4

The Fermi–Pasta–Ulam Problem and the Metastability Perspective

G. Benettin,[1] A. Carati,[2] L. Galgani[2] and A. Giorgilli[2]

[1] Università di Padova, Dipartimento di Matematica Pura e Applicata, Via G. Belzoni 7, 35131 Padova, Italy
[2] Università di Milano, Dipartimento di Matematica, Via Saldini 50, 20133 Milano, Italy

Abstract. A review is given of the works on the FPU problem that were particularly relevant in connection with the metastability perspective, proposed in the year 1982. The idea is that there exists a specific energy threshold above which the time-averages of the relevant quantities quickly agree with the predictions of classical equilibrium statistical mechanics, whereas below it there exist two time scales. First there is a quick formation of a packet of low-frequency modes which do share the energy, and this produces a metastable state that lasts for a long time; then the system attains the final equilibrium state. There are strong indications that the specific energy threshold does not vanish in the limit of infinitely many particles. The review is given for the case of a one-dimensional FPU chain.

4.1 Introduction

If one looks at the scientific literature on the FPU problem, 50 years after the original paper (or rather *report*) [1], one will find a rather large amount of papers (see, for example the recent special issue of the journal Chaos [2]). But if one tries to extract from them any clear conclusion about the mathematical status of the problem or the physical meaning of the results, one may remain rather perplexed and have the impression of a certain confusion. Or, even, one can find statements as if the problem had already been solved and there were nothing more to be said (see [3], p. 2). In this chapter, we will try to indicate, among the many papers on the subject, the ones which in our opinion played a significant role with respect to the main question we have in mind, namely that of establishing whether the FPU problem may have some relevant physical impact or not. We will try to show how the question is still completely open, although one may be confident that it may be solved in a near future.

We now give a preliminary summary of the history we are going to trace back in this chapter, in the above mentioned perspective. First of all, let us

G. Benettin et al.: *The Fermi–Pasta–Ulam Problem and the Metastability Perspective*, Lect. Notes Phys. **728**, 151–189 (2008)
DOI 10.1007/978-3-540-72995-2_4

recall that the essential result of the original FPU report was the exhibition of what we now call "the FPU paradox". Namely, numerical solutions of the equations of motion were performed for a model of a discretized string (or equivalently of a one-dimensional crystal, actually, a chain of N particles with nearest-neighbour nonlinear interactions), and it was observed that, starting from a long-wavelength initial datum (and thus very far from statistical equilibrium), there was quickly formed an apparently stationary state, extremely different from the one expected according to classical equilibrium statistical mechanics.

The FPU report had, 10 years later (1965), a great impact in mathematics, because it stimulated the well-known work [4] of Zabusky and Kruskal, in which the FPU result was interpreted in terms of solitons. In turn, this fact paved the way to the whole research on infinite-dimensional integrable systems, which quickly became a fashionable and extremely interesting mathematical field in itself, with the result that its relation to the FPU problem was somehow neglected. We will point out later how the relations between the two subjects, solitons and FPU problem, were reestablished in very recent times. In may be worth remarking that, as the soliton theory is essentially equivalent to integrability, by some naive transitivity some people may have been induced to associate FPU to integrability, which corresponds to even exalting the FPU paradox.

The next essential step, which by the way also eliminated the possible confusion just mentioned, was made one year later (1966) by Izrailev and Chirikov in the work [5] (see also [6]), with the discovery that the paradox disappears (i.e., a quick agreement with the predictions of classical equilibrium statistical mechanics is found) if initial data are taken of the FPU type (long-wavelength), but with a sufficiently high energy. In other words, there somehow exists a critical energy E_c, above which the paradox disappears. However, Izrailev and Chirikov appeared even to go beyond such a result, because they also advanced the additional conjecture (supported by some kind of analytical considerations, later adjusted by their pupil Shepelyansky in [7], with arguments subsequently critically discussed by Ponno in [8]) that the FPU paradox disappears at all in the thermodynamic limit (N tending to infinity, with positive specific energy $\epsilon = E/N$). On the other hand, a little later (1971) Bocchieri et al. (in [9]) reported numerical results that appeared to support the opposite conjecture. This fact was particularly emphasized by Galgani and Scotti and by Cercignani (see [10], [11] and the review [12]), who were pointing out that the FPU paradox, if it persists in the thermodynamic limit, may have a deep physical impact. With such papers ends the first phase of the history of the FPU problem, at least in our personal way of reconstructing it. At that moment the alternative seemed to be whether the paradox disappears in the thermodynamic limit or not, i.e., whether one has $\epsilon_c \to 0$ for $N \to \infty$ or not, where $\epsilon_c = E_c/N$ is the specific energy threshold.

But in such an alternative the mathematical (and even the physical) setting of the problem was a rather "naive" one, because it appeared that one had to

decide whether, in the terminology then used, the motions are of "ordered" or of "chaotic" type (below or above the threshold respectively), whereas the deep question of determining the "relaxation times" for the approach to equilibrium (which actually was the very question raised in the original FPU report itself) was completely overlooked. The breakthrough in this direction came from a paper of the year 1982 by Fucito et al. (see [13]), which clearly was conceived within a scientific frame, the theory of glasses (particularly studied by a group of people around Parisi in Roma), in which a special attention was naturally payed to the possibility that relaxation times of quite different orders of magnitude may show up. This actually constitutes what we now call *the metastability scenario*: The time-averages of the relevant physical quantities are expected to agree with the predictions of classical equilibrium statistical mechanics at any energy after a sufficiently long time-scale, the "final" one (which is the one described by the limit $t \to \infty$), and the paradox is interpreted as corresponding to the existence of another, shorter, time-scale (the fast scale), within which a relaxation is produced to some intermediate state. Such an intermediate state at first sight appears as an equilibrium one, although it is destined to subsequently relax, on a much longer time-scale, to the "final"equilibrium state. The way in which the existence of an energy threshold can be conceived within such a metastability scenario was understood much later (see [14]).

In an attempt to trace back, in the present days, a kind of historical review on the subject, one cannot but remain perplexed by remarking that the paper of Fucito et al. did not receive at that time the attention that would appear natural today. Indeed, apart from a bunch of papers written immediately later, for a long time the metastability scenario was essentially forgotten. In particular, no discussion was given of the relevant problem that was left open within such a scenario, namely to establish whether the formation of a metastable state is a phenomenon that persists in the thermodynamic limit or not. The idea of the metastability perspective actually reappeared only rather recently, under the stimulus of the work [15] by Carati and Galgani (see also [16]), devoted to the problem of estimating the specific heats in systems of FPU type. In such a paper, the existence of relaxation times of different orders of magnitude was reported, and a qualitative analogy with the problem of glasses was explicitly pointed out. Finally, a vivid numerical illustration of the metastability phenomenon, with a particularly impressive exhibition of two quite different relaxation times, was given by Berchialla et al. in [14]. In particular, it was found that the phenomenon of the two separated time-scales occurs only below a certain energy, which can thus be interpreted as the critical energy in the sense of Bocchieri et al.

Finally, a deep analytical understanding of the metastability scenario in the FPU problem was given in a paper by Bambusi and Ponno (see [17]), through a result holding in the thermodynamic limit. In such a paper, by the way, a bridge with the old Zabusky and Kruskal contribution was given. Indeed, a general mathematical frame (the method of resonant normal forms)

was devised in order to approximate the FPU system for not too long times, and this actually amounts to justify the use of a pair of KdV equations for not too long times, thus explaining the quick formation of the metastable state, in a way that is essentially equivalent to that of Fucito et al. The result of Bambusi and Ponno actually holds only for an extremely special class of initial conditions, but a strong indication that significant results may be obtained also for a much broader set of initial data, is afforded by a very recent result of one of us (see [18]), where for the first time it was proved, in a concrete model, that the techniques of Hamiltonian perturbation theory can be extended to the thermodynamic limit (previous results uniform in N were given, by Bambusi and Giorgilli in [19], but only for a finite energy E, namely for vanishing specific energy $\epsilon = E/N$ in the limit $N \to \infty$).

On the basis of the successes obtained with such recent results, one may be tempted to conclude that the FPU paradox should persist in the thermodynamic limit. But a deep question still remains open, namely the "question of the dimensions". Indeed, all the results previously mentioned refer to the FPU problem in its original formulation, namely in the one-dimensional case (a chain of particles), and there remains the problem of establishing whether the phenomenon of the quick formation of a metastable state persists (still in the thermodynamic limit) when one passes to the case of dimension two and especially to the "physical case" of dimension three. At the moment, a few results in dimension larger than one are available in the literature, for example [20, 21], but in our opinion they do not allow one to draw a definite conclusion. So in the present review we shall not enter the question. However, by judging from the results that were recently obtained in the one-dimensional case, we are rather confident that the problem may find a solution in the near future.

4.2 The First Phase, 1955–1972: From FPU to Izrailev and Chirikov and to Bocchieri et al.; the Suggestion of a Possible Physical Interpretation

4.2.1 The Original FPU Paper and the FPU Paradox

Fermi, Pasta and Ulam considered the simplest model of a discretized nonlinear string, which can also be interpreted as a model of a one-dimensional crystal, namely, a chain of equal particles with nearest-neighbours nonlinear interactions (nonlinear springs), and fixed ends. Denoting by x_j the displacements of the particles from their equilibrium positions and by p_j the corresponding momenta, $j = 0, \ldots, N+1$, with the boundary conditions

$$x_0 = 0 \,, \quad x_{N+1} = 0 \,,$$

the Hamiltonian of the system is then

$$H(x_1, \cdots, x_N, p_1, \cdots, p_N) = \frac{1}{2}\sum_{j=1}^{N} p_j^2 + \sum_{j=1}^{N+1} V(x_{j+1} - x_j) \tag{4.1}$$

where the potential V actually chosen was

$$V(x) = \frac{1}{2}x^2 + \frac{\alpha}{2}x^3 + \frac{\beta}{3}x^4 \ .$$

Here, the mass of the particles and the harmonic constant of the springs have been set equal to 1, while α and β are positive parameters. It is well known that the corresponding linearized system ($\alpha = \beta = 0$) can be transformed, through a linear change of variables, to a system of uncoupled linear oscillators (normal modes) with a corresponding Hamiltonian H_2 which, in terms of action-angle variables I_k, φ_k, takes the form

$$H_2 = \sum_{k=1}^{N} E_k \ ,$$

where

$$E_k = \omega_k I_k$$

are the normal-mode energies, having angular frequencies ω_k given by

$$\omega_k = 2\sin\frac{k\pi}{2(N+1)} \ .$$

According to classical equilibrium statistical mechanics, the statistical properties of an isolated system (such as the FPU one) at equilibrium at a given total energy E should be described by the microcanonical measure (the one naturally induced on the "energy surface" $H = E$ by the Lebesgue measure in the whole phase space) or equivalently (at least for sufficiently large N) by the "canonical" or Gibbs measure in the whole phase space with a suitable temperature $T = T(E)$. The fundamental result of classical equilibrium statistical mechanics is then the "equipartition theorem", according to which, in the harmonic limit ($\alpha = \beta = 0$), the expected values of the harmonic energies E_k at a given total energy E, which we denote by $< E_k >_E$, are all equal, independent of k,

$$< E_k >_E = E/N \equiv \epsilon \ , \quad k = 1, \ldots, N \ , \tag{4.2}$$

where $\epsilon = E/N$ is the specific energy, and for the common value ϵ one has the interpretation $\epsilon = k_{\mathrm{B}}T$, where k_{B} is the Boltzmann constant and T the absolute temperature. The result does not change qualitatively for a slightly anharmonic system (α and β small) and for a small temperature T (i.e., a small specific energy $\epsilon = E/N$), because the anharmonic corrections to the relations (4.2) do vanish in the limit $\alpha, \beta \to 0$ or $T \to 0$ (i.e., $\epsilon \to 0$).

Let us now come to the dynamics. In the harmonic case the system is "integrable", namely it has N integrals of motion (the harmonic actions I_k or equivalently the harmonic energies $E_k = \omega_k I_k$) which are independent and in involution (their mutual Poisson brackets vanish). Instead, the system is expected to be ergodic, i.e., to have no (measurable) integral of motion apart from the total Hamiltonian H itself, when the perturbation is present (for $\alpha \neq 0$ or $\beta \neq 0$), no matter how small the perturbation be. This was suggested by a famous theorem of Poincaré (see [22]), to which Fermi himself had contributed in one of the first works of his youth (see [23] and also [24]), and this was probably the main reason for him to come back again to such a problem near the end of his life.

Thus one meets with the problem of how is it possible to reconcile such a dichotomy (N integrals of motion in the harmonic case $\alpha = \beta = 0$, no integral of motion independent of the Hamiltonian in the perturbed case, no matter how small the perturbation be) with the continuity of the solutions of the equations of motion with respect to the parameters. A possible way of recovering continuity is by making reference to the notion of relaxation time. In order to make this point clear, let us recall what the ergodicity property is in our particular case of a Hamiltonian system with a phase space M coinciding with the energy surface Γ_E defined by $H = E$. As above, by $< f >_E$ we denote the corresponding microcanonical expectation of a dynamical variable $f : M \to \mathbb{R}$. Denoting by $\{g^t\}_{t\in \mathbb{R}}$, $g^t : M \to M$, the flow induced by the equations of motion, and by x a point in phase space, then the ergodicity of the microcanonical distribution amounts to the property

$$\overline{f}(t,x) \to < f >_E \quad \text{as} \quad t \to \infty$$

for all measurable dynamical variables f and for almost all initial data $x \in M$, where $\overline{f}(t,x)$ is the "time-average" of the function f up to time t with initial datum x:

$$\overline{f}(t,x) = \frac{1}{t}\int_0^t f(g^s x)\,\mathrm{d}s \ .$$

Now, as particularly pointed out by von Neumann (see [25]), for every significant dynamical variable f there should exist a typical relaxation time τ, defined as the first time such that the time-average essentially coincides with the "phase average" $< f >_E$ for all times larger than it.

Obviously, the elimination of the just mentioned dichotomy should correspond to the fact that the relaxation time τ does actually depend on the parameters (α, β and E), and should tend to infinity as they tend to zero. i.e., as the linear system is approached. So FPU had in mind to determine, through numerical solutions of the equations of motion, the relaxation times for the time-averages $\overline{E}_k(t,x)$ of the energies E_k, for initial data very far from equilibrium. As the equilibrium expectations of such energies are all equal (equipartition), the most significant initial datum corresponding to a situation out of equilibrium is the one in which the energy is given to just one

mode, for example the "first one" (i.e., the one with lowest frequency), namely the initial datum with $E_1 = E$, $E_k = 0$ for $k = 2, \cdots, N$, and for example all particles in their equilibrium positions, $x_j = 0$, $j = 1, \ldots, N$.

The essence of their numerical computations is well summarized by the first and the last figures of their paper (corresponding to Figs. 4.1 and 4.2 here). They considered the case $N = 32$ with the first mode initially excited (in the way just mentioned) for a certain value of the total energy E and certain values of α and β. They were expecting that the energy would soon spread over all other modes $k = 2, \cdots, N$. Instead, they found that the values of the instantaneous mode energies E_k versus time t were as in Fig. 4.1. One sees that the energy, initially given to mode 1, passes to the modes 2, 3, 4 and 5 (each of such modes entering the sharing of energy at a proper characteristic

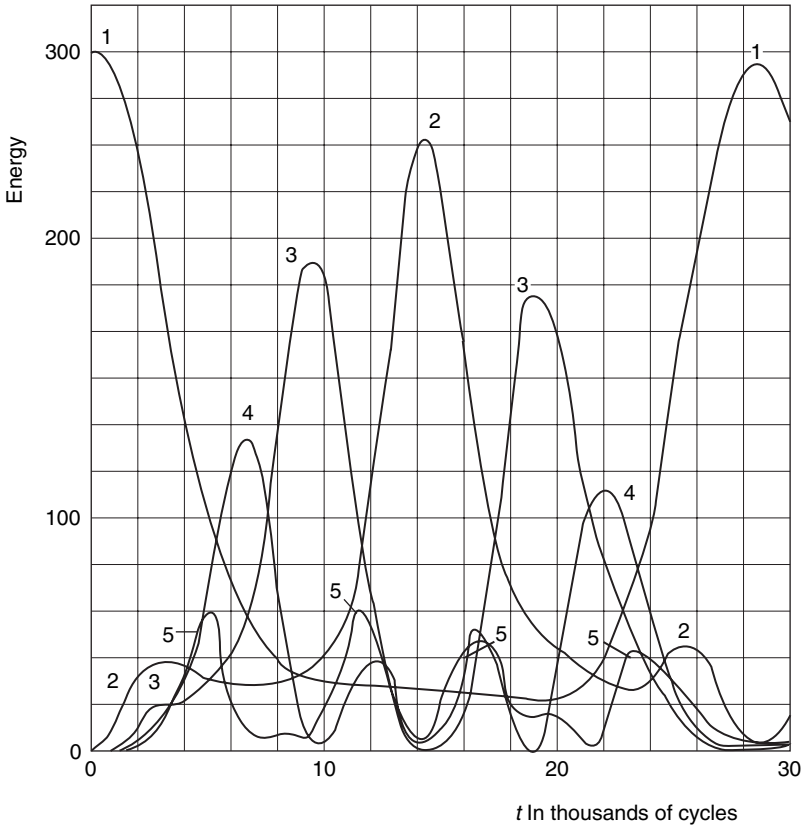

Fig. 4.1. The time evolution of the harmonic energies. The figure is a reproduction of the first one of the original FPU report. Here, $N = 32$ (with $\alpha = 1/4$, $\beta = 0$), and the energy was given initially just to the lowest frequency mode. One sees that the energy, instead of flowing to all the 32 modes, remains confined within a packet of low-frequency modes, namely modes 1 up to 5

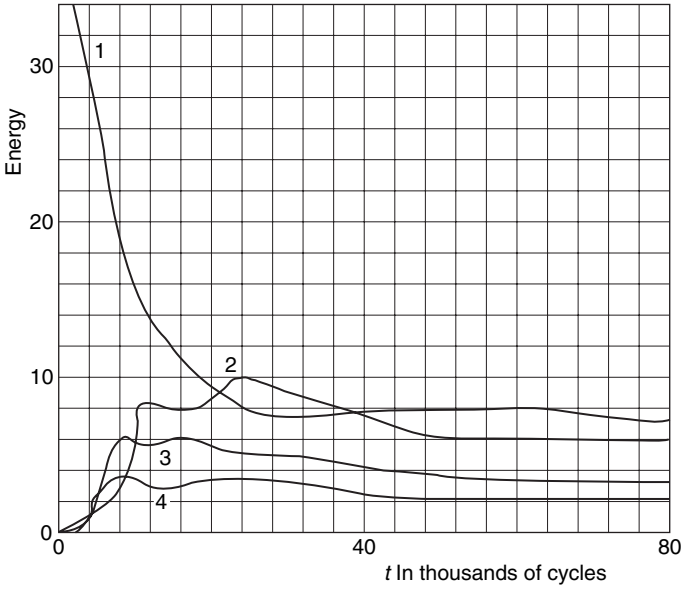

Fig. 4.2. Time-averaged harmonic energies $\overline{E}_k$ versus time. The figure is a reproduction of the last one of the original FPU report

time—this is a point to which we will come back later), and then flows back almost completely to the first mode (this is called the *recurrence property*). In any case, the energy does not appear to flow to the high-frequency modes at all (or almost at all). The most striking feature was however exhibited by the last figure of their paper (Fig. 4.2), where the time-averages $\overline{E}_k(t, x)$ of the energies E_k up to time t were plotted versus time. Indeed such a figure clearly shows not only that the final state predicted by classical equilibrium statistical mechanics was not attained, but also that a relaxation had indeed been attained to some other kind of (apparently stationary) state (after a certain time, the time-averages do not appear to change any more), which is completely different from the final expected one (equipartition). The stabilization of the averages is much more evident in Fig. 4.3, where the calculation has been pushed to a much longer time with respect to Fermi's one.

This is what we like to call *the FPU paradox*: Instead of a slow relaxation to the final equilibrium state, there is exhibited a rather quick relaxation to some kind of "nonstandard" state, in which the energy turns out to be shared only within a *packet* of low-frequency modes, having a certain well defined width, i.e., extending up to some characteristic frequency. One somehow has a kind of "partial thermalization" involving just such a packet, with the high-frequency modes essentially excluded, as if the system were composed only of an *effective number of degrees of freedom*, substantially smaller than N. This fact is well exhibited in Fig. 4.4, where we report the corresponding

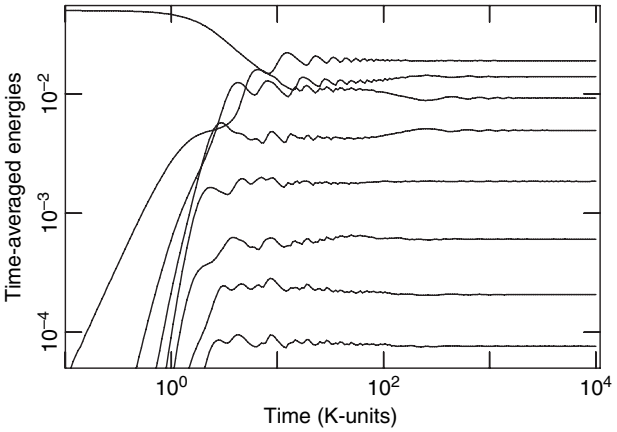

Fig. 4.3. The FPU phenomenon: exhibition of the apparent stabilization. Time-averaged harmonic energies $\overline{E}_k$ versus time in log–log scale, for a time interval much longer than in the original FPU report. The curves are drawn only for the first eight modes. Notice that the "final" value of $\overline{E}_k$ is a decreasing function of the wave-number k (which is not indicated on the corresponding curve in the figure), actually of exponential type (at least for $k > 3$). Here, $N = 32$ and $E = 0.05$ (and thus specific energy $\epsilon = E/N \simeq 0.0015$). Taken from [26]

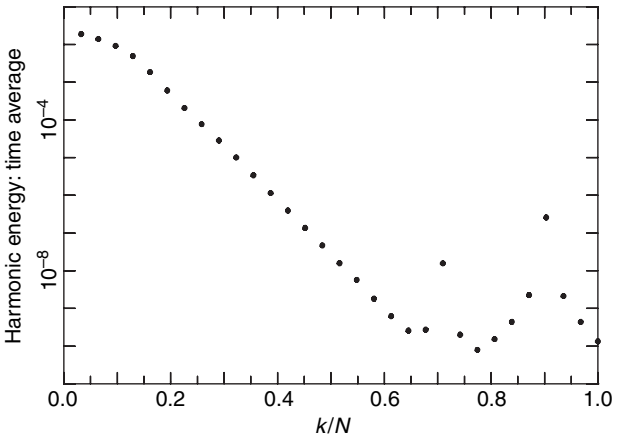

Fig. 4.4. The spectrum (namely the time-averaged energies $\overline{E}_k$ versus k/N), for the same orbit of Fig. 4.3, at the final time of the calculation. Notice the logarithmic scale for the energies. This exhibits how the modes involved in the energy sharing constitute a low-frequency packet with a tail presenting an exponential decay towards the high frequencies

"spectrum", namely the values of the time-averages $\overline{E}_k(t,x)$ versus the mode-number k (actually, versus k/N) at the final time of the calculation. Notice the exponential tail, on which we will come back later. The reaction of Fermi (who had passed away before the paper was written down) is reported by Ulam, in the preface to the reproduction of the paper in Fermi's Collected Papers, in the following terms: "The results of the calculations...were interesting and quite surprising to Fermi. He expressed to me the opinion that they really constituted a little discovery in providing intimations that the prevalent beliefs in the universality of mixing and thermalization in nonlinear systems may not be always justified".

4.2.2 The Paper of Zabusky and Kruskal, and the KdV Equation

With the paper [4] of Zabusky and Kruskal (1965), the Korteweg–de Vries (KdV) equation

$$u_t + uu_x + u_{xxx} = 0$$

entered the game. Here one thinks of a function $u = u(x,t)$ which gives, at time t, the profile of a continuous nonlinear string interpolating the FPU chain of particles. The fact that the interpolation of the FPU chain is rather well described by the KdV equation in certain situations is since then a well-known fact, and is proved in some standard way by multi-scale methods which are familiar in several fields of applied mathematics (see, for example the application given later in [27]).

From the way in which the KdV equation was associated by Kruskal and Zabusky to the FPU model, it is completely clear that the solutions of the KdV equation should provide a good approximation to those of the FPU model only for initial data corresponding to an excitation of low-frequency modes (i.e., for long-wavelength initial data). A relevant further point is however that the agreement should be expected to hold only for not too long times, as was particularly emphasized in the later "deduction" of the KdV equation that was given quite recently by Bambusi and Ponno, through a technique extending to Hamiltonian partial differential equations certain methods of perturbation theory (Birkhoff normal forms) well known in the case of a finite number N of degrees of freedom.

We give here a particular emphasis to the latter fact, because no explicit mention of it is made in the original Zabusky–Kruskal paper. Rather, just at the beginning of the paper, it is said that the KdV equation "can be used to describe the one-dimensional, long time, behavior of small, but finite amplitude,... long waves in the anharmonic crystal." Here, we are pointing out that this should be understood as meaning "long time" within the time-scale up to which the KdV equation provides a good approximation to the solutions of the FPU equations themselves. We will come back to this point later.

In any case, Zabusky and Kruskal studied the KdV equation, and were able to exhibit the existence of three time-scales (or time intervals, as they say),

namely, in their words: "(I) Initially, the first two terms (of the KdV equation) dominate and the classical overtaking phenomenon occurs; that is, u steepens in regions where it has a negative slope. (II) Second, after u has steepened sufficiently, the third term becomes important and serves to prevent the formation of a discontinuity. Instead, oscillations of small wavelength...develop on the left of the front. The amplitudes of the oscillations grow and finally each oscillation achieves an almost steady amplitude...and has a shape almost identical to that of an individual solitary-wave solution (of the KdV equation). (III) Finally, each such 'solitary-wave pulse' or 'soliton' begins to move uniformly...". For a recent numerical illustration of this description, see [28] by Lorenzoni and Paleari.[1]

So the theory of solitons had come to its modern life, and started to be pursued in itself, giving rise to the whole theory of infinite-dimensional integrable systems, while its relation to the FPU problem was somehow neglected. To such a connection we will come back later.

4.2.3 The Izrailev–Chirikov Contribution

a) The discovery of a stochasticity threshold.

The next fundamental contribution was the discovery, by Izrailev and Chirikov, of the so-called stochasticity threshold (see [5, 6]). That is, the FPU paradox disappears if the initial energy is sufficiently large, i.e., there exists a critical energy $E_c = E_c(N)$ such that one has a quick equipartition for $E > E_c$. This is illustrated in Fig. 4.5, where the time evolution of the harmonic energies is calculates for a much larger energy than in Fig. 4.3 ($E = 1$).

Actually, Izrailev and Chirikov considered initial data of a certain broader class than FPU, in that they gave the energy to a packet of modes of nearby frequencies, considering the characteristic frequency of the packet as a parameter. The analog of the FPU paradox was found to occur for any frequency of the excited packet, and in all cases the paradox disappeared above a critical energy E_c depending on the frequency of the excited packet, and on N.

Concerning the theoretical motivation behind such a discovery of the energy threshold, a reading of the Izrailev–Chirikov paper clearly indicates that they had in mind the results on Hamiltonian perturbation theory (KAM theorem) that had just been obtained in Russia by the school of Kolmogorov (see [29]). If one considers a Hamiltonian perturbation of an integrable system, for small perturbations the system resembles very much the integrable one: there exist invariant tori, near the unperturbed ones, and the relative measure of the set of such perturbed invariant tori tends to 1 as the perturbation tends

[1] This paper should be compared with [27], where the first time-scale of Zabusky and Kruskal (namely, the characteristic one for the formation of the packet) was interpreted as the time-scale for equipartition.

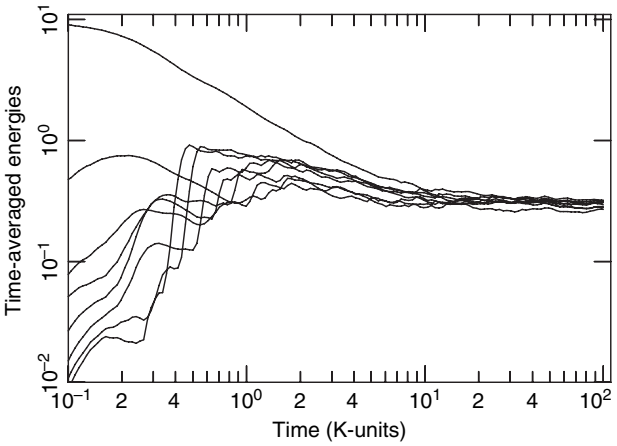

Fig. 4.5. The Izrailev–Chirikov discovery: equipartition of energy is quickly attained if energy is large enough. Time-averaged harmonic energies $\overline{E}_k$ versus time in log–log scale, still for $N = 32$ but now for $E = 1$ (i.e., $\epsilon \simeq 0.9$). Compare with Fig. 4.3, which refers to $E = 0.05$ (i.e., $\epsilon \simeq 0.0015$). Taken from [26]

to 0. Continuity is thus obtained in such a measure-theoretic sense. But the relative measure of the invariant tori is expected in general to decrease as the perturbation is increased so that, at a large enough perturbation, the resemblance of the system to the unperturbed one is essentially completely lost, and the motions are in general expected to present "chaotic" features. This is the reason why in the FPU problem one might expect that, for sufficiently large energies E, chaotic motions should dominate, and this actually led Izrailev and Chirikov to the conception of the existence of a stochasticity threshold E_c.

b) The conjecture of the disappearing of the FPU paradox in the thermodynamic limit.

The Izrailev–Chirikov discovery previously recalled, certainly constituted an extremely relevant contribution. The authors however added something more, by indicating a way in which the FPU paradox might be removed altogether, for the purposes of statistical mechanics.

Indeed, for the aims of statistical mechanics one has to consider the case of extremely large numbers N, i.e., formally, the limit $N \to \infty$. So they pointed out that one should estimate the value of the specific critical energy $\epsilon_c(N) = E_c(N)/N$ in the limit $N \to \infty$. Indeed, the FPU paradox would be completely removed if one could prove that $\epsilon_c(N) \to 0$ for $N \to \infty$. In such a way one would be guaranteed that the FPU phenomenon does not occur for large systems at any positive specific energy $\epsilon > 0$ (i.e., at any finite positive temperature $T > 0$).

The authors even indicated some kind of mathematical mechanism which should govern the vanishing of the limit-specific energy threshold. In this

connection, a relevant role should be played by resonances (i.e., relations of the type $m\omega + n\bar{\omega} = 0$ for two frequencies ω, $\bar{\omega}$, with m, n integers), because the authors had in mind that resonances would lead to stochasticity, as was familiar to them through the so-called Chirikov criterion of the overlapping or resonances. On the other hand, they pointed out that, in the limit $N \to \infty$, the FPU system presents infinitely many resonances. So they worked out some estimates based on this idea, for the case of initial data with an excitation of a few high-frequency modes, and they interpreted their considerations as suggesting that $\epsilon_c(N) \to 0$ in that case. An analogous conclusion could not be drawn for the case of initial data with excitations of low-frequency modes (the case considered in the FPU work). Quite recently, their pupil Shepelyansky, elaborating on their methods, maintained to have extended such a result to that case too (see [7]).

Serious doubts on the significance of the criterion of the overlapping of resonances may actually be raised (see [8]). In any case, however, one can say that a physical conjecture had emerged, namely, that the FPU paradox may disappear entirely in the thermodynamic limit.

4.2.4 The Result of Bocchieri et al.

Five years later (1971), in [9] Bocchieri et al. gave numerical indications in the opposite direction: The FPU paradox should persist in the limit $N \to \infty$. They actually performed computations for a slight modification of the FPU model, inasmuch as they introduced a "realistic" potential of Lennard-Jones type, namely,

$$V(r) = 4V_0 \left[(\sigma/r)^{12} - (\sigma/r)^6\right]$$

involving two parameters, the depth V_0 of the potential well and the typical distance σ, at which the potential passes from positive to negative values. They considered several types of initial data with a few nearby modes excited, of low, or of high, or of intermediate frequency, and found that in a short time equilibrium is attained (equipartition of the time-averages of the mode-energies was obtained), if the initial energy is sufficiently large, i.e., for $E > E_c(N)$ for some critical energy $E_c(N)$, in agreement with the discovery of Izrailev and Chirikov. For what concerns the dependence of the specific stochasticity threshold $\epsilon_c(N) = E_c(N)/N$ on N, they found a large dependence for small N, say for $2 < N < 10$, whereas $\epsilon_c(N)$ was found to be essentially constant for "large" N (concretely, in their computations, for $10 \le N \le 100$). They actually found for $\epsilon_c(N)$ the "limit" value $\simeq (3/100)V_0$. In their words: "When the energy of vibration per particle is equal or larger than 2 or 3 percent of the potential well and the number of particles is sufficiently large, one has, in time average, equipartition of energy among the normal modes".[2] Namely,

$$\varepsilon_c \simeq 0.03\, V_0 \ .$$

[2] It must however be added that the authors were completely aware of the possible relevance of the actual times of observation, because they also added: "We may

4.2.5 The Suggestion of a Possible Physical Interpretation

At this point the situation was as follows. Izrailev and Chirikov had with an extreme clarity indicated how one might eliminate the FPU paradox entirely, for systems of interest to statistical mechanics: One should prove that the specific energy threshold $\epsilon_c(N) = E_c(N)/N$ vanishes in the limit $N \to \infty$. Moreover, they believed to have shown that this is the case at least in the case of high-frequency excitations. On the other hand, Bocchieri et al. had given indications in the opposite direction. Thus, there was the problem of which could be a physical interpretation for the apparently stationary state exhibited by FPU, in case the indications of Bocchieri et al. were confirmed.

The idea that the relations between classical mechanics and quantum mechanics may be much subtler than usually believed was very much discussed within the group of theoretical physicists in Milano, particularly under the stimulus of Caldirola and Loinger. Thus, as one of the strongest and deepest manifestations of quantum mechanics in a statistical mechanics frame, actually the one that gave rise to quantum mechanics itself, is the lack of energy equipartition at low temperatures, the systems behaving as if the high-frequency modes were excluded from the energy sharing, quite naturally there arose the idea that the "nonstandard" apparently stationary FPU states may be a sort of classical analogs of quantum states. These are characterized by the Planck spectrum E_k^{P} given by

$$E_k^{\mathrm{P}} = \frac{\hbar\omega_k}{\exp(\beta\hbar\omega_k) - 1}$$

where $\hbar$ is Planck's constant and $\beta = 1/(k_{\mathrm{B}}T)$ the "inverse temperature".

In such a way, after many discussions with Bocchieri and Loinger, Galgani and Scotti started out an investigation in which the FPU spectrum below threshold was fitted to a Planck-like distribution E_k^{Plike}, namely,

$$E_k^{\mathrm{Plike}} = \frac{A\omega_k}{\exp(\beta A\omega_k) - 1} \ ,$$

with two free parameters A and β, having the dimensions of an action and of an inverse temperature, respectively. The fits were made to data obtained with the same computer program used by Bocchieri et al., in which the molecular parameters m (mass of the particles), V_0 and σ (the ones entering the

conclude by saying that, in the case of very low total energies, the relaxation mechanism towards the standard Boltzmann distribution of the normal modes may act so slowly that the coupling of the system with a thermal bath could be very important in determining the approach of the model towards such a distribution." This remark, by the way, actually opens another relevant problem, because it may happen that also the mechanisms of transfer of energy between a FPU system and a heat reservoir are slowed down when temperature is lowered. In fact, this actually seems to be the case.

Lennard-Jones potential) actually considered were those of Argon, as taken from standard available handbooks.

The result found was a rather striking one. Indeed, not only the qualitative fit to the Planck-like law was found to be rather good, with β behaving as expected (namely, as an inverse temperature depending only on the specific energy $\epsilon = E/N$), but it was also found that the parameter A on the one hand was pretty constant, i.e. independent of the specific energy, and on the other hand happened to have a value quite near to the Planck constant $\hbar$.

It took some time to understand how this could have happened. The simple reason is that A, being an action, has to be proportional to the natural action obtained from the dimensional parameters m, V_0, σ introduced in the model, which is $\sqrt{mV_0}\,\sigma$. So one necessarily has

$$A = \alpha\sqrt{mV_0}\,\sigma\ ,$$

where α is a pure number. On the other hand, it is well known that the molecular parameters actually met in nature do indeed contain $\hbar$, and in particular one has, for example for the noble gases,

$$\sqrt{mV_0}\,\sigma \simeq 2Z\hbar\ ,$$

where Z is the atomic number. In conclusion, Planck's constant had been introduced somehow by hands in the model through the molecular parameters. This is the way in which Galgani and Scotti came to venture the suggestion (see [10, 11]) that, if one can prove that the FPU paradox persists in the thermodynamic limit, then the apparently stationary FPU states may provide a sort of classical analog to the quantum degeneration described by Planck's law. This idea continued to be pursued up to the present days (see [30, 31]).

By the way, it may be noted that the good fit of the FPU spectrum to Planck's law (suggested by the analogy with quantum mechanics) amounts to be perhaps the first clear exhibition of the fact that, for large wavenumbers k, the energies $\overline{E}_k$ decay exponentially fast with k, i.e., with the frequency ω_k. This fact is indeed a quite general one, the nature of which was clearly understood analytically ten years later with the paper of Fucito et al., who, through the intermediary of the paper of Frisch and Morf, transported to the FPU problem general ideas of turbulence theory.

This ends the first phase of the history of the FPU problem in our personal way of reconstructing it.

4.3 A Voice in the Desert: The Paper of Fucito et al. (1982) and the Proposal of a Metastability Scenario. The Work of Parisi and the Analogy with Glasses. Relations with Turbulence Theory

In the first phase of the history we traced back in the previous section, due to the need of concentrating our attention on the papers that are most relevant

for our reconstruction, we already had to neglect a considerable amount of papers, among which stay for example several ones of the late J. Ford, the memory of whom is particularly dear to the oldest of the present authors, who exchanged with him tenths of letters on the subject. In the same way, we are going to neglect in the present section many other papers, including several ones worked out by the present authors.

The next relevant step was made with the paper [13] of Fucito et al., where both a new point of view and a new technique were introduced.

The new point of view concerns metastability and was certainly borrowed from the frame of the theory of glasses and of disordered systems, in which distinguished contributions had been given in Roma by Parisi. The idea is that the FPU state is just an apparently (rather than a true) stationary one. This is somehow at variance with the attitude of Izrailev and Chirikov, who were apparently thinking in terms of truly stationary states; indeed they were explicitly making reference to KAM theorem, which is expressed in terms of invariant surfaces (that is, surfaces on which the orbits lie for all times). Instead, in the paper of Fucito et al. reference is made to the quick formation of a state which remains essentially undisturbed for extremely long times, until it eventually precipitates through a "catastrophic mechanism" to the true "final" equilibrium state.

The new technique is simply that of relating the decay of the tail of the spectrum to the singularities of the analytic continuation of a field interpolating the positions of the FPU model. This idea was borrowed from a very interesting paper of Frisch and Morf (see [32]), the aim of which was to understand certain features of turbulence theory (see [33]) as manifestations of quite general relations between the high-frequency tail of the Fourier transform (in complex time) of a temporal signal and the singularities of the analytical continuation of the signal itself. By the way, it can be noted that the existence of a deep analogy between the problem of a dynamical justification of the Boltzmann–Gibbs equipartition principle, and the general problem of turbulence, had been clearly pointed out by von Neumann (see [34]), in the year 1949.[3]

In fact, Fucito et al. were not actually studying the FPU model itself, but rather a variant of it, namely, the so-called φ^4-model (to be presently recalled),

[3] Such an analogy between turbulence and ordinary statistical mechanics permeates the whole paper of von Neumann. See, for example p. 445, where it is said: "The $k^{-5/3}$ law calls for an interpretation akin to (although not identical with) the ultraviolet catastrophe of black-body radiation theory", and reference to "non-ergodic conservation laws" is made. See also p. 447 and finally p. 468, where it is said: "From the point of view of statistical physics, turbulence is the first clear-cut instance calling for a new form of statistical mechanics. ... The existing theories. ... suffice to show that those laws will differ essentially from those of classical (Maxwell–Boltzmann–Gibbsian) statistical mechanics. Thus it is certain that the law of equipartition of energy between all degrees of freedom, which is valid in the latter, is replaced by something altogether different in the former."

which was a very familiar one in field theory (and had also been studied a little before in the spirit of the FPU problem in [35]). In fact, it turns out that the techniques used by Fucito et al. to investigate the φ^4-model cannot be immediately transported to the FPU model itself (a subsequent attempt will be mentioned later), but the transport of the global scenario proved instead to be possible (actually in terms of the work of Zabusky and Kruskal), as shown later by Bambusi and Ponno.

Perhaps, as a preliminary introduction to the description of the paper of Fucito et al., it may be useful to illustrate the main phenomena understood by them, through the Figs. 4.6 and 4.7, which refer the FPU model (with $N=127$ and specific energies $\epsilon=10^{-4}$ and $\epsilon=5\times10^{-3}$ respectively). In each figure, the spectrum (namely, the plot of the time-averaged energies $\overline{E}_k$ versus k/N) is reported at successive times t_j (with $t_{j+1}=10\,t_j$). From Fig. 4.6 one clearly sees that, at any observation time, the spectrum consists of both a packet of low-frequency modes (having a tail which decreases exponentially fast with k/N) which essentially contains the whole available energy, and of a complementary packet of high-frequency modes displaying an essential equipartition of energy at a much smaller energy. One also observes that the slope of the main low-frequency packet decreases as time increases, until it appears to have come to a stop, remaining essentially constant (this is the phenomenon of the apparent stationarity) during a rather long-time interval (covering at least four orders of magnitude in the case of the figure). This occurs for $\epsilon = 10^{-4}$. But if one considers a larger specific energy ($\epsilon = 5\times10^{-3}$ in the case of Fig. 4.7), then the same phenomenology is speeded up, and within the same final observation time ($t = 10^8$), a further phenomenon is exhibited. This is the final attainment of global equipartition, which occurs through a quite different mechanism. Indeed, one might have imagined that the final global equipartition be attained through a successive decreasing of the slope of the tail. Instead, the approach occurs in the following way. The complementary packet of high-frequency modes continues to essentially display partial equipartition at an energy smaller than that of the main low-frequency packet, and what occurs is that the level of the energy of the complementary packet rises as time increases,[4] until global equipartition is attained. Note that, in both figures, the scale of the ordinates is not the same at the various times. In conclusion, one observes the existence of two different mechanisms: first, the quick formation of a packet of low-frequency modes with an exponential tail having an apparently stabilized slope (formation of the metastable state), and, secondly, the final approach to global equipartition through the rising of the equipartition level of the complementary packet.

After this introduction, let us finally come to an illustration of the paper of Fucito et al. They consider the one-dimensional nonlinear Klein–Gordon equation

$$\varphi_{tt} = \varphi_{xx} - m^2\varphi - g\varphi^3 \ ,$$

[4] Moreover, the complementary packet extends its size towards the left.

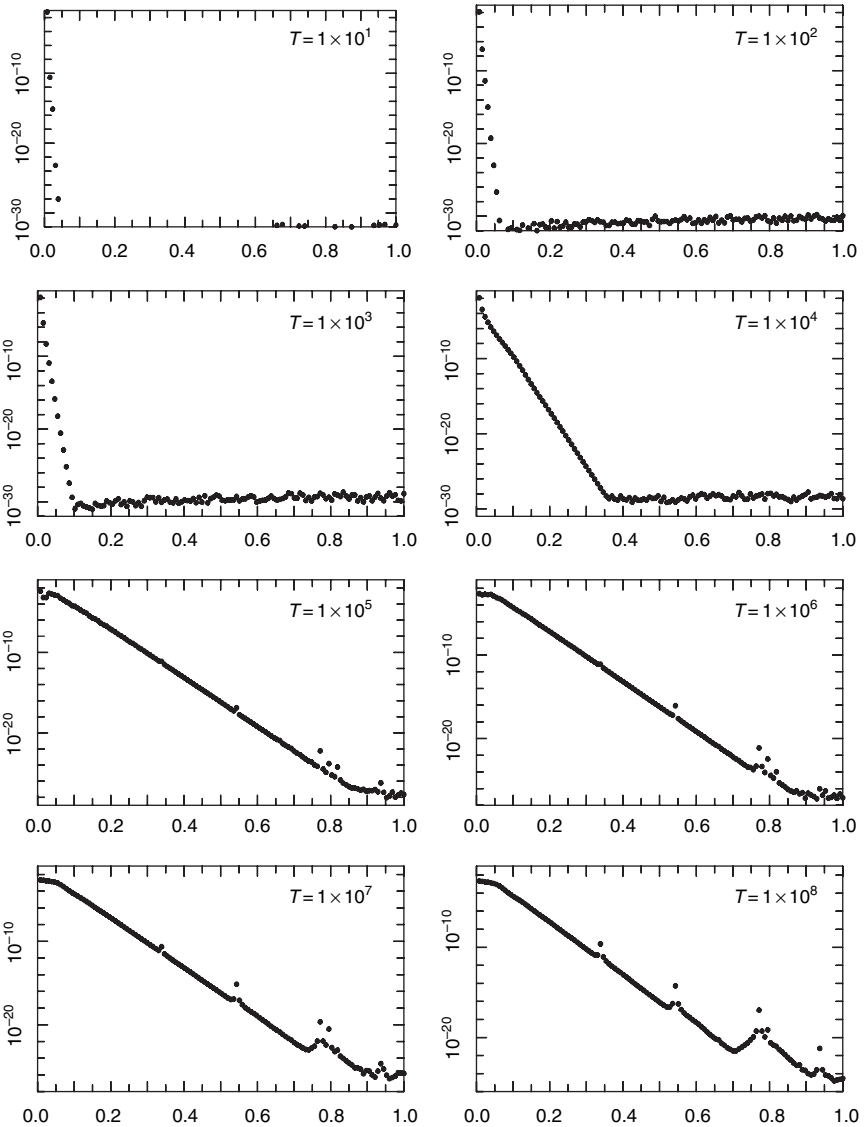

Fig. 4.6. The spectrum at several times $(10, 10^2, \ldots, 10^8)$. First phase: formation of the metastable state (note the change in the vertical scale of the figures). Here, FPU model with initial data of FPU type, $N = 127$, $\epsilon = 1 \times 10^{-4}$)

where the real, one-component, field $\varphi(x,t)$ is defined in the interval $-L/2 \leq x \leq L/2$ with periodic boundary conditions, and m and g are positive parameters. From this partial differential equation, a discretization leading to an analog of the FPU model is immediately obtained. The name φ^4-model is

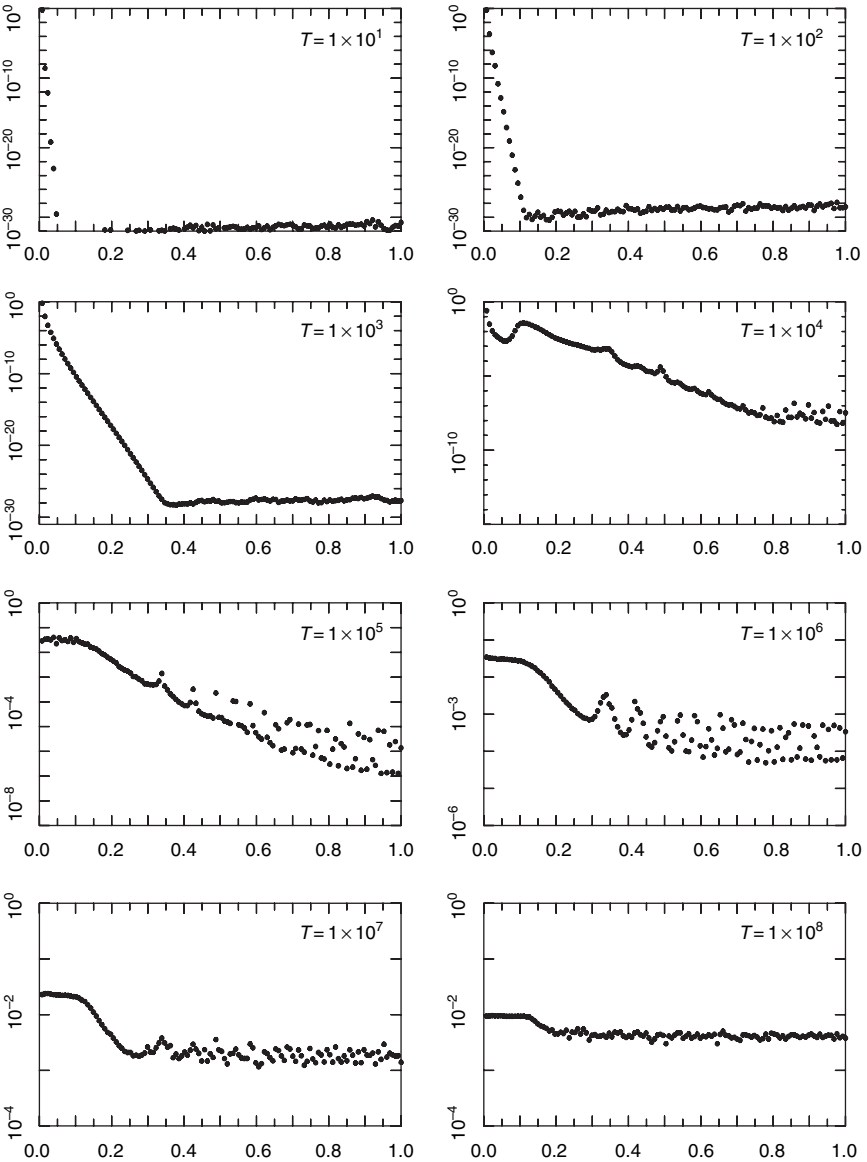

Fig. 4.7. The spectrum at several times $(10, 10^2, \ldots, 10^8)$. Illustration of the final phase following the first one, with the attainment of global equipartition (note the change in the vertical scale of the figures). Same as Fig. 4.6, but now with $\epsilon = 5 \times 10^{-3}$. The time-scale of observation is the same in both cases

due to the fact that the only nonlinearity in the model comes from a term φ^4 in the potential energy.

The quantity they are interested in is the analog of the spectrum previously discussed for the FPU model, i.e., the distribution of energy (in time-average) among the modes, as a function of time. To this end, they introduce the space Fourier transform of the field φ by

$$\hat{\varphi}(k,t) = (2\pi)^{-1/2} \int_{-L/2}^{L/2} \mathrm{d}x \; \mathrm{e}^{-ikx} \varphi(x,t)$$

and define the *power spectrum*[5] $W(k,t)$ by

$$W(k,t) = |\hat{\varphi}(k,t)|^2 \; .$$

Notice that the continuum analog of the spectrum previously defined for the FPU model would rather be twice the quantity $k^2 W(k,t)$.

Anyway, they are interested in investigating the form of the spectrum $W = W(k)$ as a function of time t for large values of k. To this end they make reference to well-known analyticity properties of Fourier transforms, and notice: "We expect the field φ to reach asymptotically a thermal equilibrium distribution, given by a Boltzmann factor $\mathrm{e}^{-\beta H}$ for some value of the inverse temperature β determined by the initial conditions. In this case, at values of the wavenumber k so large that the mass and nonlinear terms of H are negligible, one would have

$$W(k,t) \simeq \text{const.} \times k^{-2} \; . \tag{4.3}$$

This behavior of W corresponds to functions $\varphi(x,t)$ which are not differentiable with respect to x. Now it is known that, since $\varphi(x,0)$ is analytical as a function of x, the solution $\varphi(x,t)$ will remain analytical at any finite time t. Equation (4.3) can only be valid for infinite time. This means that, as time goes on, singularities of $\varphi(x,t)$ appear in the complex x plane which creep towards the real axis and accumulate onto it at infinite times. We show below that these singularities are simple poles."

Indeed, it is well known that one may relate such singularities to the large k behavior of W by means of the theorem of residues. Therefore one "obtains the following asymptotic behavior of $W(k,t)$ at large k:

$$W(k,t) \simeq \text{const.}\, \mathrm{e}^{-2k y_{\mathrm{S}}(t)} \; ,$$

where $y_{\mathrm{S}}(t)$ is the imaginary part of the location of the pole which lies nearest to the real axis."

In conclusion, "the strategy is then to evaluate the most likely value of $y_{\mathrm{S}}(t)$ by extending the approach of Frisch and Morf to a deterministic partial differential equation."

[5] They actually call it just *the spectrum.*

So, one remains with the problem of evaluating the most likely value of $y_{\mathrm{S}}(t)$. To this end, the idea was to exploit a particular feature of the φ^4-model in connection with initial data of FPU type, namely with small k and thus large wavelengths. Indeed, this means that initially, and actually up to times until which energy did not yet flow to high k modes, the profile of the field does not present large curvatures, and thus the term φ_{xx} in the equation of motion can be neglected, with the consequence that the equation of motion reduces to an ordinary one depending parametrically on the space coordinate x. So, up to not too large times, for any x one has an unknown $\varphi = \varphi(t)$ obeying the ordinary differential equation

$$\ddot{\varphi} = -m^2\varphi - g\varphi^3 \ . \tag{4.4}$$

An analytical study of such an equation is easily performed, and this leads to the result that, for initial data of the form

$$\varphi(x,0) = A\cos(k_0 x) \ , \quad \varphi_t(x,0) = 0$$

with k_0 small, the imaginary part of the nearest pole starts descending from infinity towards the real axis, approaching a point with imaginary part $(1/k_0)$ $\ln(ma/g^{1/2})$. This is illustrated in a very beautiful way in a subsequent paper by Bassetti et al. (see [36]), where the relevant poles for the φ^4-model were computed numerically by the technique of the Padé approximants.

So, there exists a first temporal phase of the dynamics, in which the spectrum presents an exponential decay towards the high wave-numbers k, with a slope decreasing as

$$|y_S(t)| = -\ln(tAg^{1/2})/k_0 \ .$$

This corresponds, for the low frequency modes, to an increase of energy as a power of t (formation of the packet). This stage, by the way, is the analog of the one in which, in the terminology of Zabusky and Kruskal, the third-derivative term u_{xxx} can be neglected, i.e.: "Initially, the first two terms (of the KdV equation) dominate and the classical overtaking phenomenon occurs; that is, u steepens in regions where it has a negative slope."

One might thus expect that equipartition will eventually occur, with the slope tending to zero (i.e., with the pole approaching the real axis). But this is not the case. As mentioned previously, the poles do not collapse onto the real axis, because the imaginary part tends to a finite positive value and so the slope stops decreasing (see [36]). At this point, according to Fucito et al. the contribution of the Laplacian starts becoming relevant, and this fact can be looked upon as the addition of a noise to the r.h.s. of (4.4). This further stage of the process is described at p. 710 of the paper by Fucito et al. in terms of a probabilistic analysis performed on the harmonic chain ($g = 0$) in

the limit of infinite length ($L \to \infty$), in which use is made of the known fact that the one-point probability distribution function of a classical harmonic field in dimension one is Gaussian. This leads to an extremely slow decrease of the slope.

The authors then turn to a qualitative discussion of the final stage of the process of approach to equipartition. The analysis is made in terms of the variance σ^2 of the Gaussian probability distribution function previously mentioned. They say: "The main effect of the nonlinear terms in this regime will be to change the value of σ^2. If σ^2 were time independent," the previous analysis "would be essentially correct. Let us distinguish between the role of short and long wavelength modes. At the times we are interested in, most of the energy is contained in the long wavelength modes, which may be assumed to be in a kind of thermal equilibrium among themselves. Their contribution to σ^2 may then be considered as essentially constant in time. The short wavelength modes will however also contribute to σ^2. As long as $W(k,t)$ is small in the large k region, their contribution is negligible. As time goes on, however, $W(k,t)$ will start increasing, what will increase the value of σ^2 and fasten therefore the transfer of energy to short wavelength modes. This triggers a catastrophic process which our analytical tools are unable to handle. We cannot therefore draw conclusions about the behavior at very long times before thermal equilibrium is reached."

In conclusion, here for the first time one finds explicitly expressed the conjecture that, for all values of the perturbation, at sufficiently long times one will attain the standard equilibrium state, and the FPU paradox is interpreted as corresponding to a preliminary stage of the process in which the energy, initially given to extremely low-frequency modes, quickly flows to a larger packet of low-frequency modes (with an exponential decay towards the high frequencies) and remains frozen there up to an extremely long time. This is what we informally call *the metastability scenario*. For what concerns the law describing the dynamical evolution towards equipartition, in the subsequent work [39] by Parisi numerical indications were given that the corresponding time scale could be a stretched exponential in terms of the inverse of the specific energy, rather than a simple exponential.

As previously pointed out, the paper of Fucito et al., with its interpretation of the FPU paradox as a metastability phenomenon, did not produce a great impact, and even was essentially forgotten for a long time. For example, if one looks at the 21 papers published quite recently on the FPU problem in a special issue that a journal devoted to it on the occasion of the 50 years from the original work, one will find out that not one of them even mentions that paper, apart from the papers [37, 38], where it is amply discussed (see also [26]).

How could this have happened? In our opinion, the main reason is that at those times the key point under discussion was the choice between the two alternatives previously mentioned about the energy threshold, namely, whether the specific energy threshold $\epsilon_c(N)$ vanishes or not in the limit $N \to \infty$; on the

other hand, no mention of a threshold at all was made in the paper of Fucito et al. According to them, equipartition should be attained at all energies. This statement, that no threshold should exist, was particularly emphasized by Parisi in [39]. So, actually, there was some misunderstanding about the sense to be attributed to the word "threshold". Indeed, Parisi was stressing that energy equipartition should be attained (after a sufficiently long time) at any specific energy, and such a conjecture is obviously opposed to the conception of a threshold, if the latter is meant as the specific energy below which equipartition is never attained. On the other hand, there is no opposition, if the threshold is understood in a softer way, namely, in the sense that for smaller energies one meets with a state which is only apparently stationary, and will later evolve, on a much longer time scale, to equipartition, i.e., to the final "true" equilibrium state.[6] However, the relation with the threshold in the sense of Bocchieri et al. was not discussed in an explicit way. In our opinion, this is the fact that generated some confusion. The situation was finally clarified in the paper of Berchialla et al. (described in a subsequent section), where a clear exhibition was given of the fact that two well distinct time-scales of relaxation exist, but only below a certain critical specific energy, which could thus be interpreted as the threshold previously discussed by Izrailev–Chirikov and by Bocchieri et al.

We mention now the very few papers in which the work of Fucito et al. was discussed.

The previously mentioned paper [36] somehow constitutes an appendix to the paper of Fucito et al., because it reported numerical computations of the relevant poles in the φ^4-model, showing a very good agreement with the theoretical predictions and some further details. Something analogous can be said of the paper [40], still devoted to numerical investigations on the φ^4 model. Here, following Frisch and Morf, the analysis concerns the statistical aspects of the field φ, which is shown to present typical non-Gaussian features. By the way, it may be worth mentioning that analogous indications of non-Gaussian behaviors were also reported much later for the process of energy exchanges of the internal degrees of freedom of diatomic molecules produced by atomic collisions (see [41]).

The suggestion that a description analogous to that of Fucito et al. for the φ^4-model could be given also for the FPU model was first advanced and

[6] In the words of Fucito et al.: "One of our main results is that the system reaches equilibrium with a logarithmic dependence on t, so that the nonequilibrium spectrum may persist for extremely long times, and may be mistaken for a stationary state if the observation time is not sufficiently long". By the way, one also finds here the words: "It is amusing to remark that the quasi-equilibrium distribution is similar to Wien's law for black body radiation with a slowly varying Planck's constant, a statement which is clearly inspired by the possible physical interpretation proposed in the work [10].

discussed in [42],[7] (see also [43])[8] in which numerical computations were performed on the FPU model itself (actually on the so called β-model, i.e., the one with $\alpha = 0$). The previously mentioned "slope" of the straight line describing, for large k, the exponential decay of $\overline{E}_k$ versus k in semi-log scale was investigated and was shown to stop decreasing, but the further evolution was not investigated. The accent was rather put on the fact that a quick approach to equilibrium occurs only for high enough energy, and the authors even added the comment (p. 3550): "The numerical results that we have described in the present section yield to the interesting conclusion that a threshold value exists, below which the equipartition of energy is never reached."

The same problem was rediscussed two years later (see [44]), still for the FPU β-model. The accent was still put on the existence of an energy threshold, trying to make a decision between the conjecture of Izrailev–Chirikov and that of Bocchieri et al. (that was there called the Galgani conjecture), and the authors said: "For the N dependence our results seem to be unquestionable and in contrast with the existing theoretical predictions" (of Izrailev and Chirikov). The point of view of Fucito et al. was mentioned in the conclusions, where they added the comment: "But as far as the time dependence is concerned we cannot conclude that the threshold does not vanish as t approaches ∞ ... The situation can be likened to the very slow relaxation behavior in disordered systems, where the evolution towards 'equilibrium' takes place through metastable states, approached at different time scales." Analogous conclusions were reached for the FPU α-model in [45].

4.4 Other Pathways

The metastability perspective, initiated with the work of Fucito et al., was finally recovered 20 years later, with the paper of Berchialla et al. that will be illustrated in the next section. Many more works were written down in the meantime by several authors (for example Kantz et al. entered the game), with an attention to several interesting problems. We cannot follow them here in detail, but it seems to us that, apparently, no reference to the metastability

[7] From the technical point of view, difficulties were met in trying to describe the motion of the poles through a direct transport of the method used by Fucito et al., which was a very special one devised for the φ^4-model. In fact the authors proposed somehow a partial differential equation which, as we now understand, can give a good agreement only for extremely short times, the ones corresponding to the first stage described by Zabusky and Kruskal, because it does not even prevent the formation of a discontinuity (in the terminology of Fucito et al., it does not prevent the falling of the poles on the real axis), and so does not lead to a blocking of the decay of the slope in the spectrum (stage II of KZ).

[8] Here, the idea is suggested that in the α–β model the β-term dominates over the α one, also at very low energies.

perspective can be found there. In the present section, we limit ourselves with a short survey of some other problems that were dealt with.

4.4.1 The Idea of Long Relaxation Times, Boltzmann and Jeans, Nekhoroshev and Landau-Teller

In the meantime, people had started becoming familiar with the fact that the relaxation times to equilibrium can actually be extremely long (see [46]). This in fact had been much discussed by Boltzmann himself and by Jeans, who had conceived of explaining by such a mechanism the observed lack of equipartition in nature (see the quotations in [47]). The same fact was later understood in terms of perturbation theory through the work of Nekhoroshev (see [48]), and through a reconsideration of the work of Landau and Teller of the years 30s on the exchanges of energy of the internal degrees of freedom in atomic collisions (see [49]). Problems of this kind actually became very popular and were much investigated, and would deserve a long discussion. Here we only remark that, while on the one hand the existence of long relaxation times was well understood, on the other hand there was no completely clear awareness of the fact that in a very short time some kind of equilibrium (or apparent equilibrium, or metaequilibrium) is attained (see however the works [50, 51]). Such a quick approach to a metaequilibrium state corresponds to what we now call the quick formation of a packet (presenting a partial thermalization), which is the one accounted for by the first two stages of Zabusky and Kruskal, and of Fucito et al.

4.4.2 The Works around Pettini

The existence of long relaxation times for the FPU problem (and also for the φ^4-model) in the spirit of Nekhoroshev's theorem was first discussed and exhibited by Pettini and Landolfi (see [52]). Indeed, already in the abstract of their paper, they make the following quite clear statement: "Below a critical value ϵ_c ... of the energy density ϵ, the relaxation time τ_R is found to follow a 'Nekhoroshev-like' law, i.e., $\tau_R = \tau_0 \exp(\epsilon_0/\epsilon)^\gamma$", and also add: "A remarkable difference with respect to Nekhoroshev's theorem (where the exponent γ scales as $1/N^2$) is the N independence of numerical experiments results. An important consequence of this fact is the existence of nonequilibrium states of arbitrary lifetimes also at large N values. On the other hand, at high-energy densities ($\epsilon > \epsilon_c$), τ_R is almost independent of ϵ". Such a scenario of Pettini and Landolfi seems to perfectly agree, actually anticipating it, with the one described in the next section along the lines of the work of Berchialla et al. However, from some subsequent works (see, for example the review paper [53]) one may have the impression that the authors rather started adhering to the scenario of Izrailev and Chirikov. We hope to come back to this point on another occasion.

In some subsequent papers (see [54]) a very ingenious method, based on certain considerations on the geometry of phase space, was devised which allowed Pettini and his collaborators to provide a semianalytical estimate of the maximal Lyapunov Characteristic Exponent as a function of the specific energy ϵ_c in the thermodynamic limit. This result, although not yet completely cleaned up from an analytic point of view, constitutes in our opinion one of the most relevant contributions to the subject. The curve of the maximal LCE versus the specific energy ϵ had been numerically investigated by Casartelli et al. in the paper[9] [56] for the FPU model with Lennard-Jones potential. Later, in their paper [52], Pettini and Landolfi found the interesting result that such a curve presents a well-marked knee at a certain value of ϵ. In fact, an analogous remark had been made three years before in a paper of Butera and Caravati (see [55]) for a plane model of rotators (the so-called O(2) planar Heisenberg model), in which the position of the knee had been associated to the presesnce of a certain phase transition (of Kosterlitz and Thouless). For previous works on the model of rotators see [57] and [58].

4.4.3 Metastability and Specific Heats

A very interesting discussion had also been started concerning estimates for the fluctuations of energy in subsystems of the FPU model. The aim was to understand whether the FPU model may be of interest in connection with the problem of the specific heats. In order to obtain some estimates through numerical studies on isolated systems, without having to make recourse to an interaction with a heat reservoir, the attention was addressed to the energy fluctuations of a subsystem of the FPU system of interest: the aim was to compare the fluctuations computed as time-averages with those expected at equilibrium, since the relation of the latter ones with the specific heat is well known. Two apparently opposite results had been obtained in the papers [59, 60] (by Livi et al. and by Perronace and Tenenbaum, respectively). The difference could be explained as due to the fact that two completely different kinds of subsystems had been considered in such papers: a spatial piece of the FPU string in [59], where the time-averages were found to agree with the equilibrium expectations, and a packet of modes of nearby frequencies in [60], where an analog of the FPU paradox was observed, because the time-averages were apparently found to tend to zero as temperature decreases.

So, there naturally arose the idea of eliminating all the problem of the good choice of the subsystem, by estimating the specific heat directly through the energy actually exchanged between the whole FPU system and a heat reservoir (see [15] by Carati and Galgani). Obviously this in turn opens the new problem which might be considered to be a good model for the energy exchanges with the reservoir, a problem we shall not discuss here. We just

[9] This, by the way, is the paper where the now familiar technique of computing the maximal LCE was first introduced.

limit ourselves to mention that in such a way some analogies between the FPU system and the glasses were pointed out, and this fact was instrumental to rediscover and recover the metastability perspective introduced by Fucito et al. Moreover, another interesting fact was observed. Namely, something analogous to the formation of a low-frequency packet (which is a standard result for long-wavelength initial data) occurs even if one starts up with initial data extracted from a Boltzmann–Gibbs distribution at a certain temperature (and so essentially with equipartition of energy among the modes). Indeed it was found (see [16]) that, if the FPU system (with initial data of the just mentioned type) is put in contact with a heat reservoir having a slightly different temperature, then only a small packet of low-frequency modes does manifest a quick reaction to the reservoir, attaining equipartition at the temperature of the latter, whereas the high-frequency modes do not manifest any reaction at all. Presumably, they too will much later attain global equipartition, in analogy with what occurs for an isolated system with long-wavelength initial data.

This fact naturally leads to expect (see [31, 60]) that metastability phenomena may show up in actual measurements of the specific heats (for example of crystals) at low temperatures, more or less in the spirit of the rationale of the time-dependent specific heats, as discussed for example by Birge and Nagel (see [61]). On this very interesting problem we plan to come back elsewhere.

4.4.4 Towards the Natural Packet through Resonance: the FPU Model with Alternating Masses

It will be shown later that, in order to understand the quick formation of an apparently stationary state, a key point is played by some relevant resonances. Such a role, already pointed out for the FPU problem in the pioneering work of Ford of the year 1961 (see [62]), became particularly evident when a modification of the original FPU model was studied (see [63]). We refer to the so-called FPU model with alternating masses which is very familiar in solid state physics, namely, the one in which the successive material points of the FPU chain have masses $m, M, m, M, \ldots$ with $m < M$. The main qualitative consequence of such a modification is that the "dispersion relation", namely, the function $\omega = \omega(k)$, presents now two branches: the "acoustical" one (emanating from near the origin) and the "optical" one, characterized by larger frequencies. The separation between the two branches becomes larger and larger (with the optical one tending to become a horizontal curve, i.e., with all frequencies equal) as the ratio M/m of the two masses is increased. So one meets here (for M/m large) with two clearly distinct subsystems, each of which can be essentially considered as completely resonant, being characterized by essentially just one frequency. Resonant systems had been previously studied in the frame of Nekhoroshev theorem (see [64]), and it had been well understood that chaotic motions in general occur within each single resonant subsystem, whereas the exchange of energy between the two subsystems is

in general extremely slow. Furthermore, the strong dependence of the results on the number of elements constituting a subsystem was almost completely eliminated. Notice that the dependence of the estimates on N in the general case is instead quite heavy, and this fact was often interpreted as indicating that "chaos should prevail" in the thermodynamic limit.

4.5 The Resurgence of the Metastability Perspective, and Its Compatibility with the Existence of a Specific Energy Threshold: The Natural Packet and the Two Relaxation Times

A very clear numerical illustration of the phenomenon of metastability, exhibiting on the one hand the existence, at low energies, of two well separated time-scales (i.e., the quick formation of a "natural packet" which persists up to very long times, when the final relaxation to equipartition occurs), and on the other hand the existence of a stochasticity threshold in the sense of Bocchieri et al., was given in [14] by Berchialla et al. The main underlying idea was to measure the width of the packet that is quickly formed by the dynamics itself when the energy is initially given to the mode of lowest frequency. While in the modified FPU model with alternating masses one was meeting with two "fixed" packets, here the packets are naturally formed by the dynamics itself (think of the first figure of the original FPU work), and their width is expected to depend on the initial energy. By the way, here too one meets with a resonance phenomenon, because the low frequencies are given in a first approximation by $\omega(k) \simeq k\pi/(N+1)$, which is just the familiar resonance relation of the continuous linear string. The idea of taking into account such a typical resonance of the low frequency modes, already indicated by Ford, was later reconsidered by Shepelyansky and by Ponno (see [7, 8]).

The first relevant result of the paper of Berchialla et al. is illustrated in Fig. 4.8 (which we familiarly refer to as "the shower"). In abscissas one has the time and in ordinates the energy. Here the results refer to a FPU system with $N = 15$, with the energy given initially to the lowest frequency mode. Having fixed the initial energy (and thus a line parallel to the axis of the abscissas), the various different symbols give an estimate of the time at which the various other modes start sharing energy with the first mode. The correspondence between the symbols and the mode numbers is not explicitly indicated in the figure, but in general it turns out that the times at which the various modes "enter the packet" are increasing with the mode number k. So one clearly sees that, for a sufficiently low energy, in a rather short time a packet of modes is formed which share the energy among themselves, then follows a rather large interval of time in which "nothing happens", until eventually the subsequent modes start entering the packet, and such an energy cascade is not interrupted until all modes did enter the packet. This is the time at which equipartition is attained.

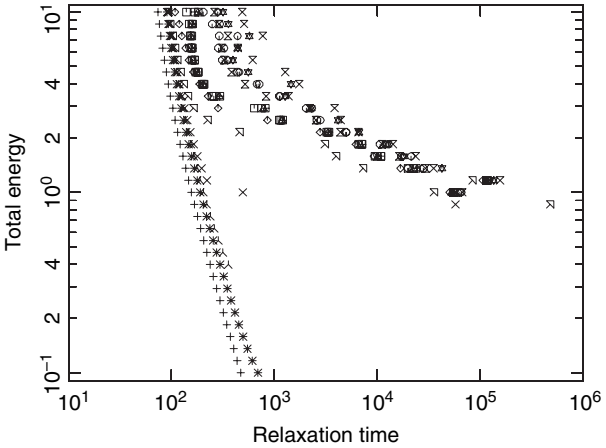

Fig. 4.8. The "shower". Here $N = 15$, and the energy was given initially just to the first mode. Having fixed an initial energy (and thus moving on a horizontal line), the various symbols give the times at which the other modes start sharing a (suitably defined) significant amount of the available energy. Such a time is found to be an increasing function of the mode number k, so that in the figure the mode number should be thought as increasing in going from left to right. The existence of two time-scales below a certain critical energy is clealy exhibited. Above the critical energy, instead, only one time-scale exists, which leads directly to equipartition. Taken from [14]

The relevant point is that such a separation of two time scales occurs only below a certain energy (namely, the one where the two inferior branches of the shower join); this just corresponds to the critical energy E_c of Bocchieri et al., because for higher energies the packet which is quickly formed covers all the available frequencies, i.e., there occurs a quick attainment of equipartition. Notice that the time needed for the quick formation of the low-frequency packet just below the critical energy is smaller than the time required for getting equipartition at larger energies.

Two more phenomena were also exhibited. The first one is that the width of the packet is a function of the specific energy, and is independent of N. The subtle point here is that such an independence with respect to N is exhibited if the width of the packet is plotted versus a quantity which is itself independent of N, and such a quantity is the frequency ω^* of the maximal mode included in the packet (or equivalently the corresponding value k^*/N). This is shown in Fig. 4.9, where the frequency ω^* defining the width of the packet (estimated in a suitable way) is plotted versus the specific energy ϵ. One very well sees that the data correspond to a curve $\omega^*(\epsilon) = c\,\epsilon^{1/4}$ (with a certain constant c), which by the way is just the law obtained later analytically. Notice that the data refer to N ranging from 8 to 1023. Notice also that the value of ϵ for which one has $\omega^* = 2$ (the maximal available frequency) provides a definition

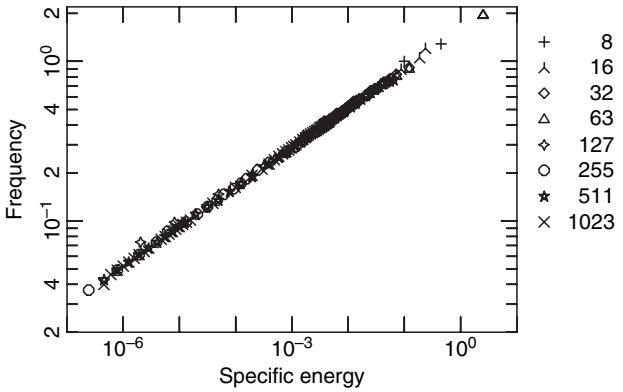

Fig. 4.9. Width of the "packet" (in frequency) versus specific energy, for N ranging from 8 to 1023. From [26] (adapted from [14])

for the critical specific energy in the sense of Bocchieri *et al.*, and that this quantity too is independent of N (i.e., pertains to the thermodynamic limit).

The second phenomenon concerns the way in which the time of formation of the packet depends on the initial conditions. One meets here with a problem that had been raised by Livi et al. (see [65]), who had pointed out that some relevant relaxation times were proportional to N if the energy was initially given to just one mode, whereas the behavior was quite different for other kinds of initial conditions. This fact is confirmed by Fig. 4.10, where the time of formation of the packet is plotted versus N for several kinds of initial conditions. Here one sees that such a time is proportional to N if the energy is given initially to the first mode. However, one also sees that the time is essentially independent of N if the energy is given initially to a packet of modes proportional to N, i.e., to a small packet extending to a maximal frequency Ω.

It should be mentioned that very interesting numerical informations on the formation of the packet had also been obtained by Biello et al. (see [66]), who were able to give quantitative estimates both of its width ω^* (namely, $\omega^* \simeq \epsilon^{1/4}$) and of its time of formation $t_{\rm f}$ (namely, $t_{\rm f} \simeq \epsilon^{-3/4}$).

We now briefly mention the results of three subsequent papers that are strictly related to the work of Berchialla et al., namely, the papers [67], [68] and [69]. In [67], the attention is addressed to the second time-scale $\tau_{\rm eq}$, namely the final time-scale to equipartition, which is shown to be of stretched exponential type, precisely, of the form $\tau_{\rm eq} \simeq \exp(\epsilon^{-1/4})$, at variance with the power law $\tau_{\rm eq} \simeq \epsilon^{-3}$ that had been suggested in [70]. Moreover, this result appears to be independent of N (for N large enough). This is clearly exhibited in Fig. 4.11, which reports the time of relaxation to the final global equipartition as a function of N, for two values of the specific energy ϵ. An analogous result was later obtained in [68], where the law $\tau_{\rm eq} \simeq \exp(\epsilon^{-1/5})$ was found for initial data with excitations of the high-frequency modes. By the way, in the latter paper

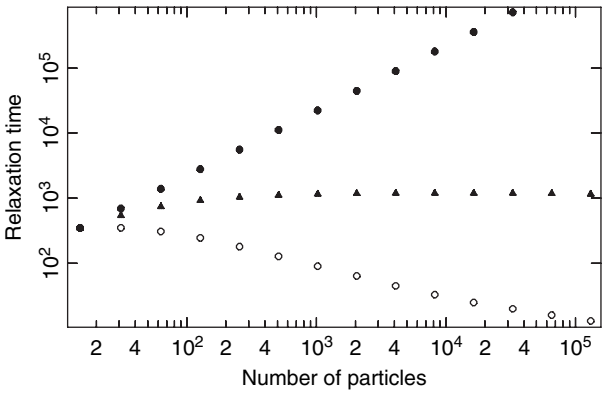

Fig. 4.10. Time of formation (relaxation time) of the packet versus N (with $N = 15, 31, 63, 127, \ldots, 32\,767$) for three types of initial conditions. *Dots*: all the energy initially on the lowest frequency mode. *Triangles* and *circles*: energy initially distributed in two different ways among the first $(N+1)/16$ modes (with zero energy to the higher frequencies). *Triangles*: energy linearly decreasing from the first mode to the last excited one. *Circles*: energy equally distributed among the initially excited modes. In all cases the specific energy is $\epsilon = 0.01$. From [14]

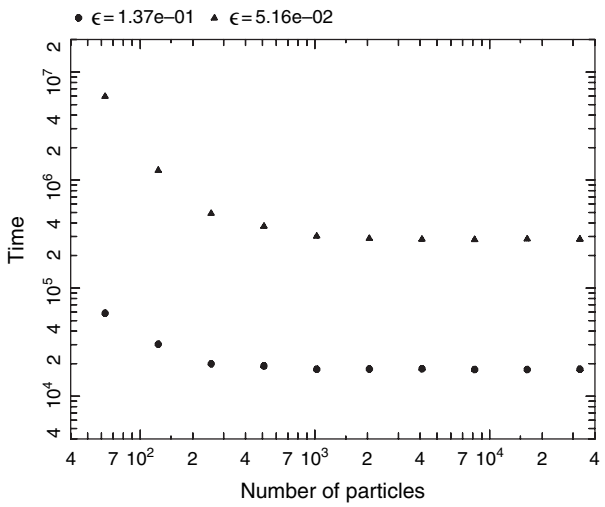

Fig. 4.11. Supporting the validity of the two-times description in the thermodynaimc limit: time of relaxation to the final global equipartition versus number N of particles, for two values of the specific energy ϵ. The relaxation times appear to tend to a definite limit (depending on ϵ) as N increases. *Circles*: $\epsilon = 0.137$, *triangles*: $\epsilon = 0.051$

an astute way was devised for exhibiting the analog of the shower when one deals with initial data of any type (and not just with excitations of the low-frequency modes). Finally, in the paper The figure showd that the relaxation time tends to a constant (depending on ϵ) as N increases, thus supporting te conjecture that the exponential law remains valid in the thermodynamic limit [69]. a vivid illustration (through projections of surfaces of section) is given of the way in which the final global equilibrium is attained. In fact, the system appears to be successively trapped into well different metastable regions, instead of finally merging, from some "ordered" region, to some "chaotic" one, as had been sometimes suggested.

4.6 New Analytical Contributions

We finally come to a brief illustration of some analytical developments that were obtained quite recently.

(a) *Solitons recovered*: The first relevant point is that soliton theory itself started to be reconsidered as a useful tool for analytical studies on the FPU problem. Indeed, in [71] it was shown that the form of the Fourier spectrum of the packet of the metastable state of the FPU system is explained in terms of KdV solitons. In particular, quantitative estimates both of its width (as $\epsilon^{1/4}$) and of its time of formation (as $\epsilon^{-3/4}$) were given. Soliton theory within the FPU problem had in fact been previously reconsidered in [72].

(b) *Shepelyansky and Ponno*: As first pointed out by Ford in his pioneering work (see [62]) of the year 1961, in order to explain the short-time dynamics of the modes (in particular, the quick formation of the packet, as we now say) one has to take into account the fact that the low-frequency modes are almost completely resonant. This idea was reconsidered by Shepelyansky (see [7]), who tried to deduce from it that the specific energy threshold tends to zero for $N \to \infty$, for initial long-wavelength excitations.[10] In [8], it was shown instead that the resonant normal form actually explains the formation of the metastable packet, and moreover that both its width and its time of formation are functions of the specific energy, exactly in the forms previously obtained (as $\epsilon^{1/4}$, and as $\epsilon^{-3/4}$, respectively) in [71] through soliton theory and in [66].

[10] In the Introduction of the paper, the result of Izrailev and Chirikov is mentioned: "According to Izrailev and Chirikov, in the case of low-mode excitation (nonlinear sound waves) the critical energy increases with the number of oscillators in the chain (or the energy per oscillator is constant)". It is then discussed how such authors had neglexted to take into account certain resonances in their semianalytical estimates, with the conclusion: "Such resonances not being considered by Izrailev and Chirikov give a sharp decrease of the chaos border in energy which goes to zero with the increase of the number of particles in the lattice. In this sense the long-wave chaos can exist for arbitrarily small nonlinearity".

(c) *Bambusi and Ponno, and the KdV equation as the resonant normal form for the FPU α-model*: In [17, 38, 73, 74] the attention was given to the resonant normal form of the FPU model for long-wavelength initial data. In [7] and [8], such a normal form had been expressed in terms of the mode coordinates, whereas in the new papers it was pointed out that, if such a normal form is read in terms of the particle coordinates in the continuum interpolation and in the thermodynamic limit, then the normal form is nothing but the KdV equation itself (actually, a pair of such equations, in agreement with the time reversal symmetry of the FPU system). In such a way, the privileged role of the KdV equation for the FPU system with long-wave initial data was recovered, with moreover an understanding of the time-scale of its validity. This happened after a previous understanding, by Bambusi et al., that the nonlinear Schroedinger (NLS) equation plays an analogous role of normal form in the FPU problem with short-wavelength initial data (see [75]).

(d) *Perturbation theory in the thermodynamic limit*: The analytical results of Bambusi and Ponno in [17] could be obtained only for an extremely special class of initial conditions, in which exactly one low-frequency mode was excited, at an energy E proportional to N, and so at a given specific energy $\epsilon = E/N$. Only in this sense does the result hold in the thermodynamic limit. It may be conjectured that such a limitation is only a technical one, to be hopefully removed in the future.

This fact rises the general problem of whether it is possible to extend the methods of classical perturbation theory of nearly integrable Hamiltonian systems to the thermodynamic limit ($N \to \infty$ with a nonvanishing specific energy ϵ). The presently available techniques do not allow it, as they apply only to finite N (or to any N, but with a bounded energy E, i.e., with a vanishing specific energy $\epsilon = E/N$ in the limit $N \to \infty$; see [19]). It was proved quite recently by one of the present authors (see [18]) that a rather simple modification of the known techniques actually allows one to do so. This is obtained at the cost of weakening the results, by renouncing to control all the orbits in phase space (a control which usually is obtained by making use, in the estimates, of the sup norm), and looking instead for results holding only in the mean. This is analogous to the way in which the von Neumann ergodic theorem can be considered as a weaker version of the Birkhoff ergodic theorem, although it still keeps all the relevant physical significance of the result (as particularly pointed out in [25]).

4.7 Conclusions

In the present review, we have illustrated the relevance of a metastability scenario for the interpretation of a large part of the results on the one-dimensional FPU model, in the thermodynamic limit. Such a scenario involves two well separated time-scales for the approach to equilibrium, below a critical specific energy.

It was also mentioned that too little information is presently available for the case of dimension two and especially for the physically significant case of dimension three. Two "simple " possibilities can be conceived. The first one is that the metastability scenario will be proved to be incorrect in the "physical case" of dimension three, in the sense that at any finite specific energy (or temperature) the time-averages of the relevant quantities present a quick relaxation to their equilibrium values. In such a case the "FPU paradox" will turn out to have been removed completely in the thermodynamic limit. This would provide a proof of the conjecture advanced long ago by Izrailev and Chirikov, at least in the way many people understand it, namely as claiming that no "FPU physical phenomenon" essentially exists.

The second "simple" possibility is that the metastability scenario as described above (with two well separated time-scales) will be proved to be correct. In such a case, in a sense the FPU paradox will still turn out to have been removed, because at any temperature the equilibrium state is finally attained. But some paradox will still remain. Indeed it will turn out that, below a certain critical specific energy (i.e., below a certain critical temperature), the FPU model predicts the existence of some metastable state which, for quite long times, may be practically indistinguishable from a true equilibrium state, although providing a statistics quite different from the standard equilibrium one (in this connection see [76, 77, 78]). So one would remain with the problem of ascertaining whether such a physical prediction is in agreement with the observations or not. We are particularly thinking of possible metastability phenomena in the measurements of specific heats at low temperatures, in the spirit of the rationale of the time-dependent specific heats (see [61]).

Naturally, other more complicated scenarios can be conceived. For example, there could exist a "cascade" of growing-time scales of different orders of magnitude as $N \to \infty$, and this, in a larger scale, could look like a continuous growth (we thank a referee for kindly pointing this out to us). To what physical phenomena would such a situation possibly correspond, is not clear to us.

References

1. E. Fermi, J. Pasta and S. Ulam, Studies of nonlinear problems, in E. Fermi: Note e Memorie (Collected Papers), Vol. II, No. 266. 977–988 (Accademia Nazionale dei Lincei, Roma, and The University of Chicago Press, Chicago 1965).
2. D.K. Campbell, P. Rosenau and G. Zaslavsky, Introduction: the Fermi-Pasta-Ulam problem—the first fifty years, Chaos **15**, 015101 (2005), and the papers following it.
3. T.P. Weissert, The genesis of simulation in dynamics; pursuing the Fermi-Pasta-Ulam problem. Springer, New York, 1997.
4. N.J. Zabusky and M.D. Kruskal, Interaction of solitons in a collisionless plasma and the recurrence of initial states, Phys. Rev. Lett. **15**, 240–243 (1965).

5. F.M. Izrailev and B.V. Chirikov, Statistical properties of a nonlinear string, Sov. Phys. Dokl. **11**, 30–34 (1966).
6. F.M. Izrailev, A.I. Khisamutdinov and B.V. Chirikov, Numerical experiments with a chain of coupled anharmonic oscillators, Report 252, Institute of Nuclear Physics, Novosibirsk, URSS, 1968 (English translation: LA 4440 TR, Los Alamos. 1970).
7. D.I. Shepelyansky, Low-energy chaos in the Fermi–Pasta–Ulam problem, Nonlinearity **10**, 1331–1338 (1997).
8. A. Ponno, The FPU problem in the thermodynamic limit: scaling laws of the energy cascade, in P. Collet et al. eds, Chaotic dynamics and transport in classical and quantum systems. Kluwer, Dordrecht, 2005, pp. 431–440.
9. P. Bocchieri, A. Scotti, B. Bearzi and A. Loinger, Anharmonic chain with Lennard-Jones interaction, Phys. Rev. A **2**, 2013–2019 (1970).
10. L. Galgani and A. Scotti, Planck-like distribution in classical nonlinear mechanics, Phys. Rev. Lett. **28**, 1173–1176 (1972).
11. C. Cercignani, L. Galgani and A. Scotti, Zero-point energy in classical nonlinear mechanics, Phys. Lett. A **38**, 403 (1972); C. Cercignani, On a nonquantum derivation of Planck's distribution law, Found. Phys. Lett. **11**, 189–199 (1998).
12. L. Galgani and A. Scotti, Recent progress in classical nonlinear dynamics, Rivista Nuovo Cim. **2**, 189–209 (1972).
13. 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. Phys., **43**, 707–713 (1982).
14. L. Berchialla, L. Galgani and A. Giorgilli, Localization of energy in FPU chains, Discr. Cont. Dyn. Syst. A **11**, 855–866 (2004).
15. A. Carati and L. Galgani, On the specific heat of the Fermi–Pasta–Ulam systems and their glassy behavior, J. Stat. Phys. **94** 859–869 (1999).
16. A. Carati, P. Cipriani and L. Galgani, On the definition of temperature in FPU systems, J. Stat. Phys. **115**, 1119–1130 (2004).
17. D. Bambusi and A. Ponno, On metastability in FPU, Comm. Math. Phys., 264(2), 539–561 (2006).
18. A. Carati, An averaging theorem for Hamiltonian dynamical systems in the thermodynamic limit, preprint.
19. D. Bambusi and A. Giorgilli, Exponential stability of states close to resonance in infinite-dimensional Hamiltonian systems, J. Stat. Phys. **71**, 569–606 (1993).
20. G. Benettin, Time scales for energy euipartition in a two-dimensional FPU model, Chaos **15**, 015108 (2004).
21. G. Marcelli and A. Tenenbaum, Quantumlike short-time behavior of a classical crystal, Phys. Rev. E **68**, 041112 (2003).
22. H. Poincaré, Méthodes nouvelles de la mécanique céleste, Vol. I, Chap. 5.
23. E. Fermi, Beweiss das ein mechanisches Normalsystem im allgemeinen quasi-ergodisch ist, Phys. Zeits. **24**, 261–265 (1923); Über die Existenz quasi-ergodischer Systeme, Phys. Zeits. **25**, 166–167 (1924), in Note e Memorie (Collected Papers), Vol. I, No. 11: 79–87 Accademia Nazionale dei Lincei, Roma, and The University of Chicago Press, Chicago, 1965.
24. G. Benettin, G. Ferrari, L. Galgani and A. Giorgilli, An extension of the Poincaré-Fermi theorem on the nonexistence of invariant manifolds in nearly integrable Hamiltonian systems, Nuovo Cim. B **72**, 137–148 (1982); G. Benettin, L. Galgani and A. Giorgilli, Poincaré's non-existence theorem and classical

perturbation theory in Nearly-Integrable Hamiltonian Systems, in Dynamics and Stochastic Processes, R. Livi and A. Politi, eds, Advances in Nonlinear World Scientific, Singapore, 1985.
25. J. von Neumann, Physical approach to the ergodic hypotesis, N.A.S. Proc. **18**, 263–266 (1932), also in J. von Neumann, Collected works. H.A. Taub, ed., Pergamon Press, Oxford, 1961. Vol. II, No. 13, pp. 274–277.
26. A. Carati, L. Galgani and A. Giorgilli, Dynamical systems and thermodynamics, in Encyclopedia of Mathematical Physics, Elsevier, Oxford (2006).
27. D. Poggi, S. Ruffo and H. Kantz, Shock waves and time scales to reach equipartition in the Fermi-Pasta-Ulam model, Phys. Rev. E **52**, 307–315 (1995).
28. P. Lorenzoni and S. Paleari, Metastability and dispersive shock waves in the FPU system, Physica D 221 (2006), 110–117.
29. A.N. Kolmogorov, Preservation of conditionally periodic movements with small change in the Hamilton function, Dokl. Akad. Nauk **98**, 527 (1954); English translation in G. Casati and G. Ford, eds, Lecture Notes in Physics No. 93, Springer Verlag, Berlin, 1979. See also G. Benettin, L. Galgani, A. Giorgilli and J.-M. Strelcyn, A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method, Nuovo Cim. B **79**, 201–223 (1984).
30. A. Carati and L. Galgani, Analog of Planck's formula and effective temperature in classical statistical mechanics far from equilibrium, Phys. Rev. E **61**, 4791–4794 (2000); A. Carati and L. Galgani, Einstein's nonconventional conception of the photon, and the modern theory of dynamical systems, in J. Bricmont et al., eds, Chance in Physics, Lecture Notes in Physics No. 574, Springer, Berlin, 2001; A. Carati and Galgani, Planck's formula and glassy behavior in classical nonequilibrium statistical mechanics, Physica A **280**, 105–114 (2000); L. Galgani, Relaxation times and the foundations of classical statistical mechanics in the light of modern perturbation theory, in G. Gallavotti and P.F. Zweifel, eds, non-linear evolution and chaotic phenomena, NATO ASI Series R71B: Vol. 176, Plenum Press, New York, 1988.
31. A. Carati and L. Galgani, The theory of dynamical systems and the relations between classical and quantum mechanics, Found. Phys. **31**, 69–87 (2001).
32. U. Frisch nad R. Morf, Intermittency in nonlinear dynamics and singularities at comlex times, Phys. Rev. A **23**, 2673–2705 (1981).
33. G.K. Batchelor, The theory of homogeneous turbulence, Cambridge U.P., Cambridge, 1960.
34. J. von Neumann, Recent theories of turbulence, in J. von Neumann, Collected works. H.A. Taub, ed., Pergamon Press, Oxford, 1961. Vol. VI, No. 32, pp. 437–472.
35. P. Butera, L. Galgani, A. Giorgilli, A. Tagliani and H. Sabata, Stochasticity thresholds in a lattice field theory, Nuovo Cim. B **59**, 81–86 (1980).
36. B. Bassetti, P. Butera, M. Raciti and M. Sparpaglione, Complex poles, spatial intermittencies, and energy transfer in a classical nonlinear string, Phys. Rev. A **30**, 1033–1039 (1984).
37. A. Carati, L. Galgani and A. Giorgilli, The Fermi-Pasta-Ulam problem as a challenge for the foundations of physics, Chaos **15**, 015105 (2004).
38. A. Ponno and D. Bambusi, Korteweg-de Vries equation and energy sharing in Fermi-Pasta-Ulam, Chaos **15**, 015107 (2004).
39. G. Parisi, On the approach to equilibrium of a Hamiltonian chain of anharmonic oscillators, Europhys. Lett. **40**, 357 (1997).

40. F. Fucito, F.Marchesoni, M. Sparpaglione and A. Vulpiani, Intermittent behaviour in non-linear-Hamiltonian systems far from equilibrium, J. Phys. A **16**, 117–124 (1983).
41. A. Carati, L. Galgani and B. Pozzi, Levy flights in the Landau-Teller model of molecular collisions, Phys. Rev. Lett. **90**, 010601 (2003).
42. R. Livi, M. Pettini, S. Ruffo, M. Sparpaglione and A. Vulpiani, Relaxation to different stationary states in the FPU model, Phys. Rev. A **28**, 3544–3552 (1983).
43. R. Livi, S. Ruffo, M. Pettini and A. Vulpiani, Short-time asymptotics in classical nonlinear wave equations, Nuovo Cim. B **89**, 120–130 (1985).
44. R. Livi, M. Pettini, S. Ruffo, M. Sparpaglione and A. Vulpiani, Equipartition threshold in nonlinear large Hamiltonian systems: the Fermi–Pasta–Ulam model, Phys. Rev. A **31**, 1039–1045 (1985).
45. R. Livi, M. Pettini, S. Ruffo, and A. Vulpiani, Further results on the equipartition threshold in large nonlinear Hamiltonian systems, Phys. Rev. A **31**, 2740–2742 (1985).
46. G. Benettin, L. Galgani and A. Giorgilli, Boltzmann's ultraviolet cutoff and Nekhoroshev's theorem on Arnold diffusion, Nature **311**, 444–445 (1984).
47. A. Carati, L. Galgani and B. Pozzi, The problem of the rate of thermalization and the relations between classical and quantum mechanics, in M. Fabrizio et al., eds, Mathematical models and methods for smart materials, Series of Advances in Mathematics no. 62, World Scientific, Singapore, 2002.
48. N.N. Nekhoroshev, Russ. Math. Surv. **32**, 1 (1977); in O.A. Oleinik, ed. Topics in modern mathematics: Petrovskii Sem. No. 5, Consultant Bureau, New York, 1985. See also G. Benettin, L. Galgani and A. Giorgilli, A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems, Celestial Mech. **37**, 1–25 (1985); A. Giorgilli, On the problem of stability for near to integrable Hamiltonian systems, Proceedings of the International Congress of Mathematicians Berlin 1998, Vol. III, Documenta Mathematica, extra volume ICM 1998, 143–152 (1998); G. Benettin, The elements of Hamiltonian perturbation theory, in D. Benest, C. Froeschle' and E. Lega eds, Hamiltonian systems and Fourier analysis, Cambridge Scientific Publisher, Cambridge (UK), 2005.
49. L.D. Landau and E. Teller, On the theory of sound dispersion, Phys. Z. Sowjet. **10**, 34–41 (1936), also in Collected Papers of L.D. Landau, ter Haar, ed., Pergamon Press, Oxford 1965, pp. 147–153; G. Benettin, A. Carati and P. Sempio, On the Landau-Teller approximation for the energy exchanges with fast degrees of freedom, J. Stat. Phys. **73**, 175 (1993); G. Benettin, A. Carati and G. Gallavotti, A rigorous implementation of the Landau-Teller approximation for adiabatic invariants, Nonlinearity **10**, 479–505 (1997); G. Benettin, P. Hjorth and P. Sempio, Exponentially long equilibrium times in a one-dimensional collisional model of a classical gas, J. Stat. Phys. **94**, 871–892 (1999); A. Carati, L. Galgani and B. Pozzi, The problem of the rate of thermalization, and the relations between classical and quantum mechanics, in M. Fabrizio, B Lazzari, A. Morro, eds, Mathematical models and methods for smart materials, World Scientific, Singapore, 2002.
50. G. Benettin, L. Galgani and A. Giorgilli, Exponential law for the equipartition times among translational and vibrational degrees of freedom, Phys. Lett. A **120**, 23–27 (1987).

51. O. Baldan and G. Benettin, Classical 'freezing' of fast rotations. A numerical test of the Boltzmann–Jeans conjecture, J. Stat. Phys. **62**, 201 (1991).
52. M. Pettini and M. Landolfi, Relaxation properties and ergodicity breaking in nonlinear Hamiltonian dynamics, Phys. Rev. A **41**, 768–783 (1990).
53. M. Pettini, L. Casetti, M. Cerruti-Sola, R. Franzosi and E.G.D. Cohen, Weak and strong chaos in Fermi-Pasta-Ulam models and beyond, Chaos **15**, 015106 (2004).
54. L. Casetti, R. Livi and M. Pettini, Gaussian model for chaotic instability of Hamiltonian flows, Phys. Rev. Lett. **74**, 375–378 (1995); L. Casetti, C. Clementi and Pettini, Riemannian theory of Hamiltonian chaos and Lyapunov exponents, Phys. Rev. E **54**, 5969–5984 (1996); L. Casetti, M. Pettini and E.G.D. Cohen, Geometric approach to Hamiltonian dynamics and statistical mechanics, Phys. Rep. **337**, 237–341 (2000).
55. P. Butera and G. Caravati, Phase transitions and Lyapunov characteristic exponents, Phys. Rev. A **36**, 962–964 (1987).
56. M. Casartelli, E. Diana, L. Galgani and A. Scotti, Numerical computations on a stochastic parameter related to the Kolmogorov entropy, Phys. Rev. A **13**, 1921–1925 (1976).
57. G. Benettin, L. Galgani and A. Giorgilli, Classical Perturbation Theory for systems of weakly coupled rotators, Nuovo Cim. **89** B, 89–102 (1985).
58. G. Benettin, L. Galgani and A. Giorgilli, Numerical investigations on a chain of weakly coupled rotators in the light of classical perturbation theory, Nuovo Cim. **89** B, 103–119 (1985).
59. R. Livi, M. Pettini, S. Ruffo and A. Vulpiani, Chaotic behavior in nonlinear Hamiltonian systems and equilibrium statistical mechanics, J. Stat. Phys. **48**, 539–559 (1987).
60. A. Perronace and A. Tenenbaum, Classical specific heat of an atomic lattice at low temperature, revisited, Phys. Rev. E **57**, 100–107 (1998); Erratum, Phys. Rev. E **57**, 6215 (1998).
61. N.O. Birge and S.R. Nagel, Specific-heat spectroscopy of the glass transition, Phys. Rev. Lett. **54**, 2674 (1985); N.O. Birge, Specific-heat spectroscopy of glycerol and propylene near the glass transition, Phys. Rev. **34**, 1631–1642 (1986).
62. J. Ford, Equipartition of energy for nonlinear systems, J. Math. Phys. **2**, 387–393 (1961).
63. L. Galgani, A. Giorgilli, A. Martinoli and S. Vanzini, On the problem of energy equipartition for large systems of the Fermi–Pasta–Ulam type: analytical and numerical estimates, Physica D **59**, 334–348 (1992).
64. G. Benettin, L. Galgani and A. Giorgilli, Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory, II, Comm. Math. Phys. **121**, 557–601 (1989); G. Benettin, J. Fröhlich and A. Giorgilli, A Nekhoroshev-type theorem for Hamiltonian systems with infinitely many degrees of freedom, Comm. Math. Phys, **119**, 95–108 (1988).
65. H. Kantz, Vanishing stability thresholds in the thermodynamic limit of nonintegrable conservative systems, Physica D **39**, 322 (1989); H. Kantz, R. Livi and S. Ruffo, Equipartition thresholds in chains of anharmonic oscillators, J. Stat. Phys. **76**, 627 (1994).
66. J.A. Biello, P.R. Kramers, and Y. Lvov, Stages of energy transfer in the FPU model, Discr. Cont. Dyn. Syst. B (Suppl.) 113–122 (2003).

67. L. Berchialla, A. Giorgilli and S. Paleari, Exponentially long times to equipartition in the thermodynamic limit, Phys. Lett. A **321**, 167–172 (2004).
68. S. Paleari and T. Penati, Equipartition times in a FPU system, Discr. Cont. Dynam. Syst. (Suppl. Volume) 1–10 (2005).
69. A. Giorgilli, S. Paleari and T. Penati, Local chaotic behavior of the FPU system, Discr. Cont. Dynam. Syst. B **5**, 1–14 (2005).
70. J. De Luca, A. Lichtenberg and S. Ruffo, Finite times to equipartition in the thermodynamic limit, Phys. Rev. E **60**, 3781–3786 (1999).
71. A. Ponno, Soliton theory and the Fermi-Pasta-Ulam problem in the thermodynamic limit, Europh. Lett. **64**, 606–612 (2003).
72. A. Ponno, L. Galgani and F. Guerra, Analytical estimate of stochasticity thresholds in Fermi-Pasta-Ulam and phi-4 models, Phys. Rev. E **61**, 7081–7086 (2000).
73. A. Ponno and D. Bambusi, Energy cascade in FPU models, in G. Gaeta, et al. eds, Symmetry and perturbation theory 2004, World Scientific, Singapore, 2005, pp. 263–270.
74. D. Bambusi and A. Ponno, Resonance, metastability and blow-up in FPU, Chap. 6 in this volume.
75. D. Bambusi, A. Carati and A. Ponno, The nonlinear Schroedinger equation as a resonant normal form, Discr. Cont. Dyn. Sys. B **2**, 109–128 (2002).
76. E. Olivieri and M.E. Vares, Large deviations and metastability, Cambridge U.P., Cambridge, 2005.
77. H. Larralde and F. Leyvraz, Metastability for Markov processes with detailed balance, Phys. Rev. Lett. **94**, 160201 (2005).
78. A. Carati, Thermodynamics and time-averages, Physica A **348**, 110–120 (2005).