THE FERMI-PASTA-ULAM PROBLEM

A. CARATI, L. GALGANI, A. PONNO
Università di Milano, Dipartimento di Matematica,
Via Saldini 50, 20133-Milano, Italy
E-mail: carati@mat.unimi.it, galgani@mat.unimi.it, ponno@mat.unimi.it

A. GIOR GILLI
Università di Milano Bicocca, Dipartimento di Matematica e Applicazioni,
Via degli Arcimboldi 8, 20126-Milano, Italy
E-mail: antonio@matapp.unimib.it

A review is given of the Fermi-Pasta-Ulam problem. Its foundational relevance in connection with the relations between classical and quantum mechanics is pointed out, and the status of the numerical and analytical results is discussed.

1 The FPU model and the FPU problem

The Fermi Pasta Ulam model is a system of \( N + 2 \) equal particles on a line with mutual interactions between adjacent particles, provided by a potential of the form

\[
V(r) = r^2/2 + \alpha r^3/3 + \beta r^4/4;
\]

\( \alpha = \beta = 0 \) the system is a linear one, and by a familiar linear transformation it can be reduced to a system of \( N \) independent harmonic oscillators (called normal modes) with certain frequencies

\[
\omega_j^2 = 2 \sin\left(\pi j / (2(N+1))\right), \quad j = 1, \ldots, N.
\]

The total energy \( E \) then reduces to the sum

\[
E = \sum_j E_j
\]

of the \( N \) normal mode energies \( E_j \), which are independent integrals of motion: \( E_j(t) = E_j(0) \). When the nonlinear interaction is active, the normal mode energies are no more integrals of motion, and a standard arguments of classical statistical mechanics suggests that their time averages \( \overline{E_j(t)} = (1/t) \int_0^t E_j(\tau) d\tau \) should tend to a common value, thus realizing what is usually called the equipartition of energy. More precisely this is expected to occur for almost all initial data with respect to the Gibbs measure, and in the thermodynamic limit, i.e. the limit of an infinite system with a finite nonvanishing specific energy \( \epsilon: N \to \infty, E \to \infty, E/N \to \epsilon > 0 \).

In such a case, the common value of the time averages of the normal mode energies is identified with the temperature \( T \) by \( \overline{E_j(t)} \to k_B T \), where \( k_B \) is the Boltzmann constant. Correspondingly, the specific heat (defined as the derivative of energy with respect to temperature) turns out to be a constant, independent of temperature.

The FPU problem consists in establishing whether the dynamics actually leads to equipartition. Typically one considers initial data with the energy given just to some low frequency modes, and one looks for the rate of thermalization, i.e. for the rate at which energy flows to the high frequency modes. Such a problem was first investigated by Fermi, Pasta and Ulam in the year 1954, through numerical solutions of the equations of motion for \( N = 64 \), using the facilities of the Los Alamos laboratory.1 The result they found is that, at least up to the actually observed times, the energy, initially given to the lowest frequency mode, did not appear to flow at all to the highest frequency modes, but was just shared among a
small group (or packet) of low frequency modes. The “final” distribution of energy also appeared to decrease more or less exponentially fast with the frequency. The results did not change qualitatively if the initial energy was given not just to the lowest frequency mode, but to a small packet about it.

2 The significance of the FPU problem: the FPU paradox

The question of the equipartition of energy has a deep foundational meaning for physics, because it is the one that gave rise to quantum mechanics. Indeed, as everyone knows, equipartition of energy (i.e. mean energy independent of frequency, and specific heat independent of temperature) is experimentally found to obtain only in the limit of high temperatures and/or low frequencies, and to completely fail in the complementary region. It is actually at this point the Planck's constant $\hbar$ enters the game, because it was found by Planck, on October 19, 1900, by fitting the experimental data of black body radiation, that the relevant dimensionless parameter is the quantity $\hbar\omega/k_BT$, and that the distribution of energy (per oscillator) $U$ versus frequency $\omega$ at temperature $T$ is

$$U(\omega, T) = \frac{\hbar\omega}{e^{\hbar\omega/k_BT} - 1} = kT \frac{x}{e^x - 1} \quad (x = \hbar\omega/k_BT).$$

Now, as shown by Planck in his second memoir and described in all textbooks, Planck’s law is obtained by the usual arguments of statistical mechanics if energy is assumed to be quantized. In particular, for a harmonic oscillator the admitted values of the energy should be $E_n = n\hbar\omega$, $n = 1, 2, \ldots$, (or rather $E_n = (n + 1/2)\hbar\omega$, which leads to the addition of the “zero-point energy” $1/2\hbar\omega$). If energy is not quantized, one instead recovers the “classical” equipartition value $U(\omega, T) = k_BT$. Thus the FPU result appeared as a paradox.

By the way, in our opinion it is not by chance that Fermi happened to study this problem. Indeed his interest for the problem of equipartition of energy goes back to his youth, as is witnessed by the work of the year 1923 in which he had given a subtle mathematical improvement to a theorem of Poincaré² (see also³). Poincaré was concerned with the number of integrals of motion for a Hamiltonian system, and had proven that “in general” there is just one integral, namely the total energy. Notice that this is a crucial point in connection with the problem of equipartition of energy, because for example in the FPU problem there are $N$ integrals of motion for the linearized system, and one should understand in which mathematical sense can one pass, with the introduction of a nonlinearity, to a situation in which one remains instead with just one integral of motion. The subtle mathematical point addressed by Fermi consisted in paying attention not to the integrals of motion themselves, but rather to single invariant surfaces in the phase space. In this connection let us recall that if there are $N$ independent integrals “in involution”, then the phase space is “foliated” by a continuous set of single invariant $N$ dimensional surfaces, which in the compact case turn out to be tori. This is actually the frame in which Kolmogorov (just in the same year 1954 of the work of Fermi, Pasta and Ulam) formulated his celebrated theorem, now usually known as the KAM theorem, and it is impressive how a physicist, as Fermi was,
might have thought of the problem in such terms. In any case, the theorem of Poincaré was universally interpreted as a dynamical support to the idea that in a generic Hamiltonian system all integrals of motion which are possibly present in an unperturbed system (such as the linearized FPU model) should disappear with the introduction of a generic perturbation; in our case, this would lead to equipartition of energy. The interest for this mathematical problem was clearly the reason for Fermi coming back to the equipartition problem when he happened to have a large computer available. To this historically documented fact, we can add a personal impression, that was formed by a conversation that one of us had several years ago with the late E. Segré (one of Fermi’s pupils) in Berkeley. What Segré said is that in private conversations with friends Fermi would admit that he did not really like the Copenhagen interpretation of quantum mechanics. So, perhaps, the opportunity of going back to the problem in which quantum mechanics itself had originated might have been particularly attractive for him.

3 The way out of Izrailev and Chirikov; the problem of the energy thresholds

All of us have learned at school that classical mechanics predicts a wrong result (equi-partition) in connection with the problem of the distribution of energy for a system of harmonic oscillators, this being exactly the point where the new quantum mechanics originated. So the result of Fermi, Pasta and Ulam appears as a paradox, and it is expected that one should be able to explain and eliminate it by a deeper scrutiny of the problem.

A fundamental contribution in this direction was given by a very deep paper of Izrailev and Chirikov of the year 1966. The main idea was that one should take into account the existence of some energy threshold. Actually, this is a familiar fact in perturbation theory, and possibly it came to the minds of the authors because they were the first physicists that made a connection at all between the FPU problem and the modern results (i.e. the KAM theorem) of perturbation theory. The point is that perturbation theory applies at all, i.e. the perturbed system is proven to be qualitatively similar to the unperturbed one (there are however some delicate points here, to which we will come back in another section), only if the perturbation is low enough, namely if it is smaller than a certain threshold, which should be suitably estimated in each particular case. In the FPU model the unperturbed system is the linearized one, with its $N$ independent integrals, and what plays the role of the perturbation is essentially the total energy, because the relative “size” of the nonlinearity tends to zero with the total energy. So the idea was that the perturbed system would be qualitatively similar to the unperturbed one only for energies smaller than a certain critical energy, say $E^c$; the result of Fermi, Pasta and Ulam should then be explained as due to the fact that they had considered small initial energies, below the threshold $E^c$. Finally, the paradox would disappear at all if one might prove that the threshold vanishes in the thermodynamic limit.

Actually, the authors even went further, because they also provided, with some heuristic arguments, an estimate for the threshold. More precisely, they also conceived the idea that the threshold should depend on the choice of the initial data:
for an initial excitation of mode $j$ one should correspondingly have a threshold $E_j^*$. So they gave their estimate for the threshold $E_j^*$, or at least two limit expressions for it for the case of low frequencies (small $j$) and for the case of high frequencies (large $j$). The key point now is: how the dependence on the number $N$ of particles because, according to their estimates, the energy thresholds would tend to zero, at least for the case of large $j$. This result almost eliminates the paradox, because, at least for initial data with excitations of the high frequency modes, in the thermodynamic limit one would always be above threshold, i.e. the system would have almost no relation to the unperturbed one and thus would be expected to lead to equipartition.

The job would be totally accomplished if one were able to produce estimates for the thresholds in the case of small $j$, presenting the same property of vanishing in the thermodynamic limit. In a recent paper by Shepelyansky\textsuperscript{5} it is stated that such an accomplishment has now been performed. In his words, the aim of his paper is the following one: “A possibility that in the FPU problem the critical energy for chaos goes to zero when the number of particles in the chain increases is discussed.” In the introduction 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 neglected 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 arbitrary small nonlinearity”.

4 The result of Bocchieri, Scotti and Loinger, and the Planck like distributions

A completely different perspective had been however imagined in the meantime. This occurred in Milano, through the work of Bocchieri, Scotti and Loinger\textsuperscript{4} of the year 1970 and a subsequent one\textsuperscript{7}. The theoretical group of Milano was then led by Caldirola, who had been among the young physicists influenced by Fermi just before the war, and happened to be particularly interested in problems concerning the foundations of theoretical physics. So in Milano people were particularly sensitive to the possible foundational relevance of the FPU problem, with its implications for the relations between classical and quantum mechanics. Thus, when the FPU problem made its way to Italy through the work of Izrailev and Chirikov, that had been discovered by Loinger, there naturally arose the idea of checking, by numerical solutions of the equations of motion, whether the specific critical energy $\epsilon^* = E^*/N$ vanishes in the thermodynamic limit or not. This was done by Bocchieri, Scotti and Loinger. Actually, the critical energy was investigated just for initial data as in the original work of Fermi, Pasta and Ulam, namely for initial excitations of the lowest mode. The energy threshold was defined in the simplest possible way, by computing the time averages $E_j^*(t)$ until they had apparently settled down to some “final” stationary value. It turned out that there was some critical energy
above which there was an apparent equipartition, while equipartition did not occur at lower energies. The main result was that the critical energy \( E^c \) appeared to be proportional to \( N \), i.e., there appeared to exist a finite nonvanishing specific critical energy \( E^c > 0 \). In other terms, it was suggested that according to classical dynamics there is a relevant set of initial data which lead to “final” states not corresponding to equipartition of energy. An interesting fact concerning this work is that the computations were performed with a realistic interatomic potential, namely a standard Lennard Jones potential \( V(r) = 4V_0 [(σ/r)^{12} − (σ/r)^{6}] \); this involves two parameters, \( V_0 \) and \( σ \), giving the depth and the width respectively of the potential well. The computations were actually performed by taking for the molecular parameters \( m \) (the mass of the particles), \( V_0 \) and \( σ \) realistic values corresponding to Argon as deduced from standard textbooks. The specific critical energy \( E^c \) turned out to have a value which is more or less 4 percent of the depth \( V_0 \) of the potential well.

Many discussions followed this striking result. Shortly later, an investigation was made\(^7\) of the distribution of energy for the ‘final’ states that are found below threshold (still with initial data of FPU type), looking for a function that gives the final values \( E^*_j \) in terms of the corresponding frequencies \( ω_j \). It was found that the curves were rather well fitted (apart from a short plateau at the very low frequencies) by Planck-like distributions of the form

\[
E^*(ω, E) = \frac{Aω}{e^{βAω} - 1},
\]

The parameter \( β \) was depending on the total energy \( E \) more or less as an inverse temperature should, while the parameter \( A \) appeared to be a constant. The most striking fact was that, with the realistic values of the molecular parameters corresponding to Argon which had been chosen in the computations, the quantity \( A \) turned out to have a value very near to that of Planck’s constant. It took some time to understand this point: in brief, Planck’s constant had been introduced, somehow by hands, through the molecular parameters. This goes as follows. One immediately checks that the natural action built up from the parameters is just \( \sqrt{mV_0} σ \), so that one can write \( A = a \sqrt{mV_0} σ \) with a pure number \( a \); on the other hand from the textbooks it turns out that one has, in an incredibly precise way, \( \sqrt{mV_0} σ ≈ 2Z\hbar \), where \( Z \) is the atomic number. Thus the numerical computations had just provided an estimate of the pure number \( a \), which turned out to be of the order of magnitude of 1/50. Shortly later an interesting contribution was also given by Cercignani, who suggested\(^8\) that there might be an analogy between energy thresholds and quantum zero point energy.

5 The problem of the relaxation times; old and modern aspects

One thus remains with the problem of deciding between two possible alternatives. Denote by “freezing” the FPU qualitative phenomenon that the “final” distribution of energy is near the initial one, so that in particular equipartition does not hold for the final distribution if the initial one corresponds to excitation of the very low frequency modes; such a freezing is expected to hold below some energy threshold.
Then the alternative is whether such a freezing persists in the thermodynamic limit or not, i.e. whether it is relevant for physics or not. At first sight, one might be tempted to say that the school of Chirikov would bet for the second alternative, and the group of Milano for the first one. But perhaps we are misinterpreting our colleagues, and their hopes might be not so dissimilar from ours.

Passing from hopes to facts, or to theorems, it turns out that the answer is not at all simple, and actually has not yet been afforded. The main difficulty resides in providing a clear definition for the freezing, especially in connection with the question of the times involved (i.e. the size of the relaxation times in relation to the observation times). Another point concerns the meaning that should be attributed to the notion of an “energy per oscillator”. These are indeed quite delicate problems on which we are presently actively working, and we limit here ourselves to some comments, mainly addressed to the first problem, i.e. that of the times involved.

The general physical problem of the dependence of the results on the observation time turns out to have a a strong counterpart in perturbation theory. Indeed, in general, in perturbation theory one aims at proving that a certain system is “similar” to another “unperturbed” one, but the mathematical implementation of such an idea requires that preliminarily a time \( t \) should be fixed up to which the similarity should hold; such a time is the counterpart of the physical observation time. On the other hand it is a general fact that the similarity can be proven to exist only if the perturbation is below a certain threshold, so that correspondingly the threshold turns out to depend on the given observation time.

Now, the KAM theorem refers to an infinite observation time, and all the available estimates indicate that it should not apply in the thermodynamic limit (we do not discuss here the problem of the existence of invariant low dimensional tori, which is now so popular in the mathematical literature, and is studied for example by Kulesin and by Bambusi). One can instead make reference to finite times, and Nekhoroshev\(^9\) has taught us\(^10\) how to deal with them in a perspicacious efficient way (see also\(^11\)). Many numerical and analytical studies have been performed from this point of view on the FPU and related models\(^12\), on which we do not have time to enter. What we want to stress here is that finally the scientific community seems to have come to agree that taking into account the observation time is a physically relevant requirement even in the FPU and related problems. Asking whether one has equipartition or not without an accurate discussion of the times involved is extremely naive and unphysical.

By the way, it has also been realized that the relevance of the observation time in this connection was actually well known since the “old times”, because it was Boltzmann himself\(^13\) that for the first time conceived that the phenomenological lack of equipartition in crystals and polyatomic molecules could be explained as corresponding to the fact that equipartition had not been achieved within the actual observation times; the relaxation times to equipartition would be much longer than the experimental observation times. This idea was pursued by Jeans\(^14\), and then discussed at the first Solvay conference\(^15\), particularly by Nernst, who declared that such long relaxation times had never been observed in experiments. The story has been described elsewhere\(^16\). In short, “long” relaxation times (even of the order of one second) have actually been observed in the phenomenon of the dispersion
and anomalous absorption of sound in diatomic molecules\textsuperscript{17}, and is rather well accounted for by the classical theory, although some delicate problems are still open. Moreover, the circumstance that one should have some "time dependent specific heat" is presently accepted even as a trivial fact\textsuperscript{18}. The situation is however rather delicate, especially in connection with the problem of understanding from this point of view the standard static measurements of the specific heats.

The discussion would become here really intricate, and at the moment we do not have a clear answer available. We have however a main qualitative perspective. The general idea is that in classical models of crystals and of polyatomic molecules one might meet with situations qualitatively analogous to those which are met in the phenomenology of glasses, spin glasses and polymers, where an essential role is played by the fact that there exist relaxation times differing from each other by huge orders of magnitude. So, one might have an essentially rapid relaxation to some kind of metaequilibrium state, which should last for an extremely long time; the final relaxation to a standard Maxwell Boltzmann equilibrium, and thus to equipartition, might then occur only over such huge time scales. Something like this was suggested for the first time in a work\textsuperscript{19}, where the interaction of a FPU system with a heat reservoir was studied numerically. Such an idea had also been pursued for the case of polyatomic molecules\textsuperscript{20}. A special attention had there been given to the so called Landau Teller model\textsuperscript{21,22}, which takes there the role of the FPU model. In particular, it was recently observed that in the Landau Teller model of molecular collisions the energy of the internal vibrations performs a kind of random walk in which there occur rare conspicuous jumps, somehow analogous to those occurring in Lévy processes\textsuperscript{23}. One should then meet there with the phenomenon of the anomalous diffusion, which might thus be expected to occur also in the FPU model.

In connection with the physical necessity of taking into account the observation times, one also meets with a quite delicate problem of interpretation pointed out by Boltzmann. The problem concerns the identification, which is usually made in statistical mechanics, between thermodynamic energy and mechanical energy. According to Boltzmann, one should declare in advance which is the chosen observation time. Then, considering a system which possesses a certain mechanical energy in virtue of some initial conditions, its thermodynamic energy should be identified with the fraction of the mechanical energy that the system can actually exchange with the measurement instrument up to the given observation time. As pointed out by Nernst\textsuperscript{24} (see also\textsuperscript{25}), in such a way one might have a situation in which there is equipartition of energy for the mechanical energy, just in virtue of the choice of the initial data (according to the Maxwell Boltzmann distribution), and instead a Planck distribution for the exchangeable energy (see also\textsuperscript{26}). This is a crucial point if one wants to interpret the phenomenon of the freezing of the high frequencies modes when one considers initial conditions of a generic type (i.e., according to Maxwell Boltzmann), and not just of the special type corresponding to an excitation of the low frequency modes.
6 Some recent results

We now quickly describe some recent results, which are mostly still unpublished. The first one is of numerical type, and is already in print \(^{27}\). For initial data of the FPU type, a strong evidence is given of the fact that the results depend on the specific energy \(\epsilon\) in the following way. There exists a critical specific energy \(\epsilon^c\) such that for \(\epsilon > \epsilon^c\) equipartition is obtained within a time that increases as an inverse power of \(\epsilon\) with decreasing \(\epsilon\). Instead, below threshold, i.e. for \(\epsilon < \epsilon^c\), one meets with two time scales: in a short time there is formed a "natural packet" that extends up to a maximal frequency \(\omega(\epsilon)\) proportional to \(\epsilon^{1/4}\). Only on a much longer time scale would one get equipartition. Just in these days indications are being found that such a large time scale might increase as a stretched exponential of \(1/\epsilon\).

The natural packet mentioned above is presumably to be identified with what in the year 1972 was considered to be the "final state" providing a Planck-like distribution. So it is of particular interest to obtain any possible analytical information about it. This has now been afforded, and will in a short time be written down. By arguments related to the description of the FPU model in terms of solitons, along the lines of the celebrated work of Zabusky and Kruskal \(^{28}\) (see also \(^{29}\)), it is shown that the results actually depend on the specific energy, and an explicit analytical formula is given for the natural packet, which is confirmed to extend up to a maximal frequency \(\omega(\epsilon)\) proportional to \(\epsilon^{1/4}\). The analytical form of the packet is found to fit in an extremely good way the numerical data.

Some progress was also made in the direction of getting rigorous analytical results in the thermodynamic limit, because for the first time it has been possible to perform a finite number of perturbative steps in that limit. This required the establishment of a suitable measure theoretic framework for perturbation theory itself, and in particular a clarification of what should be meant by "energy per oscillator", in that limit.

7 Final comments

So, let us come back to the problem of deciding between the two alternatives, namely whether the original FPU result is relevant for physics or not. In the light of the recent results just mentioned, we are confident that the FPU paradox cannot be eliminated and that it has a deep physical meaning.

The general perspective mentioned above naturally leads to the following interpretation of the paradox. Before Fermi, Pasta and Ulam the alternative was between classical mechanics, which should be wrong, and quantum mechanics, which is correct. But this makes no reference to times. We would instead suggest: up to "short" times classical mechanics might qualitatively agree with quantum mechanics, and only later on might they differentiate. Indeed, according to quantum mechanics Planck’s law is the final equilibrium distribution, while, apparently, according to classical mechanics it might just describe a metaequilibrium distribution which only over much longer glassy like time scales would finally evolve to the "classical" Maxwell Boltzmann equilibrium.
By the way, such a perspective seems to be in the way of becoming a rather popular one. Indeed it is presently often stated that there should exist some characteristic Ehrenfest time up to which classical and quantum mechanics agree in predicting motions of “ordered” type; later on they would instead differentiate, because nothing would happen according to quantum mechanics, while “chaotic” motions would occur according to classical mechanics. This is actually qualitatively analogous to the perspective proposed here.

So much for what concerns the possible logical relations between classical and quantum mechanics, in connection with the problem of equipartition of energy. It would be very interesting to know what is the actual status of the experiments concerning measurements of the specific heats of crystals and of polyatomic molecules over extremely long times.

References

2. E. Fermi, Nuovo Cim. 25, 267 (1923); Phys. Z. 24, 261 (1923)
14. J.H. Jeans, Phil. Mag. 35, 279 (1903)
15. La théorie du rayonnement et les quanta, ed. P. Langevin, M. de Broglie (Gauthier Villar, Paris, 1912)