\mathcal{E}_k = \omega_k \mathcal{I}_k$ in place of $\overline{E}_k$. The $\mathcal{E}_k$'s and $\mathcal{M}$ are constant in time, and can be computed at any time, including the initial state. This means *The scaling laws characterizing the FPU state are contained in the initial state, and stay in the nontrivial correspondence between harmonic actions and Toda actions.* This is a somehow innovative perspective, in which dynamics (integration of Hamilton equations) does not play a role.

According to (9) and (10), the conjecture is expected to hold for states in which one or a few harmonic actions are different from zero, or also a number proportional to N, with however random phases; it is instead expected to fail for coherent phases. This is precisely what we shall check in the next paragraphs A and B, devoted respectively to states including a single travelling wave and to states including a number of waves proportional to N, with either random or coherent phases. Concerning the algorithm to compute the actions, a quick account is provided in paragraph C.

A. States with a single travelling wave

Here we restrict the attention to FPU-like initial conditions, precisely to states including a single travelling wave with $k = 1$ (a preliminary account of such results, limited to smaller N, can be found in [14]). Figure 11a shows $\mathcal{M}$ vs. ε, in log-log scale, for different N ranging from 32 to $32,768$. The computed slopes, reported in the figure, indicate with *great* evidence that, for large N, $\mathcal{M}$ is indeed proportional to $\varepsilon^{1/4}$. The proportionality to N could be similarly checked, but the best way to check the conjecture is to directly plot $\mathcal{M}$ vs. $N\varepsilon^{1/4}$: if we do, see figure 11b, curves for different N exactly superimpose, and for large N the computed slope, see the data in the figure, is virtually 1. (One might observe that $N - 1$, rather than N, enters the

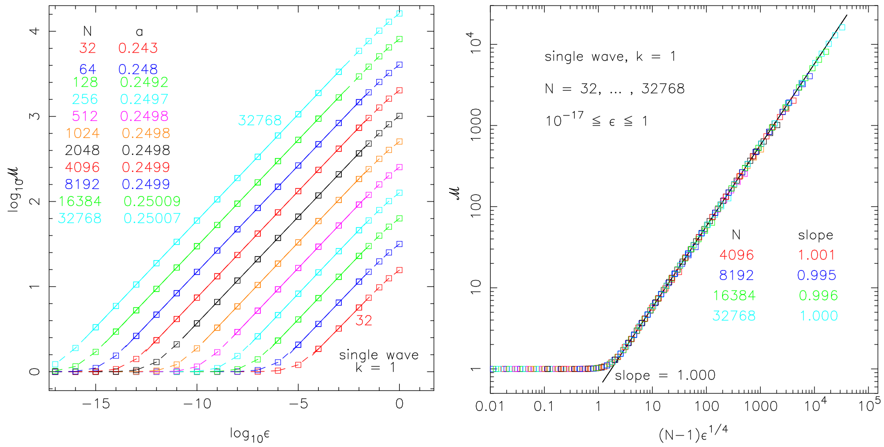
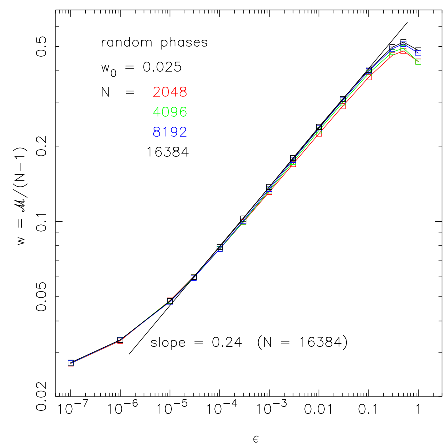


Fig. 11 Left: the effective dimensionality $\mathcal{M}$ of the Toda torus vs. ε, log-log scale, for different N. Right: the same quantity, reported as a function of $(N-1)\varepsilon^{1/4}$. Single travelling wave, $k = 1$.

Fig. 12 The effective width $\mathcal{M}/N$ of the Toda torus, when energy is equipartized among travelling waves with $0 < k/N \leq w_0 = 0.025$ and phases are random. Average on 128 random extractions of the phases.



abscissa of figure 11b. Indeed, the barycenter being at rest, the number of degrees of freedom is $N - 1$; the difference is very minor, but is visible at small N, and slightly improves the figure.)

B. States with many travelling waves

We studied FPU-like states with energy equipartized among a number of waves proportional to N, namely travelling waves with $0 < k/N \leq w_0$, small w_0. The phases of the waves are chosen randomly; more precisely, before computing $\mathcal{M}$, the Toda spectrum $\mathcal{E}_k$ is averaged on several different random extractions of the phases, actually 128 of them.

Results here are encouraging, although not as satisfying as in the above case of single wave. Figure 12 shows $\mathcal{M}/(N-1)$ vs. ε, for $w_0 = 0.025$. The expected slope

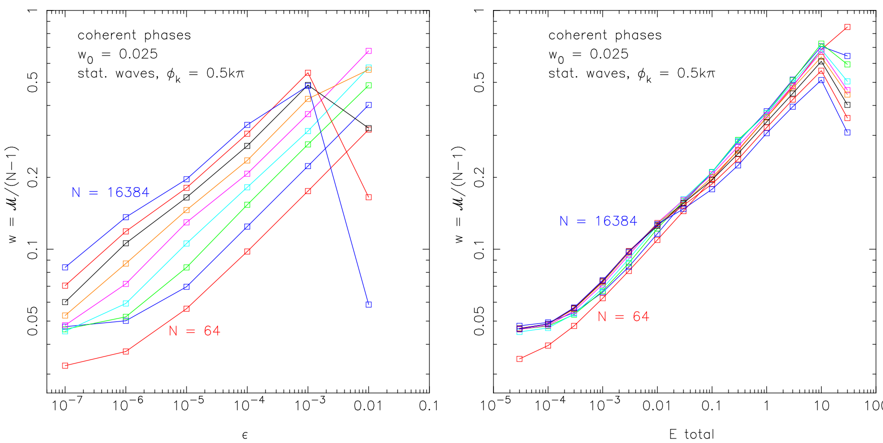


Fig. 13 Left: the fraction $\mathcal{M}/N$ vs. the specific energy ε, for $N = 64, \ldots, 16384$; stationary waves, phases $\theta_k = k\pi/2$; $w_0 = 0.05$. Right: same quantity vs. the total energy E.

was $1/4$, the computed slope, for large N, is 0.24. Preliminary computations show that by decreasing w_0 the slope gets closer to $1/4$, but a systematic study (which would require larger N) has not yet been done.

Quite remarkably, however, if we pass to coherent phases θ_k, results drastically change, reflecting the difference between (10) and (9). Indeed, the independence of $\mathcal{M}/N$ on N, strongly evident in the superposition of curves in figure 12, gets lost; see figure 13a, where $\mathcal{M}/N$ is plotted vs. ε for different N (also observe the crazy behavior at large N). The choice of phases is here $\theta_k = k\pi/2$. The stability in N is however roughly recovered if, according to (10), $\mathcal{M}/N$ is plotted vs. the total energy E; see figure 13b. Changing the way coherent phases are chosen, for example equal to each other, or following a different "easy" pattern, changes the details of the curves, but not the phenomenon.

C. On the algorithm to compute Toda actions

The algorithm to compute Toda actions is essentially as follows:

– Consider a periodic Toda chain with N particles, and let

$$L = \begin{pmatrix} b_1 & a_1 & & & & a_N \\ a_1 & b_2 & a_2 & & & \\ & a_2 & b_3 & a_3 & & \\ & & & \ddots & & \\ & & & & b_{N-1} & a_{N-1} \\ a_N & & & & a_{N-1} & b_N \end{pmatrix}, \qquad \begin{aligned} a_i &= e^{\lambda(q_{i+1}-q_i)/2} \\ b_i &= -\lambda p_i \end{aligned}$$

be the associated Lax matrix (tridiagonal periodic). Let $P(x) = \det(xI - L)$ be its characteristic polynomial, and $\Delta(x) = P(x) + 2$ be the so-called discriminant. When the system is at rest, Δ oscillates between -2 and 2 [8], as in figure 14a;

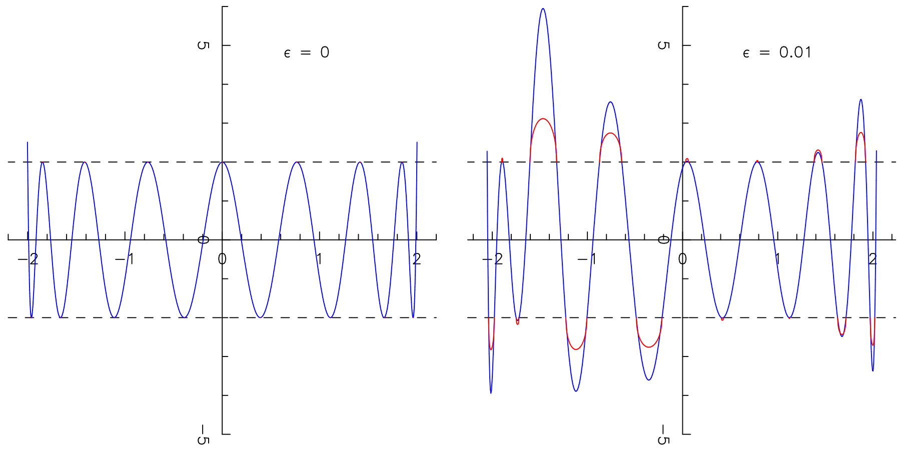


Fig. 14 Left: the discriminant $\Delta(x)$ for $N = 16$, chain at rest. Right: the discriminant (blue) and the arcs $\pm(2+\rho)$ (red), for $\varepsilon = 10^{-2}$, in a typical situation.

for positive energies instead the shape looks as in figure 14b, blue curve, with "gaps", that is intervals where maxima and minima exceed ± 2. For clarity, the figure refers to quite small $N = 16$ (observe the number of extremals is precisely $N-1$).

– Let g_k be the k-th gap. The recipe to compute the actions is:

$$\mathcal{I}_k = \frac{1}{\pi}\int_{g_k} \rho(x)\,\mathrm{d}x\,, \qquad \rho(x) = \mathrm{acosh}\,\frac{|\Delta(x)|}{2}\,.$$

Red arcs in figure 14b are the curves $\pm(2+\rho(x))$; the actions are precisely, up to a trivial factor, the areas between such curves and the lines ± 2. Notice that if $|\Delta|$ is large, then $\rho \simeq \log|\Delta|$.

Practically, applying the algorithm is not as simple. A main difficulty is that *for large N, unbelievably large numbers enter the game.* Indeed, assume tentatively that $\mathcal{I}_k$ is not far from the corresponding linear action I_k, which in turn, when energy is shared only by a few modes, is of the order E/ω_k. If k is small, then $\omega_k \sim N^{-1}$, so that $\mathcal{I}_k$ is of order N, and correspondingly *the peaks of* $\Delta(x)$ *are of the order of the exponential of* N. Large numbers are in a sense virtual, since then ρ_k and I_k are not as large, but they unavoidably enter the algorithm. Such a singularity for large N also reflects the irreducible difference between Toda actions and linear actions, unless, as in [34, 35, 36], $\varepsilon \ll N^{-4}$ is assumed. How to proceed numerically in such conditions is a little technical, and we haven't the possibility to further discuss the question here.

3.3 A conclusion?

The only conclusion we feel confident to draw, is that considering FPU as a perturbed Toda model, rather than a perturbed linear model as is more commonly done, is very fruitful, and allows to acheive a unitary view of the FPU behavior. This includes both the first time scale, with the formation of a state common to FPU and Toda, trajectories staying (almost) confined to a torus, and the second time scale, where diffusion across tori becomes important. The main point, highly non obvious, is that *FPU appears to stay close to Toda uniformly in* N; on the contrary, *both FPU and Toda look distant from the harmonic chain, no matter how small is* ε*, if* $N\varepsilon^{1/4}$ *is not small.*

Investigating the relation between normal modes, or harmonic actions, and Toda nonlinear actions, seems to us particularly important: indeed on the one hand normal modes are the elementary bricks of statistical mechanics, which play a key role in the equipartition teorem, and for ergodic-like systems are known to have a simple statistical behavior; on the other hand, Toda actions are very essential elements of the dynamics. As is not easy, the two points of view should be kept together. We are making some effort to continue this investigation that we consider promising.

References

1. E. Fermi, J. Pasta, and S. Ulam: *Studies of Non Linear Problems,* Los-Alamos Internal Report, Document LA-1940 (1955), in: *Enrico Fermi Collected Papers*, Vol. II, The University of Chicago Press, Chicago, and Accademia Nazionale dei Lincei, Roma, 1965, pp. 977-988. Later reproduced in *Lect. Appl. Math.* **15**, 143-156 (1974) and in ref. [3] below.
2. Chaos focus issue: *The "Fermi-Pasta-Ulam" problem–the first 50 years.* Chaos **15**, 2005.
3. G. Gallavotti (Ed.): *The Fermi-Pasta-Ulam Problem: A Status Report*, Lect. Notes Phys. **728**, Springer, Berlin-Heidelberg, 2008.
4. M. Toda, *Vibration of a Chain with Nonlinear Interaction*, Journ. Phys. Soc. Japan **22**, 431-436 (1967); *Wave Propagation in Anharmonic Lattices*, Journ. Phys. Soc. Japan **23**, 501-506 (1967); *Mechanics and Statistical Mechanics of Nonlinear Chains*, Journ. Phys. Soc. Japan Suppl. **26**, 109-111 (1969); *Waves in nonlinear lattice*, Progr. Teor. Phys. Suppl. **45**, 174-200 (1970).
5. S.V. Manakov, *Complete integrability and stochastization of discrete dynamical systems*, Sov. Phys. JEPT **40**, 269-274 (1974).
6. M. Hénon, *Integrals of the Toda lattice,* Phys. Rev. B **9**, 1921-1923 (1974).
7. H. Flaschka, *The Toda Lattice. II. Existence of integrals*, Phys. Rev. B **9**, 1924-1925 (1974).
8. W.E. Ferguson, H. Flaschka and D.W. McLaughlin, *Nonlinear Toda Modes for the Toda Chain,* Journ. Comput. Phys. **45**, 157-209 (1982).
9. G. Benettin and A. Ponno, *Time-scales to equipartition in the Fermi-Pasta-Ulam problem: finite-size effects and thermodynamic limit,* Journ. Stat. Phys. **144**, 793-812 (2011).
10. G. Benettin, H. Christodoulidi and A. Ponno, *The Fermi-Pasta-Ulam problem and its underlying integrable dynamics*, Journ. Stat. Phys. **152**, 195-212 (2013).
11. G. Benettin, S. Pasquali e A. Ponno, *The Fermi-Pasta-Ulam problem and its underlying integrable dynamics: an approach through Lyapunov Exponents*, J. Stat. Phys. 171, 521-542 (2018).
12. H. Christodoulidi and C. Efthymiopoulos, *Stages of dynamics in the Fermi-Pasta-Ulam system as probed by the first Toda integraly*, Mathematics in Engineering **1**, 359-377 (2019).

13. T. Goldfriend and J. Kurchan, *Equilibration of quasi-integrable systems*, Phys. Rev. E **99**, 022146-1-9 (2019).
14. G. Benettin and A. Ponno, *Understanding the FPU state in FPU-like models*, Mathematics in Engineering **3**, 1-22 (2021).
15. C. Cercignani, L. Galgani and A. Scotti, *Zero-point energy in classical non-linear mechanics*, Phys. Lett. A **38**, 403-404 (1972).
16. C. Cercignani, *Solitons. Theory and application*, Riv. Nuovo Cim. **7**, 429-469 (1977).
17. J.L. Tuck, M.T. Menzell, *The superperiod of the nonlinear weighted string (FPU) problem,*, Adv. Math. **9**, 399-407 (1972). In spite of the late publication date, results go back to 1961 (see Ulam's foreward to the FPU paper in the *Enrico Fermi Collected Papers* [1]).
18. N.J. Zabusky and M.D. Kruskal, *Interaction of solitons in a collisionless plasma and the recurrence of initial states,* Phys. Rev. Lett. **15**, 240-245 (1965).
19. C.S. Gardner, J.M. Green, M.D. Kruskal and R.M Miura, *Method for solving the Korteweg-de Vries equation*, Phys. Rev. Lett. **19**,1095-1097 (1967).
20. P.D. Lax, *Integrals of Nonlinear Equations of Evolution and Solitary Waves*, Comm. Pure Appl. Math. **21**, 467-490 (1968).
21. R.M Miura, C.S. Gardner and M.D. Kruskal, *Korteweg-de Vries equation and generalization, II. Existence of conservation laws and constants of motion*, J. Math. Phys. **9**, 1204-1209 (1968).
22. V.E. Zakharov and L.D. Feddeev, *Korteweg-de Vries Equation: a completely integrable Hamiltonian system*, Funct. Analysis Appl. **5**, 280-286 (1971).
23. V.E. Zakharov, *On stochastization of one dimensional chains of nonlinear oscillators*, Sov. Phys. JETP **38**, 108-110 (1973).
24. F.M. Izrailev and B.V. Chirikov, *Statistical properties of a nonlinear string*, Sov. Phys. Dokl. **11**, 30-34 (1966).
25. E. Fucito, F. Marchesoni, E. Marinari, G. Parisi, L. Peliti, S. Ruffo and A. Vulpiani, *Approach to equilibrium in a chain of nonlinear oscillators,* J. de Physique **43**, 707-713 (1982).
26. R. Livi, M. Pettini, S. Ruffo, M. Sparpaglione and A. Vulpiani, *Relaxation to different stationary states in the Fermi-Pasta-Ulam model,* Phys. Rev. A **28**, 3544-3552 (1983).
27. A. Carati, L. Galgani, A. Giorgilli and S. Paleari, *FPU phenomenon for generic initial data,* Phys. Rev. E **76**, 022104/1-4 (2007).
28. D.L. Shepelyansky, *Low-energy chaos in the Fermi-Pasta-Ulam Problem,* Nonlinearity **10**, 1331-1338 (1997).
29. J.A. Biello, P.R. Kramer, Y.V. L'vov, *Stages of energy transfer in the FPU model,* Proceedings of the Fourth International Conference on Dynamical Systems and Differential Equations (May 24-27, 2002, Wilmington, NC, USA), AIMS Conference Publications 2003 (special), 113-122 (2003).
30. L. Berchialla, L. Galgani and A. Giorgilli, *Localization of energy in FPU chains,* DCDS - A **11**, 855-866 (2004).
31. G. Benettin, R. Livi and A. Ponno, *The Fermi-Pasta-Ulam problem: scaling laws vs. initial conditions*, Journ. Stat. Phys. **135**, 873-893 (2009).
32. T. Nishida, *A note on an existence of conditionally periodic oscillation in a one-dimensional anharmonic lattice*, Mem. Fac. Engrg. Kyoto Univ. **33**, 27-44 (1971).
33. B. Rink, *Proof of Nishida's conjecture on anharmonic lattices,* Comm. Math. Phys. 261, 613-627 (2006).
34. A. Henrici and T. Kappeler, *Global action-angle variables for the periodic Toda lattice*, Int. Math. Res. Not. 2008, article ID rnn031, 1-52 (2008).
35. A. Henrici and T. Kappeler, *Global Birkhoff coordinates for the periodic Toda lattice*, Nonlinearity 21 2731-2758 (2008).
36. D. Bambusi and A. Maspero, *Birkhoff coordinates for the Toda Lattice in the limit of infinitely many particles with an application to FPU*, Journ. Funct. Anal. **270**, 1818-1887 (2016).