The Fermi-Pasta-Ulam problem in the thermodynamic limit: Scaling laws of the energy cascade

Antonio Ponno
Univerisità degli Studi di Milano
Dipartimento di Matematica "F. Enriques"
Via Saldini 50, 20133 Milano, Italy

November 1, 2004

Abstract

In the present contribution we justify and discuss the scaling laws characterizing the first phase of the energy transfer from large to small spatial scales in a chain of nonlinear oscillators (the so-called Fermi-Pasta-Ulam α-model). By means of qualitative estimates, we show that large scale initial excitations (long wavelength Fourier modes) produce injection of energy into smaller scales on times $t > \tau_c \sim \varepsilon^{-3/4}$ and up to a cutoff spatial scale $\ell_c \sim \varepsilon^{-1/4}$, where ε is the energy per degree of freedom of the system.

Introduction

In the present work we will discuss some aspects characterizing the dynamics of the oscillator chain defined by the Hamiltonian function

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

where $q = (q_0, q_1, \ldots, q_{N+1})$ and $p = (p_0, p_1, \ldots, p_{N+1})$ are the two sets of canonically conjugate coordinates, while $\alpha > 0$ is the nonlinear coupling

constant. We will consider a chain with fixed ends, namely $q_0 = q_{N+1} = 0$ and $p_0 = p_{N+1} = 0$; as a consequence, the Hamiltonian system (1) has N degrees of freedom. One can regard the system as modeling a nonlinear string as well as a one dimensional crystal with a nonlinear interaction between nearby particles.

Such a model was introduced in 1954 by Fermi, Pasta and Ulam (FPU) [1], with the explicit purpose of measuring (numerically) the time rate of approach to equilibrium of the system, starting with energy initially given to the fundamental (longest wavelength) Fourier mode. In agreement with the prescriptions of classical statistical mechanics, FPU used as an indicator of equilibrium the (approximate) equipartition of energy among the Fourier modes of the system. What they got was essentially a negative result: energy was observed to be shared by a small fraction of modes only, within the times accessible to numerical computation.

Since then, the explanation of such a lack of complete equipartition observed in systems of the kind (1), known as the FPU paradox, has been the object of thorough investigations. We skip here most of the story (see for example [2, 3, 4] and references therein), and come directly to what we consider to be the most significant problem in the field, namely that of understanding whether the FPU paradox survives the thermodynamic limit: $N \to \infty$, $H = E \to \infty$, the specific energy $\varepsilon = E/N$ being constant. Were the answer negative, the phenomenology described by FPU would not be relevant to classical statistical mechanics.

The way to an affermative answer to the above fundamental question was paved by the numerical work of Bocchieri et al. in 1970 [5]. Since then, other results have been confirming the thesis that the FPU paradox might be of actual relevance to statistical mechanics (see e.g. [4] and references therein).

In particular, we will refer here to some recent numerical findings [6, 7] which show that large scale (long wavelength) initial excitations in system (1) inject energy to small scales (short wavelengths) up to a cutoff spatial scale $\ell_c \sim \alpha^{-1/2}\varepsilon^{-1/4}$, and that a state of partial equipartition involving modes of wavelength $\lambda > \ell_c$ sets in on a time scale larger than $\tau_c \sim \ell_c^3 \sim \alpha^{-3/2}\varepsilon^{-3/4}$. This means that, if the specific energy of the system is low enough, the FPU paradox persists in the thermodynamic limit. An analytical explanation of these results based on soliton theory and inspired to the work of Zabusky and Kruskal [8] can be found in [9].

The aim of the present work is to give an elementary justification of the above fundamental scaling laws by using arguments of canonical perturbation

theory and statistics. We start by exploiting an idea of Shepelyansky [10] which consists in taking into account the fact that long wavelength Fourier modes are nearly resonant in order to obtain a suitable averaged system. Anyway, we do not follow the author in using the criterion of resonance overlapping [11, 12] (the reason is briefly explained in *Remark 1* in the last section of the paper). We proceed instead by analogy with turbulence theory [13], as follows. The flow of energy from long wavelength modes (the integral scale) to shorter wavelength modes is determined by the nonlinearity, responsible for mode-coupling. Thus, the dynamics of the modes involved in the cascade process and sharing their energy has to be dominated by the nonlinear term appearing in the equations of motion (the interval of spatial scales corresponding to such strongly coupled modes is the analog of the inertial range in turbulence). For such a reason, by assuming that at a given time the set of interacting modes is in a state of energy equipartition, we show that the energy cascade is interrupted at scales corresponding to modes for which the linear term appearing in the equations of motion, responsible for mode-dispersion, is of the same order of magnitude of the nonlinear one. Here dispersion plays a role analogous to that of dissipation in turbulence (though, of course, energy is conserved in Hamiltonian systems). The time-scale characterizing the dynamics of the interacting modes is also estimated, showing that high modes (close to the border of the inertial range) reach the equipartition level more quickly than the low ones (close to the integral scale). We thus produce estimates of all the relevant quantities.

1 Model and setting

For the present purposes, it is convenient to rewrite the Hamiltonian (1) in the usual complex canonical coordinates z, z^* (the star denotes complex conjugation). The canonical change of variables $(q, p) \mapsto (z, z^*)$ is given by the formulæ

$$\begin{aligned} q_n &= \sqrt{\frac{2}{N+1}} \sum_{k=1}^{N} \left[\frac{z_k - z_k^*}{i\sqrt{2\omega_k}} \right] \sin\left(\frac{\pi k n}{N+1} \right) \\ p_n &= \sqrt{\frac{2}{N+1}} \sum_{k=1}^{N} \left[\sqrt{\frac{\omega_k}{2}} \, (z_k + z_k^*) \right] \sin\left(\frac{\pi k n}{N+1} \right) \, , \end{aligned} \tag{2}$$

from which one sees that the k-th Fourier (harmonic) component of q_n and p_n is characterized by a wavelength $\lambda_k = 2(N+1)/k$. It has to be kept in mind that if one regards z_k as a coordinate ($k = 1, \ldots, N$), its conjugate canonical momentum turns out to be $-iz_k^*$ ($i = \sqrt{-1}$). In terms of the z-variables the Hamiltonian of the system becomes

$$H(z, z^*) = \sum_{k=1}^{N} \omega_k |z_k|^2 + H_3(z, z^*) \ , \tag{3}$$

where the dispersion relation (frequency spectrum) of the linearized system

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

has been introduced, while the cubic part of the Hamiltonian reads

$$H_3(z, z^*) = \frac{i\alpha}{12\sqrt{N+1}} \sum_{k_1,k_2,k_3=1}^{N} S_{k_1 k_2 k_3} \prod_{j=1}^{3} \sqrt{\omega_{k_j}} (z_{k_j} - z_{k_j}^*) \ . \tag{5}$$

The coefficient

$$S_{k_1 k_2 k_3} \equiv \delta_{k_1+k_2-k_3,0} + \delta_{k_1-k_2+k_3,0} + \delta_{k_1-k_2-k_3,0} - \delta_{k_1+k_2+k_3,2N+2} \tag{6}$$

appearing in (5) defines the exact structure of mode-coupling due to the nonlinearity.

The z_k variables determine the amplitude of the k-th Fourier mode of the system. In the linear approximation ($H_3 = 0$), the modes of the system are uncoupled and their energies

$$E_k = \omega_k |z_k|^2 \quad , \quad k = 1, \ldots, N \tag{7}$$

are constants of motion. Instead, for the full (nonlinear) system a variation in time of the modal energies E_k due to nonlinear mode-coupling is expected, with a consequent transfer of energy to modes which were initially switched off. In the next section we try to quantify the effectiveness of such a transfer in the case of long wavelength initial excitations.

2 Nearly resonant averaging

As pointed out first by Ford [14], energy sharing in the FPU problem is mainly due to the almost resonant character of low modes ($k \ll N$). Indeed,

by expanding (4) in powers of $k/(N+1)$ one gets $\omega_k \simeq \pi k/(N+1)$, which implies $\omega_k \simeq k\omega_1$ up to an error increasing with k. The first to take into account this fact in a smart way (at least within the context of canonical perturbation theory) was Shepelyansky [10], who performed an averaging of Hamiltonian (3) over the unperturbed flow of low modes with frequencies almost linearly dependent on k. Quasi-resonance of low modes can be taken into account by performing the time dependent canonical change of variables $(z, z^*) \mapsto (\zeta, \zeta^*)$ given by

$$z_k = \exp\left(i\frac{\pi k}{N+1}\ t\right)\zeta_k \quad , \quad z_k^* = \exp\left(-i\frac{\pi k}{N+1}\ t\right)\zeta_k^* \ . \tag{8}$$

This is easily seen to transform the Hamiltonian (3) into

$$K(\zeta,\zeta^*) = \sum_{k=1}^{N} \Omega_k |\zeta|^2 + \overline{H}_3(\zeta,\zeta^*) + \widehat{H}_3(\zeta,\zeta^*,t) \tag{9}$$

where

$$\Omega_k = \omega_k - \frac{\pi k}{N+1} \simeq -\frac{1}{24}\left(\frac{\pi k}{N+1}\right)^3 \ , \tag{10}$$

$$\overline{H}_3 = \sum_{1\le k_1+k_2\le N} \frac{i\sqrt{\omega_{k_1}\omega_{k_2}\omega_{k_1+k_2}}}{4\sqrt{N+1}}\ (\zeta_{k_1}^*\zeta_{k_2}^*\zeta_{k_1+k_2} - \zeta_{k_1}\zeta_{k_2}\zeta_{k_1+k_2}^*) \ , \tag{11}$$

while

$$\widehat{H}_3(\zeta,\zeta^*,t) \equiv H_3 - \overline{H}_3 \quad .$$

This is the part of the coupling Hamiltonian that explicitly depends on time and has the property that its time average, with all the ζ_k's kept constant, is zero. At this stage, one simply neglects the contribution of the oscillating part $\widehat{H}_3$, and conjectures that the dynamics of the system is correctly described by the averaged Hamiltonian

$$\overline{K} = \sum_{k=1}^{N} \Omega_k |\zeta_k|^2 + \overline{H}_3(\zeta,\zeta^*) \ . \tag{12}$$

By construction, the quantity

$$J = \sum_{k=1}^{N} \left(\frac{\pi k}{N+1}\right) |\zeta_k|^2 \tag{13}$$

is a constant of motion for the averaged system (12). Moreover, if energy is localized mainly over large spatial scales (low k's), it represents approximately the sum of the harmonic energies of the excited modes and its value is almost equal to that of the total energy E of the original system: $J \simeq \sum_k E_k \approx E \equiv N\varepsilon$, where ε denotes the specific energy of the system.

3 Estimates

Now we come to the main point. The equations of motion of system (12) read

$$\dot{\zeta}_k = i\Omega_k \zeta_k + i\frac{\partial \overline{H}_3}{\partial \zeta_k^*} \quad . \tag{14}$$

The first term on the right hand side represents the effect of dispersion which, alone, would cause the mode variable ζ_k to oscillate at the frequency Ω_k defined in (10). The second term represents instead the effect of the nonlinearity and couples the motion of the k-th mode to that of the other ones. Were dispersion (first term) dominant, then the mode would oscillate almost freely and would not exchange energy with the other ones; in particular, were such a mode initially switched off, it would remain so afterwards too. In the opposite case, namely if the nonlinearity (second term) prevails, the mode is strongly coupled to other modes and shares its energy with them. Thus, what we have to estimate, for a given k, is the ratio (in a sense to be specified below) of $\Omega_k z_k$ to $\partial \overline{H}_3/\partial z_k^*$. In order to do this, we suppose the system to be in a "typical" configuration, characterized by partial equipartition between the first k_c modes, all of them having approximately the same value of modal energy E_c, while the other ones (for $k > k_c$) are essentially at rest. Further, we suppose k_c/N small, so that $\omega_k \simeq \pi k/(N+1)$ for all $k \leq k_c$. Such a configuration is given by

$$\begin{aligned} \zeta_k &= \sqrt{\frac{E_c}{\omega_k}}\, e^{i\phi_k} \quad , \quad k = 1, \ldots, k_c \\ \zeta_k &= 0 \quad , \quad k = k_c + 1, \ldots, N \ , \end{aligned} \tag{15}$$

where the phases $\phi_1, \ldots, \phi_{k_c}$ are considered as independent random variables, each of them being uniformly distributed over $[0, 2\pi]$. We come back to the hypothesis (15) in *Remark 2* in the last section. For the moment one can

accept it as the simplest possible one. Now, by inserting (15) into (13) we get a first relation linking E_c and k_c, namely

$$k_c E_c = J \approx N\varepsilon \ . \tag{16}$$

As an estimate for the ratio of the dispersive to the nonlinear term appearing in the right hand side of equation (14) we use the ratio

$$R(k) = \frac{\sqrt{\left\langle \left|\Omega_k \zeta_k\right|^2 \right\rangle_\phi}}{\sqrt{\left\langle \left|\frac{\partial \overline{H}_3}{\partial \zeta_k^*}\right|^2 \right\rangle_\phi}} \ , \tag{17}$$

where both the numerator and the denominator are calculated on the configuration (15), $\langle\ \rangle_\phi$ denoting average over the random phase distribution. Notice that the simple phase average of both $\Omega_k \zeta_k$ and $\partial \overline{H}_3 / \partial \zeta_k^*$ is zero. Of course, the function R is defined for values of k smaller than k_c. An easy computation yields

$$R(k) = \frac{\pi^2}{6} \frac{\left(\frac{k}{N+1}\right)^2}{\alpha \sqrt{E_h} \sqrt{4\frac{k_c}{N+1} - 3\frac{k}{N+1}}} \ , \tag{18}$$

showing that $R(k)$ is in fact a monotonically increasing function of $k/(N+1)$. Thus, the higher $k/(N+1)$ is, the greater is the effect of dispersion with respect to that of nonlinearity for the corresponding mode k. Our criterion to determine k_c consists in requiring $R(k_c) = 1$, which, by exploiting (13), for large N, yields

$$\frac{k_c}{N+1} = \frac{\sqrt{6}}{\pi} \sqrt{\alpha}\ \varepsilon^{1/4} \ . \tag{19}$$

The meaning of the relation $R(k_c) = 1$ is almost obvious: it means that k_c is the first mode for which dispersion compensates nonlinearity, so that k_c represents the highest mode of the inertial (nonlinear) scale. Remembering that the wavelength of the k-th mode is $\lambda_k = 2(N+1)/k$, one then has that energy is injected on spatial scales larger than $\ell_c = \lambda_{k_c} \sim \alpha^{-1/2} \varepsilon^{-1/4}$. By inserting (19) into (13) one gets (for large N)

$$E_c = \frac{\pi}{\sqrt{6}}\ \frac{\varepsilon^{3/4}}{\sqrt{\alpha}} \ , \tag{20}$$

which gives the scaling of the partial equipartition level.

Finally, the characteristic time scale over which such a state of partial relaxation sets in can be estimated through the formula

$$T(k) = \sqrt{\frac{\langle|\zeta_k|^2\rangle_\phi}{\langle|\dot{\zeta}_k|^2\rangle_\phi}} \ , \tag{21}$$

where $\dot{\zeta}_k$ stands for the right hand side of equation (14). The dimensional meaning of the above formula is obvious: $T(k)$ represents the time scale characterizing the dynamics of the k-th mode. Here too, both the numerator and the denominator have to be calculated on the configuration (15). Noting that $\langle|\dot{\zeta}_k|^2\rangle_\phi = \langle|\Omega_k\zeta_k|^2\rangle_\phi + \langle|\partial\overline{H}_3/\partial\zeta_k^*|^2\rangle_\phi$, an easy computation yields for (21)

$$T(k) = \frac{24}{\left[\left(\frac{\pi k}{N+1}\right)^3 + 36\alpha^2 E_c \left(\frac{\pi k}{N+1}\right)^2 \left(4\frac{k_c}{N+1} - 3\frac{k}{N+1}\right)\right]^{1/2}} \ , \tag{22}$$

showing that $T(k)$ too depends in fact on $k/(N+1)$. Anyway, at variance with $R(k)$, the quantity $T(k)$ is not a monotonic function of its argument, at least if the specific energy is low enough. The time scale of interest here corresponds to $\tau_c \equiv \lim_{1\leq k\leq k_c} T(k) = T(\hat{k})$; the mode $\hat{k}$ thus determined is the quickest one to reach the "plateau" of partial equipartition at E_c. It can be easily shown that τ_c has the same order of magnitude of $T(k_c)$, which can be easily computed, yielding

$$\tau_c \approx T(k_c) = \frac{6}{3^{3/2}} \ \frac{1}{\alpha^{3/2}\varepsilon^{3/4}} \ . \tag{23}$$

It can also be immediately checked that for those low modes such that $k/(N+1) \to 0$ as $N \to \infty$, one has $T(k) \sim (N+1)/(k\alpha\varepsilon^{1/2})$, a time which is much longer than τ_c and diverges with N. This represents the characteristic time scale of those first few modes that approach the partially relaxed state on times diverging with the size of the system.

What we can conclude is the following: if the specific energy of the system is low enough, say $\alpha^2\varepsilon \ll 1$, the system is expected to relax to a state of partial equipartition involving only a small fraction of long wavelength modes. Such a state is reached in times larger than a first characteristic time scale which depends only on intensive quantities and, in practical (numerical) computations, might be quite large.

One can expect that this first stage of the cascade be followed by a second stage, which takes place on much lomger times $t \gg \tau_c$, possibly leading to full equipartition. The existence of extremely long time scales of such a type (stretched exponentials of $1/\varepsilon$) is supported by numerical evidence (see e.g. [15, 16]). An analytical description of such phase of the relaxation process is beyond the scope of the present work.

4 Concluding remarks

Remark 1 - In the quoted paper of Shepelyansky [10] use is made of the resonance overlap criterion along the lines of Chirikov [11, 12]. In applying such a principle, one compares the so called nonlinear frequency shift of mode k with the difference of frequencies of modes k and $k+1$. In our opinion, such a criterion cannot be applied to the FPU models for two reasons. The first one is that no resonance of the form $\Omega_{k+1} - \Omega_k$ appears in the FPU system (12). For such a system dangerous (quasi-)resonances (appearing as small denominators in the first step of canonical perturbation theory) have the form $\Omega_{p+q} - \Omega_p - \Omega_q$. The second problem is due to the use of action-angle variables (I, φ) in the theory, which involves quantities such as the frequency of a given mode $(\dot{\varphi}_k)$ and its nonlinear correction. When one considers a nearly unexcited high mode k, the value of its action I_k almost vanishes and the action-angle coordinates become singular. This reflects in the fact that, the cubic part of the Hamiltonian being dependent on $\sqrt{I_k}$, its derivative with respect to I_k (which yields the nonlinear frequency correction) becomes artificially large. That is why, we think, the results of Shepelyansky seem not to persist in the thermodynamic limit. But in our opinion this is an artifact of the method, and not a property of the system.

Remark 2 - Our result is essentially based on the hypothesis that the state of the system at a given time $t > 0$ has the form (15). The assumption of complete equipartition between the first k_c modes has been made here just for the sake of simplicity. One could relax it and obtain the ratio $R(k)$ (17) as a functional of the unknown energies $E_1, \ldots, E_{k_c}$. The minimization of R constarined to the conservation of the second integral $J \simeq \sum_k E_k$ would then yield a more realistic distribution of the modal energies (E_k vs. k). This is deserved to further research. At present, an estimate of the spectral distribution of modal energies, based on soliton theory, is given in [9]. The connection between the present approach and that given in [9], namely the

rigorous justification of the Korteweg-de Vries equation as a resonant Hamiltonian normal form of system (1) is the object of a joint work in progress with Dario Bambusi.

Remark 3 - It is sometimes objected that the cubic potential in the Hamiltonian (1) is not bounded from below and the chain could, in principle, break-down (blow-up of some coordinates and momenta in a finite time). Anyway, break-down is caused by possible highly localized excitations on the chain, which requires very short wavelengths having a consistent amount of energy, i.e. the energy cascade to be really effective. The results of the present paper confirm that, under suitable conditions, small scale motions are frozen over rather long times, which implies absence of break-down. In references [2, 17] numerical runs over large times display no pathology, even when the system approaches equipartition.

Acknowledgments

The present contribution is the result of many discussions made on the problem within the Mathematical Physics Group, University of Milano, to which I presently belong. Special thanks go to Dario Bambusi, for useful discussions on the subject, and to Luigi Galgani, who introduced me to the "core" of the FPU problem as meant here.

References

[1] Fermi E., Pasta J., Ulam S. (1954), reprinted in "Non Linear Wave Motion", ed. by A. Newell, Lectures in Applied Mathematics **15**, 1974, pp. 143-156.

[2] Benettin G. (1986), in "Molecular dynamics simulation in classical statistical mechanical systems", Proceedings of the E. Fermi school of Varenna, 1986, ed. by G. Ciccotti and W. G. Hoover.

[3] Ford J. (1992), Phys. Rep. **213**, 271-310.

[4] Carati A., Galgani L., Ponno A., Giorgilli A. (2002), Il Nuovo Cimento B **117**, 1017-1026.

[5] Bocchieri P., Scotti A., Bearzi B., Loinger A. (1970), Phys. Rev. A **2**, 2013-2019.

[6] Biello J. A., Kramer P. R., Lvov Y. (2002), in "Proceedings of the 4th International Conference on Dynamical Systems and Differential Equations", Wilmington, NC, USA, 2002.

[7] Berchialla L., Galgani L., Giorgilli A. (2003), DCDS-B, in print.

[8] Zabusky N. J., Kruskal M. D. (1965), Phys. Rev. Lett. **15**, 240-243.

[9] Ponno A. (2003), Europhys. Lett. **64**, in print.

[10] Shepelyansky D. L. (1997), Nonlinearity **10**, 1331-1338.

[11] Izrailev F. M., Chirikov B. V. (1966), Sov. Phys. Dokl. **11**, 30-32.

[12] Chirikov B. V. (1969), reprinted and translated in CERN Trans. 71-40, Geneva, 1971, pp. 1-237.

[13] Frisch U. (1995), *Turbulence*, Cambridge University Press, 1995.

[14] Ford J. (1961), J. Math. Phys. **2**, 387-393.

[15] Pettini M., Landolfi M. (1990), Phys. Rev. A **41**, 768-783.

[16] Berchialla L., Giorgilli A., Paleari S. (2003), preprint.

[17] Livi R., Pettini M., Ruffo S., Vulpiani A. (1985), Phys. Rev. A **31**, 2740-2742.