The two-stage dynamics in the Fermi-Pasta-Ulam problem: From regular to diffusive behavior

A. Ponno,[1,a)] H. Christodoulidi,[1,b)] Ch. Skokos,[2,c)] and S. Flach[2,d)]

[1]*Università degli Studi di Padova, Dipartimento di Matematica Pura e Applicata, Via Trieste 63, 35121 Padova, Italy*
[2]*Max Planck Institut für Physik komplexer Systeme, Nöthnitzer Str. 38, D-01187 Dresden, Germany*

A numerical and analytical study of the relaxation to equilibrium of both the Fermi-Pasta-Ulam (FPU) α-model and the integrable Toda model, when the fundamental mode is initially excited, is reported. We show that the dynamics of both systems is almost identical on the short term, when the energies of the initially unexcited modes grow in geometric progression with time, through a secular avalanche process. At the end of this first stage of the dynamics, the time-averaged modal energy spectrum of the Toda system stabilizes to its final profile, well described, at low energy, by the spectrum of a *q*-breather. The Toda equilibrium state is clearly shown to describe well the long-living quasi-state of the FPU system. On the long term, the modal energy spectrum of the FPU system slowly detaches from the Toda one by a diffusive-like rising of the tail modes, and eventually reaches the equilibrium flat shape. We find a simple law describing the growth of tail modes, which enables us to estimate the time-scale to equipartition of the FPU system, even when, at small energies, it becomes unobservable. © *2011 American Institute of Physics*. [doi:10.1063/1.3658620]

The present paper reports on a comparative study of the dynamics of both the Fermi-Pasta-Ulam (FPU) α-model and of the integrable Toda model. The fundamental (longest wavelength) Fourier mode of the chain with fixed ends is initially excited in both cases, with the same energy, and a fixed system size $N = 32$ is considered. Our purpose is to clearly distinguish those aspects of the dynamics of the FPU system that are regular, i.e., of integrable origin, from those that are of chaotic nature. We find that the dynamics of both systems is almost identical during a first time interval, when energy is pumped into the initially unexcited modes. This happens through a *secular avalanche* process, i.e., an almost resonant transfer of energy from the fundamental mode to the other ones, characterized by a growth of the modal energies E_k in geometric progression with squared time: $E_k \propto t^{2(k-1)}, k = 1, \dots, 31$. At the end of this first stage, the time-averaged modal energy spectrum of the Toda system stabilizes to its final profile, characterized by an exponentially decreasing tail in mode space. Such an exponential localization is well described, at low energy, by the spectrum of a *q*-breather, namely the Lyapunov continuation of the fundamental mode initially excited. The energy spectrum of the FPU system almost superposes and stays close to the Toda one during this first time stage, which is the longer, the lower the energy is. During a second time stage, the modal energy spectrum of such a long-living *quasi-state* of the FPU system slowly detaches from the Toda one and eventually reaches the flat shape predicted by the equipartition theorem of statistical

mechanics. We find that the growth of tail modes is well described, on a certain time-window, by a *diffusive-like rising of the tail modes*: $\sum_{k=22}^{31} E_k \propto t^\gamma$, with an energy dependent exponent γ. Exploiting the latter law we obtain, by extrapolation, a numerical estimate (from below) of the time-scale to equipartition of the FPU system, even when, at small energies, its actual value becomes unobservable.

I. INTRODUCTION

The present work deals with the so-called Fermi-Pasta-Ulam (FPU) problem, which consists in understanding and characterizing the process of dynamical relaxation to the micro-canonical equilibrium of nonintegrable Hamiltonian systems with many degrees of freedom, when initial conditions far from equilibrium are chosen. This issue, under investigation for more than 50 yr, is relevant to equilibrium and out-of-equilibrium statistical mechanics, as well as to many problems in condensed matter physics.

Following the original work of FPU,[1] one of the most studied systems has been, and presently is, the so-called α-model, defined by the Hamiltonian

$$H_\alpha(q,p) = \sum_{n=0}^{N-1} \left[\frac{p_n^2}{2} + \frac{(q_{n+1} - q_n)^2}{2} + \alpha \frac{(q_{n+1} - q_n)^3}{3} \right], \quad (1)$$

where α is a parameter and fixed ends conditions are chosen: $q_0 = q_N = p_0 = p_N = 0$. Such a Hamiltonian model with $N - 1$ degrees of freedom can be regarded either as a space-discretized version of a nonlinear string (as FPU did) or as the expansion around the equilibrium configuration of a one-dimensional system of pairwise interacting identical particles and is commonly referred to as a nonlinear oscillator chain.

---

a)Electronic mail: ponno@math.unipd.it.
b)Electronic mail: echrist@math.unipd.it.
c)Electronic mail: hskokos@pks.mpg.de.
d)Electronic mail: flach@pks.mpg.de.

FPU numerically integrated the equations of motion of system (1), given by $\dot{q}_n = \partial H_\alpha/\partial p_n$, $\dot{p}_n = -\partial H_\alpha/\partial q_n$, and they chose initial conditions of the form

$$q_n(0) = A \sin\left(\frac{\pi n}{N}\right), p_n(0) = 0, \qquad (2)$$

A being a parameter, thus initially exciting the fundamental (longest wavelength) Fourier mode of the chain. FPU considered system sizes $N = 16, 32, 64$. The surprising and for a long time debated result they got, namely the absence of any complete relaxation of the system to a state of equilibrium compatible with the laws of statistical mechanics (at least up to the observation times then available) was named, after them, *the FPU paradox*. As first conjectured in Refs. [2] and [3] and then clearly shown in Ref. [4], the paradox is due to a separation of the relevant time-scales of the problem, which in turn depends on the value of the total energy E and on the size N of the system. More precisely, it has been understood that energy is first transferred to shorter wavelength modes, which, within a first characteristic time-scale $T^{qs}(E,N)$, gives rise to a *quasi-state* (as named by FPU, later also referred to as "metastable state"[2,3] or "natural packet"[4] in the literature). This state is characterized by a modal energy spectrum displaying an exponentially decreasing tail. On a second, larger time-scale $T^{eq}(E,N)$, the modal energy spectrum approaches an almost flat shape, i.e., the system does reach the equilibrium state characterized by modal energy equipartition, as predicted by the laws of equilibrium statistical mechanics. The FPU paradox appears then when, depending on E and N, a strong inequality $T^{qs} \ll T^{eq}$ holds, such that the approach to equilibrium becomes unobservable on the available computation times.

Up to now, the short term dynamics of the system has been thoroughly investigated, both from a numerical and an analytical point of view, also for initial conditions more general than (2), and both the time-scale T^{qs} and the slope of the exponential tail of the energy spectrum have been more or less quantified. In particular, the localization of the energy in Fourier space, i.e., the existence of the quasi-state, has been explained essentially in two ways. One approach consists in looking for a particular stable solution of the equations of motion that is close to the actual solution of the initial value problem and displays a localized energy spectrum. One way to get such a solution consists in looking for the Lyapunov continuation of the initially excited mode or *q*-breather.[5–9] The other approach consists instead in looking for a suitable integrable system whose dynamics is close, on a certain time-scale, to that of the actual system and leads to the Korteweg-de Vries equation and its truncations in mode space[9–15] (Ref. [9] contains a first tentative comparison between the two approaches). On the other hand, for the α-model (1), there exists an approximating integrable system given *a priori*, namely the Toda model (see Hamiltonian (3) below), and a comparative study of the two models has been performed, e.g., in Refs. [16–20].

Notwithstanding the results just quoted, both a systematic study of the dynamics of problems (1) and (2) on short times and a deeper understanding of the possible existing links between the aforementioned theoretical approaches are still lacking. Such a gap is partially covered in the present paper. We explain the detailed resonance mechanisms ruling the energy

cascade on short term together with a simple characterization of the quasi-state. Moreover, we compare, for the same initial conditions of the form (2), the dynamics of the FPU chain with that of the integrable Toda model[21] defined by the Hamiltonian

$$H_T(q,p) = \sum_{n=0}^{N-1}\left[\frac{p_n^2}{2} + \frac{e^{2\alpha(q_{n+1}-q_n)} - 1}{4\alpha^2}\right], \qquad (3)$$

with fixed ends at $n = 0$ and $n = N$. The integrability of the latter Hamiltonian system was first suggested in Ref. [22] and then proved in Refs. [23] and [24]. Expanding the exponential in Eq. (3) and taking into account the boundary conditions, one can write

$$H_\alpha(q,p) = H_T(q,p) - \sum_{n=0}^{N-1}\sum_{r\geq4}(2\alpha)^{r-2}\frac{(q_{n+1}-q_n)^r}{r!}, \qquad (4)$$

i.e., the FPU Hamiltonian can be regarded as a perturbation of the Toda one. The call for such a comparative study for the evolution towards an equilibrium comes from the necessity to distinguish those phenomena that are of integrable, i.e., nonchaotic, nature, from those that are instead signatures of nonintegrability, i.e., chaos. In the present paper, we show that the FPU quasi-state coincides well with the Toda equilibrium state (see below, Sec. II and III; see also Ref. [16]) and is therefore a manifestation of closeness to integrability. Chaos becomes relevant in the α-model only *after* the saturation to the quasi-state. We also show that the quasi-state, i.e., the Toda equilibrium state, is well described, at low energy, by a time-periodic orbit or one-dimensional torus, namely the q-breather associated to the first mode.

For what concerns the second time-scale T^{eq}, its numerical estimate has been given, e.g., in Ref. [18] for the α-model, and in Refs. [25–27] for the so-called β-model; the two models are treated together and compared in the more recent Ref. [28], where accurate numerical estimates of T^{eq} are given. Yet, a detailed description of how the tail of the modal energy spectrum raises from zero to the equipartition level has not been provided so far. In the present paper, we perform a fit of the numerical data relative to the modal energy spectrum of the α-model during the second part of its evolution, which yields a simple law describing the motion of the tail. This allows us to measure the second time-scale T^{eq} on times which can be much shorter than T^{eq} (notice that a direct measurement of the equipartition time-scale T^{eq} becomes impossible when it exceeds the available computation times).

We find that at low energy (when the quasi-state is a one-dimensional Toda torus) the estimated time-scale T^{eq} undergoes a transition from power law to stretched exponential of the inverse of the energy. As a consequence, the quasi-state of the α-model becomes effectively stable: in practice, the approach to equipartition becomes unobservable, which explains the FPU paradox.

Let us stress that the existing literature on the FPU problem is huge, encompassing more than 50 yr of research on an unsolved problem. We thus defer the reader to some of the existing high-quality reviews on the subject.[29–31]

The paper is organized as follows. In Sec. II, after introducing some notations and a few relevant quantities, we

state our main results on the evolution towards equilibrium of systems (1) and (3) with initial conditions (2). In Sec. III, the phenomenology of the problem is illustrated by the numerical results and the fits with the analytical predictions. In Sec. IV, the theory leading to the analytical estimates is reported. Finally, Sec. V is devoted to concluding remarks.

II. MAIN RESULTS

We integrated the equations of motion associated to the Hamiltonians (1) and (3) by means of a symplectic algorithm, namely the Yoshida kinetic-potential splitting algorithm of fourth order,[32,33] with a time-step kept fixed to 0.05 all over the explored energy range.[34] The size N of the system considered in the present paper is kept fixed to $N = 32$, and no systematic exploration of size dependent phenomena is performed. The chosen value of the nonlinearity parameter is $\alpha = 0.33$.[35]

We follow the evolution of the harmonic energies E_k of the Fourier modes of the system, also coined *modal energies*. These quantities are defined as

$$E_k(t) \equiv \frac{P_k^2(t) + \omega_k^2 Q_k^2(t)}{2}, (k = 1, ..., N-1), \quad (5)$$

where $(Q_k, P_k) = \sqrt{2/N} \sum_{n=1}^{N-} (q_n, p_n) \sin(\pi k n/N)$ are the Fourier coordinates and

$$\omega_k = 2 \sin\left(\frac{\pi k}{2N}\right) \quad (6)$$

is the dispersion relation of the linearized problem ($\alpha = 0$). For the FPU model (1), the total energy $E_\alpha = H_\alpha(q(0),p(0))$ in terms of the initial condition (2) reads

$$E_\alpha(A) = N\left(\frac{A\omega_1}{2}\right)^2 \simeq \frac{\pi^2 A^2}{4N}, \quad (7)$$

whereas, for the same initial condition, the total energy $E_T = H_T(q(0),p(0))$ of the Toda model (3) reads[36]

$$E_T(A) = \frac{N[I_0(2\alpha A\omega_1) - 1]}{4\alpha^2}, \quad (8)$$

where $I_0(x)$ is the first modified Bessel function of order zero.[37] All the comparative runs were performed at the same value of the total energy E for the two models, which required to consistently choose different initial amplitudes A and A' in the two cases: $E = E_\alpha(A) = E_T(A')$.[38]

For $\alpha = 0$, the modal energies E_k do not evolve in time both for the FPU and the Toda model, since they are constants of motion: $E_1 = E$ and $E_{s \geq 2} = 0$, for all times. When $\alpha > 0$, the modal energies are no longer constants of motion for both systems, and evolve in time. Now, for the nonintegrable FPU system, one expects that each time-averaged modal energy

$$\bar{E}_k(t) \equiv \frac{1}{t} \int_0^t E_k(s)ds \quad (9)$$

converges, as $t \to \infty$, to its final expectation value,[39] which is approximately equal to the specific energy E/N of the system.

On the other hand, for the integrable Toda model, the modal energies are quasi-periodic functions of time (with a maximum number $N-1$ of rationally independent

frequencies), which ensures that their time-averages (9) converge to some limit as $t \to \infty$. Such a limit defines what we refer to as the *equilibrium state of the Toda system*. In the sequel, we use and plot the normalized modal energies

$$e_k(t) \equiv \frac{E_k(t)}{E} \quad (10)$$

and their time-averages $\bar{e}_k(t) = \bar{E}_k(t)/E$, and we refer to the plots of e_k vs. k and of $\bar{e}_k$ vs. k as the instantaneous and averaged modal energy spectrum (m.e.s.), respectively. The initial values of the normalized modal energies are then $e_1(0) = 1$ and $e_1(0) \simeq 1 - \alpha^2 E/N$, for the FPU-$\alpha$ and Toda model respectively, whereas $e_k(0) = 0$ for $k = 2, ..., N-1$. The expected equipartition value of the $\bar{e}_k$'s is approximately $1/N$.

We report in Fig. 1 nine snapshots of the FPU and Toda averaged m.e.s. at times $t = 0, 10, ..., 10^8$, all referring to the same run at $E = 1$. We note that up to time $t = 10^3$, the two spectra almost perfectly superpose, whereas beginning with $t = 10^4$ the tail of the FPU m.e.s. starts to rise while the whole Toda m.e.s. no longer evolves. Finally, at $t = 10^8$, the FPU m.e.s. displays an almost perfect flat shape, a signature of statistical mechanical equilibrium.

A. Stage I: Secular avalanche

On a very short time-scale, the initially excited first mode ($k = 1$) looses its energy, while the energies of all the other modes increase with a mode-dependent power of time. For both the FPU and the Toda models, the time-evolution of the instantaneous m.e.s. is described by the following law:

$$e_1(t) = 1 - \frac{\omega_1\omega_2(\mu t)^2}{8}, \quad (11)$$

$$e_k(t) = (\mu t)^{2(k-1)} c_k^2 (2 \leq k \leq N-1), \quad (12)$$

where

$$\mu \equiv \alpha\sqrt{\frac{E}{N}} \quad (13)$$

is a small parameter and the time-independent coefficients c_k are recursively computed by solving the nonlinear equations

$$c_1 = 1; gt c_k = \frac{\omega_k}{4(k-1)} \sum_{q=1}^{k-1} c_{k-q} c_q (2 \leq k \leq N-1). \quad (14)$$

These results follow from the smallness of the difference $\omega_k - k\omega_1$, and essentially consist of a resonant transfer of energy, where the secular (proportional to t^2) growth of the modal energy e_2 in turn implies the law $e_k \propto t^{2(k-1)} (k \geq 2)$. We call this stage of the cascade *the secular avalanche stage* (see Sect. III A). The relevance of the acoustic resonance as a driving mechanism of the cascade in the FPU problem was first stressed in Ref. 40.

B. Toda equilibrium and quasi-state

Until a characteristic, energy-dependent time-scale

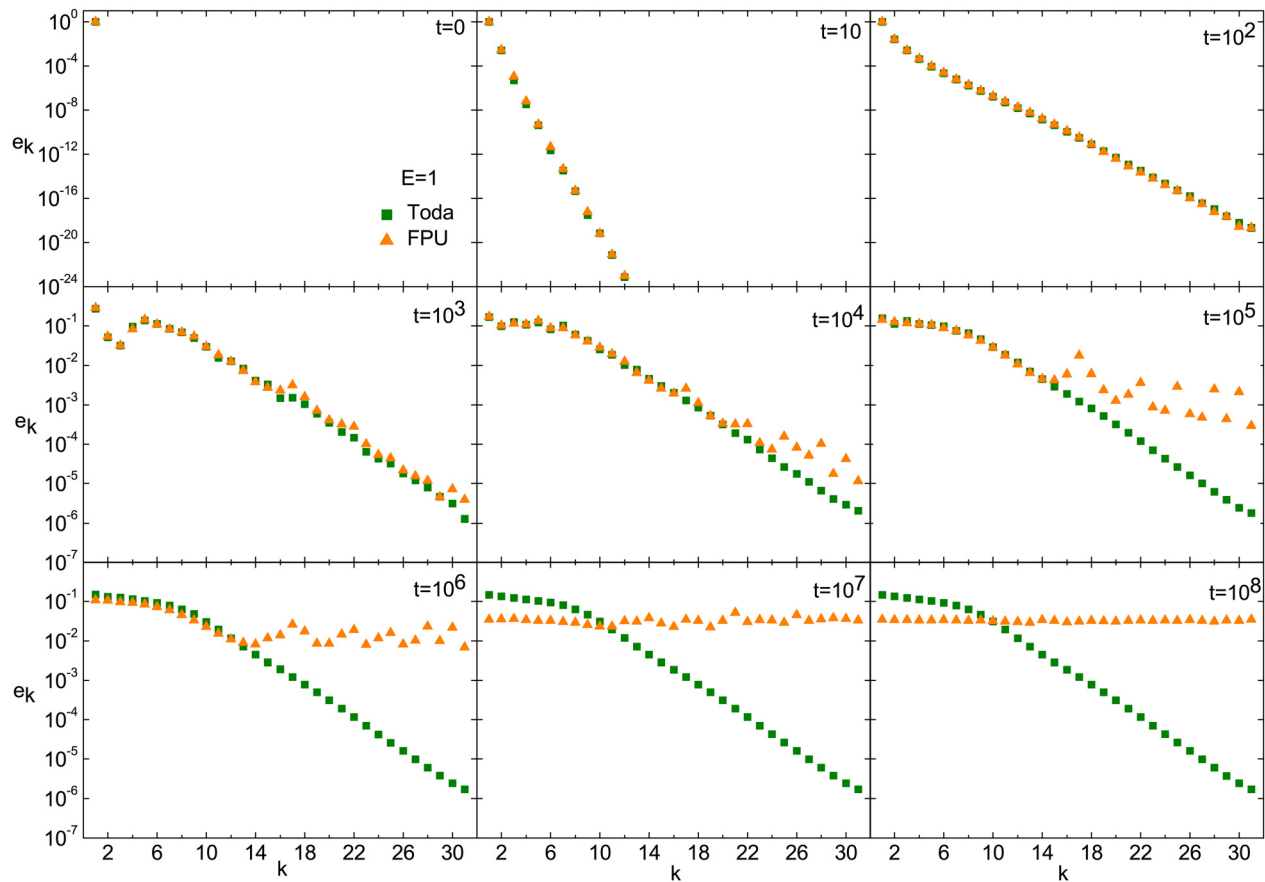$$T^{qs} \approx \frac{N}{\mu}, \quad (15)$$

FIG. 1. (Color online) $\overline{e}_k$ vs. k (log-linear scale) for FPU-α (Eq. (1)) [orange triangles] and Toda (Eq. (3)) [green squares] models, $E = 1$, $N = 32$, $\alpha = 0.33$. From the upper left corner to the lower right one $t = 0, 10, \ldots, 10^8$. Note the change of vertical scale from $t = 10^3$ on last six panels.

the decrease of e_1 and the resonant growth of $e_{k \geq 2}$ due to the secular avalanche effectively slow down in the FPU case and actually stop in the Toda case. The time-averaged m.e.s. of both the FPU and Toda models converge to an exponentially localized profile, which, for energy values up to about 10^{-2}, is well described by the expression

$$\overline{e}_k \simeq k^2 \left(\frac{\alpha^2 E N^3 \overline{e}_1}{\pi^4} \right)^{k-1} \overline{e}_1, \tag{16}$$

with $\overline{e}_1$ determined by the normalization condition

$$\sum_{k=1}^{N-1} \overline{e}_k = 1. \tag{17}$$

The m.e.s. profile (16) is that of the q-breather associated to mode $k = 1$, i.e., the Lyapunov continuation of the first mode to the nonlinear system.[5] The single energy-dependent frequency of this periodic orbit, or one-dimensional torus, is

$$\Omega_1(\mu) \simeq \omega_1 + \mu^2 \overline{e}_1 \frac{\omega_1 \omega_2}{8(2\omega_1 - \omega_2)} + \ldots. \tag{18}$$

Though the agreement between formula (16) and the numerical m.e.s. becomes the less accurate the higher is the energy, we numerically observe that the m.e.s. of both models keep on saturating to (almost) the same exponentially localized profile for all the explored energy values up to $E = 1$. Such

an agreement suggests a simple interpretation to the nature of the quasi-state guessed by FPU: *the quasi-state of the α-model coincides with the final equilibrium state of the Toda model*. It must be stressed that strong indications to such a conclusion can be found already in Refs. 16 and 20.

The analytical predictions (11) and (12) and (16) were obtained by computing a particular resonant normal form Hamiltonian of the system to leading order. The same normal form is the starting point of heuristic considerations concerning the observed stability of the q-breather m.e.s. (16) on long times.

C. Stage II: Diffusion of tail modes

After saturation, the FPU m.e.s. slowly continues to evolve and detaches from the reference, stationary Toda m.e.s. This behavior becomes actually unobservable below some energy threshold E_c ($E_c \simeq 0.1$ for $N = 32$), at least up to the available observation times (10^8 to 10^9). Above this threshold, there is an evident raising of the m.e.s. tail. We consider the evolution of the normalized total energy residing in the last third of the modes, namely

$$\eta(t) \equiv \sum_{k=22}^{31} e_k(t). \tag{19}$$

The choice of the last third of modes in defining the tail of the m.e.s. is of empirical nature and agrees, when the quasi-

state is well described by the q-breather, with the definition of tail given in Ref. 8 and based on resonance arguments. We find that, after a transient time approximately coinciding with the time-scale T^{qs}, the quantity $\bar{\eta}(t)$ starts to increase with time as a power law

$$\bar{\eta}(t) \sim Dt^{\gamma}, \qquad (20)$$

within an energy depending time-window. Both the exponent γ and the coefficient D depend on the energy. The diffusion exponent $\gamma(E)$ roughly displays a step-like behavior: $\gamma(E) \simeq 0$ for $E \ll E_c$ and $\gamma(E) \simeq 1$ as $E \gg E_c$. The energy dependence between the two limits is not smooth: the exponent $\gamma(E)$ strongly fluctuates around E_c, the latter value being more properly identified as the center of a transition interval rather than an actual sharp threshold. This is possibly a signature of anomalous diffusion processes, characterizing the dynamics of tail modes (recall that $e_k = E_k/E$ depends quadratically on the modal coordinates P_k and Q_k). Depending on the value of the energy E, one can observe either subdiffusion ($\gamma < 1$) or super-diffusion ($\gamma > 1$). An analytical derivation of the law (20) is currently missing.

We then extrapolated the time-scale to equipartition T^{eq}, defined as the time necessary for its time average $\bar{\eta}$ to reach its approximate equipartition value $1/3$,

$$T^{eq} \equiv (3D)^{-1/\gamma}. \qquad (21)$$

Both D and γ can be therefore obtained on relatively short times. These values lead to lower bounds for the actual equipartition time, which, for example, can reach the unconceivable value of 10^{1000} at $E = 0.01$. We find that $T^{eq} \propto E^{-a}$ at higher energies (from $E = 0.5$ up to 2), with a close to 3, whereas $T^{eq} \sim \exp(c/E^b)$ for low energies (in the range 0.01 to 0.1), with $b \simeq 0.7$.

In this second stage of the dynamics, we also measured the *erosion* of one of the first integrals of the Toda model independent of the Hamiltonian, computing it along the FPU dynamics. It turns out that, above the energy threshold E_c, when the diffusive-like raising of the tail modes becomes observable, such a quantity undergoes a drift starting at times of the order of T^{qs}. This behavior indicates that the relaxation to equilibrium of the FPU system is characterized by the drift of the Toda integrals, which become quasi-invariants of the FPU system at low energy.

III. PHENOMENOLOGY AND NUMERICS

We now come to a detailed description of the phenomenology of the problem, presenting our numerical results and their comparison with the theoretical predictions.

A. Stage I: Secular avalanche

The secular avalanche stage is shown in Fig. 2, where the log-log scale plot of $e_k(t)$ vs. t at $E = 0.01$ is reported, both for the Toda and the FPU-α system, together with the lines corresponding to the theoretical laws (11) and (12). The values of the mode number are restricted to $k = 1, \ldots, 10$ just to improve the readability of some details in the figure. One
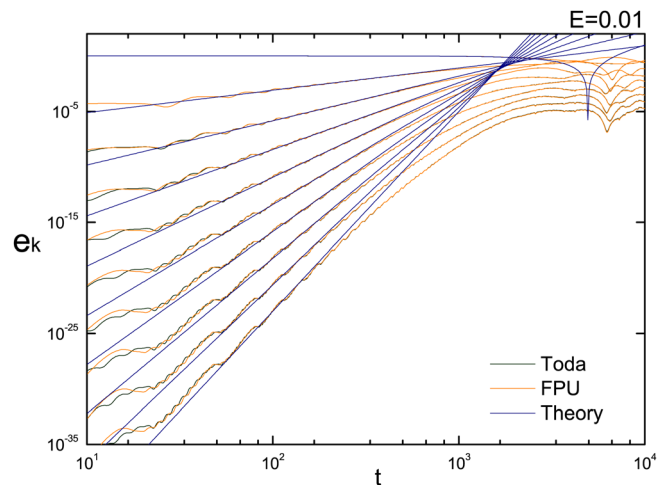

FIG. 2. (Color online) $e_k(t)$ vs. t (log-log scale) for $k = 1, \ldots, 10$ at $E = 0.01$; black and orange curves refer, respectively, to the Toda and the FPU numerically computed energies, while blue lines correspond to the theoretical laws (11) and (12).

can note the almost perfect superposition, mode by mode, of the Toda and FPU numerical modal energies, which are both well described, on the very short term, by the laws (11) and (12). The same good agreement of the latter theoretical law with the numerical data is found all over the energy range 10^{-2} to 1, though further figures are not reported.

The law $e_k(t) \propto t^{2(k-1)}$ for low modes is due to the smallness of the difference $k\omega_1 - \omega_k$, i.e., to their almost complete resonance. Indeed, mode $k = 2$ is forced in resonance by the second harmonic of mode $k = 1$, so that its amplitude grows linearly and its energy grows quadratically with time: $e_2 \propto t^2$. Mode $k = 3$ is in turn forced in resonance by the product of amplitudes of modes $k = 1$ and $k = 2$, which has a factor t in front, so that $e_k \propto (t^2)^2 = t^4$, and so on. In this way, a resonant cascade of energy transfer sets in, which causes the growth of the initially unexcited modes one after the other. Notice that for high modes, the difference $\omega_k - k\omega_1$ is not small, but their amplitude is initially zero, and this (as shown in Sec. IV) ensures the validity, also for them, of the approximation leading to the secular avalanche laws (11) and (12).

Let us denote by $T_k^{sa}(E, N)$ the mode-dependent characteristic time of validity of the secular avalanche laws (11) and (12). For the first mode, this time can be estimated by simply requiring that the leading order estimate (11) yields $e_1 \geq 0$. Setting $e_1 = 0$ in Eq. (11), one gets $T_1^{sa} = 2\sqrt{2}/(\mu\sqrt{\omega_1\omega_2}) \simeq 2N/(\pi\mu)$. Moreover, in the regime of weak mode-coupling explored here, almost all the energy is retained by the harmonic part of the Hamiltonian, so that the sum $\sum_{k=1}^{N-1} e_k$ is close to its initial value (one) for all times, which in turn implies $e_k \leq 1$ for all k's. Setting $e_k = 1$ in Eq. (12), one gets $T_k^{sa} = c_k^{\frac{1}{k-1}}/\mu$ for $k \geq 2$. We finally notice that $T_k^{sa} \sim T^{sa} \equiv 0.2N/\mu$ for large values of k. By exploiting such asymptotic behavior, the law (12) reads, for $k \gg 1$, $e_k(t) = e^{-2(k-1)\sigma(t)}$, where $\sigma(t) = -\ln(t/T^{sa})$. A logarithmic time-dependence of the slope $-\sigma(t)$ of the m.e.s on the short term was first predicted, with methods based on continuum approximations, in Refs. 2 and 41. It is interesting

to notice that the time-scale (15), derived below (in Sec. IV E) from the point of view of mode dynamics, was first determined as the break-down time of the solution of the nonlinear wave-equation approximating the FPU dynamics in the continuum, zero dispersion limit.[42,43]

B. Toda equilibrium and quasi-state

At the end of the secular avalanche, the time-averaged m.e.s. ($\bar{e}_k$ vs. k) of both the FPU-α and the Toda models saturate to one and the same exponentially localized profile—the FPU quasi-state. For larger times, the Toda time-averaged m.e.s. no longer evolves, whereas the FPU one keeps on evolving, the more slowly the lower is the value of the energy E. In order to illustrate this, in Fig. 3 we report both the time-averaged Toda and FPU m.e.s. and the FPU instantaneous one (log-linear scale). The nine panels refer to three values of the energy $E = 0.01$, 0.1, 1 at times $t = 10^4$, 10^6, 10^8. The value of the energy E is fixed in each row, whereas that of the time is fixed in each column. At $E = 0.01$ (first row), the Toda and FPU spectra are almost perfectly superposed at any time, the only observable difference being the small fluctuations $\bar{e}_k - e_k$ of the FPU energies of the last (highest) few modes. At $E = 0.1$ (second row), the FPU spectral fluctuations $\bar{e}_k - e_k$ become observable for any k, and increase with time. Also at $t = 10^6$, the tail of the FPU m.e.s. starts to rise detaching from that of the Toda. Finally, at

$E = 1$ (third row) (see also Fig. 1) one observes, in the FPU m.e.s., even larger fluctuations, the raising of the tail, and the relaxation to equipartition.

Fig. 3 also shows that at low energy, when the FPU model takes a long time to relax to equilibrium, its motion stays close to the invariant torus of the Toda model corresponding to (almost) the same initial condition. The slow variation along the FPU actual trajectory of the Toda invariants is at the basis of the FPU paradox. This was first conjectured in Ref. 16, where a phenomenological picture of the Toda dynamics based on inverse scattering is presented. In particular, in the latter reference, it is shown that at small energy, the Toda dynamics corresponding to the initial excitation of a single mode takes place on a one-dimensional torus, and that such a periodic orbit must be the Lyapunov continuation to the nonlinear problem of the linear normal mode solution.

With this in mind, in Fig. 4, the time-averaged m.e.s. (in log-linear scale) of the Toda model at $E = 0.01$ is reported for three different values of the time $t = 10^4$, 10^6, 10^8, using different symbols. Notice that the three numerical spectra at different times almost perfectly superpose, which shows that the time-averaged m.e.s. we are looking at is stationary. The theoretical m.e.s. (16) (with $\bar{e}_1$ computed from Eq. (17)) is also plotted in Fig. 4 (asterisks). The agreement with the numerical data is very good and holds for energies lower than $E = 0.01$. On the other hand, in Ref. 16, it is also shown that the dimension of the Toda torus corresponding to a single
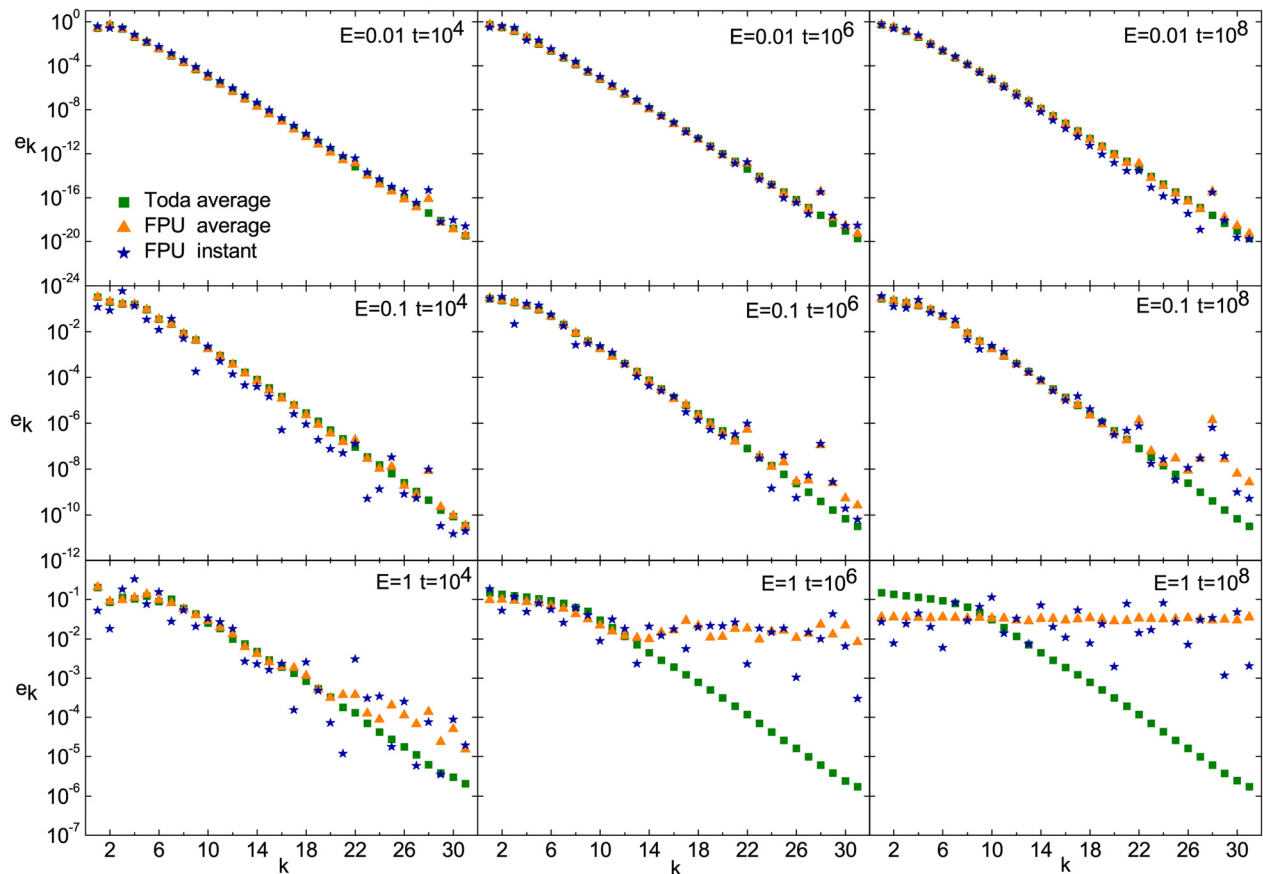

FIG. 3. (Color online) $\bar{e}_k(t)$ vs. k (log-linear scale) for the Toda (green squares) and the FPU-α (orange triangles) systems, and $\bar{e}_k(t)$ vs. k for the FPU-α model (blue stars). $E = 0.01$, 0.1, 1 from top to bottom, $t = 10^4$, 10^6, 10^8 from left to right, i.e., the energy is constant along the rows and the time is constant along the columns. Notice the change of vertical scale from top to bottom.
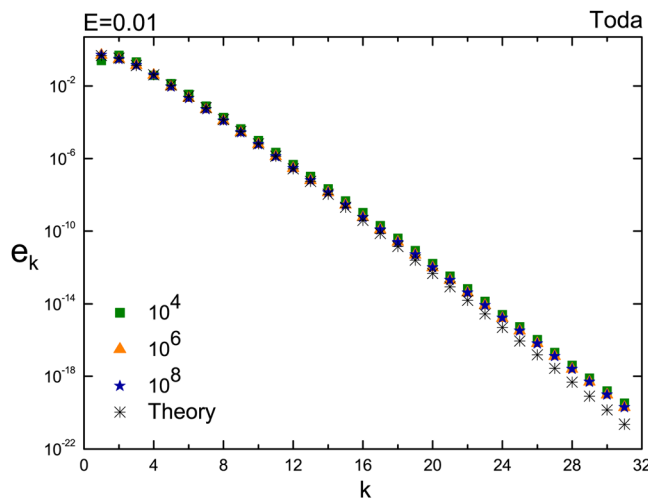
FIG. 4. (Color online) Time-averaged Toda m.e.s. at three different times (10^4, 10^6, 10^8), $E = 0.01$. The asterisks correspond to the theoretical prediction (16).

mode excitation grows with the energy, so that the (Toda) motion becomes quasi periodic. This explains why the q-breather m.e.s. (16) fails to describe well, at larger energy values, the numerical spectra of the Toda and FPU systems, although the latter spectra go on to almost superpose on short times. The analytic description of the Toda or FPU m.e.s. corresponding to the full quasi-periodic case is beyond the scope of the present paper and is left to future investigation (see however Refs. 44 and 45 for interesting results in this direction).

C. Stage II: Diffusion of tail modes

In order to describe the approach to equipartition in the FPU system, in Fig. 5, we plot (in log-log scale) the instantaneous normalized tail energy $\eta(t) = \sum_{k=22}^{31} e_k(t)$ and the corresponding time-averaged values $\bar{\eta}(t)$ vs. t, for three values of the energy $E = 0.01$, 0.1, 1 (bottom to top). Note that the instantaneous values evolve by fluctuating around the corresponding time-averages. We observe a growth of both
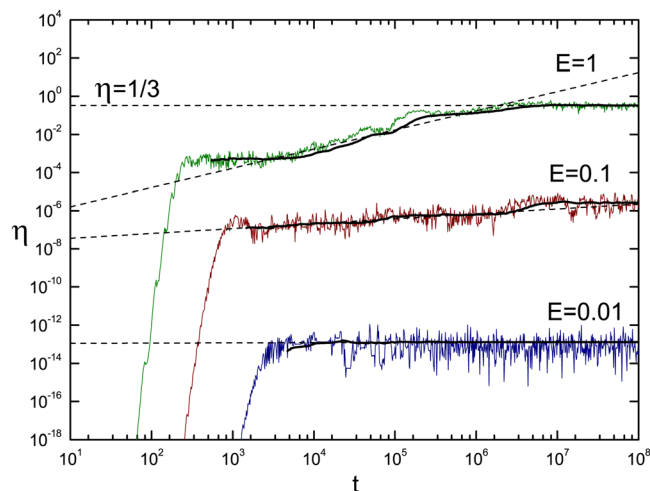

FIG. 5. (Color online) Bottom to top: $\eta(t)$ vs. t (log-log scale) for $E = 0.01$ (blue), 0.1 (magenta), and 1 (green) together with the corresponding time-averages $\bar{\eta}(t)$ vs. t (black curves) and the fitting lines Dt^γ vs. t (dashed lines).
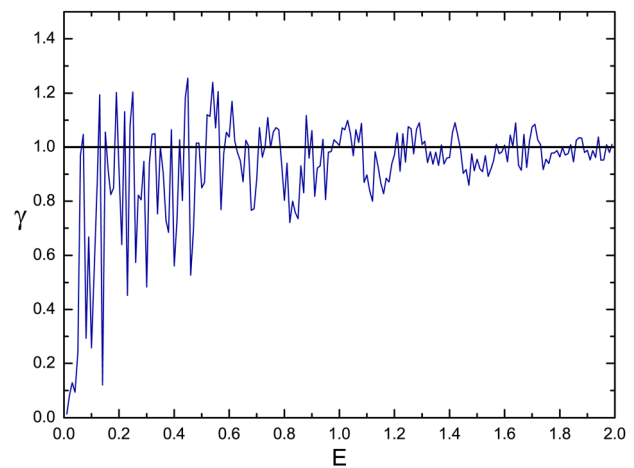

FIG. 6. (Color online) γ vs. E corresponding to $\bar{\eta}$.

quantities, well fitted by a power-law $\bar{\eta}(t) \sim Dt^\gamma$ within some time-interval $[t_0, t_1]$. It follows the linear behavior $\log \bar{\eta}(t) \simeq \log D + \gamma \log t$ in log-log scale, as reported in Fig. 5. Notice that when $E = 1$, η, and $\bar{\eta}$ asymptotically reach the equipartition value 1/3 and do not grow further, as expected.

By a systematic and detailed numerical investigation, we first determined the time intervals $[t_0(E), t_1(E)]$ and then computed $\gamma(E)$ and $D(E)$ by a least squares fit (200 values of E were considered, in the energy interval 0.01 to 2, with an energy step $\Delta E = 0.01$). The errors on both quantities turn out to be small (a few percents of the value) except for very low energy values, close to 10^{-2}, where large errors affect the almost vanishing exponent γ. The function $\gamma(E)$ is plotted in Fig. 6. Despite strong fluctuations, we find that γ changes from a value very close to zero at energies lower than 0.1 to values oscillating around $\gamma = 1$ for energies larger than 0.1. Moreover, the absolute value of the fluctuations decreases with increasing energy.

We conjecture that the fluctuations in the exponent γ are intrinsic to the dynamics and are due to the complicated phase space structure of the FPU system. To show that, we plot $\eta(t)$ in Fig. 7 for two values of the energy which differ
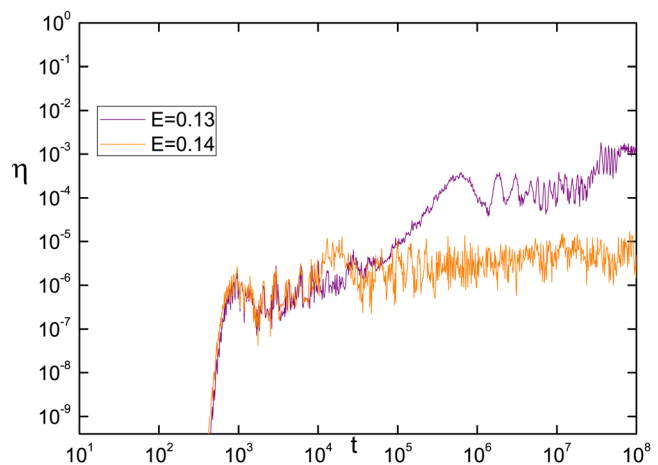

FIG. 7. (Color online) $\eta(t)$ vs. t for two values of the energy differing by $\Delta E = 10^{-2}$: the upper and lower curves correspond to the energy values $E = 0.13$ and $E = 0.14$, respectively.
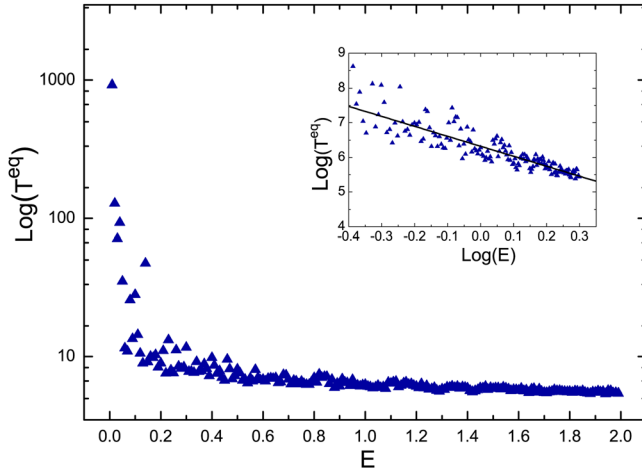
FIG. 8. (Color online) Estimated $\log(T^{eq})$ vs. E from average $\bar{\eta}$ data. Inset: $\log(T^{eq})$ vs. $\log(E)$. The line drawn has slope -2.9.

by $\Delta E = 10^{-2}$. Indeed, the growth rates differ substantially: the values of the exponent γ corresponding to these two energies are $\gamma(0.14) = 0.03$ and $\gamma(0.13) = 1.1$.

1. Estimate of the time to equipartition

The observed power-law growth $\bar{\eta} \sim Dt^{\gamma}$ allows us to define the extrapolated equipartition time such that Dt^{γ} equals $1/3$, i.e., the equipartition value of $\bar{\eta}$, which yields $T^{eq} \equiv (3D)^{-1/\gamma}$. In Fig. 8, we plot $\log(T^{eq})$ vs. E. We observe a sharp crossover around $E = 0.1$, below which T^{eq} increases dramatically, up to values $T^{eq} \approx 10^{1000}$. Such large times strongly suggest that below the reference energy threshold $E_c \simeq 0.1$, the quasi-state of the FPU system (i.e., the equilibrium state of the Toda model) might become stable in the sense of Nekhoroshev, i.e., over exponentially long times of the order $10^{c/E^b}$, with suitable constants c and b. We get a rough estimate for the stretching exponent $b \simeq 0.7 \pm 0.2$. Existing numerical results in the literature[18,26–28] suggest one that for larger values of the energy, the time to equipartition grows inverse proportional to a power of the energy E. In the inset of Fig. 8, we plot $\log(T^{eq})$ vs. $\log(E)$ in the energy interval 0.5 to 2. We indeed find some evidence for a power law, with exponent $a \simeq 2.9$. This number is close to the value $a = 3$ first reported in Ref. 18 for $N = 32$, and there obtained with other methods.

2. Evolution of a Toda integral

We have studied the evolution of one of the additional constants of motion of the Toda model. The integral $J(q,p)$ we considered is[23]

$$J \equiv \sum_{n=0}^{N-1} \left[\frac{p_n^4}{2} + \frac{(q_n^2 + q_{n+1}q_n + q_{n+1}^2)f_n}{\alpha} + \frac{(f_{n+1} + f_n + f_{n-1})f_n}{4\alpha^2} \right],$$
(22)

where $f_n \equiv e^{2\alpha(q_{n+1} - q_n)}/(2\alpha)$ $(n = 0, \ldots, N-1)$ are the Flaschka variables,[24] with boundary conditions $f_{-1} = f_0$, $f_N = f_{N-1}$, inherited by the restriction of the $2N$-periodic case
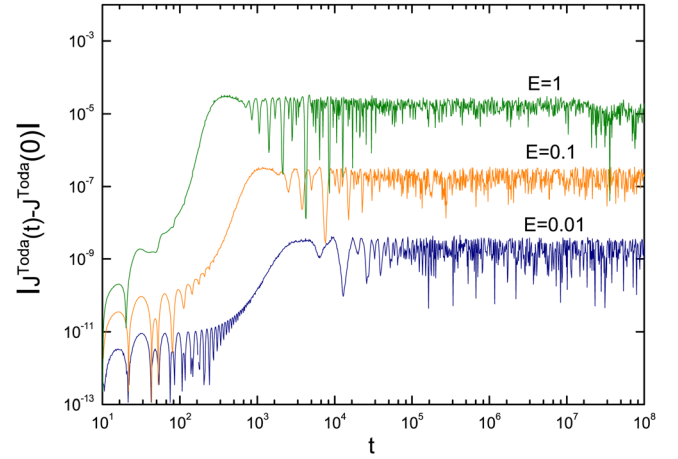
to that of fixed ends. Since J is an integral of the Toda model, $\dot{J} = \{J, H_T\} = 0$ along the Toda flow. Along the FPU flow, the quantity $\dot{J} = \{J, H_\alpha\} = \{J, H_\alpha - H_T\}$ is expected to be small in the perturbative regime where the two models can be considered to be close to each other. Let us denote by $J^{FPU}(t)$ and by $J^{Toda}(t)$ the evolution of the quantity (22) along the *numerical* FPU and Toda trajectories, respectively.

In Fig. 9, the absolute error $|J^{Toda}(t) - J^{Toda}(0)|$ vs. t is reported, for the three different values of the energy $E = 0.01, 0.1, 1$ (bottom to top), in log-log scale. After a transient time, the error stabilizes quickly at a practically constant value ranging from 10^{-9} at $E = 0.01$ to 10^{-5} at $E = 1$. Thus, the numerical algorithm reproduces the integrable Toda dynamics with very good accuracy.

In Fig. 10, we plot the evolution of $|J^{FPU}(t) - J^{Toda}(t)|$, for the same three values of the energy $E = 0.01, 0.1, 1$ (bottom to top), in log-log scale. For $E = 0.01$ or 0.1, this difference stays almost constant (up to the final observation time 10^8), while for $E = 1$, it starts to grow at about $t = 10^4$ to eventually saturate at a constant value. Therefore at low energies, the FPU system can be regarded as a perturbation

FIG. 9. (Color online) $|J^{Toda}(t) - J^{Toda}(0)|$ vs. t (log-log scale) at $E = 0.01$ (blue), $E = 0.1$ (orange), and $E = 1$ (green).
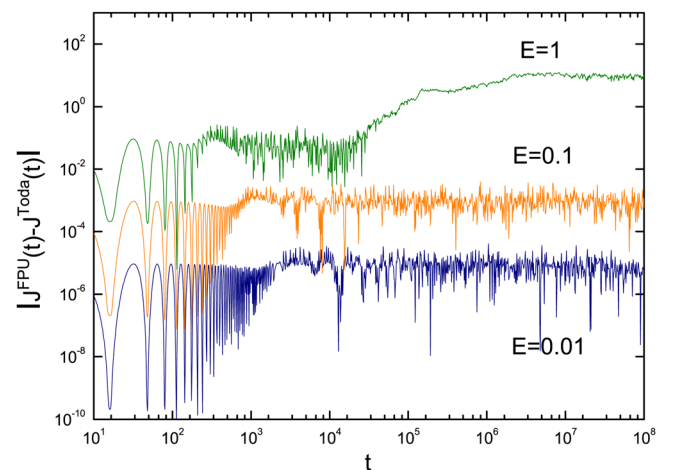

FIG. 10. (Color online) $|J^{FPU}(t) - J^{Toda}(t)|$ vs. t (log-log scale) at $E = 0.01$ (blue), $E = 0.1$ (orange), and $E = 1$ (green).

of the Toda one, the exact integrals of the latter being adiabatic invariants, or quasi-integrals, of the former. Note that at $E = 0.01$ and 0.1, no effective tendency to equipartition was observed up to $t = 10^8$. On the other hand, for $E = 1$ the tail of the FPU m.e.s. starts to effectively raise at $t = 10^4$, when the drift of the Toda integral starts up.

IV. THEORY

Our theoretical approach is based on Hamiltonian perturbation theory. We construct a resonant normal form Hamiltonian of the FPU and Toda systems. The laws (11) and (12) are then derived by solving the normal form equations, to leading order, for short times. The law (16) is instead obtained as the m.e.s. of a time-periodic solution of the normal form equations. The theoretical explanation of the diffusion law (20) goes beyond our present capabilities.

A. Complex modal variables

Let us consider the Hamiltonian (1) of the FPU α-model. We first perform a change of variables passing from the real space coordinates $(q, p) \in R^{2(N-1)}$ to the complex modal variables $u \in C^{N-1}$. The complex modal variables u_k are defined in terms of the real ones by

$$u_k \equiv \frac{\omega_k Q_k + i P_k}{\sqrt{2E}}, \quad |u_k(t)|^2 = \frac{E_k(t)}{E} = e_k(t), \quad (23)$$

whose utility is self-explanatory. The FPU Hamiltonian (1), when expressed in the new complex variables u_k, reads

$$\mathcal{H}(u, u^*) = \underbrace{\sum_{k=1}^{N-1} |u_k|^2}_{\mathcal{H}_2} + \underbrace{\frac{\mu}{12} \sum_{k_1, k_2, k_3 = 1}^{N-1} \Delta_{k_1, k_2, k_3} \prod_{j=1}^{3} (u_{k_j} + u_{k_j}^*)}_{\mathcal{H}_3},$$

$$(24)$$

where a superscript asterisk denotes complex conjugation, and the mode-coupling coefficient

$$\Delta_{k_1, k_2, k_3} \equiv \delta_{k_1 + k_2, k_3} + \delta_{k_2 + k_3, k_1} + \delta_{k_3 + k_1, k_2} - \delta_{k_1 + k_2 + k_3, 2N} \quad (25)$$

has been introduced ($\delta_{n,m}$ is the usual Kronecker symbol). We also recall that

$$\mu \equiv \alpha \sqrt{\frac{E}{N}} \quad (26)$$

denotes a small parameter in the theory. In Eq. (24), we have explicitly pointed out the quadratic part $\mathcal{H}_2$ and the cubic part $\mathcal{H}_3$ of the Hamiltonian. The Hamilton equations of motion $\dot{u}_k = \{u_k, \mathcal{H}\} = -i\omega_k \partial \mathcal{H}/\partial u_k^*$ associated to the Hamiltonian $\mathcal{H}$ explicitly read

$$\dot{u}_k = -i\omega_k \left[u_k + \frac{\mu}{4} \sum_{p, q = 1}^{N-1} \Delta_{k, p, q} (u_p + u_p^*)(u_q + u_q^*) \right]. \quad (27)$$

Finally, the initial condition (2) expressed in the new variables reads

$$u_k(0) = \delta_{k,1}. \quad (28)$$

B. Averaging

The initial condition (28), at low energy, is expected to excite an oscillation of the first mode at a frequency Ω_1 close to the unperturbed one, the difference $\Omega_1 - \omega_1$ vanishing as $\mu \to 0$. Indeed, in the latter limit, Eqs. (27) become linear and are solved by $u_k = e^{-i\omega_1 t} \delta_{k,1}$. Now, by substituting the leading order *Ansatz* $u_k = e^{-i\Omega_1 t}\delta_{k,1}$ in the right hand side of equations (27), and making use of Eq. (25), one realizes by a direct inspection that the mode $k = 2$ is forced, among others, by a term oscillating at a frequency $2\Omega_1 \simeq 2\omega_1 \simeq \omega_2$. Thus, mode $k = 2$ responds close to resonance and its amplitude grows, which requires to modify the starting *Ansatz* to $u_k = Ae^{-i\Omega_1 t}\delta_{k,1} + Be^{-i2\Omega_1 t}\delta_{k,2}$. Re-inserting the latter in the right hand side of (27), one finds that mode $k = 3$ is almost resonantly pumped at a frequency $3\Omega_1 \simeq 3\omega_1 \simeq \omega_3$, and so on. Such a process leads to an almost resonant transfer of energy from the fundamental mode to higher frequency modes, with mode k oscillating at the frequency $k\Omega_1$. For this reason, we perform the time-dependent change of variables $(u, u^*) \mapsto (v, v^*)$

$$(u, u^*) = \Phi^{\Omega_1 t}(v, v^*) : u_k = e^{-ik\Omega_1 t} v_k, \quad (29)$$

for $k = 1, \dots, N - 1$. The above defined map $\Phi^{\Omega_1 t}$ is actually the $2\pi/\Omega_1$-periodic flow of the Hamiltonian

$$h_{\Omega_1}(u, u^*) = \sum_{k=1}^{N-1} \left(\frac{k\Omega_1}{\omega_k} \right) |u_k|^2. \quad (30)$$

Notice also that all the functions of the modal energies, such as $\mathcal{H}_2$ and h_{Ω_1}, are invariant under the flow $\Phi^{\Omega_1 t}$, i.e., Poisson-commute with h_{Ω_1}. The new variables v evolve according to the Hamilton equations $\dot{v}_k = \{v_k, \mathcal{A}\} = -i\omega_k \partial \mathcal{A}/\partial v_k^*$ associated to the explicitly time-dependent, $2\pi/\Omega_1$-periodic Hamiltonian $\mathcal{A}$:

$$\mathcal{A}(v, v^*, \Omega_1 t) \equiv \mathcal{H}_2(v, v^*) - h_{\Omega_1}(v, v^*) + \mathcal{H}_3\big(\Phi^{\Omega_1 t}(v, v^*)\big). \quad (31)$$

According to the usual averaging theory,[46,47] we replace the explicitly time-dependent Hamiltonian $\mathcal{A}$ with its time average $\overline{\mathcal{A}}$. This amounts to replacing $\mathcal{H}_3^\circ \Phi^{\Omega_1 t}$, on the right hand side of (31), with its time average

$$\overline{\mathcal{H}}_3(v, v^*) = \frac{\Omega_1}{2\pi} \int_0^{\frac{2\pi}{\Omega_1}} \mathcal{H}_3\big(\Phi^{\Omega_1 s}(v, v^*)\big) ds. \quad (32)$$

We then obtain an averaged system, whose solution stays close to that of the original system over times inverse proportional to the small parameter. The averaged Hamiltonian $\overline{\mathcal{A}} = \mathcal{H}_2 - h_{\Omega_1} + \overline{\mathcal{H}}_3$ of the system then reads

$$\overline{\mathcal{A}}(v, v^*) = \sum_{k=1}^{N-1} \left(\frac{\omega_k - k\Omega_1}{\omega_k} \right) |v_k|^2$$
$$+ \frac{\mu}{4} \sum_{p, q=1}^{N-1} (v_{p+q}^* v_p v_q + v_{p+q} v_p^* v_q^*), \quad (33)$$

and its associated Hamilton equations $\dot{v}_k = \{v_k, \overline{\mathcal{A}}\}$ $= -i\omega_k \partial \overline{\mathcal{A}}/\partial v_k^*$ read

$$\dot{v}_k = -i(\omega_k - k\Omega_1)v_k$$
$$- i\frac{\omega_k \mu}{4}\left(\sum_{q=1}^{k-1} v_{k-q}v_q + \sum_{q=1}^{N-k-1} 2v_{k+q}v_q^*\right), \quad (34)$$

where, if $k=1$ or $k=N-1$, the first or the second sum on the right hand side are, respectively, absent. Such equations must be solved, for $k=1,\dots,N-1$, with the initial condition

$$v_k(0) = \delta_{k,1}. \quad (35)$$

Note that since the averaging (32) is performed along the flow of the Hamiltonian h_{Ω_1}, defined in (30), then h_{Ω_1} is a constant of motion of the averaged system: $\{h_{\Omega_1}, \overline{\mathcal{A}}\} = 0$. Thus also the quantity

$$h_{\omega_1}(v, v^*) = \sum_{k=1}^{N-1}\left(\frac{k\omega_1}{\omega_k}\right)|v_k|^2 \quad (36)$$

is a constant of motion of the averaged system.

C. Solution of the averaged system

The averaged equations (34) contain the small parameter μ. We set

$$\Omega_1(\mu) = \omega_1 + \delta\omega_1(\mu), \quad (37)$$

the dependence on μ being left in the *nonlinear frequency shift* $\delta\omega_1 = \Omega_1 - \omega_1$. We require that $\delta\omega_1 \to 0$ as $\mu \to 0$, so that in absence of nonlinearity the initially excited mode will oscillate with its unperturbed frequency ω_1. It follows:[48]

$$v_k(t; \mu) = \mu^{k-1}\left(v_k^{(0)}(t) + \sum_{j\geq 1}\mu^{2j}v_k^{(j)}(t)\right), \quad (38)$$

$$\delta\omega_1(\mu) = \Omega_1(\mu) - \omega_1 = \sum_{j\geq 1}\mu^{2j}\nu^{(j)}. \quad (39)$$

The initial condition $v_k(0) = \delta_{k,1}$ implies

$$v_k^{(j)}(0) = \delta_{k,1}\delta_{j,0}. \quad (40)$$

Upon substitution of Eqs. (38) and (39) in Eq. (34) and balancing order by order, one gets a nonlinear system of $N-1$ coupled equations at any order $j \in N$. The first two orders $(j=0,1)$ explicitly read

$$\dot{v}_k^{(0)} = -i(\omega_k - k\omega_1)v_k^{(0)} - i\frac{\omega_k}{4}\sum_{q=1}^{k-1}v_{k-q}^{(0)}v_q^{(0)}gt, \quad (41)$$

$$\dot{v}_k^{(1)} = -i(\omega_k - k\omega_1)v_k^{(1)} + ik\nu^{(1)}v_k^{(0)} +$$
$$- i\frac{\omega_k}{4}\left[\sum_{q=1}^{k-1}\left(v_{k-q}^{(0)}v_q^{(1)} + v_{k-q}^{(1)}v_q^{(0)}\right) + 2v_{k+1}^{(0)}\left(v_1^{(0)}\right)^*\right]. \quad (42)$$

Again, a sum is absent if the upper limit of summation is less than one. The latter two systems, in which the unknowns are

$v_k^{(0)}$, $v_k^{(1)}$, and $v^{(1)}$, have a triangular structure, i.e., the equations can be solved one after the other starting from the first down to the last one, with the initial condition (40). The expansion of the frequency shift (39) (i.e., the coefficients $v^{(j)}$) is determined by requesting the vanishing of possible secular terms in the expansion of v_1. One thus arrives at the following leading order expansions:

$$\Omega_1(\mu) = \omega_1 + \frac{\mu^2\omega_1\omega_2}{8(2\omega_1 - \omega_2)} + \dots, \quad (43)$$

$$v_1(t) = 1 - \frac{\mu^2\omega_1\omega_2}{8(2\omega_1 - \omega_2)^2}\left(1 - e^{i(2\omega_1 - \omega_2)t}\right) + \dots, \quad (44)$$

$$v_2(t) = \frac{\mu\omega_2}{4(2\omega_1 - \omega_2)}\left(1 - e^{i(2\omega_1 - \omega_2)t}\right) + \dots, \quad (45)$$

$$v_3(t) = \frac{\mu^2\omega_2\omega_3}{8(\omega_2 - 2\omega_1)}\left(\frac{e^{i(3\omega_1 - \omega_3)t} - e^{i(2\omega_1 - \omega_2)t}}{\omega_3 - \omega_2 - \omega_1} + \frac{1 - e^{i(3\omega_1 - \omega_3)t}}{\omega_3 - 3\omega_1}\right)$$
$$+ \dots, \quad (46)$$

the dots standing for terms of higher order.

In Fig. 11, the first four modal energies $e_k(t)$ of the α-model are plotted vs. time t, in log-log scale, at the energy value $E = 10^{-4}$. The red (pale) curves represent the numerical values, while the four black (dark) lines correspond to the three analytic formulae (44)-(46) and to that of mode $k=4$, which is not explicitly reported here, since it has a very long, uninteresting expression. Note that $e_k = |v_k|^2$. Note how the agreement is excellent up to the final computation time $t = 10^7$.

Fig. 11 shows that the perturbative scheme used here works very well at low energies. The regime of validity of the perturbative solution can be estimated by requiring that in the expansions (43) and (44) of Ω_1 and of the v_1, respectively, the leading order is dominant with respect to the correction proportional to μ^2. The latter condition turns out to be more restrictive than the former and, by taking into account that $2\omega_1 - \omega_2 = \pi^3/(4N^3) + O(N^{-5})$, yields

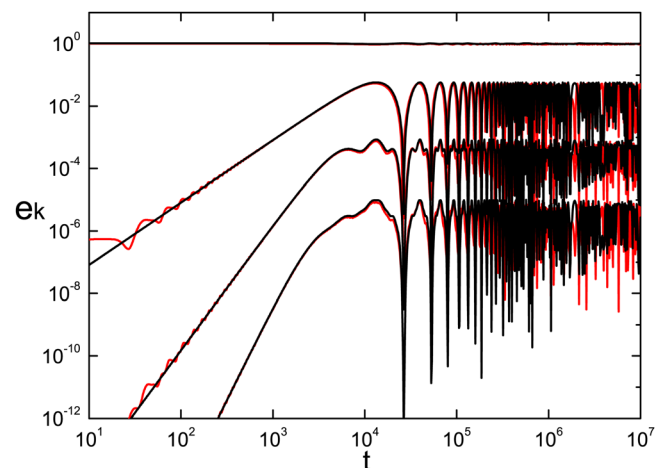$$\alpha^2 E N^3 < \frac{\pi^4}{8} \simeq 10, \quad (47)$$

FIG. 11. (Color online) e_k vs. t (log-log scale), $k=1, 2, 3, 4$, for FPU, $E = 10^{-4}$. Red curves: numerical data. Black curves: analytic prediction.

i.e., an upper bound to the energy which decreases with increasing N. For $N = 32$ and $\alpha = 0.33$, one gets $E < 3 \times 10^{-3}$. Using Eq. (43), condition (47) can be transformed into

$$\frac{\delta\omega_1}{2\omega_1 - \omega_2} < \frac{1}{2}. \qquad (48)$$

Thus, the ratio of the nonlinear frequency shift $\delta\omega_1 = \Omega_1 - \omega_1$ to the *resonance width* $2\omega_1 - \omega_2$ must be smaller than a constant, which is akin to the so-called resonance overlap criterion first formulated in Ref. 49 (notice that in the latter paper, dealing with the FPU-β model, the considered resonance width was different; see also Refs. 9 and 11 for comments on this point).

For values of the energy larger than about 10^{-3}, the agreement of the expressions (44)-(46) with the numerical data starts to worsen, and one would really need to include higher order corrections. However, some leading order analytic predictions concerning the behavior of the v_k's can still be obtained, within the same perturbative scheme, either on very short times, or at saturation, as follows.

D. Secular avalanche

Since by the initial condition (40), all the $v_k^{(0)}$ are zero at $t = 0$ if $k \geq 2$ and $\omega_k - k\omega_1$ is small for low modes (and exactly vanishes for $k = 1$), the contribution to the short term dynamics coming from the linear term $(\omega_k - k\omega_1)v_k^{(0)}$ on the right hand side of Eqs. (41) is expected to be small on short time. By neglecting such a term, the leading order equations (41) simplify to

$$\dot{v}_k^{(0)} = -i\frac{\omega_k}{4}\sum_{q=1}^{k-1} v_{k-q}^{(0)}v_q^{(0)}, \qquad (49)$$

where again the sum on the right hand side is absent for $k = 1$. Equations (49) with the initial condition $v_k^{(0)} = \delta_{k,1}$ admit the exact (unique) solution

$$v_k^{(0)}(t) = (-it)^{k-1}c_k (k = 1, \ldots, N-1), \qquad (50)$$

where the coefficients c_k are computed by solving the nonlinear map

$$c_1 = 1; gtc_k = \frac{\omega_k}{4(k-1)}\sum_{q=1}^{k-1} c_{k-q}c_q \qquad (51)$$

for $2 \leq k \leq N-1$. Notice that to leading order, the first mode does not change from the initial condition: $v_1^{(0)}(t) = 1$. The evolution equation for the first mode at second order can be obtained by solving the second order equations (42) for $k = 1$ with the *choice* $v^{(1)} = 0$.[50] Taking into account that $v_1^{(0)} = 1$, and making use of Eqs. (50) and (51) to get $v_2^{(0)} = -itc_2$ and $c_2 = \omega_2/4$, one gets a differential equation for $v_1^{(1)}$, which can be integrated with the initial condition $v_1^{(1)} = 0$. The result is

$$v_1^{(1)}(t) = -\frac{\omega_1\omega_2 t^2}{16}. \qquad (52)$$

Finally, the normalized modal energies $e_k = |v_k|^2$ can be computed,

$$e_1(t) = |v_1^{(0)} + \mu^2 v_1^{(1)} + \ldots|^2 \simeq 1 - \frac{\omega_1\omega_2(\mu t)^2}{8}gt, \qquad (53)$$

$$e_k(t) = |\mu^{k-1}v_k^{(0)} + \ldots|^2 = (\mu t)^{2(k-1)}c_k^2 (k \geq 2), \qquad (54)$$

which explains the derivation of the laws (11) and (12) reported in Sec. II and tested in Fig. 2 of Sec. III.

E. Role of resonance detuning and energy conservation

Let us discuss the validity of the approximation obtained by neglecting the linear term on the right hand side of Eqs. (41). This is done by substituting on the right hand side of the same equation the solution (50) and by checking whether and how long the linear term stays smaller than the quadratic term. One thus finds that a first necessary condition for the approximation to be valid is that

$$t < T_k^{det} \equiv \frac{k-1}{k\omega_1 - \omega_k}(k \geq 2). \qquad (55)$$

The *detuning time* T_k^{det} is the characteristic time within which mode k is resonantly pumped, the term proportional to the resonance amplitude $k\omega_1 - \omega_k$ being negligible. Now, T_k^{det} is a monotonically decreasing function of k, and for high mode numbers $T_k^{det} \approx N$, whereas for low modes $T_k^{det} \approx N^3/k^2$, which can be as large as $O(N^3)$ for the first few modes, actually too large to be compatible with the observed timescale of formation of the quasi-state. When checking the validity of the approximation made by neglecting the linear term in Eq. (41), one has to take into account that the quantity $h_{\omega 1}$, defined in Eq. (36), is constant and equal to one. The latter condition implies that $|v_k|^2 = e_k < 1$ for any $k \geq 2$, yielding, when the (approximate) solution (50) is taken into account, a second necessary condition for the validity of the approximation made, namely

$$t < T_k^{sa} \equiv \frac{c_k^{-\frac{1}{k-1}}}{\mu}. \qquad (56)$$

T_k^{sa} is the mode dependent characteristic time of the secular avalanche process, i.e., the time within which mode k can grow, when resonantly pumped by the cascade process, without violating the conservation of the energy $h_{\omega 1}$. We find that T_k^{sa} is a monotonically decreasing function of k and, by direct inspection of the map (51), that $T_k^{sa} \sim 0.2N/\mu$, if k is not too small. Note that from Eq. (53), it follows that $e_1 > 0$ if $t < T_1^{sa} \approx 2N/(\pi\mu)$, while there is obviously no detuning time for the first mode.

Having two characteristic times for each mode $k \geq 2$, namely T_k^{det} and T_k^{sa}, the actual characteristic time T_k at which e_k stops its secular growth is defined as the minimum of the two for each $k : T_k \equiv \min\{T_k^{sa}, T_k^{det}\}$. For small values of μ, $T_k = T_k^{sa}$ for $k < k_c$ and $T_k = T_k^{det}$ for $k > k_c$. An order of magnitude estimate of k_c is obtained by observing that $T_k^{sa} \approx N/\mu$ for almost all modes, whereas $T_k^{det} \sim N^3/k^2$ if

$k/N \ll 1$ (numerical factors are neglected). Then, the cross-over $T_k^{sa} = T_k^{det}$ takes place at the critical mode number

$$\frac{k_c}{N} \approx \sqrt{\mu}, \qquad (57)$$

valid if $k_c/N \ll 1$, i.e., if $\mu \ll 1$. The time-scale for the formation of the quasi-state is defined as the crossover time at k_c,

$$T^{qs} \equiv T_{k_c}^{det} = T_{k_c}^{sa} \approx \frac{N}{\mu}, \qquad (58)$$

which is the estimate (15).

From the above reasonings, it follows that modes with $k < k_c$ are always inside a resonance layer and their modal energies stop to grow due to energy conservation, whereas modal energies e_k with $k > k_c$ stop to grow due to resonance detuning. We thus distinguish core modes with $k < k_c$ from those with $k > k_c$. Core modes share most of the energy in the quasi-state, and the resonant transfer of energy among them is effective. Higher modes outside the core are resonantly pumped for a certain time and are tuned out of resonance rather quickly. The quasi-state is characterized by a few core modes sharing the energy and by an exponential localization of the m.e.s. for high modes. Notice that for $N = 32$ and $\alpha = 0.33$, Eq. (57) gives k_c varying from 2 to 8 as the energy varies from 0.01 to 1, in agreement with the number of modes that define the plateau of partial equipartition observable in the numerical spectra reported in the left column of Fig. 3.

F. Fixed point of the averaged system: Quasi-state

Let us consider the averaged equations (34). In the limit $\mu \to 0$ one has $\overline{\mathcal{H}}_3 \to 0$, we require $\Omega_1 \to \omega_1$, and the equations simplify to

$$\dot{v}_k = -i(\omega_k - k\omega_1)v_k, \qquad (59)$$

whose solution with initial condition (35) is $v_k(t) = v_k(0) = \delta_{k,1}$. In other words, the system is placed at $t = 0$ on a fixed point and stays there forever. Notice that system (59) admits a one-parameter family of fixed points

$$C_0 = \{|v_1| = 1gt; |v_k| = 0\, k \geq 2\}, \qquad (60)$$

which is a unit circle in the v_1-plane of C^{N-1}. It can be easily checked that the set C_0 is the maximum of the Hamiltonian $\mathcal{H}_2 - h_{\omega_1}$ (limit of the Hamiltonian $\overline{\mathcal{A}}$ as $\mu \to 0$) constrained to the set $h_{\omega_1} = 1$. If $\mu > 0$ but small enough, the set C_0 is slightly deformed to the set C_μ of stationary points of the full averaged system (34) on the set $h_{\omega_1} = 1$, i.e., the set of critical points of the Hamiltonian $\overline{\mathcal{A}} = \mathcal{H}_2 - h_{\Omega_1} + \overline{\mathcal{H}}_3$ constrained to the set $h_{\omega_1} = 1$. Moreover, if μ is small enough, the critical set C_μ preserves its maximum property (i.e., C_μ is the maximum of $\overline{\mathcal{A}}$ on $h_{\omega_1} = 1$) and turns out to be Lyapunov stable: if v is initially close to C_μ it stays close to it forever. On the other hand, the initial condition $v_k(0) = \delta_{k,1}$ means that one starts on C_0, so that if C_0 and C_μ are close, i.e., if μ is small enough, the orbit of the averaged system stays forever close to C_μ. We make use of this qualitative

(but exact) reasoning to identify the quasi-state orbit of the α-model with the set C_μ.

The fixed points of system (34) can be perturbatively computed by means of the expansions (38) and (39). To leading order, one has to find the fixed point of system (41), setting $\dot{v}_k^{(0)}$, which yields

$$v_1^{(0)} = \lambda; gtv_k^{(0)} = \frac{\omega_k}{4(k\omega_1 - \omega_k)}\sum_{q=1}^{k-1} v_{k-1}^{(0)}v_q^{(0)}, (k \geq 2), \quad (61)$$

where λ, at this stage, is an arbitrary complex number such that $|\lambda|^2 = e_1$, to leading order. Isolating the dependence of $v_k^{(0)}$ on λ, one gets

$$v_k^{(0)} = \lambda^k f_k, \qquad (62)$$

where the real coefficients f_k are found by solving the nonlinear map problem

$$f_1 = 1; gtf_k = \frac{\omega_k}{4(k\omega_1 - \omega_k)}\sum_{q=1}^{k-1} f_{k-q}f_q, \qquad (63)$$

for $2 \leq k \leq N - 1$. Recalling that, to leading order, $v_k = \mu^{k-1}v_k^{(0)}$, one can compute the m.e.s.

$$e_k = |v_k|^2 = f_k^2 (\mu^2 e_1)^{k-1} e_1. \qquad (64)$$

Such a spectrum depends on the unknown $e_1 = |\lambda|^2$, which is determined by the normalization condition $h_{\omega_1} = 1$, i.e.,

$$h_{\omega_1} = \sum_{k=1}^{N-1} \left(\frac{k\omega_1}{\omega_k}\right)e_k = \sum_{k=1}^{N-1} \left(\frac{k\omega_1}{\omega_k}\right)f_k^2 (\mu^2 e_1)^{k-1} e_1 = 1. \quad (65)$$

The latter condition determines the modulus of λ, so that an overall phase remains undetermined. For this reason, the set C_μ is composed by a one-parameter family of stationary points of the averaged system.

We now observe that to a stationary solution of the averaged system, there corresponds a periodic orbit, or one-dimensional torus, of the original system: recall that, before averaging, we performed the change of variables $u_k = e^{-ik\Omega_1 t}v_k$. The fundamental frequency Ω_1 of the quasi-state periodic orbit is easily determined by means of Eqs. (42). Setting there $k = 1$ and $\dot{v}_1^{(1)} = 0$, and making use of Eqs. (62) and (63), one gets

$$\nu^{(1)} = \frac{\omega_1 v_2^{(0)}(v_1^{(0)})^*}{2v_1^{(0)}} = \frac{e_1\omega_1\omega_2}{8(2\omega_1 - \omega_2)}, \qquad (66)$$

which implies

$$\Omega_1 = \omega_1 + \mu^2 \nu^{(1)} = \omega_1 + \frac{\mu^2 e_1\omega_1\omega_2}{8(2\omega_1 - \omega_2)}. \qquad (67)$$

Notice that in the regime of strong localization, when $e_1 \simeq 1$, the above expression for the corrected frequency coincides with that given in Eq. (43).

As already mentioned, the quasi-state m.e.s. (64) fits the numerically obtained one at low energy, when the exponential

localization is so strong that only a few low modes contribute to the total energy. In this case, the map (63) can be approximately solved, as follows. First of all, notice that if only low (acoustic) modes are relevant, one can substitute the approximate expressions $\omega_k \simeq \pi k/N$, $k\omega_1 - \omega_k \simeq \pi^3 k(k^2-1)/(24N^3)$ in the numerator and denominator of Eq. (63), respectively. Second, we look for a solution of the map in the form $f_k = k\, g^{k-1}$, which fits the condition $f_1 = 1$. Upon substitution of the latter expression into Eq. (63), and after some simple calculation, one gets $g = \pi^2/N^2$. Thus, in the regime of strong localization, the approximate solution of the map (63) is given by $f_k \simeq k(N^2/\pi^2)^{k-1}$. As a consequence, recalling that $\mu^2 = \alpha^2 E/N$ and that $e_1 = E_1/E$, for the spectrum (64), one gets the approximate expression

$$e_k = k^2 \left(\frac{\alpha^2 E N^3 e_1}{\pi^4} \right)^{k-1} e_1. \qquad (68)$$

Under the same acoustic approximation ($\omega_k \simeq k\omega_1$) leading to the latter formula, condition (65) becomes

$$\sum_{k=1}^{N-1} e_k = \sum_{k=1}^{N-1} k^2 \left(\frac{\alpha^2 E N^3 e_1}{\pi^4} \right)^{k-1} e_1 = 1. \qquad (69)$$

This explains formulas (16)-(18) given in Sec. II. We stress again that the m.e.s. (68) corresponds to the leading order spectrum of the Lyapunov continuation to the nonlinear case of the first mode, first reported in Ref. 5 and there named q-breather. We also recall here that the normalization condition (69) can be simply solved. Indeed, by defining $x \equiv \alpha^2 E N^3/\pi^4$ and $y \equiv \alpha^2 E_1 N^3/\pi^4$, it is easily found that Eq. (69), in the limit of large N and for $y < 1$, yields the simple implicit equation $y(1+y)/(1-y)^3 = x$. This is explicitly solved, for $x \geq 0$ (and $0 \leq y < 1$), to give $y = F(x)$, for a suitable function F, which determines E_1 (and thus e_1) in terms of the total energy E, of α and of the number of particles N. In particular, having determined $F(x)$, it turns out that, at any fixed x, $e_1 = F(x)/x$, which can then be used to finally determine the m.e.s. (68).

To conclude this subsection, we observe that all the above construction was performed to leading order, and applied to the Toda model, would give the same result. This is a consequence of the coincidence to third order of the FPU-α and the Toda models expressed by relation (4).

G. Long term stability of the quasi-state

The argument used to guess that the one parameter family C_μ of stationary points of the Hamiltonian (33) constrained to the set $h_{\omega_1} = 1$ is the candidate quasi-state orbit of the FPU-α and Toda models is a qualitative one. We did not specify how small the parameter μ has to be to ensure that C_μ is a maximum of the constrained averaged Hamiltonian and moreover the unperturbed orbit C_0 (which contains the initial datum) is in the domain of concavity of C_μ. In this respect, we do not provide any precise estimate on μ, which would require technical arguments that are beyond the aim of the present paper. We only mention that, from the discussion made in Sec. IV C above, one may expect that a

condition of the form $E = O(N^{-3})$ should hold to ensure the Lyapunov stability of C_μ.

Apart from the precise energy-size dependence of the stability domain of the one-dimensional torus C_μ, it is important to understand that the construction made above is a leading order one, which ensures, for the real system, confinement close to the set C_μ on a relatively short time-scale, typically an inverse power law of the small parameter μ. In order to explain the stability on much longer times observed numerically, one should go beyond the leading order construction, which modifies the averaged Hamiltonian (33) with the addition of higher order terms and increases the length of the confinement time. Such an approach was used in Refs. 51–53 to prove that in the case of initial excitation of the first mode in a suitable class of string equations, the confinement time grows as the exponential of the inverse of the small parameter. However, it must be stressed that importing those techniques to the problem at hand is nontrivial at all, essentially due to the lack of complete resonance in the FPU problem (explicit reference to the FPU problem, in this perspective, has been made in Ref. 52).

V. CONCLUDING REMARKS

We have shown that the classical FPU paradox can be easily understood if one compares the behavior of the (presumably) nonintegrable FPU-α-model with that of the integrable Toda model, for the same initial condition, at the same energy. One thus finds that the dynamics of the α-model on a short initial time interval is very close to that of the Toda model. The quasi-state guessed by FPU is nothing but the quasi-periodic Toda torus on which the dynamics is initiated for the same initial condition. At very low energy, such a torus becomes one-dimensional and the FPU-α system evolves close to it up to extremely long times. As the energy increases, the dimension of the reference torus grows, while the time of confinement close to it (or stability time) for the FPU-α system decreases. Finally, the approach to a state of energy equipartition of the FPU-α model takes place through a diffusive-like detachment from the quasi-state, a phenomenon that is chaotic in character, and consequently, absent in the integrable model.

Regarding the FPU α-model as a perturbation of the Toda one, as first suggested in Ref. 16, and strongly motivated here, seems to be quite promising. For example, in Ref. 54, it is shown that the Toda Hamiltonian fits the hypotheses requested to the unperturbed Hamiltonian in the Nekhoroshev theorem, with the purpose to explain the phenomenology of the FPU-α model in terms of the exponential stability of all the Toda integrals (or actions). However, precise/optimal estimates of the regime of validity (energy-size domain) of such a conclusion are still lacking. Of course, if the energy is small enough, one cannot exclude the possibility to enter a KAM regime, i.e., to have stability over infinitely long times. Here also, we know that this is possible in principle,[55] but again, precise/optimal estimates of the regime of validity are not available yet. We finally notice that it is not presently clear whether this approach (i.e., considering the FPU system as a perturbation of the

Toda one) is valid for initial conditions involving also high frequency modes or even for generic excitations.

One interesting question is whether and how the energy region for the long-term stability of the quasi-state collapses or not to zero by increasing N. In the light of the purely numerical results contained in Ref. 28, a first answer seems to be that exponentially long times to equipartition might be, for the α-model, a finite-size effect, observable only in an unspecified regime $E \to 0$ as $N \to \infty$. In the same paper, it is shown that for particular low frequency initial data, the β-model seems to exhibit exponentially long times to equipartition in the thermodynamic limit. On the other hand, in Ref. 56, an analytic lower bound to the canonical mixing time of a weakly coupled φ^4 model is provided, showing that such time is longer than a stretched exponential of the inverse of the temperature. The full scenario is quite complicated and deserves a lot of further study.

ACKNOWLEDGMENTS

[1] E. Fermi, J. Pasta, and S. Ulam, LA-1940 internal report, 1955. Reprinted in the *Collect Papers of E. Fermi*, Vol. II, University of Chicago Press and Academia Nazionale dei Lincei, 1965, 978-988.
[2] E. Fucito, F. Marchesoni, E. Marinari, G. Parisi, L. Peliti, S. Ruffo, and A. Vulpiani, J. Phys. **43**, 707 (1982).
[3] R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani, Phys. Rev. A **31**, 2740 (1985).
[4] L. Berchialla, L. Galgani, and A. Giorgilli, DCDS-A **11**, 855 (2004).
[5] S. Flach, M. V. Ivanchenko, and O. I. Kanakov, Phys. Rev. Lett. **95**, 064102 (2005).
[6] S. Flach, M. V. Ivanchenko, and O. I. Kanakov, Phys. Rev. E **73**, 036618 (2006).
[7] S. Flach, O. I. Kanakov, M. V. Ivanchenko, and K. G. Mishagin, Int. J. Mod. Phys. B **21**, 3925 (2007).
[8] T. Penati and S. Flach, Chaos **17**, 023102 (2007).
[9] S. Flach and A. Ponno, Physica D **237**, 908 (2008).
[10] N. J. Zabusky and M. D. Kruskal, Phys. Rev. Lett. **15**, 240 (1965).
[11] D. L. Shepelyansky, Nonlinearity **10**, 1331 (1997).
[12] A. Ponno, Europhys. Lett. **64**, 606 (2003).
[13] A. Ponno, in *Chaotic Dynamics and Transport in Classical and Quantum Systems*, edited by P. Collet, M. Courbage, S. Metens, A. Neishtadt, and G. Zaslavsky (Kluwer Academic, Dordrecht, 2005), pp. 431-440.
[14] D. Bambusi and A. Ponno, Commun. Math. Phys. **264**, 539 (2006).
[15] G. Benettin, R. Livi, and A. Ponno, J. Stat. Phys. **135**, 873 (2009).
[16] W. E. Ferguson, Jr., H. Flaschka, and D. W. McLaughlin, J. Comp. Phys. **45**, 157 (1982).
[17] S. Isola, R. Livi, S. Ruffo, and A. Vulpiani, Phys. Rev. A **33**, 1163 (1986).
[18] L. Casetti, M. Cerruti-Sola, M. Pettini, and E. G. D. Cohen, Phys. Rev. E **55**, 6566 (1997).
[19] A. Giorgilli, S. Paleari, and T. Penati, DCDS-B **5**, 991 (2005).
[20] N. J. Zabusky, Z. Sun, and G. Peng, Chaos **16**, 013130 (2006).
[21] M. Toda, Phys. Rep. **18**, 1 (1975).
[22] J. Ford, S. D. Stoddard, and J. S. Turner, Prog. Theor. Phys. **50**, 1547 (1973).
[23] M. Hénon, Phys. Rev. B **9**, 1921 (1974).
[24] H. Flaschka, Phys. Rev. B **9**, 1924 (1974).
[25] M. Pettini and M. Landolfi, Phys. Rev. A **41**, 768 (1990).
[26] J. De Luca, A. J. Lichtenberg, and S. Ruffo, Phys. Rev. E **60**, 3781 (1999).
[27] L. Berchialla, A. Giorgilli, and S. Paleari, Phys. Lett. A **321**, 167 (2004).
[28] G. Benettin and A. Ponno, J. Stat. Phys. **144**, 793 (2011).
[29] J. Ford, Phys. Rep. **213**, 271 (1992).
[30] Focus Issue *The "Fermi-Pasta-Ulam"-Problem: The First 50 Years*, Chaos **15** (2005).
[31] *The Fermi-Pasta-Ulam Problem: A Status Report*, edited by G. Gallavotti, Lect. Notes Phys. Vol. 728 (Springer-Verlag, Berlin Heidelberg, 2008).
[32] H. Yoshida, Phys. Lett. A **150**, 262 (1990).
[33] H. Yoshida, Celest. Mech. Dyn. Astron. **56**, 27 (1993).
[34] The time-step τ must be much smaller than π, the latter being the minimum normal mode period (the maximum frequency is 2). The fourth order algorithm ensures energy conservation with an absolute error of the order $\tau^4 E^2$. For details on the numerical integration of FPU systems see, e.g., G. Benettin and A. Ponno, Physica D **240**, 568 (2011).
[35] The value of α is irrelevant, since it can be set to one by the canonical transformation $(q, p, H, t) \mapsto (\alpha q, \alpha p, \alpha^2 H, t)$, which also shows that the actual parameter of nonlinearity in the problem is the product $\alpha^2 E$.
[36] Formula (8) can be obtained by first computing $q_{n+1} - q_n$ for the initial condition (2), expanding the exponential in the Toda potential (3), and then summing first over n.
[37] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).
[38] In the regime explored in the present paper $A \simeq A'$, or, which is the same, $E_\alpha(A) \simeq E_T(A)$. Indeed, since $I_0(x) = 1 + x^2/4 + x^4/64 + \dots$, from Eqs. (7) and (8), it follows that $E_T(A) = E_\alpha(A)[1 + \alpha^2 E_\alpha(A)/N + \dots]$, and the quantity $\alpha^2 E_\alpha(A)/N$ is always small in our runs, being at most 10^{-2} down to 10^{-5}, and even lower in some cases.
[39] Since the cubic potential is not lower bounded, the micro-canonical measure of the α-model does not exist and the system breaks-down in a finite time when a sufficient amount of energy localizes on a few particles. As a matter of fact, in the quasi-linear regime considered here, the break-down takes place after equipartition has been reached, the latter state being well described by the micro-canonical measure of the harmonic chain. In all the numerical runs reported in the sequel, the break-down was never reached.
[40] J. Ford, J. Math. Phys. **2**, 387 (1961).
[41] R. Livi, S. Ruffo, M. Pettini, and A. Vulpiani, II, Nuovo Cimento B **89**, 120 (1985).
[42] N. J. Zabusky, J. Math. Phys. **3**, 1028 (1962).
[43] M. D. Kruskal and N. J. Zabusky, J. Math. Phys. **5**, 231 (1964).
[44] H. Christodoulidi, C. Efthymiopoulos, and T. Bountis, Phys. Rev. E **81**, 016210 (2010).
[45] H. Christodoulidi, C. Efthymiopoulos, and T. Bountis, preprint (2011).
[46] N. N. Bogolyubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Non-linear Oscillations* (Gordon & Breach, New York, 1961).
[47] J. A. Sanders, F. Verhulst, and J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems* (Springer, New York, 2007).
[48] As can be proved by induction, Eqs. (38) and (39) are the only possible formal power series expansions compatible with Eqs. (34) and with the initial condition (35).
[49] F. M. Izrailev and B. V. Chirikov, Sov. Phys. Dokl. **11**, 30 (1966).
[50] The quantity $v(1)$ is an unknown of the problem and there is more than one way to determine its value; by setting the latter value to zero means giving up to control secular divergencies, which still yields an approximate solution, though valid over very short times.
[51] D. Bambusi and N. N. Nekhoroshev, Physica D **122**, 73 (1998).
[52] N. N. Nekhoroshev, in *The Proceedings of the Porquerolles School 2001 on Hamiltonian Systems and Fourier Analysis*, edited by D. Benest, C. Froeschlé, and E. Lega (Cambridge Scientific, Cambridge, UK, 2005), pp. 289-302.
[53] N. N. Nekhoroshev, Trans. Mosc. Math. Soc. **2002**, 151 (2002).
[54] A. Henrici and T. Kappeler, Chaos **19**, 033120 (2009).
[55] B. Rink, Commun. Math. Phys. **261**, 613 (2006).
[56] A. Carati and A. Maiocchi, preprint on http://arxiv.org/abs/1011.5846, Commun. Math. Phys. (to be published).