Università degli Studi di Torino
Scuola di Dottorato

Wave Turbulence dynamics in one-dimensional nonlinear lattices with nearest-neighbour interactions

Lorenzo Pistone

Università degli Studi di Torino
Scuola di Dottorato in Scienza ed Alta Tecnologia

Indirizzo di Fisica ed Astrofisica

Wave Turbulence dynamics in one-dimensional nonlinear lattices with nearest-neighbour interactions

Lorenzo Pistone

Supervisor: Miguel Onorato
Co-supervisor: Sergio Chibbaro

Grazie ai miei maestri,
Antonio,
Tito,
Miguel,
Sergio,
Alireza.

Abstract

The thermalization properties of several nonlinear Hamiltonian lattices are studied. The studied models are the α and β-Fermi-Pasta-Ulam-Tsingou models, the nonlinear Klein-Gordon lattice, and its disordered variant. We find evidence that the dynamics are dominated by resonant and quasiresonant interactions of the eigenmodes of the linearized Hamiltonians. The existence of resonances makes it possible to use Wave Turbulence theory in order to derive quantitative estimates over the equipartition dynamics of the lattices. Wave Turbulence is a statistical theory of dispersive wave systems with a weak nonlinearity, and it is of perturbative nature. These estimates concern the route to thermalization from a non-equilibrium initial conditions. In particular, no energy threshold is found for the system to thermalize. The statistical dynamics are predicted to show one possible additional conserved quantity that has no counterpart in the deterministic dynamics. This statistically conserved quantity could provide insight in the study of metastable states in the Fermi-Pasta-Ulam-Tsingou model and similar. The time to equipartition is found to scale as a power-law of the nonlinearity. The role of quasiresonances is investigated, and it is found to be particularly important in the disordered Klein-Gordon lattice. The disordered Klein-Gordon lattice is characterized by random but fixed-in-time parameters. Our analytical results are validated with extensive numerical simulations. The code for these simulations is also presented and made available to the scientific community.

Contents

Introduction

In statistical physics, the concept of equilibrium plays a central role. In the broadest possible way, we can say that equilibrium is a state where the macroscopic observables of the system are essentially constant over a prolonged period of time. In general, this state is also the one where the system is most disordered, as in, there are no recognizable patterns in the state at microscopic level. In this sense, the evolution of a system from an initial configuration, is its route to equilibrium, and we observe generation of disorder.

This is also a common experience. Whenever we set up something in a particular way, we decide the initial state of a system with possibly a large number of symmetries, but we also find that, if we just let the system evolve, the symmetries cease to exist after a rather short time. For example, we could slap our hand on the water surface of a full bathtub: the initial excitation would be very much localized in space, its parts would be highly correlated in space, etc., but then after some time the ripples on the surface would contain all the initial energy in a much more sparse way. Arguably, the final distribution of energy is the most disordered.

This kind of common experience is so solid that often the concept of equilibrium and its implications are not deduced from the characteristics of the thermodynamic system, but they rather pose the foundation of any successive reasoning. It is a fair assumption.

My doctoral thesis has its roots in the famous FPUT (Fermi, Pasta, Ulam, Tsingou) experiment [1], which is one case where equilibrium was expected, but it failed to show incontrovertibly. This must have been of much interest to Fermi, who penned in young age a demonstration that all mechanical systems are quasi-ergodic [2]. The authors chose to study a toy model: a chain of equal masses coupled with nearest-neighbour interactions (each mass pushes or pulls the mass in front and behind). The number of masses was not very large, and the simulated time was rather short due to the limitations in computing power of the time. They tried to observe any sign of equilibrium in the system from an initial arbitrary distribution of energy, but they failed to do so.

The puzzling results of the FPUT model are a popular topic in mathematical physics. The dynamics are rich even though the description of the system is simple and short. Explaining in a satisfactory way the FPUT experiment would proba-

bly teach us something on the connection from microscopic to macroscopic state, because we try to derive macroscopic observables from simply the specification of the mechanical system, a connection which is valuable in itself. Additionally, the simplicity of the model allowed for countless modifications. To name a few, one could consider multi-neighbour interactions, dimensionality, all different form of interaction, number of masses, as well as studying the conductivity of the lattice, or introducing random variations of its parameter over time or space, etc.

When I began my PhD course, my supervisor Prof. Miguel Onorato and collaborators had just published a paper [3] on the original FPUT problem which applied techniques borrowed from his work on ocean waves. The prediction was on the time the FPUT chain needed to evolve to reach equilibrium from an initial arbitrary excitation. Their contribute was in my opinion interesting because it appeared to be quite robust to the change of details of the system, of course at the expense of some precision in the predictions.

We first investigated whether the results of [3] could be applied to another well-known model, the discrete nonlinear Klein-Gordon lattice [4]. We were confident enough then that the result could have some significance across the details of the lattice dynamics. We published a comparative view of multiple models in [5], where we also explored the thermodynamic limit of the α-FPUT, β-FPUT lattices. We are now in the phase of editing the material on the disordered Klein-Gordon lattice (a lattice with a random variation of the masses along the chain). We have also published the software [6] that I developed to run the extensive simulations needed to support our ideas, because I believe that a better accountability of the software engineering process is urgently needed in our scientific community. In this PhD thesis I am going to present these publications with an introduction in order to help the reader understand the message of the papers.

In the course of my PhD my understanding of the problem evolved and has changed. I believe that in our latest effort in cooperation with Prof. Sergio Chibbaro, several aspects of our approach where challenged. This difficulty has led in my opinion to a better understanding of some aspects of our earlier works. I wanted to make use of this accumulated knowledge to communicate to the reader the work of three years, which is disseminated with corrections and new questions arising during the way. For this reason, I will first give in Chapter 1 a very high-level overview of the topic of my PhD. Then, the published papers [4, 5, 6] are presented in chronological order in Chapters 2, 3 and 4, with a short introduction to highlight the takeaways of each. This should allow the reader to track the evolution of the ideas of the PhD. In Chapter 5 I will present the latest work, in collaboration with S. Chibbaro and M. Onorato, on disordered nonlinear lattices. Finally, in Chapter 6 I will present my conclusions and point the reader to some possible further direction of research.

Chapter 1

A primer on the FPUT problem, Wave Turbulence and their relation

1.1 The FPUT model

In the original FPUT experiment [1], Fermi and collaborators came up with a toy model and a simple problem of statistical physics, in order to test one of the first computers available. They devised a simple model made of point particles in motion, and they tried to observe a typical trait of ergodicity, that is equipartition of energy among the system. Let us consider an Hamiltonian system with N particles, described by their coordinates and momenta q_i, $p_j = \dot{q}_j$, $0 \leq i < N$ (in the original work, $N \leq 64$). The original expression of the FPUT Hamiltonian is

$$H = H_{\text{lin}} + H_{\text{nlin}} = \sum_j^N \left(\frac{p_j^2}{2} + \frac{1}{2}(q_{i+1} - q_j)^2 \right) + H_{\text{nlin}}, \tag{1.1}$$

where H_{nlin} contains all terms that are not quadratic in the Hamiltonian. If $H_{\text{nlin}} = 0$, the system can be cast in an eigenvalue problem, an analytical or numerical solution can always be extracted, and the system is integrable. Fermi and collaborators considered the cubic and quadratic potentials

$$H_\alpha = \sum_j^N \frac{\alpha}{3} \left(q_{j+1} - q_j\right)^3, \; H_\beta = \sum_j^N \frac{\beta}{4} \left(q_{j+1} - q_j\right)^4 \tag{1.2}$$

and a piecewise quadratic potential, which it is little studied as it can be approximated by the cubic potential and it is significantly harder to treat analytically. The full Hamiltonians complemented by eq. (1.2) are widely known as the α-FPUT and β-FPUT models. The parameters α and β control the deviation of the system from the linear dynamics ($H_{\text{nlin}} = 0$). In the original experiment, and throughout

this thesis, the deviation is small. This is quantified by the nonlinearity parameter

$$\epsilon = \frac{\langle |H_{\text{nlin}}| \rangle}{\langle H_{\text{lin}} \rangle} \ll 1, \tag{1.3}$$

where the average is intended either over time or over an ensemble. Note that there is to date no proof that that the FPUT system is ergodic.

The original experiment went as follows. Energy was injected in a few of the eigenmodes of the linearized eq. (1.1). The eigenmodes of H_{lin} in this case are sinusoidal waves. Then the time was advanced in the simulation, and the energy per eigenmode was tracked. The expectation was to see a rather quick radiation of energy from the initially excited modes into the others, so that the energy would be on average equipartitioned among all the degrees of freedom. Rather, it was observed that the system followed an almost periodic motion in phase-space, up to the time that they could reach with the simulations of the time. In other words, equipartition was not observed.

While initially forgotten, this problem proved to be quite tricky, and it attracted a strong interest from the physical community, especially from the dynamical systems community. A very good review of the historic evolution around the FPUT problem can be found in [7]. About the history of the FPUT problem, I will comment on a couple of milestones that can help put this thesis into context.

The next logical step was to identify the possibly different regimes of the FPUT dynamics. Eq. (1.1) can be rescaled [3] so that its dynamics depend only on the nonlinear parameters α or β (in the present notation they are proportional to ϵ). It was in fact observed [8] that the FPUT chain showed mostly regular motion at low nonlinearity, and then above a certain threshold dependent on N it showed an abrupt transition to chaotic motion. This threshold appears to decrease quickly with N. The existence of a kind of threshold is generally acknowledged in all branches of research around the FPUT problem. It is observed [9] also that this threshold appears to exist at all N, even though coherent nonlinear structures can be observed at large N and times.

1.1.1 The proximity of the FPUT model to integrable dynamics

Initially, the efforts around the FPUT problem were spent in the direction of finding an application of the Kolmogorov-Arnold-Moser (KAM) theorem, which roughly states that perturbations of the regular tori in the phase-space of a system can remain essentially stable for exponentially long periods of time. This interpretation was also able to explain in some sense the threshold mentioned above. In fact, KAM theory can only be applied if the perturbation remains small. In practice, if the nonlinearity is small enough, there would be exponentially long times to escape from the trajectory of the linearized dynamics. For the fact that the threshold quickly decreases with N, KAM theory is practically relevant to very short FPUT chains.

The idea that the FPU model shows long thermalization times because it is close to integrable dynamics has a long history. It was quickly realized that the long-wave approximation of the FPUT problem (large N, slow variation of position and momentum across near-by particles) is the well-known Korteweg–de Vries equation [10], which is known to have stable solutions, the solitons. This is a concept that is found also in other models. For example, different variations of the nonlinear Schrodinger equation are integrable [11], but their discrete counterparts are not necessarily integrable. Another important integrable model that is close to the FPUT problem is the Toda lattice [12], a nonlinear chain with an exponential nonlinearity. It is widely known in fact that the FPUT problem can be seen as a first order approximation of the Toda lattice, which is integrable. Analysis in the base of the Toda constants of motion is possible [13], and can be extended to higher-order models such as the $\alpha - \beta$ FPUT model (where both the cubic and quartic nonlinearities are present).

In many dynamical systems, the leading dynamic of the transfer of energy is due to resonances. Resonances are a coupling of the modes that are phase-locked. Imagine for example two strings tuned to the same note in a musical instrument: if we strike one string, also the second will start to vibrate and emit sound almost immediately. This is because the two strings share the same base harmonic, hence the transfer of energy is very efficient. Surprisingly enough, in the FPUT system the existence of first-order resonances actually leads to integrable dynamics [14, 15, 16]. However, we found that higher-order resonances, or quasi-resonances can explain the route to thermalization of these lattices.

A significant advancement in the understanding of how an Hamiltonian system can transit to chaos is given in [8]. The authors suggested that the nonlinearity could make the frequency of the modes stochastic, to the point that new resonances appear in the system. This is what will be referred to in this thesis as the Chirikov argument, introduced in [17]. This is a concept that allows me introduce a few useful properties of the linearized FPUT problem, which obviously apply also to the other dynamical systems considered in this thesis, therefore I will introduce it in some detail.

1.1.2 The linearized FPUT dynamics, and the nonlinear broadening

A linear mechanical system can be cast in matrix form and diagonalized. This means that the dynamics can be decomposed into a linear sum of modes, that is configurations of the chains, characterized by an amplitude (how much energy is in that mode) and a phase ϕ_j. These modes evolve independently, that is their, amplitude does not change and the modes do not interact with each other. The evolution of the dynamics that can be observed in physical space (the q_j and p_j variables) is the result only of the evolution of the phase, which is periodic with a

specific frequency ω_j,

$$\phi_j(t) = -\omega_j t + \phi(0) \mod (2\pi), \tag{1.4}$$

where $\phi(0)$ is an arbitrary initial condition. In the original FPUT problem, the modes are plane waves in complex space characterized by a momentum $0 \leq k < N$, and frequencies are given by

$$\omega_k = 2\left|\sin\left(\pi\frac{k}{N}\right)\right|. \tag{1.5}$$

Since we generally study the FPUT problem in the weakly nonlinear regime, it is natural to use this decomposition into modes even if there is a small nonlinearity, and the exchange of energies is intended as an exchange of energies between modes which are now coupled by the nonlinearity.

The core of the argument of Chirikov's proposition is the following. Note that it is not a formal argument, but it is of qualitative nature. It is observed that when a nonlinearity is introduced in a linear system, the frequency ω_j becomes stochastic in time. That is, since the modes are now coupled, eq. (1.4) as an additional, small contribute which depends on all the other modes in the system. This contribution, however, is a function of several *weakly* coupled degrees of freedom, so it is highly uncorrelated in time. This kind of incoherent interaction is referred sometimes as heating [18].

When the stochasticization of the frequencies is activated through nonlinearity, a mode could excite other modes that are in quasi-resonance. This mechanism, and its measure σ_{ω_j} is generally referred to as "nonlinear broadening". In [8], it is proposed that the transition to chaotic dynamics happens when the typical broadening σ_ω becomes comparable with the typical frequency separation $\Delta\omega$,

$$\sigma_\omega \sim \Delta\omega. \tag{1.6}$$

The modes that are close to resonance form a quasiresonance, and they will be central to our analysis.

1.1.3 The route to equipartition, and initial conditions

Still after more than sixty years, there is no clear picture of the thermalization properties of the FPUT chain. Besides the intrinsic difficulty of the subject, among the possible reason is that the dynamics are very sensitive to the initial conditions. The experimental strong dependence on the initial condition makes theoretical approaches quite limited in scope. I will briefly outline here some relevant points regarding this issue, so that the reader can understand some choices adopted in this thesis.

In the original paper, the chain was initialized with the energy in the lowest mode, that is $k = 1$ in eq. (1.5). Note that the mode $k = 0$ formally exists but it

corresponds to a solid translation of the chain which is always uncoupled to the other degrees of freedom. There was no formal measurement of the equipartition, because it was simply not observed even qualitatively. It was soon found that the time to equipartition depends strongly on the initial conditions. Since the original FPUT experiment initialized energy in the lowest modes, other experiments focused on different inputs of energies, such as the highest modes [19]. The distinction between low-frequency and high-frequency initial conditions is quite ubiquitous in the literature around the FPUT problem. Several multi-mode initial excitations were also investigated, and showed remarkably different results from single-mode excitations [20].

The definition of the measure of thermalization is also not uniform across literature on the FPUT problem. While this is a technical matter and often measures are equivalent, it is important to be aware of the implications. Equipartition can be measured essentially in two ways. One can look at the average potential and kinetic energies of the individual sites, that is a measure in physical space, or at the energies of the modes of the basis chosen to decompose the nonlinear problem, that is in the eigenvectors space. As a broad indication, the first method is generally used when the initial conditions are localized in space (e.g. we observe the spread of the excitation across the chain after it is initially spatially localized), while the second way is used when the initial energy is initially localized in some modes. It should be noted that a system at equipartition according to the first measure is not necessarily equipartitioned according to the second measure, because the modes may cover the whole physical space, and so all the sites carry some energy. The physical space variant is in fact especially used in the context of "disordered" lattices where the there are no travelling wave solutions to the linearized lattice (see Chapter 5). It is evident that if the excitation is initially localized, then its action on the dynamics evidently change over time when the initial excitation attains relaxation. This can make numerical and theoretical investigations much harder, because different dynamical regimes may be crossed during time evolution.

Another possible complication in the measure of thermalization is that the FPUT chain appears to often evolve across different timescales. That is, some states are metastable, and may appear stable if one observes a specific time span, but they actually evolve at a much larger timescale (see for example [21]). The definition of equipartition time then becomes difficult. Lastly, the existence of these metastable states in some numerical experiment makes it difficult to determine whether equipartition is reached or not, because an extensive numerical validation would require large simulation times.

Because of the problems outlined above, it should be clear to the reader why it is important to state what kind of initial conditions one intends to investigate when approaching the FPUT model and other nonlinear lattices. I will refer to this section later when the next necessary concepts are introduced.

1.2 A briefing on Wave Turbulence

In the previous section we saw how the idea of resonances in Hamiltonian lattices has some application. The existence of a structure of resonances is a widely studied subject [22, 23] in the FPUT community. Here, I briefly introduce Wave Turbulence theory, where the concept of resonance plays a central role, and as such it appears as a great candidate to explain some of the thermalization properties of nonlinear lattices.

Wave Turbulence (WT) is a statistical theory for dispersive wave systems with a weak nonlinearity. It is relevant in a number of physical systems that span disparate fields, from oceanography to quantum condensates [24, 25, 26]. Traditionally, WT is developed only on continuous systems, or on the thermodynamic limit ($N \to \infty$) of a discrete system. To be clear, we intend a infinite lattice, so continuity is in Fourier space, but the physical space is still discrete. To the best of my knowledge, any attempt in the direction of removing this limitation has failed to reach the level of insight that WT provides on weakly nonlinear wave systems. This does not mean that WT deals only with continuous systems. In a discrete system resonance still exist, though the finiteness can make them more scarce [27, 14, 28]. Throughout this thesis we will assume that a continuous limit can be still relevant in discrete case with not too few degrees of freedom. Obviously, the lower the number of degrees of freedom, the more the system is expected to approach quasi-integrable dynamics, hence the WT predictions eventually do not apply any more. The final verdict on whether WT is applicable is then obviously numerical experimentation.

I will use in the following an example model to clarify some of the propositions. This model is the disordered nonlinear Klein-Gordon lattice. It is an extension of the ordered nonlinear Klein-Gordon lattice, that we studied in [4]. Here the word disordered means that some parameters of the Hamiltonian are random, but do not vary in time. This kind of systems is studied in solid state physics; we will provide a motivation for studying this model Chapter 5. For now, let us consider this system just for explanatory purposes. It is particularly instructive because, since random parameters are present, the diagonalization procedure has to be performed symbolically, and as such, it is quite generic and can cover several nonlinear lattices.

1.2.1 Normal modes

The Hamiltonian of the disordered nonlinear Klein-Gordon lattice is

$$H = \sum_{j}^{N} \frac{1}{2}p_j^2 + \frac{1}{2}(q_{j+1} - q_j)^2 + \frac{1}{2}m_j q_j^2 + \frac{1}{4}\beta q_j^4. \tag{1.7}$$

This is the model that we are going to use as a primary example in this section. The variables m_j are random positive numbers, but the details are not important for now. β controls the nonlinearity, and it is restricted to be positive so that the Hamiltonian is bounded from below. The equation of motion reads

$$\ddot{q}_j = -2q_j + q_{j-1} + q_{j+1} - m_j q_j - \beta q_j^3. \tag{1.8}$$

In standard WT methodology [26, 29], one must first diagonalize the quadratic part of the Hamiltonian of the system. The linear eigenvalue problem can be formulated by inserting the plane wave solution $q_j = v_j e^{-i\omega t}$ in eq. (1.8),

$$-\omega^2 v_j = -m_j v_j + v_{j-1} - 2v_j + v_{j+1}, \tag{1.9}$$

which can be solved as an eigenvalue problem. This is a feature common to all systems which only have a nearest-neighbour interaction, such as the ones considered in this thesis. After diagonalizing, we obtain an orthonormal basis v_j^k with eigenvalues $\lambda_k = \omega_k^2$ (k is the eigenvalue index and j the site index in physical space). The solutions to this eigenvalue problem can be found analytically or numerically. In the case of ordered media, it may be possible to obtain a dispersion relation, that is a function $\omega(k)$. In disordered media, the index k can in general only be used as a label of the eigenstates.

The physical space variables are projected onto the diagonal basis to obtain the new variables Q_k and P_k,

$$Q_k = \sum_i^N v_i^k q_i,\ P_k = \sum_i^N v_i^k p_i. \tag{1.10}$$

It is easy to see that in the diagonal variables (1.10) the energy of the linearized system is constant per mode $\frac{1}{2}\left(P_k^2 + \omega_k^2 Q_k^2\right)$, and it is continuously shuffled between kinetic and potential energy. This is better shown by using the normal modes,

$$a_k = \frac{1}{\sqrt{2\omega_k}}(P_k - i\omega_k Q_k), \tag{1.11}$$

so that the original Hamiltonian finally reads (we use the shorthand notation for indices $i_1, i_2, i_3, ... \to 1, 2, 3, ...$)

$$H = \sum_1 \omega_1 |a_1|^2 + \frac{\beta}{4} \sum_{1,2,3,4} W_{1234} \left(a_1 a_2 a_3 a_4 - 4a_1 a_2 a_3 a_4^* + 3a_1 a_2 a_3^* a_4^* + c.c.\right) \tag{1.12}$$

with interaction coefficient

$$W_{1234} = \frac{1}{4\sqrt{\omega_1\omega_2\omega_3\omega_4}} \sum_i^N v_i^1 v_i^2 v_i^3 v_i^4, \tag{1.13}$$

and the equation of motion reads

$$i\dot{a}_1 = a_1\omega_1 - \beta \sum_{2,3,4} W_{1234} \left(a_2 a_3 a_4 - 3a_2 a_3 a_4^* + 3a_2 a_3^* a_4^* - a_2^* a_3^* a_4^* \right) \tag{1.14}$$

We observe that eq. (1.14) is very similar to the case when disorder is not present in the Klein-Gordon model [4]. This is to be expected, because the features of the linearized dyamics are all inside the particular solution to the eigenvalue problem, which shows in normal mode space simply as the specific set of frequencies ω_j.

1.2.2 Resonances and quasiresonances

Let us now focus on the nonlinear terms in eq. (1.12). We will show how these terms identify resonances, which are primary building objects of WT. The following considerations apply to quite a lot of nonlinear lattices, as long as the nonlinearity is a positive integer power-law of the variables.

From eq. (1.14) consider the action-angle coordinates,

$$a_j = \sqrt{I_j} e^{-i\phi_j} \tag{1.15}$$

with I_j and ϕ_j real. It is easy to see that the evolution equation for the action I_j is

$$\begin{aligned} \dot{I}_1 = 2\beta \sum_{2,3,4} \sqrt{I_1 I_2 I_3 I_4} W_{1234} \Im \big[-e^{i(\phi_1 - \phi_2 - \phi_3 - \phi_4)} + \\ + 3e^{i(\phi_1 - \phi_2 - \phi_3 + \phi_4)} - 3e^{i(\phi_1 - \phi_2 + \phi_3 + \phi_4)} + e^{i(\phi_1 + \phi_2 + \phi_3 + \phi_4)} \big] , \end{aligned} \tag{1.16}$$

and for the phase θ_i it is

$$\begin{aligned} \dot{\theta}_1 = \omega_i + \beta \sum_{2,3,4} \sqrt{\frac{I_2 I_3 I_4}{I_1}} W_{1234} \Re \big[-e^{i(\phi_1 - \phi_2 - \phi_3 - \phi_4)} + \\ + 3e^{i(\phi_1 - \phi_2 - \phi_3 + \phi_4)} - 3e^{i(\phi_1 - \phi_2 + \phi_3 + \phi_4)} + e^{i(\phi_1 + \phi_2 + \phi_3 + \phi_4)} \big] , \end{aligned} \tag{1.17}$$

with $\Re$ and $\Im$ the real and imaginary part operators. When the nonlinearity is weak, the angle coordinates ϕ_j that appear in the exponentials in eq. (1.16) can be approximated with $\phi_j \simeq \omega_j t$, and the evolution of the amplitudes I_j happens on a timescale that is much larger than the linear timescale ω_j: consequently, the terms in the sum oscillate as $e^{i(\omega_1 \pm \omega_2 \pm \omega_3 \pm \omega_4)t}$. It is evident that most of these terms oscillate quickly, leading to a large cancellation of the overall effect on the evolution of I_1. Consequently, the amplitude I_1 remains essentially constant. On longer timescales, the terms that effectively contribute in eq. (1.16) are those whose argument of the exponentials is zero, or very close to zero, that is, the amplitudes change due to quadruplets of modes such that

$$\Delta\omega = \omega_1 \pm \omega_2 \pm \omega_3 \pm \omega_4 \tag{1.18}$$

is zero, or close to zero (and obviously the overlap sum W_{1234} is not exponentially small). The parameter $\Delta\omega$, quantifying the detuning of a resonance, is essentially a beat frequency. The plus-minus signs account for the different possible resonances in eq. (1.16). These quadruplets are called resonances when the sum of the frequencies is zero, or quasiresonances if the sum is very close to zero. The different signs in the sums of frequency are linked to distinct resonant processes, that is the resonance linked with the term $\Delta\omega = \omega_1 + \omega_2 + \omega_3 + \omega_4$ is a $4 \to 0$ process (four-to-zero), while the term $\Delta\omega = \omega_1 + \omega_2 - \omega_3 - \omega_4$ is a $2 \to 2$ process, etc.

1.2.3 The random phase approximation

In textbook WT derivation (see [29], and [30]) the evolution equation for the wave spectrum,

$$n_j = \langle a_j a_j^* \rangle, \tag{1.19}$$

is obtained. The average is intended over an ensemble of realizations of the same system, with similar total energy. The variable n_j is also called the number of particles of a mode j. The evolution equation for n_j is called the kinetic equation.

It turns out that in order to calculate the particle number, which is a second order correlator, one has to find the properties of higher-order correlators. While formal WT is developed only for continuous systems, we will show how these higher order correlators appear in the equations using our example model. In eq. (1.14): multiply the sides for a_1^*, then sum to it its complex conjugate, to obtain

$$\dot{n}_1 = 2\beta\, \Im \sum_{2,3,4} W_{1234} \left\langle -a_1 a_2 a_3 a_4 + 2a_1^* a_2 a_3 a_4 - 3a_1^* a_2^* a_3 a_4 \right\rangle . \tag{1.20}$$

An analytical solution of eq. (1.20) can in principle be obtained by calculating the evolution in time of the fourth order correlators. We attempt a perturbative approach in the first order. To remain in a first order perturbation theory of n_j, one calculates the evolution of $\dot{n}_j$ as an integral of the first derivative. However, it can be easily checked that this results in higher order correlators appearing in the system, which in turn depends on even higher order correlators. This is generally called a BBGKY hierarchy [31]. For example, let us take the fourth-order correlator in eq. (1.20) that corresponds to $2 \to 2$ interactions,

$$\begin{aligned} \frac{\partial \langle a_1 a_2 a_3^* a_4^* \rangle}{\partial t} &= \langle \dot{a}_1 a_2 a_3^* a_4^* \rangle + \langle a_1 \dot{a}_2 a_3^* a_4^* \rangle + \ldots = \\ &- (\omega_1 + \omega_2 - \omega_3 - \omega_4) \langle a_1 a_2 a_3^* a_4^* \rangle - 3\beta W_{1567} \langle a_2 a_5 a_6 a_3^* a_4^* a_7^* \rangle + \ldots \end{aligned} \tag{1.21}$$

In eq. (1.21) we show only one of the many six order correlators that appear [32].

WT provides a closure to the BBGKY hierarchy at the first order. This closure is the random phase approximation, that is we assume that the phase of

every mode is randomly distributed between $[0, 2\pi)$, and that there are no strong correlations between different modes. As already introduced in Section 1.1.2, this is a reasonable assumption as long as the nonlinearity is small. Since phases are random, only correlators that do not depend on the phase are non-zero or, in other words, where it is possible to match every normal mode to its conjugate, e.g.

$$\begin{aligned} \langle a_j \rangle &= \left\langle |a_j| e^{i\phi_j} \right\rangle = 0 \\ \langle a_j a_k \rangle &= \left\langle |a_j a_k| e^{i(\phi_j + \phi_k)} \right\rangle = 0 \\ \langle a_j a_j^* \rangle &= \langle |a_j^2| \rangle = n_j. \end{aligned} \tag{1.22}$$

Following this assumption, higher-order correlators can be decomposed into pairwise matches of a mode and its conjugate. Consequently, only correlators with an equal number of a_j and a_j^* are non-zero, and they can be expressed as a function of the number of particles. This procedure is also known as Wick's decomposition rule.

1.2.4 The kinetic equation and its properties

As we mentioned, the kinetic equation, with its properties that we are going to show in this Section, cannot be derived for a discrete system. A thermodynamic limit is also meaningless for the disordered Klein-Gordon lattice, as we will argument in Chapter 5. We assume that details of the dynamics of discrete systems with certain order of wave-interactions can be obtained from the corresponding (same order) WT kinetic equation [3, 33, 4, 5, 34, 35]. In the case of the disordered Klein-Gordon lattice, we will argue that this is a four-wave system where only the $2 \to 2$ interactions are active. The corresponding kinetic equation reads [26, 29]

$$\dot{n}_1 = \beta^2 \int |W_{1234}|^2 \, \delta(\omega_1 + \omega_2 - \omega_3 - \omega_4) f(n_1, n_2, n_3, n_4) \mathrm{d}k_1 k_2 k_3 k_4, \tag{1.23}$$

where $\delta(x)$ is the Dirac's delta, and we omitted the expression of $f(n_1, n_2, n_3, n_4)$ for clarity since it is not important to the discussion.

In the kinetic equation, the condition of resonance is embedded in the argument to the Dirac's delta. A reader familiar with WT will notice the absence of the usual selection rule on the wave numbers k. This should come to no surprise, because such condition comes from the fact that very often WT deals with systems of plane waves, and that is why this condition arises. However, in systems where the potential term in the Hamiltonian is not homogeneous in space, such condition does not arise [32].

The resonance condition arises from an integration in time of the closure for the high-order correlators, see eq. (1.21). We see in fact that the beat frequency

$\Delta\omega$ appears explicitly in eq. (1.21). The kinetic equation is valid for times much longer than the than the linear timescale,

$$\tau_{\text{lin}} \sim \frac{2\pi}{\omega}, \tag{1.24}$$

because the fast oscillations of the non-resonant terms ($\Delta\omega \neq 0$) are cancelled out only if one considers a timescale larger than the linear timescale. The prefactor β^2 arises from the fact that WT is a second order perturbation theory in H_{nlin} [26]. For some systems a second-order perturbation theory is not enough to explain the dynamics of the chain, which means that higher-order wave interaction must be considered. We will discuss this case in Chapter 2 and 3.

The kinetic equation implicitly defines a number of important statistical quantities. In particular, it can be shown that the total energy of the linear system is conserved,

$$E = \int \omega_1 n_1 \mathrm{d}k_1, \tag{1.25}$$

where the integral is intended over all the modes of a continuous system (or a continuous approximation of a discrete system). Additionally, for systems where only four-wave interactions of the type $2 \to 2$ are present ($\Delta\omega = \omega_1 + \omega_2 - \omega_3 - \omega_4$), it can be shown that the number of particles is conserved as well,

$$M = \int n_1 \mathrm{d}k_1. \tag{1.26}$$

It is important to note that the quantities E and M are conserved only in a statistical sense, and in particular they can, and do, show fluctuations in actual computations. The total linear energy E has its deterministic counterpart in the total energy of the original Hamiltonian system, but notably the total number of particles M may not have a deterministic counterpart. We will show later how this fact can be used to show that in the disordered KG lattice the effective mechanism for equipartition is indeed four-wave resonances of the type $2 \to 2$, because M is conserved.

Another very important aspect of the kinetic equation is that it implicitly defines an entropy function,

$$S = \int \log(n_1) \mathrm{d}k_1, \tag{1.27}$$

The entropy (1.27) satisfies

$$\frac{\mathrm{d}S}{\mathrm{d}t} \geq 0, \tag{1.28}$$

where the equality holds only at equilibrium. The equilibrium is found to be a Rayleigh-Jeans distribution

$$n_j|_{\text{equilibrium}} = \frac{T}{\omega_j + \mu}, \tag{1.29}$$

where μ and T are constants linked to the conservation of M and E respectively ($\mu = 0$ if M is not conserved), and they are called chemical potential and temperature. Equations (1.28) and (1.29) contain an important statement: they tell us that the system of waves does reach equilibrium, that is in principle there is no threshold to activate the spreading of energy across modes, in contrast to KAM theory. This should be of no surprise because KAM theory is expected to work in system with very few degrees of freedom, while textbook WT requires a large number of modes active in the system.

1.2.5 Measuring the equipartition time $T_{\mathbf{eq}}$

The last point of interest of the kinetic equation in our study is that it is possible to extract a timescale of the route to equilibrium from the kinetic equation. The kinetic equation in fact defines a scale of the rate of change of the number of particle n_j. This scale must be inversely proportional to the timescale to equipartition, that is the evolution time from an out-of-equilibrium initial condition, to the Rayleigh-Jeans equilibrium distribution eq. (1.29). In eq. (1.23) the kinetic equation shows a prefactor β^2, that controls the overall speed of the evolution of the amplitudes. The timescale to equipartition should then be proportional to it,

$$T_{\text{eq}} \sim \beta^{-2}. \tag{1.30}$$

Note that some studies tuned the nonlinearity strength by setting the total energy rather than β, but the two approaches are equivalent because a scaling exist between the two [34]. It is also easy to see that this quadratic scaling appears also in the discrete equations (1.20) and (1.21) of our example system. With some specific variations for each model, this is the core idea stated in [3, 33, 4, 5, 34, 35]. The specific differences are due to the lattice model details, but they can be encompassed in this idea.

It is very important to remark the following point that an alert reader may have noticed. The interaction coefficient between modes in resonance is present from the Hamiltonian equation to the final kinetic equation. Yet, we claim to have an unique thermalization time that does not depend on which mode we are observing in its route to equipartition. This is a bold statement because, as we mentioned above, in most literature of the FPUT-like models, there is a clear swing in conductivity between different modes. In practice, in writing eq. (1.30), we implicitly assume that the modes have roughly the same connectivity (borrowing the term from graph theory), that is the sum of the squared interaction coefficients $|W_{1234}|^2$ for all the resonances that a mode is part of, is roughly the same for all modes. Otherwise, the dynamics to thermalization could be dominated by some poorly connected modes, or by the fact that the dispersion relation shows distinct and far away branches, as in the case of the coexistence of an acoustic and optic branch [36]. This aspect is commented in the following section.

1.3 The choice of initial conditions

As we explained in Section 1.1.3, it is very important in the context of FPUT thermalization to clarify what kind of initial conditions we consider. The necessary concepts have been introduced, so it is possible now to clarify on this point.

As we mentioned in the previous section, we do not look for a characterization of the equipartition time as a function of the initial conditions. We are trying to establish that the methods of WT can work on nonlinear lattices to predict the scaling of the equipartition time, at least in a semi-quantitative way. If we choose an initial condition that favours initially a specific class of modes (such as high-frequency or low-frequency, or spatially localized), then there would be a difficult-to-estimate transient time where the dynamics are dominated by the specific interaction coefficients linked to the initially excited modes. A similar effect is the sandpiling behavior (see for example [37]), which is a condition where transmission of energies is burst-like because of bottlenecks in spectrum space. It is an interesting phenomenon, but it is likely to disturb any experiment where energy is injected in a specific class of modes. Our aim is to understand whether the global, long-time dynamics of the system can be explained with the methodology proposed so far, for a large class of nonlinear lattices. Correlations should then be avoided in the initial conditions that we choose.

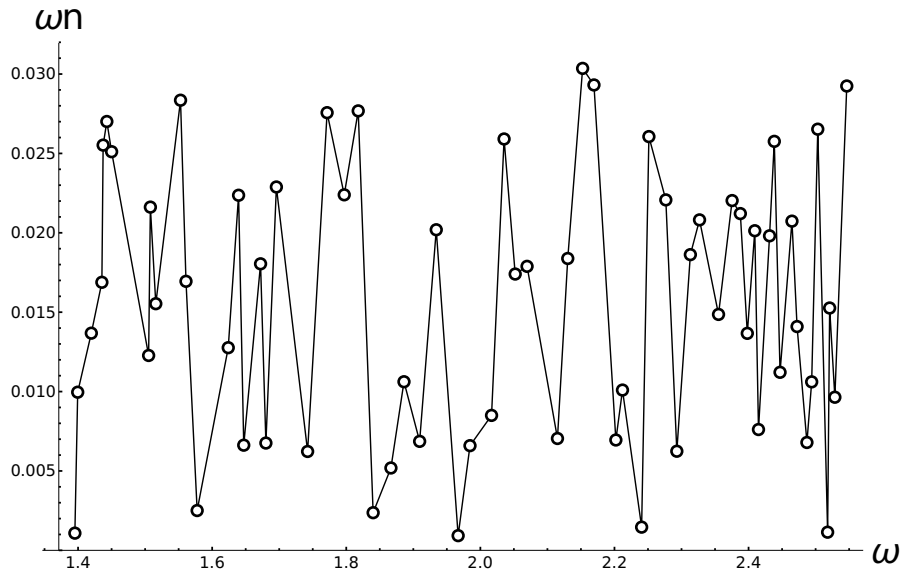


Figure 1.1: An example of an initial condition of the energies $\omega_j n_j$ as a function of the normal mode frequencies ω_j, for a lattice with N, from an actual simulation shown later in this thesis (disordered nonlinear Klein-Gordon lattice with $m_{\max} = 3$ and $\mu = 0$). The total sum of the energy is $E = 1$, and the expected value for the energy at thermalization is $E/N = 0.016$.

Since we want to measure the wave spectrum n_j, ensemble averages with random phases should be taken. This is also useful in avoiding to rely on initial conditions with arbitrary correlations. The only constraints in designing the initial conditions are that the total linear energy $E = H_{\text{lin}}$ must be the same throughout the ensemble, and obviously that the system should be out of equilibrium, that is the initial energies are different than E/N. A set of N random numbers are drawn from an uniform distribution, then they are rescaled such that the sum equals a specific, arbitrary energy value E. These values are the initial energy of each mode in the model. Each realization in the ensemble will differ in the initial phase of the modes, which is chosen uniformly between $[0, 2\pi)$. In Figure 1.1 a typical initial condition for the energies $n\omega$ is show.

With this scheme, we are observing the route to equipartition from a rather uncorrelated initial condition, which is however still out of equilibrium. Because of the absence of particular initial symmetries in physical and normal mode space, and also due to the large ensemble size employed, our numerical experiments are not expected to be "brittle" to the change of details in initial conditions.

There is one obviously correlation in the initial condition, that is all the realizations in the ensemble have the exact same profile of energies. A less correlated condition would be to have for all modes a probabilistic distribution of energies *across* the realizations of the chain. While this is a valid point, it is difficult to integrate with the constraint of the same initial energy E across the ensemble. We have adopted the previously explained schema for simplicity. We have tested across all the models presented in this thesis that the energy of a mode across the realizations quickly decorrelates across the ensemble, much earlier than equipartition. The energy distribution of a single mode can be fitted very well to an exponential distribution, a result that is expected [29]. This was observed however only as a confirmation that the realizations decorrelate, and no further probing was done.

1.4 The measure of equipartition

Since we want to observe the route to the relaxation distribution of eq. (1.29), we will directly inspect the entropy (1.27). In Chapter 5 we will show some experiments where the chemical potential μ is set to values that are non-zero. This means that the final state at relaxation is not a strict equipartition of energy $e_j = \omega_j n_j$. The methodology involved in this thesis cannot answer if and how complete equipartition is attained in these cases at much larger timescales. In this sense, there could be situations in which we actually observe only a metastable state. This is a subject that surely deserves further research. The initial experiments on the FPUT models with just a few low or high-frequency initial excited models are extreme examples of this category.

In all the experiments aimed to determine the equipartition time, the system

is let evolve until a specific threshold of the entropy is crossed. This threshold is chosen by first inspecting a simulation run for long enough to observe a plateau. The plateau depends in fact on the size of the ensemble, because of the unavoidable random fluctuations of macroscopic functions for finite ensembles, but it is kept fixed for a given initial condition and parametrization of the linear part of the Hamiltonian. In this way it is possible to run a number of simulations with a different value for ϵ, and since $\epsilon \propto H_{\text{nlin}}$, then it is also proportional to the nonlinear control parameter.

This approach is limited to the case of nonlinear lattices where there is a single control parameter of the nonlinearity, and that is proportional to ϵ. In the case for example of the $\alpha - \beta$-FPUT system, wave processes of multiple orders are present. We have not tried to apply this methodology to these case. The most likely scenario is the coexistence of different regimes linked to the two kinds of wave processes, with a cross-over in the relative importance of the different nonlinear terms. This is also an area that deserves more attention.

Chapter 2

Thermalization in the discrete nonlinear Klein-Gordon chain in the wave-turbulence framework

A commonly studied nonlinear lattice in the class of FPUT-like problems is the Klein-Gordon (KG) model [38, 39],

$$H = \sum_{j}^{N} \frac{1}{2} p_j^2 + \frac{1}{2} \left(q_{j+1} - q_j \right)^2 + \frac{1}{2} m q_j^2 + \frac{1}{4} \beta q_j^4, \tag{2.1}$$

with $m \geq 0$. The KG lattice is similar yet different in some details to the original FPUT problem, hence it can be used to check the robustness of the arguments introduced in the previous Chapter and introduced in [3]. This is the reason for which we considered this model in [4], alongside the β-FPUT model which is treated in [33]. We will comment on some of the results of the paper, in order to put them in the broader context of successive research.

The first important difference in eq. (2.1) compared to the FPUT model is the presence of a new parameter m which appears in the quadratic part of the Hamiltonian. Consequently, the dispersion relation of the linearized dynamics,

$$\omega_k = \sqrt{m + 4 \sin\left(\pi \frac{k}{N} \right)} \tag{2.2}$$

is different from the class of FPUT models. The parameter m is often called the mass parameter, but as a matter of fact it acts as an on-site potential. The principal difference from the FPUT dispersion relation is that it does not go to zero for $k = 0$. This means that the zeroth mode participates in the dynamics, in contrast to the FPUT models where a solid translation of the chain is a constant of motion uncoupled from the other degrees of freedom. This actually makes the KG

model easier to study with the tools of WT, because the linear timescale eq. (1.24) does not diverge for the low-frequency modes.

The nonlinearity is quartic, which makes the KG model more similar to the β-FPUT model. While the expression in real space is different, that is q_j^4 in contrast to $(q_j - q_{j-i})^4$, this difference is embedded in the interaction coefficient W_{1234} in normal mode space, but the order of wave interactions is the same.

In [3] it is found that the route to equipartition of the α-FPUT problem is determined by six-wave interactions, despite the fact that the α-FPUT model contains three-wave interactions. In the KG model, we found that despite the different dispersion relation the resonance manifold has very similar features. The first order of wave interactions that can explain the thermalization of the chain is again the sixth order. Low-order wave interactions can in fact be removed with a quasi-linear change of variables, at the price of obtaining higher-order interactions [32].

The principal novelty of our study on the KG lattice is that we considered the thermodynamic limit (N large), and the possible effects of quasiresonances, and we checked whether the outlined WT arguments were consistent. We studied the range of nonlinearity ϵ from $5 * 10^{-4}$ to $3 * 10^{-2}$ for N between 31 and 1024. In observing the thermalization time as a function of the nonlinearity ϵ, we observed that there is an evident cross-over between regimes at a threshold $\epsilon_c(N)$, that is dependent on N. For nonlinearities $\epsilon < \epsilon_c$, the behaviour is essentially the same as the α-FPUT model, in the sense that the wave-interaction order is six. This is because the resonant manifold is very similar to that of the FPUT models. For larger nonlinearity values $\epsilon > \epsilon_c$ the behavior appears consistent with four-wave dynamics.

As we mentioned, the resonances in the KG model are essentially the same of the FPUT. The wave interaction that appear in the KG Hamiltonian are of the fourth order, but they are ineffective to bring the system to equipartition, since they are all disconnected. This fact is already known, from results that showed that the truncated FPUT Hamiltonian can be cast in integrable form [14, 16]. Four-wave dynamics, since integrable, should not drive the system to equipartition.

Our proposed explanation is that at larger nonlinearities, the frequency broadening activates quasiresonances linked with four-wave dynamics. This is explained through Chirikov's criterion, that we introduced in Section 1.1.2. It should be noted that five-wave interactions are not relevant to the KG system: in fact, the evolution of the fourth-order correlators in eq.(1.21) depend on the sixth order correlators only. Hence, if new quasiresonances appear due to the broadening, they may bridge the gaps between the fourth order resonances, which drive a faster thermalization time. Fourth-order quasiresonances, when active, overwhelm sixth-order resonances because their interaction coefficient is much larger. This is because four-wave interactions are a lower order in the perturbative expansion operated in WT.

The fact that $\epsilon_c(N)$ is a decreasing function in N is consistent with the previous argument. In fact, in the thermodynamic limit $N \to \infty$ the frequency spacing is very small, according to eq. (2.2), and smaller nonlinear frequency broadening is required to activate quasiresonances. The frequency spacing also depends on m, being smaller at larger m, and a sign of this effect was also observed in numerics.

Unfortunately, at the time of the publication we did not have a clearer, possibly quantitative explanation of details of this proposed mechanism. We could only observe the broadening on an arbitrary beat frequency $\Delta\omega$ of a quasiresonance very close to exact resonance, $\Delta\omega \ll 1$. This result was however obtained with a technique that I found inadequate in later research (see Chapter 5 for an in-depth explanation applied to the disordered KG lattice). Similar results of two regimes of interaction were also observed in the β-FPUT model [33]. However, the large nonlinearity regime does not show in the β-FPUT problem the same four-wave dynamics, as it would be expected from the previous reasoning.

The result of this paper can be stated stated as follows. While at very low nonlinearities quasiresonances are scarse, and the dynamics are driven by exact resonances, for larger nonlinearities quasiresonances appear to take over due to the nonlinear broadening. The fact that the frequency broadening can activate new resonances is a well-known phenomenon [40]. In a way, this crossover has a correspondence with the sandpile behaviour observed in discrete turbulence [37], where the nonlinear broadening is localized in normal mode space. Our numerical evidence calls for further investigation of the role of quasiresonances in WT, because the dependence of the threshold $\epsilon_c(N)$ on the number of particles is consistent with the activation of quasi-resonances.

In the following pages, we propose the paper [4] where the results on the KG lattice were presented, and this commentary should be useful to understand the significance of the paper.

A Letters Journal Exploring the Frontiers of Physics

February 2018

EPL, **121** (2018) 44003
doi: 10.1209/0295-5075/121/44003

www.epljournal.org

Thermalization in the discrete nonlinear Klein-Gordon chain in the wave-turbulence framework

L. Pistone[1], M. Onorato[1,2] and S. Chibbaro[3]

[1] *Dipartimento di Fisica, Università degli Studi di Torino - Via P. Giuria, 1, Torino, 10125, Italy*
[2] *INFN, Sezione di Torino - Via P. Giuria, 1, Torino, 10125, Italy*
[3] *Sorbonne Université, CNRS, Institut Jean Le Rond d'Alembert - F-75005 Paris, France*

received 15 January 2018; accepted in final form 26 March 2018
published online 17 April 2018

PACS 47.52.+j – Chaos in fluid dynamics

Abstract – We study the time of equipartition, T_{eq}, of energy in the one-dimensional Discrete Nonlinear Klein-Gordon (DNKG) equation in the framework of the Wave Turbulence (WT) theory. We discuss the applicability of the WT theory and show how this approach can explain qualitatively the route to thermalization and the scaling of the equipartition time as a function of the nonlinear parameter ϵ, defined as the ratio between the nonlinear and linear part of the Hamiltonian. Two scaling laws, $T_{eq} \propto \epsilon^{-2}$ and $T_{eq} \propto \epsilon^{-4}$, for different degrees of nonlinearity are explained in terms of four-wave or six-wave processes in the WT theory. The predictions are verified with extensive numerical simulations varying the system size and the degree of nonlinearity.

Introduction. – The behavior of the small perturbation of an integrable system is one of the fundamental problems of mechanics [1]. In particular, an important question is whether a system reaches thermalization (*e.g.*, equal distribution of energy among the degrees of freedom) given some initial distribution of energy. In this framework, Fermi, Pasta and Ulam (FPU) with the help of Tsingou [2] performed the very first numerical simulations of a nonlinear chain, obtaining puzzling results and then opening up the way to the modern nonlinear physics.

Almost at the same time, the same problem was tackled from a mathematical point of view by the Kolmogorov school, obtaining the fundamental result of the KAM (Kolmogorov-Arnold-Moser) theorem which is at the basis of modern perturbation theories [3–6]. A remarkable work was performed some years later by Zabusky and Kruskal [7] who showed that in the continuous limit the FPU system reduces to the celebrated Korteweg-de Vries equation which is known to be integrable via the inverse scattering transform. Interestingly, at the same time of the work of Zabusky and Kruskal, important advances in the statistical mechanics theory of weakly nonlinear interacting waves were developed [8,9]: this theory is called Wave Turbulence (WT) and has been applied to many physical situation since then [10,11], see also [12] for the concept of resonant interactions in discrete systems. Nevertheless, the understanding of the problem of ergodicity in nonlinear chains has been tackled mostly in terms of chaos tools, other than direct numerical simulations, as can be seen in the good reviews [13,14] and in more recent important results [15,16]. Only more recently, the wave turbulence approach has been used in relation with one-dimensional anharmonic chains [17–19].

In this letter, we deal with the problem of the dynamics of the Discrete Nonlinear Klein-Gordon (DNKG) or ϕ^4 model from generic initial data, using the tools of wave turbulence. The ϕ^4 chain has been already considered in an original way to shed light on the problem of thermalization in [20]. From extensive numerical simulations, it appears that thermalization takes place for very small nonlinearity in agreement with our predictions: in different regimes $T_{eq}(\epsilon)$ displays a power-law dependence on ϵ.

The model. – The equation of motion of a ϕ^4 chain reads

$$\ddot{\phi}_i = (\phi_{i+1} + \phi_{i-1} - 2\phi_i) - m\phi_i - g\phi_i^3, \qquad (1)$$

where $\phi_i \in \mathbb{R}$ and $i = 1, 2, \ldots, N$; m and g are two positive constants, the latter controls the strength of the nonlinearity. The Hamiltonian associated to eq. (1) is

$$H = \sum_i \frac{1}{2}\pi_i^2 + \frac{1}{2}(\phi_{i+1} - \phi_i)^2 + \frac{1}{2}m\phi_i^2 + \frac{1}{4}g\phi_i^4, \qquad (2)$$

with $\pi_i \equiv \dot{\phi}_i$. In order to characterize the degree of nonlinearity of the system, we introduce the following parameter:

$$\epsilon \equiv H_{nl}/H_l, \tag{3}$$

where H_l accounts for first three terms in eq. (2) and H_{nl} for the last one. For a given initial condition, a change of g leads to a change of the degree of nonlinearity ϵ of the simulation ($\epsilon \propto g$). We mention that ϵ is not constant throughout the evolution of the system. Nevertheless, for the weakly nonlinear regimes considered in our simulations, it will be shown that it is essentially equivalent for scaling considerations to use as a controlling parameter of the nonlinearity either g, or ϵ computed at the equipartition time, or ϵ computed at the initial conditions (that is, $\epsilon(t=0) \simeq \epsilon(t=T_{eq})$).

The linear version ($g = 0$) of eq. (2) can be diagonalized in the normal modes. Assuming periodic boundary conditions, we introduce the discrete Fourier transform of the ϕ_i and π_i,

$$\hat{\phi}_k \equiv \sum_{j=1}^{N} \phi_j e^{-i2\pi kj/N}, \quad \hat{\pi}_k \equiv \sum_{j=1}^{N} \pi_j e^{-i2\pi kj/N} \tag{4}$$

(note $\hat{\phi}_k^* = \hat{\phi}_{-k}$ and $\hat{\pi}_k^* = \hat{\pi}_{-k}$), and then normal modes are given by

$$a_k \equiv \left(\omega_k \hat{\phi}_k + i\hat{\pi}_k\right)/\sqrt{2\omega_k}, \tag{5}$$

where ω_k is the positive branch of the linear dispersion relation,

$$\omega_k \equiv \sqrt{m + 4\sin\left(k\pi/N\right)^2}. \tag{6}$$

In normal variables, the equation of motion reads

$$\begin{aligned} i\dot{a}_{k_1} = \omega_{k_1} a_{k_1} + g \sum_{k_2,k_3,k_4} V_{k_1k_2k_3k_4} (a_{k_2} a_{k_3} a_{k_4} \delta_1^{234} \\ + 3a_{k_2}^* a_{k_3} a_{k_4} \delta_{12}^{34} + 3a_{k_2}^* a_{k_3}^* a_{k_4} \delta_{123}^{4} + a_{k_2}^* a_{k_3}^* a_{k_4}^* \delta_{1234}), \end{aligned} \tag{7}$$

with the interaction coefficient $V_{1234} = 1/(4\sqrt{\omega_1\omega_2\omega_3\omega_4})$ and δ is the generalized Kronecker's delta,

$$\delta_{ab\ldots}^{\alpha\beta\ldots} \equiv \begin{cases} 1, & k_a + k_b + \ldots = k_\alpha + k_\beta + \ldots \mod (N), \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Equation (7) describes a wave dynamics with a third-order nonlinearity [10,21].

The large-box limit and the weak WT theory. – In the framework of WT [8,10,21,22] one considers the evolution of a statistical ensemble of realisations, each characterized by random phases and amplitudes, in the limit of a large-box ($N \to \infty$) and small nonlinearity (small wave amplitude).

In the standard formulation of the WT, the physical space is considered continuous while our system will be considered always discrete (in physical space). Here, we consider the large-box limit of the system (1), assuming that, as $N \to \infty$, also the length of the chain, L, goes to infinity with their ratio constant, $\Delta x = L/N = 1$; this implies that the Fourier space becomes dense and the wave numbers are not discrete anymore, yet the dynamics in the physical space is intrinsically discrete, and therefore $k_{\max} = \pi$. In such limit, it is possible to derive the evolution for the observable $\langle a_{k_i} a_{k_j}^* \rangle = n_i \delta_i^j$, that is a *kinetic equation* [18,19,23],

$$\begin{aligned} \dot{n}_1 = g^2 \int_{-\pi}^{\pi} \delta_{2\pi}(k_1 + k_2 - k_3 - k_4)\delta(\omega_1 + \omega_2 - \omega_3 - \omega_4) \\ \times \left(V_{1234}\right)^2 n_1 n_2 n_3 n_4 \left(\frac{1}{n_1} + \frac{1}{n_2} - \frac{1}{n_3} - \frac{1}{n_4}\right) \mathrm{d}k_2 \mathrm{d}k_3 \mathrm{d}k_4, \end{aligned} \tag{9}$$

where δ is now the Dirac δ-function, $\delta_{2\pi}(k) \equiv \delta(2\pi j + k)$ with $j \in \mathbb{Z}$, and $\omega_k = \sqrt{m + 4\sin(k/2)^2}$, with now $k \in [-\pi, \pi)$. It is interesting to note that the time scale of the evolution of the spectrum via the collision integral (9) is proportional to $1/g^2$. The δ-functions in eq. (9) define the conditions under which the collision integral does not vanish; these *resonance conditions* are in the following form:

$$k_1 + k_2 - k_3 - k_4 = 0, \mod 2\pi, \quad \omega_1 + \omega_2 - \omega_3 - \omega_4 = 0. \tag{10}$$

As can be shown numerically, nontrivial solutions of the above equations for $k \in [-\pi, \pi)$ and $\omega_k = \sqrt{m + 4\sin(k/2)^2}$ can be found for each value of m, while other scattering processes are forbidden because of the shape of the dispersion relation (6). For this reason, only $2 \to 2$ processes have been explicitly included in eq. (9).

The importance of eq. (9) relies on the fact that it is possible to define an entropy function,

$$S = \int_{-\pi}^{\pi} \log(n_k) \mathrm{d}k, \tag{11}$$

such that $\mathrm{d}S/\mathrm{d}t \geq 0$, *i.e.*, the resonant interactions lead to irreversible dynamics. Moreover, the entropy is maximized by the Rayleigh-Jeans (RJ) distribution,

$$n_k = \frac{T}{\omega_k + \mu}, \tag{12}$$

where T and μ are constants (that can be calculated from initial conditions) associated to the conserved quantities of the kinetic equation: energy and number of particles. Essentially, the prediction in the large-box limit is that n_k relaxes to the RJ distribution, eq. (12), and if $\mu \sim 0$ then the energy per mode $e_k \equiv n_k \omega_k = \text{const}$. Note that with respect to continuous wave turbulence, because of the presence of periodic Dirac delta functions the total momentum is not conserved.

The picture outlined above is valid in the limit of infinitely many modes and small nonlinearity. We expect

that in regimes close to this limit, the predictions should apply with some degree of accuracy. In any numerical simulation, the number of modes is always finite, therefore, the resonances in principle may take place only for integer values of wave numbers. However, in nonlinear dispersive wave systems, another effect comes into play, that is the broadening of frequencies [24]: the frequency of the modes becomes stochastic around the value described by the dispersion relation. The implication of this phenomenon is that if N or ϵ are sufficiently large, the resonance conditions, eqs. (10), do not need to be satisfied exactly in the computational grid and *quasi-resonances* may become important [10]. On the other hand, in the weakly nonlinear regime and when the number of modes is low, we assume that exact resonant interactions in a discrete system may lead on average to an irreversible process just like in the large-box limit, even though a statistical description, *i.e.*, a kinetic equation, with discrete wave numbers has not been developed. This case is considered hereafter.

Exact resonances in the DNKG for small number of masses. –

Four-wave exact resonances. We now consider the finite-size system and study whether eqs. (10) can hold for the dispersion relation eq. (6). We will indicate the l.h.s. of the equation for the resonance condition on frequencies for a process $X \to Y$ as $\Delta\omega_X^Y$. Processes of the kind $4 \to 0$ are obviously excluded because $\Delta\omega_X^0 > 0$ for $m > 0$. Processes of the type $3 \to 1$ can be decomposed into the sum of two $2 \to 1$ processes, $\omega_1 + \omega_2 - \omega_{1+2} + (\omega_{1+2} + \omega_3 - \omega_{1+2+3}) = 0$. For $m = 0$ it can be shown that $\Delta\omega_{12}^{1+2} > 0$ (except for $k = 0$, but the mode a_0 is not well defined for $m = 0$ and it is effectively excluded from the dynamics): furthermore, assuming $\Delta\omega_{12}^{1+2} = 0$, then $\partial\Delta\omega_{12}^{1+2}/\partial m > 0$, and, therefore, for continuity in m (for $m > 0$) of both $\Delta\omega_{12}^{1+2}$ and $\partial\Delta\omega_{12}^{1+2}/\partial m$ no new solution can appear by increasing m. Consequently, no processes of the type $3 \to 1$ are present in the system. Let us consider now the resonance condition on frequencies for a process $2 \to 2$, $\Delta\omega_{12}^{34}$. After successively squaring the equation and renaming $s_i \equiv \sin(k_i\pi/N)$ one gets to

$$m \prod \left(s_1^2 \pm s_2^2 \pm s_3^2 \pm s_4^2\right) = \frac{1}{2} \prod \left(s_1 \pm s_2 \pm s_3 \pm s_4\right), \tag{13}$$

where the first product is intended between all combinations of the plus-minus signs with a total of two minus signs (*e.g.*, $s_1^2 + s_2^2 - s_3^2 - s_4^2$), while the second product is intended between all possible combinations of plus-minus signs without any constraint. Equation (13) can be used to find solutions for $\Delta\omega_{12}^{34} = 0$ both dependent on and independent of m. The independent solutions can be found by requiring that the product on the left, that is the coefficient of m, is zero, and then inserting back the resulting solutions in the definition of $\Delta\omega_{12}^{34}$ with $m = 0$. Using $k_4 = k_1 + k_2 - k_3 + jN$ with $j \in \mathbb{Z}$ arbitrary, after some algebra, one finds that the only solutions are the trivial process $\{k_1, k_2, k_1, k_2\}$, and the so-called Umklapp scattering process

$$\{k_1, k_2, -k_1, -k_2\} \quad \text{with} \quad k_1 + k_2 = N/2, \tag{14}$$

which is only possible when N is even; such solutions are related to the periodicity of the dispersion relation and are the same for the standard α and β-FPU chain [17]. The trivial processes do not contribute to the dynamics but for a nonlinear correction to the frequencies, [10]. The Umklapp scattering process, on the other hand, does contribute to the dynamics, but since all quartets are disconnected (*i.e.*, there is no common mode between the quartets), their dynamics cannot lead the system to equipartition.

Six-wave exact resonant interactions. Since the only m-independent four-waves resonances permitted are due to the Umklapp scattering mechanism, and they are not connected, it is necessary to perform a quasi-linear change of variables and remove the nonresonant four-waves terms in eq. (7) and obtain a Hamiltonian with higher-order wave-wave interactions to predict the evolution of the system at large times (see, for example, [8]). This change of variables is canonical, up to the order in consideration, and in our case the transformation reads

$$\begin{aligned} a_1 \;=\; b_1 - g \sum_{234} V_{1234} \big(& 3b_2^* b_3^* b_4 / \Delta\omega_{123}^4 + 3b_2^* b_3 b_4 / \Delta\omega_{12}^{34} \\ & + b_2 b_3 b_4 / \Delta\omega_1^{234} + b_2^* b_3^* b_4^* / \Delta\omega^{1234} \big) + (\text{h.o.t}), \end{aligned} \tag{15}$$

where h.o.t means higher-order terms. Due to this transformation the equation of motion in the new variables b_k gains new terms corresponding to six-wave interactions, and after examination they turn out to be all proportional to factors of the type $g^2 V_{1234} V_{5678}$. Note also that, for interactions of the type $2 \to 2$, the sum is made only over nonresonant terms because otherwise such transformation would be singular. In the new variables it would be formally possible to derive a six-wave kinetic equation [10], and obtain resonance conditions analogous to eqs. (10), but with six wave numbers and frequencies. However, in doing so one would need to take again the thermodynamic limit first, in which, as explained later, the m-dependent exact four-wave interactions would dominate the dynamics. In the following we conjecture that, in practice, there exists a regime with finite N and small nonlinearity ϵ such that six-wave resonance dynamics are dominant, even though, as far as we are aware, a WT formalism with finite N has not been developed yet.

Six-waves resonances. As for the $4 \to 0$ and $3 \to 1$ cases, also $6 \to 0$ and $5 \to 1$ resonances are not allowed. Resonances of the type $3 \to 3$ can be trivially constructed by analogy from eq. (14) using an arbitrary wave number k_3,

$$\{k_1, k_2, k_3, -k_1, -k_2, -k_3\},\ k_1 + k_2 + k_3 = 0 \mod (N). \tag{16}$$

These resonant sextuplets are interconnected, and they provide a way to exchange energy among all the modes

of the system. Note that eq. (16) is valid also for odd values N. For even N, eq. (16) is also valid for $k_1+k_2+k_3 = 0$ mod $(N/2)$, and finally it is also possible to construct again from eq. (14) these additional resonances,

$$\{k_1, k_2, k_3, -k_1, -k_2, k_3\}, \quad k_1 + k_2 = N/2. \tag{17}$$

As in the case of four-waves interactions, there also exist the trivial resonances $\{k_1, k_2, k_3, k_1, k_2, k_3\}$ but they only contribute to the nonlinear frequency shift.

The previous analysis of the resonances in the KG system is limited to resonances that are independent of the value of m, *i.e.*, they are valid for any value of m. One may ask if, for specific values of m, new resonances appear. Such analysis if far more difficult. For the four-waves case, it is possible to solve for m in eq. (13), and then to find values of m that connect arbitrary wave numbers. These solutions however appear to have little significance in the dynamics: in fact, we could not find numerically for $N = 32$ and $N = 64$ cases in which for a specific value of m there are more than two connected quartets. For the six-waves resonances it is possible, just as in the case of four-waves resonances, to successively square $\Delta\omega^{456}_{123}$ and obtain an equation for m. However, this turns out to be of the sixth order in m. Furthermore, it is possible to show that resonances of the type $4 \to 2$ can only exist if $m < 4/3$, in particular we have obtained the relation

$$m \leq \frac{4Y^2}{X^2 - Y^2} \tag{18}$$

which leads to $m < 4/3$ for a $4 \to 2$ resonance.

No special solutions of the exact resonances conditions for the four- and six-waves resonances were numerically found for the chosen value of $m = 1$ with N finite, and only the resonances of the Umklapp type were satisfied exactly in our simulations. Numerical simulations of the equation of motion show that, for the specific values of m for which new resonances appear, there is no appreciable effect on the dynamics, as will be shown shortly.

Estimation of the scaling of $T_{eq}(\epsilon)$. – From the previous discussion we understand the following. A WT formalism can be developed from the equation of motion (7) in the weakly nonlinear regime and in the large-box limits. The latter implies that wave numbers are dense in the domain $[-\pi, \pi)$. In these limits the expected time scale of thermalization is $T_{eq} \sim \epsilon^{-2}$. When the number of modes is not sufficiently large for the system to be considered as in the thermodynamic limit, wave numbers are discrete; in such condition four-wave exact resonant interactions exist but cannot bring the system to equipartition because they are not interconnected. However, if the frequency gaps between the modes are sufficiently low (a condition attained by having either N or ϵ sufficiently large), we conjecture that the frequency broadening (Chirikov's argument) can allow "quasi-resonaces" effectively bridging between the disconnected Umklapp-type four-waves resonances in the system, and hence bringing the system to equipartition with typical times $T_{eq} \sim \epsilon^{-2}$, as per eq. (9). If both N and ϵ are small, the dynamics is intrinsically discrete both in physical and spectral space and we assume that in the weakly nonlinear regime discrete six-wave exact resonances lead, as in the continuous case, to an irreversible dynamics, although the WT has not been developed for such discrete case.

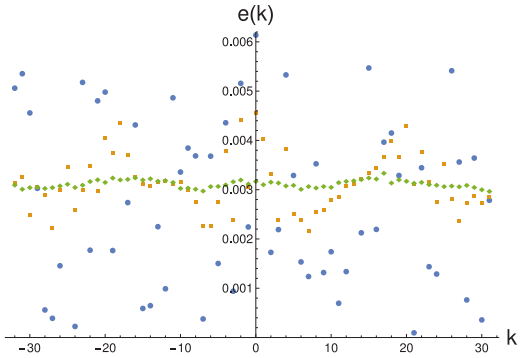


Fig. 1: (Colour online) A typical process of thermalization, for $N = 64$, $m = 1$, $E = 0.2$, $g = 5$ ($\epsilon \simeq 0.002$). The three datasets are $s(0) \simeq 14$ (●), $s(100000) \simeq 1$ (■), $s(2000000) \simeq 0$ (♦).

The leading-order nontrivial wave-wave interaction process displayed by the DNKG equation is the six-wave one, as it turns out from using the standard procedure of removing the nonresonant four-waves interactions through eq. (15). The time scale associated to the dynamical equation for the six-wave interactions is ϵ^{-2} and, even though there is no formal derivation of the wave kinetic equation, we expect, just as for the standard WT, that the spectrum evolves on a time scale which is given by the square of the coupling coefficient, *i.e.*, we expect to observe $T_{eq} \sim \epsilon^{-4}$ (see also [17]).

Numerical simulations. – In order to verify the predictions, a number of simulations have been run with different values of N and g. The numerical integration of eq. (1) has been implemented with a symplectic integrator of the sixth order [25], and the marching timestep was set to $\delta t = 0.1$, a value that has been found to conserve the Hamiltonian at least up to the sixth digit in all simulations. The number of realizations in the ensemble is 4096 for all the following simulations. The initial ensemble of eq. (1) consists of a fixed random choice of amplitudes of the normal modes, all with random and different phases for each realization. The scheme ensures that all the realizations have the same initial linear energy. The initial random amplitudes are also adjusted in order to cancel the chemical potential in eq. (12), in order to better observe the equipartition. The equations of motion are then integrated, and the approach to equipartition is monitored with the following indicator function [26]:

$$s(t) \equiv \sum_k f_k \log(f_k), \qquad f_k = \frac{N}{E} e_k(t), \tag{19}$$

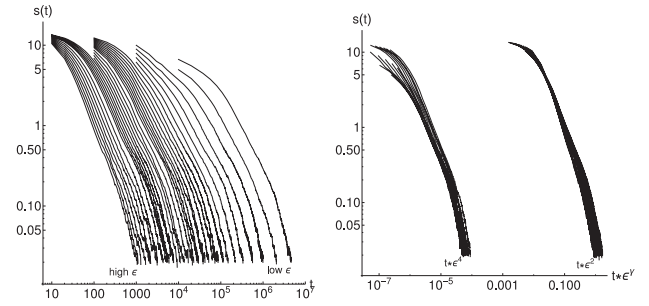


Fig. 2: Left: the entropy curves for the same parameters as fig. 1. The entropy curves are naturally ordered from left to right for decreasing ϵ, with values in the range 0.002–0.03 for $N = 64$. Right: the same curves with rescaled time: the two regimes $T_{eq} \propto \epsilon^{-2}$ and $T_{eq} \propto \epsilon^{-4}$ are highlighted by scaling the time either by ϵ^2 or by ϵ^4 (transition occurs at $\epsilon \simeq 0.007$).

with $e_k(t) = n_k(t)\omega_k$, the energy per mode, and $E = \sum e_k$. While the fundamental entropy function in WT is defined in eq. (11), it is common to refer to $s(t)$ just as "entropy", (because it is proportional to a Shannon entropy [14]). At equipartition, $s(t) = 0$, otherwise $s(t) > 0$. During evolution, $s(t)$ will decrease, and T_{eq} is defined as the time when $s(t)$ goes under some threshold value that depends in general on the number of realizations in the ensemble. The parameter ϵ is calculated *a posteriori*, that is inspecting the Hamiltonian at thermalization. A typical initial condition and its thermalization can be seen in fig. 1. The parameters are $N = 64$, $m = 1$, $E = 0.2$, $g = 5$ ($\epsilon \simeq 0.002$). The system goes to thermalization, although approximatively, as the size of the ensemble is finite. The difference between ϵ calculated from the initial and final (equipartition) states was found to be in the few percent points for the lowest values of ϵ, and larger (up to 20%) for large values of ϵ. In the left panel of fig. 2 we show the approach to equipartition for different values of $\epsilon = 0.002$–0.03 and $N = 64$. For large times the indicator function $s(t)$ settles and fluctuates to some value right below the chosen threshold (not shown here). The two predicted scaling laws can be highlighted by rescaling the time properly (fig. 2, right). We see that all the results now collapse into two separated curves with ϵ^{-2} and ϵ^{-4}, respectively. This indicates that the whole dynamics depends only on ϵ. Furthermore, two distinct regimes appear for ϵ lower and higher of a given critical ϵ_c. In fig. 3 we show $T_{eq}(\epsilon)$ in a log-log plot for $N = 64$ for the same simulations of fig. 2. The change of the dynamics between the ϵ^{-4} and ϵ^{-2} scalings is evident. In the same figure we also show the dependence of the thermalization dynamics on different values of $m = \{0.1, 0.5, 1\}$. There is no substantial difference between the runs, as the two different scalings are clearly visible, and the threshold of ϵ that determines the crossover between them is roughly the same for the cases $m = \{0.5, 1\}$, and slightly displaced to higher nonlinearity values for $m = 0.1$. The larger span of the steeper ϵ^{-4} scaling for the lower values of m is due to the larger average spacing between frequencies, and

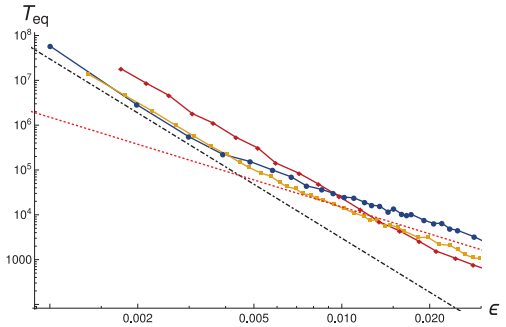


Fig. 3: (Colour online) The scaling of T_{eq} on ϵ for $N = 64$ and $m = 1$ (■), $m = 0.5$ (●) and $m = 0.1$ (♦). Scaling laws ϵ^{-2} and ϵ^{-4} in red dotted and black dash-dotted lines.

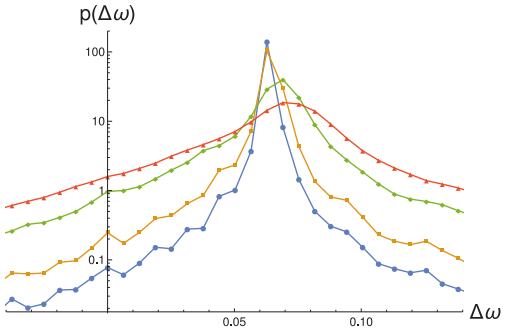


Fig. 4: (Colour online) The dispersion of the frequency mismatch $\Delta\omega$, for a $2 \to 2$ resonance with $N = 32$ and $k = \{1, -15, -11, -3\}$, calculated numerically with an accuracy of $\delta\omega = 2\pi/1000 \simeq 0.00628$. The product $a_1 a_{-15} a^*_{-11} a^*_{-3}$ is sampled in time, and the square amplitude of the Fourier coefficients of this time series is plotted after averaging and normalizing as a probability. $\epsilon \simeq 0.0026$ (●), 0.0052 (■), 0.0144 (♦), 0.023 (▲).

consequently the suppression of quasi-resonances. Indeed, as mentioned, the nonlinearity causes a broadening of the frequencies which, for a sufficiently strong nonlinearity, can lead to a frequency overlap phenomenon. In this case quasi-resonances can take place. This mechanism is qualitatively supported in fig. 4 for $N = 32$, where the probability density function of the mismatch frequency $\Delta\omega^{34}_{12}$ for a generic $k = \{1, -15, -11, -3\}$ (this is not an exact resonance) is computed from the simulations. If one considers frequencies from the very weakly nonlinear regime, one will obtain a function peaked at $\Delta^{34}_{12} \simeq 0.0626$, that is the waves are not in resonance ($\Delta^{34}_{12} \neq 0$); because of the nonlinearity, the value of Δ^{34}_{12} is actually stochastic, and a numerical evaluation of its probability distribution leads to the broad functions displayed in fig. 4. It is seen that with the increase of the nonlinearity over the range of the numerical studies reported in this letter, the probability of $\Delta^{34}_{12} \simeq 0$ increases by an order of magnitude, showing that four-waves quasi-resonances are not possible for small nonlinearities but become probable for larger nonlinearity.

We have then investigated the robustness of the predictions in the thermodynamic limit increasing N. The dependence of T_{eq} on N is shown in fig. 5. The transition

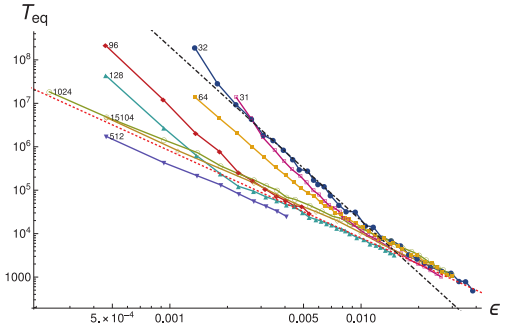


Fig. 5: (Colour online) The scaling of T_{eq} on ϵ for multiple values of N, with $m = 1$ and $E = 0.1N/32$. Scaling laws ϵ^{-2} and ϵ^{-4} are shown in red dotted and black dash-dotted lines.

between the two scalings is still well visible for the cases $N = 32$, $N = 64$ and $N = 96$, $N = 128$. As expected, there is no qualitative difference between even and odd N (see dataset with $N = 31$). For larger N the transition turns out to manifest always later until it disappears for $N \geq 512$, meaning that four-waves dominate the thermodynamic limit. It has been checked that no new dynamics appear even for a larger number of modes, up to $N = 15104$. The behavior in the thermodynamics limit is consistent with the WT analysis.

We conclude this section with a short discussion about the issue of averages in statistical mechanics [27,28]. While we have performed ensemble averages in order to evaluate accurately $T_{eq}(\epsilon)$ suppressing the entropy fluctuations, the behavior of the single realization appears always "typical", in the sense that eventually it fluctuates around the thermalized irreversible final state. We have indeed verified numerically that thermalization is observed if a suitable time average is performed over the energy per mode of a single system. It is remarkable that it is true even for a rather small number of degrees of freedom.

Conclusions and discussion. – In this letter we have studied the problem of the thermalization of the discrete nonlinear Klein-Gordon chain. In this framework, we understand the dynamics toward equipartition of energy through the mechanism of exchange of energy between resonant wave-wave interactions. Our prediction are the following: i) for generic weakly nonlinear initial conditions and a large enough number of modes (large-box limit), the system always thermalizes according to the WT theory (resonant four-wave interactions) on a time scale ϵ^{-2}; ii) when the number of modes and the nonlinearity are small, then exact resonances over a discrete regular grid of wave numbers become relevant. The six-wave interactions are dominant and the time scale for equipartition is ϵ^{-4} (this is the same scaling observed for the α-FPU, [17] and β-FPU [29]; iii) by increasing the nonlinearity in the discrete regime a crossover from ϵ^{-4} to a scaling ϵ^{-2} is observed. The transition takes place at different threshold values of ϵ depending on the number of modes; this is due to the phenomenon of frequency overlap. Interestingly, the scaling is the same as the one predicted in the large-box limit, implying that the frequency overlap has the role of reactivating exact four-wave resonances.

All the results have been verified by extensive numerical simulations, for different values of N, m and ϵ.

* * *

We thank M. Bustamante, A. Vulpiani, S. Caruso and B. Giulinico for discussions. MO would like to thank Y. Lvov with whom many of the ideas here discussed have been developed. Simulations were run on GPU hardware hosted at the OCCAM supercomputer, Università di Torino.

REFERENCES

[1] Poincaré H., *Les méthodes nouvelles de la mécanique céleste*, Vol. **1-VIII** (Gauthier-Villars et fils) 1899.
[2] Fermi E., Pasta J. and Ulam S., Technical Report I, Los Alamos Scientific Laboratory Report No. LA-1940 (1955).
[3] Kolmogorov A., *The general theory of dynamical systems and classical mechanics*, in *Proceedings of the International Congress of Mathematics, Amsterdam*, Vol. **1**, 1954, pp. 315–333.
[4] Arnol'd V. I., *Izv. Ross. Akad. Nauk. Ser. Mat.*, **25** (1961) 21.
[5] Arnold V. I., *Russ. Math. Surv.*, **18** (1963) 85.
[6] Möser J., *Nachr. Akad. Wiss. Göttingen*, **II** (1962) 1.
[7] Zabusky N. J. and Kruskal M. D., *Phys. Rev. Lett.*, **15** (1965) 240.
[8] Zakharov V. E., L'vov V. S. and Falkovich G., *Kolmogorov Spectra of Turbulence I: Wave Turbulence* (Springer Science & Business Media) 2012.
[9] Benney D. and Newell A., *Stud. Appl. Math.*, **46** (1967) 133.
[10] Nazarenko S., *Wave Turbulence*, *Lect. Notes Phys.*, Vol. **825** (Springer) 2011.
[11] Newell A. C. and Rumpf B., *Annu. Rev. Fluid Mech.*, **43** (2011) 59.
[12] Kartashova E., *Nonlinear Resonance Analysis* (Cambridge University Press) 2010.
[13] Zaslavsky G., *Chaos*, **15** (2005) 015103.
[14] Gallavotti G., *The Fermi-Pasta-Ulam Problem: A Status Report*, *Lect. Notes Phys.*, Vol. **728** (Springer) 2007.
[15] Benettin G. and Ponno A., *J. Stat. Phys.*, **144** (2011) 793.
[16] Benettin G., Christodoulidi H. and Ponno A., *J. Stat. Phys.*, **152** (2013) 195.
[17] Onorato M., Vozella L., Proment D. and Lvov Y. V., *Proc. Natl. Acad. Sci. U.S.A.*, **112** (2015) 4208.
[18] Spohn H., *J. Stat. Phys.*, **124** (2006) 1041.
[19] Lukkarinen J., *Kinetic theory of phonons in weakly anharmonic particle chains*, in *Thermal Transport in Low Dimensions* (Springer) 2016, pp. 159–214.
[20] Fucito F., Marchesoni F., Marinari E., Parisi G., Peliti L., Ruffo S. and Vulpiani A., *J. Phys. France*, **43** (1982) 707.

[21] Chibbaro S., Dematteis G. and Rondoni L., *Physica D: Nonlinear Phenom.*, **362** (2018) 24.
[22] Eyink G. L. and Shi Y.-K., *Physica D: Nonlinear Phenom.*, **241** (2012) 1487.
[23] Aoki K., Lukkarinen J. and Spohn H., *J. Stat. Phys.*, **124** (2006) 1105.
[24] Chirikov B. V., *Phys. Rep.*, **52** (1979) 263.
[25] Yoshida H., *Phys. Lett. A*, **150** (1990) 262.
[26] Izrailev F. and Chirikov B., *Sov. Phys. Dokl.*, **11** (1966) 30.
[27] Chibbaro S., Rondoni L. and Vulpiani A., *Reductionism, Emergence and Levels of Reality* (Springer) 2014.
[28] Cerino L., Cecconi F., Cencini M. and Vulpiani A., *Physica A: Stat. Mech. Appl.*, **442** (2016) 486.
[29] Lvov Y. V. and Onorato M., *Phys. Rev. Lett.*, **120** (2018) 144301.

Chapter 3

Universal route to thermalization in weakly-nonlinear one-dimensional chains

It has already been mentioned that our analysis of the thermalization dynamics of nonlinear lattices with WT arguments appears to be quite independent of the specific model details. Of course, different models have different types of wave interactions, and that causes a different exponent α in the power law for the thermalization time,

$$T_{\text{eq}} \sim \epsilon^{-\alpha}. \tag{3.1}$$

In the next paper that we published [5], we decided to propose this view in a review of three models, the α-FPUT, β-FPUT and KG lattice. Our objective was to show that a large class of lattices can be explained through WT concepts. Concurrently to our research, another group found a whole class of large N FPUT problems with different nonlinearities [34], and the perturbed Toda lattice [35] also can be understood in terms of WT dynamics.

We also investigated the thermodynamic limit of the FPUT models, which were previously unexplored in [3] and [33]. We found that in all the models considered, the leading dynamics when N is large is compatible with four-waves dynamics. The numerics show a quite good agreement with the predictions of WT. The least convincing dataset is that of the thermodynamic limit of the α-FPUT system, which appears to have a steeper power-law than the predicted $\alpha = -4$. Our attempt to explain this deviation at the time were quite inconsistent. A better understanding of this issue is surely provided by an analysis of the structure of quasiresonances in the α-FPUT model, however that was unavailable at the time.

It should be noted that we decided in this issue not to claim that the cross-over at larger nonlinearities observed in [4] is an universal trait. That is, we do not claim that in all lattice models for large enough nonlinearities a specific kind of quasiresonances is activated. Little is known in fact quantitatively on how these

quasiresonances arise, and how abundant they are. In the models studied in this thesis, the frequencies ω_j are mostly uniformly distributed from a minimum and a maximum value. For a larger number of modes (N large), this causes crowding in the spectrum, and so the dependence $\epsilon_c(N)$ found in [4] is explained. However, at fixed N, the appearance of quasiresonances must depend strongly on the details of nonlinearity, that is the interaction coefficient W_{1234} of the wave interactions, and the order of wave-interactions. Furthermore, W_{1234} is strongly dependent on the features of the eigenvalue problem (cp. eq. (1.13)). For this reason, we do not expect an universal behaviour across different FPUT-like models on the transition to lower-order wave interactions at larger nonlinearities.

http://www.aimspress.com/journal/mine

Mathematics in Engineering, 1(4): 672–698.
DOI:10.3934/mine.2019.4.672
Received: 12 February 2019
Accepted: 07 July 2019
Published: 08 August 2019

Research article

Universal route to thermalization in weakly-nonlinear one-dimensional chains†

Lorenzo Pistone[1], Sergio Chibbaro[2], Miguel D. Bustamante[3], Yuri V. Lvov[4] and Miguel Onorato[1,5,*]

[1] Dipartimento di Fisica, Università di Torino, via Pietro Giuria 1, 10125, Torino, Italy

[2] Sorbonne Université, CNRS, Institut Jean Le Rond D'Alembert, F-75005 Paris, France

[3] School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland

[4] Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180

[5] INFN, Sezione di Torino, Via P. Giuria, 1-Torino, 10125, Italy

† **This contribution is part of the Special Issue:** Hamiltonian Lattice Dynamics
Guest Editors: Simone Paleari; Tiziano Penati
Link: https://www.aimspress.com/newsinfo/1165.html

* **Correspondence:** Email: miguel.onorato@unito.it.

Abstract: Assuming that resonances play a major role in the transfer of energy among the Fourier modes, we apply the Wave Turbulence theory to describe the dynamics on nonlinear one-dimensional chains. We consider α and β Fermi-Pasta-Ulam-Tsingou (FPUT) systems, and the discrete nonlinear Klein-Gordon chain. We consider both the thermodynamic limit and the discrete regime and we conjecture that all the systems thermalize for large times, and that the equipartition time scales as a power-law of the strength of the nonlinearity, at least for a range of values of the nonlinear parameter. We perform state of the art numerical simulations and show that the results are mostly consistent with theoretical predictions. Some observed discrepancies are discussed. We suggest that the route to thermalization, based on the presence of exact resonance, has universal features. Moreover, a by-product of our analysis is the asymptotic integrability, up to four wave interactions, of the discrete nonlinear Klein-Gordon chain.

Keywords: FPUT; Klein-Gordon; nonlinear interactions; Wave Turbulence; thermalization

1. Introduction

The numerical experiment of Fermi, Pasta, Ulam and Tsingou (FPUT) in 1955 [18] has been of great influence to physics in many respects. The idea of the experiment was related to the role of chaos on the foundations of statistical mechanics. As chaotic systems are highly irregular, like stochastic ones, they lose memory of initial conditions rapidly, in agreement with ergodic hypothesis. In this perspective, the result of Poincaré about the non existence in Hamiltonian systems of first integrals of motion, other than the energy or those due to a particular symmetry, seems positive. In his early theoretical activity, Fermi generalised the result of Poincaré showing that in Hamiltonian systems with $N > 2$ degrees of freedom no smooth surface can divide the phase space into two regions containing invariant sets. From this correct result, Fermi deduced that non-integrable Hamiltonian systems are generically ergodic and, even in absence of a rigorous proof, the ergodic problem was considered basically solved. However, Fermi came back to the problem after the war with this numerical experiment, probably too much absorbed by the development of the just born quantum mechanics in the meanwhile.

With one of the first available computers, they investigated the dynamics of a simple system of springs and masses, where the force is only between nearest neighbours. The novelty consisted in the methodology (it was one of the first numerical studies), but also in the fact that the force was not linear. In the case of linear forces, it is known that the Hamiltonian of the system can be diagonalized, the eigenstates are linear waves, and there is no interaction between them, namely the initial energy distribution among the modes does not change over time. Fermi and collaborators wanted to test whether introducing a small nonlinear term in the equations of motion (they used quadratic, cubic and split-linear terms) would allow energy to spread among the linear modes, and attain equipartition, confirming the earlier idea of Fermi that ergodic hypothesis can be considered true, giving a dynamical justification to statistical mechanics. However, this was not the case and instead quasi-periodic motion was observed or, at most, just a few modes close to the ones initially excited were active for the whole simulation time.

The surprising results of FPUT generated a lot of interest, and sparked research that resulted in a large body of work, as witnessed by an excellent review [23]. In particular, even though ergodicity and chaos have been clarified to be largely irrelevant for statistical mechanics of macroscopic systems, that is when $N \gg 1$ [30, 33, 35], the existence of solitons arise from a possible explanation of the FPUT phenomenology [46]. Moreover, the FPUT experiment was the first to show that numerical simulations are a powerful instrument to get physical insights of complex phenomena, and therefore an indispensable tool for theoretical progress. In parallel, at the same time, it was shown by mathematicians via the KAM theorem [25, 28, 31] that generic non-integrable Hamiltonian systems are not ergodic for small perturbations. This fundamental result showed that the first guess by Fermi was incorrect, even though in high dimensional systems ($N > 2$) the presence of KAM tori does not preclude chaotic orbits because of the Arnold diffusion mechanism.

Despite these outstanding developments and considerable efforts, a clear and exhaustive explanation of the FPUT problem is still lacking. In particular, while there is some consensus concerning the metastability picture [3, 22], the precise mechanisms driving to equipartition on very long time-scales is still elusive. One crucial point is that studies have focused on peculiar initial conditions and the specific phenomena that emerge from them in the short or medium timescales. For example, in the original experiments and in most of the successive studies the initial conditions consisted in exciting the

low-frequency modes [2], which leads to the formation of coherent nonlinear structures, that is solitons, at least in the short-time. Some attention has been focused also on high-frequency initial conditions and this led to the study of breathers. However, it is difficult to understand all these peculiar directions in one larger vision of the FPUT problem. The problem with generic initial conditions seems to be yet more relevant from a general perspective, as recognised in an important recent contribution [8].

In the present work, we propose a more unifying picture; therefore, we seek universality traits that can explain the dynamics in a way that is only weakly dependent on the particular microscopic processes of the nonlinear chain in consideration. For this reason, we will consider generic initial conditions that are randomly out of equilibrium, involving low and high modes. Moreover, we will use tools and methods of the so-called Wave Turbulence (WT) theory. Originally introduced in fluid dynamics, WT is a statistical mechanics theory of weakly interacting waves, and it has been recently applied in relation to one dimensional anharmonic chains. In some sense, our approach is in line with the recent results that have clearly showed that the α-FPUT chain is a perturbation of Toda lattice, which is an integrable system [1]. Our main claim is that the irreversible transfer of energy in the spectrum in a weakly nonlinear system is achieved by exact resonant wave-wave interactions. In the context of the FPUT system, this idea was around already in the sixties, see [19]. Such resonances are the base for the WT theory and are responsible for the phenomenon of thermalization.

In this paper we give an overview on resonances and WT in a comprehensive manner as well as present new results regarding the limit of a large number of modes, providing some insight in the fundamental mechanics of anharmonic chains. Our main result is the establishment of the power-law scaling of the equipartition time T_{eq} as a function of the nonlinearity strength (to be defined precisely later). In varying the number of elements in the chain, we show the crossover between two different scaling laws (lower exponent for large systems, that is the thermodynamic limit, steeper for smaller ones).

We remark that some of the ideas presented here have partially appeared in a series of papers by the same group of authors [5, 36, 41, 42]. All the simulations presented on the α and β FPUT problem are performed with a different class of initial conditions with respect to our previous work. Moreover, the α and β FPUT simulations for large values of the number of particles are completely new.

2. The models

We consider Hamiltonian systems of the form

$$H = H_{\text{lin}} + H_{\text{nlin}}, \tag{2.1}$$

where H_{lin} is the integrable Hamiltonian, corresponding to a linear dispersive dynamics, and H_{nonlin} that contains the anharmonicity of the potential. More in particular we will deal with the α, β-FPUT and the discrete nonlinear Klein-Gordon (DNKG) (also called the ϕ^4) systems, all characterized by the quadratic Hamiltonian

$$H_{\text{lin}} = \sum_{j=0}^{N-1} \frac{1}{2} p_j^2 + \frac{1}{2}(q_j - q_{j+1})^2 + \frac{1}{2} m q_j^2, \tag{2.2}$$

where N is the number of elements in the chain, q_j is the j-th coordinate and p_j its conjugate momentum and m models the linear part of the site-potential. For the α and β-FPUT systems $m = 0$. The nonlinear

part of the Hamiltonian for the three models is given by:

$$H_{\text{nlin}}^{(\alpha)} = \frac{\alpha}{3}\sum_{j=0}^{N-1}(q_j - q_{j+1})^3, \quad H_{\text{nlin}}^{(\beta)} = \frac{\beta}{4}\sum_{j=0}^{N-1}(q_j - q_{j+1})^4, \quad H_{\text{nlin}}^{(KG)} = \frac{\beta}{4}\sum_{j=0}^{N-1} q_j^4. \tag{2.3}$$

Without loss of generality, we use the same parameter β in front for nonlinearity in the DNKG and the β-FPUT. We will use periodic boundary conditions. It should be noted, given an initial energy, that there is a finite probability that the α-FPUT chain breaks up, especially for large N, since the nonlinear part of the Hamiltonian is not bounded from below. While being an inherently ill-posed problem, it has been shown that for low enough energies the α-FPUT chain can be stable for long times [7]. For the other models, we restrict ourselves to the case where the quartic terms are positive, that is $\beta > 0$, so that no blow-ups can occur. The parameter m is constrained to be greater than zero in the DNKG model, and we can anticipate that it will play a secondary role since we are interested in the effects of the nonlinear interactions. The presence of the quadratic finite difference terms in the linear Hamiltonian is important, because without that the eigenstates of the linear system would not be waves.

Our approach is of perturbative nature, hence we need a way to quantify the nonlinearity strength. In general, the linear and nonlinear parts of the Hamiltionan are not conserved separately, but only the total Hamiltonian is. However, in typical situations where the nonlinear interactions are not too strong, operatively, we define the following nondimensional parameter for the α-FPUT

$$\epsilon_\alpha = \left(\frac{\sum_{j=0}^{N-1}\left|h_{\text{nlin}}^{(\alpha)}(j)\right|}{H_{\text{lin}}}\right)^2 \propto \alpha^2, \tag{2.4}$$

and for the β-FPUT and DNKG

$$\epsilon_{\beta,\text{KG}} = \frac{H_{\text{nlin}}^{(\beta,\text{KG})}}{H_{\text{lin}}} \propto \beta. \tag{2.5}$$

$h_{\text{nlin}}^{(\alpha)}(j)$ is the nonlinear part of the Hamiltonian density of the j-th particle in the for α-FPUT. The absolute value in Eq (2.4) is necessary because $h_{\text{nlin}}(j)$ can be negative and cancellations in the sum would cause an incorrect accounting of the nonlinearity strength. The different scalings in Eqs (2.4) and (2.5) are due to the different degrees of nonlinearity between the α-FPUT and the other two models. Definitions (2.4) and (2.5) allows us to refine our statement that the nonlinearity must be small, that is, we require that the parameters are much less than one. In general, they fluctuate over time, so an appropriate averaging is needed to have a meaningful estimate of the nonlinearity strength. The parameters are further discussed when presenting the numerical results.

2.1. Fourier space

The eigenstates of the quadratic Hamiltonian are waves, hence it is useful to work in Fourier space. We define the direct and inverse discrete Fourier transform of the q_j variables,

$$\hat{q}_k = \frac{1}{N}\sum_{j=0}^{N-1} q_j e^{-i2\pi kj/N}, \; q_j = \sum_{k=0}^{N-1} \hat{q}_k e^{i2\pi kj/N} \tag{2.6}$$

and similar definitions for $\hat{p}_k$. We will use the convention that $0 \le k < N$, and we note $\hat{q}_k^* = \hat{q}_{N-k}$ and $\hat{p}_k^* = \hat{p}_{N-k}$. After this change of variables, and using $\sum_{j=0}^{N-1} e^{i2\pi(k_1-k_2)j/N} = N\delta_{k_1-k_2}$ with $\delta_{k_1+k_2+\ldots}^{(N)} =$

$\delta(k_1 + k_2 + ... \mod N)$ that is the Kronecker's delta modulo N, the linear part of the Hamiltonians can be written as

$$\frac{H_{\text{lin}}}{N} = \frac{1}{2}\sum_{k=0}^{N-1}\left(\hat{p}_k^2 + \omega_k^2\left|\hat{q}_k^2\right|\right), \tag{2.7}$$

where we have defined the linear dispersion relation

$$\omega_k = \sqrt{m + 4\sin^2\left(\pi\frac{k}{N}\right)} \tag{2.8}$$

(we recall that $m = 0$ for the FPUT models). The nonlinear part of the Hamiltonians couple the modes and can be interpreted as n-wave collision terms. For the sake of brevity, we denote $\hat{q}_j = \hat{q}_{k_j}$ and $\delta^{(N)}_{1+2+...} = \delta(k_1 + k_2 + ... \mod N)$. For the FPUT models we obtain

$$\frac{H^{(\alpha)}_{\text{nlin}}}{N} = \frac{\alpha}{3}\sum_{k_1,k_2,k_3=0}^{N-1}\widetilde{A}_{1,2,3}\hat{q}_1\hat{q}_2\hat{q}_3\delta^{(N)}_{1+2+3}, \quad \frac{H^{(\beta)}_{\text{nlin}}}{N} = \frac{\beta}{4}\sum_{k_1,k_2,k_3,k_4=0}^{N-1}\widetilde{B}_{1,2,3,4}\hat{q}_1\hat{q}_2\hat{q}_3\hat{q}_4\delta^{(N)}_{1+2+3+4}, \tag{2.9}$$

where the collision matrices $\widetilde{A}_{1,2,3}$ and $\widetilde{B}_{1,2,3}$ are (after symmetrization)

$$\widetilde{A}_{1,2,3} = 8ie^{i\pi(k_1+k_2+k_3)/N}\sin\left(\pi\frac{k_1}{N}\right)\sin\left(\pi\frac{k_2}{N}\right)\sin\left(\pi\frac{k_3}{N}\right), \tag{2.10}$$

$$\widetilde{B}_{1,2,3,4} = 16e^{i\pi(k_1+k_2+k_3+k_4)/N}\sin\left(\pi\frac{k_1}{N}\right)\sin\left(\pi\frac{k_2}{N}\right)\sin\left(\pi\frac{k_3}{N}\right)\sin\left(\pi\frac{k_4}{N}\right). \tag{2.11}$$

The complex exponential is relevant only when crossing Brillouin zones in the sum of wave numbers. For the DNKG model, it turns out that the interaction matrix is equal to unity, hence the nonlinear part of the Hamiltonian is simply

$$\frac{H^{(KG)}_{\text{nlin}}}{N} = \frac{\beta}{4}\sum_{k_1,k_2,k_3,k_4=0}^{N-1}\hat{q}_1\hat{q}_2\hat{q}_3\hat{q}_4\delta^{(N)}_{1+2+3+4}. \tag{2.12}$$

From Eq (2.7), we can further simplify the problem using the normal modes representation,

$$a_k = \frac{1}{\sqrt{2\omega_k}}\left(\hat{p}_k - i\omega_k\hat{q}_k\right), \; a^*_{N-k} = \frac{1}{\sqrt{2\omega_k}}\left(\hat{p}_k + i\omega_k\hat{q}_k\right) \tag{2.13}$$

as the linear part of the Hamiltonians becomes simply

$$\frac{H_{\text{lin}}}{N} = \sum_{j=0}^{N-1}\omega_k\left|a_k\right|^2, \tag{2.14}$$

whereas the nonlinear terms in the FPUT models and in the DNKG become (after renaming $N - k_i = k'_i$ where needed and dropping the prime)

$$\frac{H^{(\alpha)}_{\text{nlin}}}{N} = \alpha\sum_{k_1,k_2,k_3=0}^{N-1}\left[\frac{1}{3}(A_{1,2,3}a_1a_2a_3 + c.c.)\delta^{(N)}_{1+2+3} + (A_{-1,2,3}a^*_1a_2a_3 + c.c.)\delta^{(N)}_{1-2-3}\right], \tag{2.15}$$

$$\frac{H^{(\beta,\mathrm{KG})}_{\mathrm{nlin}}}{N} = \beta \sum_{k_1,k_2,k_3,k_4=0}^{N-1} \Bigg[\frac{1}{4}(B^{(\beta,\mathrm{KG})}_{1,2,3,4} a_1 a_2 a_3 a_4 + c.c.)\delta^{(N)}_{1+2+3+4} + (B^{(\beta,\mathrm{KG})}_{-1,2,3,4} a_1^* a_2 a_3 a_4 + c.c.)\delta^{(N)}_{1-2-3-4} + + \frac{3}{2} B^{(\beta,\mathrm{KG})}_{-1,-2,3,4} a_1^* a_2^* a_3 a_4 \delta^{(N)}_{1+2-3-4} \Bigg], \tag{2.16}$$

with the interaction matrices given by

$$A_{1,2,3} = \frac{\widetilde{A}_{1,2,3}}{2^{3/2}\sqrt{\omega_1\omega_2\omega_3}}, \quad B^{(\beta)}_{1,2,3,4} = \frac{\widetilde{B}_{1,2,3,4}}{4\sqrt{\omega_1\omega_2\omega_3\omega_4}}, \quad B^{(\mathrm{KG})}_{1,2,3,4} = \frac{1}{4\sqrt{\omega_1\omega_2\omega_3\omega_4}}. \tag{2.17}$$

The nonlinear terms show that a nonlinearity in physical space turns into n-mode collision terms such as $a_1 a_2 a_3 \delta^{(N)}_{1+2-3}$. From the Hamilton equations, we can obtain the dynamical equations $\dot{a}_k = -i(1/N)\delta H/\delta a_k^*$, which read for the α-FPUT

$$i\dot{a}_1 = \omega_1 a_1 + \alpha \sum_{k_1,k_2,k_3=0}^{N-1} \left(A_{-1,2,3} a_2 a_3 \delta^{(N)}_{1-2-3} - 2A_{-3,1,2} a_2^* a_3 \delta^{(N)}_{1+2-3} - A_{1,2,3} a_2^* a_3^* \delta^{(N)}_{1+2+3} \right) \tag{2.18}$$

and for the β-FPUT and DNKG

$$i\dot{a}_1 = \omega_1 a_1 + \beta \sum_{k_1,k_2,k_3,k_4=0}^{N-1} (B^{(\beta,\mathrm{KG})}_{-1,2,3,4} a_2 a_3 a_4 \delta^{(N)}_{1-2-3-4} + 3B^{(\beta,\mathrm{KG})}_{-1,-2,3,4} a_2^* a_3 a_4 \delta^{(N)}_{1+2-3-4} + + B^{(\beta,\mathrm{KG})}_{-4,1,2,3} 3a_2^* a_3^* a_4 \delta^{(N)}_{1+2+3-4} + B^{(\beta,\mathrm{KG})}_{1,2,3,4} a_2^* a_3^* a_4^* \delta^{(N)}_{1+2+3+4}). \tag{2.19}$$

These types of Hamiltonians are the canonical form of Hamiltonians for system composed of interacting particles or waves. If nonlinearity is small in the above mentioned sense, then there is a well developed theory called Wave Turbulence theory, which is described in detail in [17, 38] and briefly in the subsequent section.

The number of modes in the nonlinear terms defines the collision order, that is the α-FPUT contains 3-wave collisions, while β-FPUT DNKG 4-wave collisions (we will see that the effective collision order in the α-FPUT is also of the 4-wave type). The number of non-conjugate and conjugate variables (which matches the number of positive vs. negative terms in the Kronecker's deltas) defines different collision processes, e.g., $a_1 a_2 a_3 \delta^{(N)}_{1+2+3}$ is a $3 \to 0$ collision process (annihilation), while $a_1 a_2 a_3^* a_4^* \delta^{(N)}_{1+2-3-4}$ is a $2 \to 2$ collision process (scattering).

The WT approach starts from the equations of motion in the canonical variables a_k. Note that so far we have not taken any approximation; we only made the choice of using periodic boundary conditions.

3. Wave-Turbulence theory in a nutshell

When the number of modes is large, the microscopic dynamics given by Eqs (2.18) and (2.19) do not provide much analytical insight and a statistical approach is needed. WT is the general statistical theory valid in the weak-nonlinear regime. While the theory has been recently developed for higher-order statistical observables [10, 16, 38], the core of the WT is the kinetic equation developed for the prediction of the energy spectrum. The basic statistical observable is the two point correlator $\langle a_1 a_2^* \rangle$ which, under the hypothesis of homogeneity, it is given by

$$\langle a_1 a_2^* \rangle = n_1 \delta(k_1 - k_2), \tag{3.1}$$

with $n_k = n(k,t)$ the wave actions spectral density, i.e., the Fourier transform of the autocorrelation function of $a(x,t)$. The average is over an ensemble of realizations of the system with the same initial linear energy $H_{\rm lin}$. Then, the energy spectrum that gives the mean energy per mode is

$$e_k = \omega_k n_k. \tag{3.2}$$

The evolution equation for the wave action spectral density can be obtained by various techniques [10, 16, 17, 38–40]. The main issue is due to the fact that calculating the evolution of n_k from the equation of motion (2.18) or (2.19), one encounters the well know problem of the BBGKY hierarchy [14], that is, the evolution of lower order correlators depends on higher order correlators and a closure problem arises. Different approaches have been used in the past, see for example [40] for a recent discussion. Accordingly to [39], a natural asymptotic closure arises because of the smallness of higher order cummulants. A closure was also obtained with a quasi-gaussian approximation which allowed to express higher cumulants in terms of lower ones [17]. It is sufficient to make a random-phase approximation and to consider initial conditions which have also random amplitudes [10, 12, 13, 16, 38]. The random phase assumption is generally deemed to be solid because the linear waves decorrelate quickly even in the linear regime, hence it is expected that such property still holds true in the weakly nonlinear regime, as corroborated by numerical experiments [9]. The resulting kinetic equation can be obtained in the thermodynamic limit $N \to \infty$. Generally also the continuum limit is taken at the same time, and both physical $\mathbf{x}$ and momentum space $\mathbf{k}$ are continuous.

The main concept underneath the kinetic equation is the existence of conservation laws associated to the wave scattering processes. Indeed, each nonlinear term in the Eqs (2.18) and (2.19) contains an appropriate Kroneker δ function over wave numbers. Depending on the shape of the dispersion relation, it may be possible to associate also a related conservation of energy. To be more specific, let us consider for example the scattering processes,

$$k_1 = k_2 + k_3, \tag{3.3}$$

contained in Eq (2.18), then the main question is to verify if a similar relation holds for frequencies, i.e.,

$$\omega_1 = \omega_2 + \omega_3. \tag{3.4}$$

Eqs (3.3) and (3.4) define a resonant manifold in wave number space, whose existence, as anticipated, depends only the analytical form of the dispersion curve ω_k. If Eqs (3.3) and (3.4) are satisfied for the same wave numbers, then we are dealing with a *resonant* process which, according to the kinetic equation, will lead to an irreversible transfer of energy.

There are two main types of wave systems that are typically considered in the framework of WT: The one dominated by three wave resonances and by four wave resonances (higher order are also possible). Examples of the systems dominated by three wave resonances include capillary waves on a surface of fluid and internal waves in the ocean. Examples of four wave resonant systems include gravity waves on deep water, as well as the celebrated 2D-Nonlinear Schrödinger equation.

For capillary waves on a surface three wave resonant interactions are possible and a kinetic equation can be derived [17]:

$$\frac{\partial n_1}{\partial t} = \int_{-\infty}^{+\infty} |A_{123}|^2 n_1 n_2 n_3 \left(\frac{1}{n_1} - \frac{1}{n_2} - \frac{1}{n_3}\right) \delta(\mathbf{k}_1 - \mathbf{k}_2 - \mathbf{k}_3)\delta(\omega_1 - \omega_2 - \omega_3) \mathrm{d}\mathbf{k}_{23} + 2\left\{(\mathbf{k}_1 \leftrightarrow \mathbf{k}_2)\right\}, \tag{3.5}$$

where $\mathbf{k}$ is a two dimensional wave vector, $\mathrm{d}\mathbf{k}_{23} = \mathrm{d}\mathbf{k}_2\mathrm{d}\mathbf{k}_3$ and $\{(\mathbf{k}_1 \leftrightarrow \mathbf{k}_2)\}$ implies the same integral with $\mathbf{k}_1$ and $\mathbf{k}_2$ exchanged. A_{123} is a matrix that can be derived directly from the dynamical equations for capillary waves.

For surface gravity waves in deep water, it can be show that the leading order resonant process is the four-wave one, characterized by the following resonant process:

$$\begin{aligned} &\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{k}_4 \\ &\omega_1 + \omega_2 = \omega_3 + \omega_4. \end{aligned} \tag{3.6}$$

The appropriate kinetic equation describing such process is

$$\frac{\partial n_1}{\partial t} = \int_{-\infty}^{+\infty} |B_{1234}|^2 n_1 n_2 n_3 n_4 \left(\frac{1}{n_1} + \frac{1}{n_2} - \frac{1}{n_3} - \frac{1}{n_4}\right) \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4)\delta(\omega_1 + \omega_2 - \omega_3 - \omega_4)\mathrm{d}\mathbf{k}_{234}. \tag{3.7}$$

Note that in both kinetic equation the interaction matrix is squared. The latter equation, with a transport term, a forcing term due to the wind and term mimicking the dissipation, is integrated daily for operationally ocean wave forecasting pourposes, see [27].

The kinetic Eqs (3.5) and (3.7) have a number of important properties. First of all, it can be shown that both equations conserves the total energy and momentum:

$$E = \int \omega_k n_{\mathbf{k}} \mathrm{d}\mathbf{k}, \qquad \mathbf{M} = \int \mathbf{k} n_{\mathbf{k}} \mathrm{d}\mathbf{k}. \tag{3.8}$$

The kinetic equation for four-wave interactions preserves also the number of waves (particles) in the scattering:

$$\mathcal{N} = \int n_{\mathbf{k}} \mathrm{d}\mathbf{k}. \tag{3.9}$$

The kinetic Eq (3.7) is reminiscent of the Boltzmann equation for hard spheres [44] and there exists an entropy function

$$S = \int \mathrm{d}\mathbf{k} \log(n_{\mathbf{k}}) \tag{3.10}$$

such that $\mathrm{d}S/\mathrm{d}t \geq 0$ (the same entropy can be defined for the three wave kinetic equation). The equilibrium is reached when $\mathrm{d}S/\mathrm{d}t = 0$, i.e., at the Rayleigh-Jeans distribution:

$$n_k|_{\text{equilibrium}} = \frac{T}{\omega_k + \mu + \mathbf{u} \cdot \mathbf{k}}, \tag{3.11}$$

where T is some normalization constant linked to the total energy and μ is a chemical potential linked to the conservation of the total number of particles (3.9) and it is zero in the case of a three wave process; $\mathbf{u}$ is a constant vector associated to momentum conservation. We observe that in isotropic conditions $\mathbf{u} = 0$, and when $\mu = 0$ then WT predicts a relaxation towards equipartition of energy. In this framework, exact resonances bring the system to thermalization with a time-scale which is proportional to the coupling coefficient in front of the collision integral.

Some remarks are in order.

- Given a physical dispersive waves system, it has to be checked if exact resonances are allowed. In some cases, only trivial collisions or processes which are not able to transfer energy are found. In those cases, it is necessary to pursue the perturbative procedure and derive a kinetic equation with higher-order processes to check if at some point resonances are found. This issue is related to the dispersion relation, and in physical phenomena 3-, 4-, 5- and 6-wave processes have been recognised [38]. In other cases, it can be shown that no exact resonances exist and therefore no exchange of energy is possible, as rigorously shown to be the case for integrable systems [47]. Other systems may show some resonances which are not yet able to transfer energy among modes because resonating waves belong to isolated clusters [6].
- The WT framework is an asymptotic theory valid in the weakly nonlinear regime. If the nonlinearity is not small, different effects can play a role. In particular, the broadening of the frequencies due to the finite value of the nonlinear coupling, may trigger some resonances which were not present in the weak asymptotic limit, therefore allowing a more efficient transfer of energy.

In the present paper, we estimate the time scale of equipartition; therefore, it is of major importance to be able to derive a kinetic equation for our chain models. One should note that chains are intrinsically discrete in physical space and, for a finite number of masses, also the Fourier space is discrete. This issue makes the derivation of the kinetic equation not straightforward and, only in the thermodynamic limit, a formal derivation of the kinetic equation is possible. We shall deal with these points in the following section in some detail.

4. Resonance and effective Hamiltonians in the thermodynamic limit

We consider the thermodynamic limit which consists in taking the number of particles going to infinity, $N \to \infty$, keeping constant the mass linear density. The general linear dispersion relation becomes

$$\omega_k = \sqrt{m + 4\sin^2\left(\frac{k}{2}\right)} \tag{4.1}$$

with k real in the interval $0 \le k < 2\pi$. We consider interaction processes from X to Y waves for which the following equations are satisfied:

$$\begin{aligned} &k_1 + k_2 + ... + k_X = k_{X+1} + k_{X+2}... + k_{X+Y} \qquad \text{mod } 2\pi, \\ &\omega_1 + \omega_2 + ... + \omega_X = \omega_{X+1} + \omega_{X+2}... + \omega_{X+Y}. \end{aligned} \tag{4.2}$$

Our goal is to find resonances. Once the leading order resonance (lowest value of X and Y) has been found, using tools from Hamiltonian mechanics, an effective Hamiltonian can be established and a kinetic equation can be built.

The first trivial observation is that resonances of the type X to 0 are excluded for all three models; this is because $\omega_k \ge 0$, and it is equal to zero only for $m = 0$ for mode $k = 0$ which does not play any role in the dynamics (for $m = 0$).

4.1. Three-wave resonant interactions in the α-FPUT

The leading order nonlinearity in the evolution equation for the α-FPUT is cubic; this implies three-wave interactions of the form $2 \to 1$, $1 \to 2$ or $3 \to 0$ (the latter have been already excluded). It is

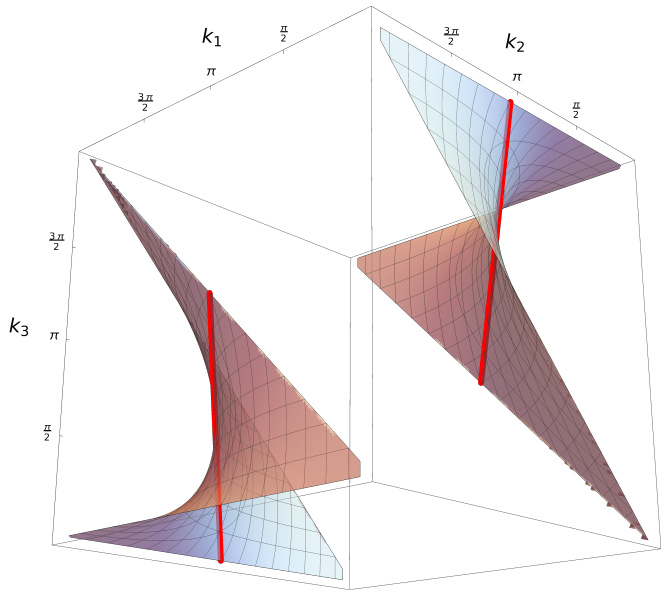


Figure 1. The continuous resonant manifold for $2 \to 2$ resonances with $m = 0$, that is $\omega_1 + \omega_2 = \omega_1 + \omega_{1+2-3}$. The trivial solutions $k_1 = k_3$ or $k_2 = k_3$ are not shown for clarity. The pairing-off resonances, to be discussed later, of Eq (5.3) are highlighted with the red line.

known, see for example [5, 37], that the function $f_k = 2|\sin(k/2)|$ is a subadditive function, i.e.,

$$|\sin(k_1/2)| + |\sin(k_2/2)| \geq |\sin((k_1 + k_2)/2)|. \tag{4.3}$$

The equality holds only for k_1 or $k_2 = l2\pi$, with $l \in \mathbb{Z}$; those wave numbers do not enter into the dynamics. The subadditivity implies the non existence of three-wave resonant interactions for $m = 0$.

4.2. Four-wave resonant interactions in the α, β-FPUT and DNKG models

The β-FPUT and DNKG models include processes involving four waves, while, at first sight, the α-FPUT model does not include them; however, as it will be shown in the following, the absence of three-wave resonances allows for a perturbative approach on the α-FPUT model that leads to four and higher order wave-wave interaction processes. In [5] it has been shown that resonant processes $3 \to 1$ or $1 \to 3$ are excluded for the FPUT models, i.e., in the case $m = 0$. In the Appendix we show that this implies that also for the case of $m \neq 0$ there are no such processes. It turns out that in the thermodynamic limit there exist, apart from trivial resonances $k_1 = k_2 = k_3 = k_4$ or $k_1 = k_3$, $k_2 = k_4$, a nontirvial resonant manifold for interactions of the type $2 \to 2$ (see Figure 1 for a visualization of the case $m = 0$, similar result apply for $m > 0$). The result has been obtained using the software Mathematica Wolfram.

4.3. The effective Hamiltonian and the kinetic equation for α, β-FPUT and DNKG models

The effective Hamiltonian is obtained by removing all non resonant interactions from the original one. As standard in Hamiltonian systems, this can be carried out through a canonical change of variables (quasi-identity transformation), from the normal modes a_k to some other variables b_k, such

that in the new variables the nonresonant terms are removed from the Hamiltonian, generating higher order nonlinearities. It is worth stressing that one has to ensure the canonicity of the transformation, at least up to the order of the new interaction terms that appear in the new variables. For example, to remove all the three-wave terms in the α-FPUT model one uses

$$a_1 = b_1 + \int_0^{2\pi} \left(X^{(1)}_{1,2,3} a_2 a_3 \delta^{(2\pi)}_{1-2-3} + X^{(2)}_{1,2,3} a^*_2 a_3 \delta^{(2\pi)}_{1+2-3} + X^{(3)}_{1,2,3} a^*_2 a^*_3 \delta^{(2\pi)}_{1+2+3} \right) \mathrm{d}k_2 \mathrm{d}k_3 + ..., \tag{4.4}$$

where the integral corresponds to three integrals from 0 to 2π and not from $-\infty$ to $+\infty$ and $\delta_{2\pi}$ account for periodicity of the Fourier space; this is because of the discreteness of our system in physical space. In Eq (4.4) we recognize terms equivalent to the nonlinear interactions in Eq (2.18). The matrices $X^{(i)}_{1,2,3}$ are determined by inserting Eq (4.4) into the α-FPUT Hamiltonian and by grouping the terms corresponding to different wave processes. The transformation generates four-wave interactions (and higher) and again one has to look for resonances and remove the non resonant ones. The calculation leads to the following effective Hamiltonian for the α-FPUT

$$H^{(\alpha)} = \int_0^{2\pi} \omega_k |b_k|^2 \, \mathrm{d}k + \frac{\alpha^2}{2} \int_0^{2\pi} \bar{B}^{(\alpha)}_{1,2,3,4} b^*_1 b^*_2 b_3 b_4 \delta^{(2\pi)}_{1+2-3-4} \mathrm{d}k_2 \mathrm{d}k_3 \mathrm{d}k_4 \tag{4.5}$$

For the β-FPUT and for the DNKG a transformation is used to remove the $3 \to 1$, $1 \to 3$ and $4 \to 0$ terms, so that the effective Hamiltonian takes the form:

$$H^{(\beta,KG)} = \int_0^{2\pi} \omega_k |b_k|^2 \, \mathrm{d}k + \frac{\beta}{2} \int_0^{2\pi} \bar{B}^{(\beta,KG)}_{1,2,3,4} b^*_1 b^*_2 b_3 b_4 \delta^{(2\pi)}_{1+2-3-4} \mathrm{d}k_2 \mathrm{d}k_3 \mathrm{d}k_4. \tag{4.6}$$

For an interested reader, a comprehensive procedure for removing three and four-wave interactions, is presented in [15, 32, 34]. All three models are dynamically described by the Zakharov equation:

$$i \frac{\partial b_1}{\partial t} = \omega_1 b_1 + \int_0^{2\pi} W_{1,2,3,4} b^*_2 b_3 b_4 \delta^{(2\pi)}_{1+2-3-4} \mathrm{d}k_2 \mathrm{d}k_3 \mathrm{d}k_4, \tag{4.7}$$

where the coefficient $W_{1,2,3,4}$ takes different form, depending on the particular system under consideration. Its actual analytical form is irrelevant in our discussion (except in those cases for which the coefficient is zero on the resonant manifold [48, 49]); it is just sufficient to mention that it is proportional to $\alpha^2 (A_{1,2,3})^2$ for the α-FPUT model and to $\beta B^{(\beta,KG)}_{1,2,3,4}$ for the β-FPUT and DNKG models.

The Zakharov equation is the starting point for developing the statistical theory, i.e., the wave kinetic equation, which takes the following form (see for details [10, 17, 38]):

$$\frac{\partial n_1}{\partial t} = \int_0^{2\pi} \mathrm{d}k_{2,3,4} |W_{1,2,3,4}|^2 \delta^{(2\pi)}(k_1 + k_2 - k_3 - k_4) \delta(\omega_1 + \omega_2 - \omega_3 - \omega_4) n_1 n_2 n_3 n_4 \left(\frac{1}{n_1} + \frac{1}{n_2} - \frac{1}{n_3} - \frac{1}{n_4} \right). \tag{4.8}$$

Note that, because of the $\delta^{(2\pi)}$, instead of δ, the momentum is not conserved. The time scales, T_{eq}, of four wave resonant interaction, which lead to a thermalized state, are:

$$T_{\mathrm{eq}} \propto \alpha^{-4} \propto \epsilon_\alpha^{-2} \tag{4.9}$$

for the α-FPUT model and

$$T_{\mathrm{eq}} \propto \beta^{-2} \propto \epsilon_{\beta,KG}^{-2} \tag{4.10}$$

for the DNKG and β-FPUT models.

We stress that the theory developed in this section is asymptotic and valid when the thermodynamic limit and the weak non-linear limit are taken, in the given order. It is important to wonder if the scaling laws obtained can be observed in actual finite-size systems and in particular in numerical experiments. We expect a positive answer, at least for sufficient large N and small nonlinearity. Indeed, it is well known that a small nonlinearity cause a frequency shift of the linear modes, and also a stocasticization of the frequencies, or in other words, a broadening [11, 26]. It is reasonable to assume that resonances do not need to be satisfied exactly in practical applications [29]. It is sufficient that broadening of frequencies becomes comparable with the spacing of the frequencies, which decreases with the number of modes, so that ω_k becomes continuous in the thermodynamic limit. In this sense, the discrete representation should converge to the continuous theory, and the higher the number of modes the smaller the nonlinearity required to be in agreement with the theory. We shall check this argument with extensive numerical simulations in the following.

5. Discrete exact resonances

In many interesting cases, the number of modes N is not too large, like for instance in the original FPU numerical experiments. In this case, it is not possible to consider the system a good approximation of the continuous one, and it has to be studied in its discrete form. As mentioned, one should recall that in discrete systems the condition in the Kroneker δ over wave numbers should be intended mod N.

5.1. Four-wave resonant interactions, effective Hamiltonian and itegrability

The first problem is to find out if discrete exact resonances exist and at which order. The previous discussion on three-wave resonance still applies here, hence three-wave resonances are always excluded from the α-FPUT model. Concerning 4-wave resonances, the only process that is potentially active for the dispersion relation (2.8) is the scattering process $2 \to 2$. These resonances have been considered extensively in [5, 41, 42], and we will recap here the results. Obviously, even discrete processes of the type $X \to X$ admit trivial resonances of the type

$$k_1 + k_2 = k_1 + k_2 \mod N,\ \omega_1 + \omega_2 = \omega_1 + \omega_2 \tag{5.1}$$

or $k_1 = k_2 = k_3 = k_4$. These processes, however, do not result into an exchange of energy, but rather cause a frequency shift of the linear modes [38]. Because the system is periodic and the dispersion relation is symmetric,

$$\omega_k = \omega_{N-k} = \omega_{-k}, \tag{5.2}$$

it turns out that nontrivial resonances of the type $2 \to 2$ are possible with the crossing of Brillouin zones (Umklapp scatterings). For N even, these resonances take the form

$$k_1 + k_2 = -k_1 - k_2 \mod N, \qquad k_1 + k_2 = N/2, \tag{5.3}$$

and are known as *pairing-off* resonances [5]. However, one can easily check that these resonances are all disconnected, and they actually give rise to integrable dynamics [24, 43]. In [42] other special resonances have been considered for some very specific values of $m \neq 0$, but they are limited in number

and they do not appear in general to be able to cover the whole Fourier space (a detailed study should be performed for such specific cases).

The effective integrable Hamiltonian, up to four-wave interactions for the α, β and DNKG equations (with m different for those special values for which other resonant quartets exists), can be recast as follows:

$$\begin{aligned}\frac{H_{\text{integrable}}}{N} = \sum_k \omega_k |b_k|^2 + \frac{1}{2}\sum_k W_{k,k,k,k}\left(|b_k|^2\right)^2 + \sum_{k_1\neq k_2} W_{k_1,k_2,k_1,k_2}|b_{k_1}|^2|b_{k_2}|^2 + \\ + \sum_{k=1}^{\lfloor N/4\rfloor} 2W_{k,\frac{N}{2}-k,-k,-\frac{N}{2}+k}\left(b_k^* b^*_{\frac{N}{2}-k} b_{-k} b_{-\frac{N}{2}+k} + c.c.\right),\end{aligned} \tag{5.4}$$

where $c.c.$ denotes complex conjugate. The first three terms depend on the moduli of the amplitudes only and underline an integrable dynamics (Birkhoff normal form); interestingly, the last term does not break integrability (resonant Birkhoff normal form) [24,43], due to the relations (5.2) and the following symmetries:

$$\begin{aligned}&W_{k_1,k_2,k_1,k_2} = W_{k_2,k_1,k_2,k_1} = W_{-k_1,k_2,-k_1,k_2},\\ &W_{k_1,k_2,k_1,k_2} + W_{\frac{N}{2}-k_1,k_2,\frac{N}{2}-k_1,k_2} - W_{-k_1,k_2,-k_1,k_2} - W_{-\frac{N}{2}+k_1,k_2,-\frac{N}{2}+k_1,k_2} = 0\,,\end{aligned} \tag{5.5}$$

valid for all admissible values of k_1, k_2. It can be proven that $N-1$ functionally independent invariants exist and are in involution (for α and β-FPUT see the proof of integrability in [24, 43]). We present these invariants in a way that highlights the fact that the resonant quartets are disconnected:

For each $k = 1, \ldots, \lfloor N/4\rfloor$, define an irreducible quartet as the set $Q_k = \left\{b_k, b_{-k}, b_{\frac{N}{2}-k}, b_{-\frac{N}{2}+k}\right\}$. There are four different modes in Q_k, except for the degenerate case $k = N/4$, valid when $N/4$ is integer, where $Q_{N/4}$ has two different modes. In the non-degenerate case, the following four invariants depend on the four modes in Q_k:

- 3 quadratic invariants:

$$|b_k|^2 + |b_{-k}|^2, \qquad |b_k|^2 + |b_{-\frac{N}{2}+k}|^2, \qquad |b_{\frac{N}{2}-k}|^2 + |b_{-\frac{N}{2}+k}|^2.$$

- 1 quartic invariant:

$$2W_{k,\frac{N}{2}-k,-k,-\frac{N}{2}+k}\left(b_k^* b^*_{\frac{N}{2}-k} b_{-k} b_{-\frac{N}{2}+k} + c.c.\right) + W_{k,k,k,k}|b_k|^2|b_{-k}|^2 + W_{\frac{N}{2}-k,\frac{N}{2}-k,\frac{N}{2}-k,\frac{N}{2}-k}|b_{\frac{N}{2}-k}|^2|b_{-\frac{N}{2}+k}|^2.$$

In the degenerate case, valid when $N/4$ is integer, we have $Q_{N/4} = \left\{b_{\frac{N}{4}}, b_{-\frac{N}{4}}\right\}$ and the following two invariants depend on the two modes in $Q_{N/4}$:

- 1 quadratic invariant:

$$|b_{\frac{N}{4}}|^2 + |b_{-\frac{N}{4}}|^2\,.$$

- 1 quartic invariant:

$$W_{\frac{N}{4},\frac{N}{4},-\frac{N}{4},-\frac{N}{4}}\left[\left(b^*_{\frac{N}{4}} b_{-\frac{N}{4}}\right)^2 + c.c.\right] + W_{\frac{N}{4},\frac{N}{4},\frac{N}{4},\frac{N}{4}}|b_{\frac{N}{4}}|^2|b_{-\frac{N}{4}}|^2.$$

Thus, in terms of counting, any irreducible quartet (degenerate or not) contributes with a number of invariants that is equal to the number of modes in the quartet. Finally, notice that the mode $b_{\frac{N}{2}}$ is not in any irreducible quartet. In fact, the following is an invariant: $|b_{\frac{N}{2}}|^2$.

It is thus easy to show by simple counting of the modes in the irreducible quartets that the system has a total of $N-1$ functionally independent invariants:

1. When $N/4$ is not integer, the irreducible quartets give a total of $4\times\lfloor N/4\rfloor = N-2$ invariants. The missing invariant is $|b_{\frac{N}{2}}|^2$.
2. When $N/4$ is integer, we get a total of $4\times(N/4-1) = N-4$ invariants from the non-degenerate quartets, plus 2 invariants from the degenerate quartet $Q_{N/4}$, totalling again $N-2$ invariants. The missing invariant is $|b_{\frac{N}{2}}|^2$.

The result shows the asymptotic integrability of the DNKG model up to four wave interactions and justifies the metastable states observed in numerical simulations previously performed [22].

5.2. Higher order resonances and break-down of integrability

For the discussion above, we see that the 3- and 4-wave collision terms cannot be effective in bringing the system to equipartition. The isolated resonant quartets of the 4-wave integrable Hamiltonian do not bring the system to a thermalized state and resonances at higher order are to be investigated. For all discrete models considered here we have to perform an extra canonical change of variables, from the normal modes b_k to some other variables c_k, such that in the new variables the nonresonant terms are removed from the Hamiltonian. The question is now what is the leading order resonant wave interaction. In [41], only power-of-two values of N were investigated (akin to the original paper on the α-FPUT problem) and five-wave interactions were excluded on numerical grounds. From a more recent investigation [5], it turned out that five-wave resonant interactions exist only if N is divisible by 3 when $m=0$. When $N=2^a3^b$ with $a,b>1$, the resulting five-wave clusters are connected and, in principle, a thermalized state might be reached, but this requires further study, to be discussed in a subsequent paper.

Excluding such specific values of N, six-wave resonant interactions are the leading order processes: It is always possible to find resonant six-wave tuples that are all interconnected and cover the whole Fourier space. These resonances are due to the same symmetries of ω_k and are of the form:

$$k_1+k_2+k_3 = -k_1-k_2-k_3 \mod N,\ k_1+k_2+k_3 = 0 \mod N. \tag{5.6}$$

Such resonances are valid for even and odd values of N. Additional resonances can be found both in pairing-off form for even N, and recently also other resonances not in pairing-off form have been found [5], but they are not necessarily to cover the whole Fourier space and we will not discuss them. Another six-wave processes is theoretically possible, that is $4\to 2$. Explicit formulas for $4\to 2$ resonances have been found recently in [5], at least for N divisible by 3.

Excluding those specific values of N for which five wave interactions and $4\to 2$ (or $4\to 2$) exist, the six wave interaction Hamiltonian for the α,β-FPUT and DNKG equation can be written as

$$\frac{H}{N} = \frac{H_{\text{integrable}}}{N} + \sum_{1,2,3,4,5,6} Z_{1,2,3,4,5,6} c_1 c_2 c_3 c_4^* c_5^* c_6^* \delta^{(N}_{1+2+3-4-5-6}, \tag{5.7}$$

where δ_N is the Kronecker modulo N. The new matrix interaction can be calculated [34], but its precise form is not relevant to our scope.

5.3. The time scale of thermalization

The next step to be taken in order to evaluate the thermalisation time-scale should be the derivation of the corresponding kinetic equation, as done in the continuous case. Unfortunately, the kinetic equation can be rigorously derived only in the thermodynamic limit, that is $N \to \infty$. Deriving a *discrete* version (N finite) of the kinetic equation poses significant mathematical problems [16].

Here, we make the conjecture that the time scale for thermalization in the discrete dynamics corresponding to the 6-wave interaction given by Eq (5.7) is at the leading order equivalent to the corresponding continuous 6-wave process. We also assume that the integrable part of the dynamics does not lead to any irreversible transfer of energy and the irreversible dynamics is fully contained in the six-wave interaction term. Our conjecture is therefore tantamount to saying that the thermalisation time-scale in the discrete case will be always inversely proportional to the square of the coupling coefficient $Z_{1,2,3,4,5,6}$ in the six by Eq (5.7), the latter coefficient being proportional to α^4 in the α-FPUT model and β^2 in the β-FPUT and DNKG system. We get the following estimates for the thermalisation time-scale for the α-FPUT model

$$T_{\text{eq}} \propto \alpha^{-8} \propto \epsilon_{\alpha}^{-4} \tag{5.8}$$

and

$$T_{\text{eq}} \propto \beta^{-4} \propto \epsilon_{\beta,KG}^{-4} \tag{5.9}$$

for the β-FPUT, DNKG models. The argument used in the discrete case is always made for the limit of vanishing nonlinearity, where only exact resonances are important. However, the presence of a small but finite nonlinearity is expected to broaden the frequencies. In the case of a small number of modes N, that is in the present discrete case, an important difference with respect to the continuous case is expected. Since the wave-space is discrete, a large enough nonlinearity may cause a sufficient broadening to trigger resonances which are not in the exact-resonance manifolds and are found at a lower-order than the exact ones. Therefore, we can expect that at small N and sufficient nonlinearity the transfer of energy can be made through lower-order processes on smaller time-scales. A definite answer on the role of quasi-resonance is beyond the scope of the present manuscript.

5.4. Final remarks on the Wave Turbulent approach

Concluding the theoretical analysis, the WT formalism applied to the different anharmonic chains leads to consider the problem of transfer of energy as related to the collisions between waves in the system. In this framework, the energy is redistributed among modes if resonances exist which connect the whole wave-space. In the thermodynamic limit, the theory predicts that 4-wave processes are the leading order process in all chains and this fact allows a transparent estimate of the equipartition time-scale. In the case of a finite small N, we have considered the system intrinsically discrete and we have searched for the exact discrete resonances. We have seen that in this case, 4-waves processes are not able to transfer energy among modes, nor 5-wave collisions. The leading order is found to be 6-waves. Although we are not able to write the corresponding discrete kinetic equation, we have made the hypothesis that the same reasoning as in the continuous case can be followed, at least when looking at statistical observables like equipartion time-scale. In this way, an estimate of this time-scale is obtained always proportional to the square of the nonlinear coupling coefficient. Moreover, also when N is small, nonlinear effects can be important. We expect that the broadening of the frequencies

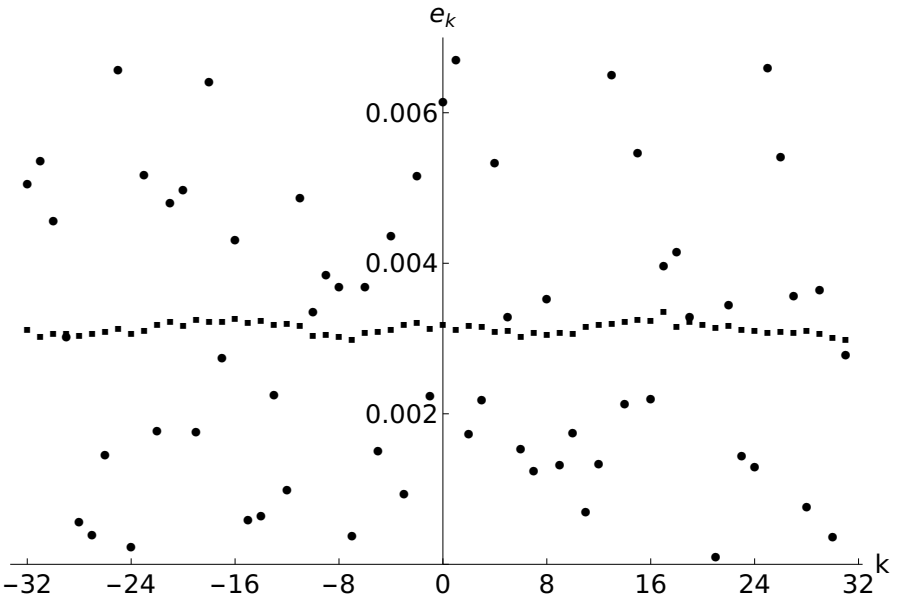


Figure 2. The initial distribution of e_k for the DNKG model with $m = 1$, $N = 64$ (circles), with $E = 0.2$ and $\mathcal{N} \simeq 0.129$. The thermalized final state is shown with the squares, and it approximately corresponds to a Rayleigh-Jeans distribution (3.11) with $\mu = 0$.

due to nonlinearity can trigger lower-order process to transfer energy among modes when it becomes of the same order of the spacing in the frequency space, which is inversely proportional to N. The discrete case stands therefore on a less rigorous ground and the conjectures proposed can be only verified numerically *a posteriori*.

6. Numerical simulations

We now present the result of numerical simulations in support of the previous discussion. The Eqs (2.18) and (2.19) have been implemented in physical space, with a symplectic algorithm of the sixth order [45]. It is important to use a high-order integration scheme that preserves accurately the energy of the system, because simulation times are very long compared to the typical wave periods. The initial random energy per mode, e_k, was drawn from an uniform distribution, and then scaled in order to obtain the desired total linear energy E and an initial number of particles $\mathcal{N}$ compatible with a final relaxation distribution with $\mu = 0$ in Eq (3.11), in order to better observe equipartition. Each wave number has the same energy per mode e_k for different members of the ensemble. The initial condition is out of equilibrium. As an example, we show the initial distribution of e_k for the simulations of the DNKG system with $N = 64$ in Figure 2, together with its final thermalized state. The linear energy E and the wave action $\mathcal{N}$ change only for one part in 10^4 across the whole simulation time, being quasi-conserved quantities.

Each realization in the ensemble shares the same initial energy per mode, but phases are randomized. This scheme of initialization ensures that all the realization share the same initial linear energy. The time step was set to 0.1 in all simulations, and it was checked that from beginning to the

end of the simulation the total energy is conserved up to one part in 10^7. The computation is very easy to parallelize, since ensemble averages need to be computed. For this reason we implemented the code on GPU hardware, which is powerful when the code can be parallelized.

The parameter α and β are chosen in such a way that the nonlinear energy is small compared to the linear one and the parameters ϵ_α and $\epsilon_{\beta,KG}$ are small. We will present the results of T_{eq} as a function of ϵ_α and $\epsilon_{\beta,KG}$.

In our numerical simulations we monitor the approach to equipartition using the entropy

$$S'_{\text{WT}} = -\sum_{k=0}^{N-1} \log(e'_k), \tag{6.1}$$

where

$$e_k = \omega_k \langle |a_k|^2 \rangle \quad \text{and} \quad E = H_{\text{lin}} = \sum_{k=0}^{N-1} e_k. \tag{6.2}$$

Such entropy is a discretized version of Eq (3.10) and has the useful property that at equipartition then $S'_{\text{WT}} = 0$ and it is greater than zero in any other state. This expression preserves the monotonicity of the WT entropy with an inverted sign ($\mathrm{d}S'_{\text{WT}}/\mathrm{d}t \leq 0$, $S'_{\text{WT}} = 0$ at equipartition).

In numerical simulations, the entropy shows some fluctuations and never reaches exactly zero, because the simulation time is finite, but most importantly because the ensemble size is finite. In the following, we will determine T_{eq} looking at the time when the entropy crosses a threshold value from above. The selection of the threshold depends on the ensemble size and the number of particles, but it does not depend on the degree of nonlinearity. Generally we chose a value of S'_{WT} two orders of magnitude smaller than that of the initial conditions.

6.1. Small number of particles, discrete resonances

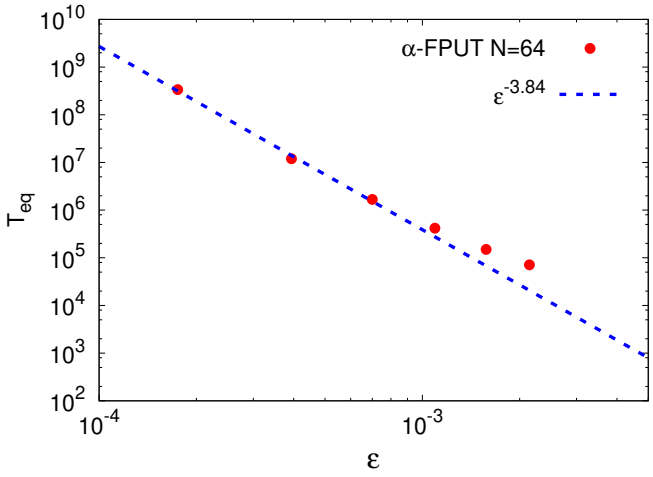
We first present the results on the case of a limited number of particles, $N = 64$, so that discrete resonances need to be considered. To recap, we expect for the α-FPUT system the scaling

$$T_{\text{eq}} \propto \alpha^{-8} \propto \epsilon_\alpha^{-4} \tag{6.3}$$

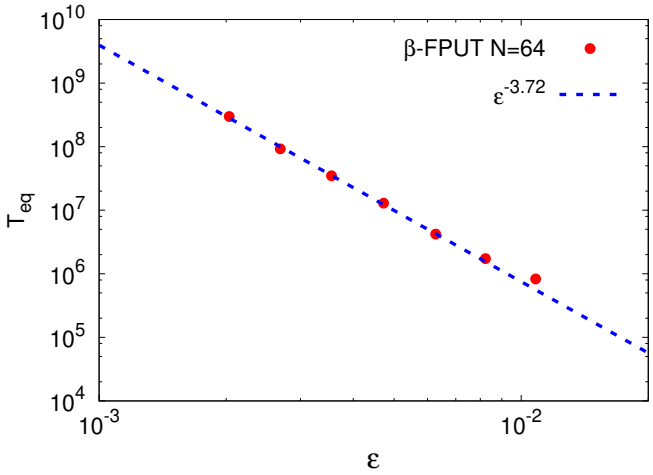
and for β-FPUT and DNKG systems the scaling

$$T_{\text{eq}} \propto \beta^{-4} \propto \epsilon_{\beta,KG}^{-4}. \tag{6.4}$$

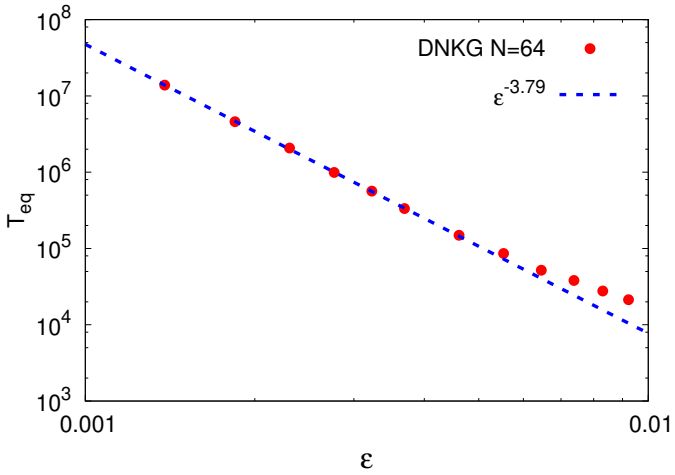
We show the results of the simulations for $N = 64$ in Figure 3. A line obtained from a fitting procedure has also been included. We see a good alignment with the expected power-laws for all the three models. We chose $m = 1$ in the DNKG model arbitrarily: In [42] it is argued that since the resonances in the discrete case do not depend on m, then in the scope of the thermalization dynamics the value of m is not relevant. However, we should add here a comment for the limit of m very small. In such cases, the contribute to H_{nlin} increases for the low frequency modes. On the other hand, for the FPUT systems the nonlinearity depends on a power of the finite differences, which are approximately proportional to k, hence the contribute to the nonlinear part of the Hamiltonians for modes with small k does not grow. As a consequence, in the case of the DNKG system with $m = 0$, a complete thermalization of the lower modes would invalid the weak-nonlinear assumption, because the nonlinearity would be too high. What is observed actually (see Figure 4) is a partial thermalization of the system, with the lower modes stationary at a lower energy than equipartition.



(a) α-FPUT, $E = 1$, ensemble size 4096.



(b) β-FPUT, $E = 1$, ensemble size 4096.



(c) DNKG, $m = 1$, $E = 0.2$, ensemble size 4096.

Figure 3. T_{eq} as a function of ϵ for the α, β-FPUT and DNKG ($m = 1$) systems with $N = 64$.

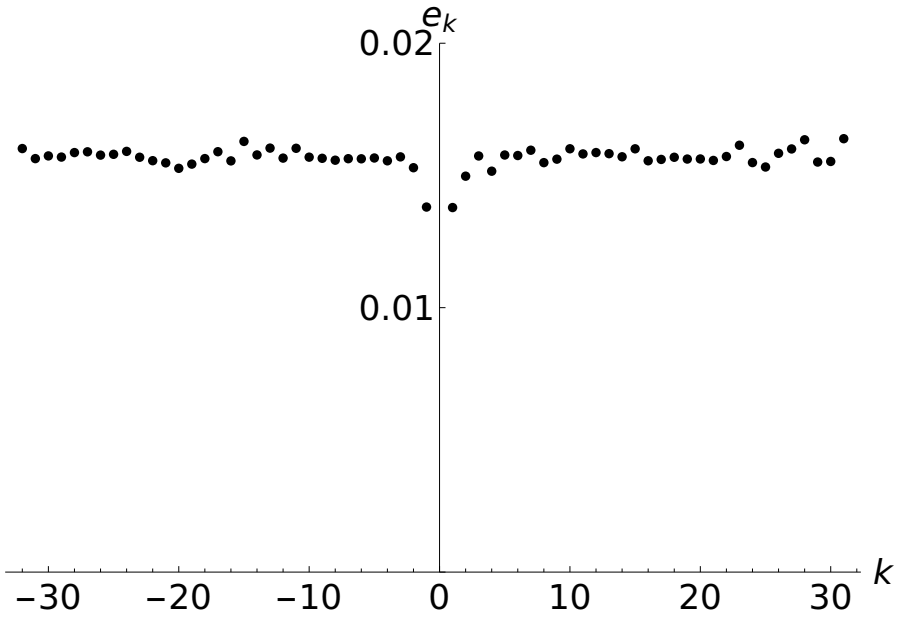


Figure 4. The equilibrium linear energy of the DNKG model with $N = 64$, $m = 0$, $E = 1$. The modes close to $k = 0$ do not thermalize as the other modes, because their contribute to H_{nlin} is significantly higher than the other modes.

6.2. Large number of particles, continuous resonances

We now present the results of numerical simulations for the continuous resonance manifold. The expected results are

$$T_{\text{eq}} \propto \alpha^{-4} \propto \epsilon_{\alpha}^{-2} \tag{6.5}$$

for the α-FPUT system, and

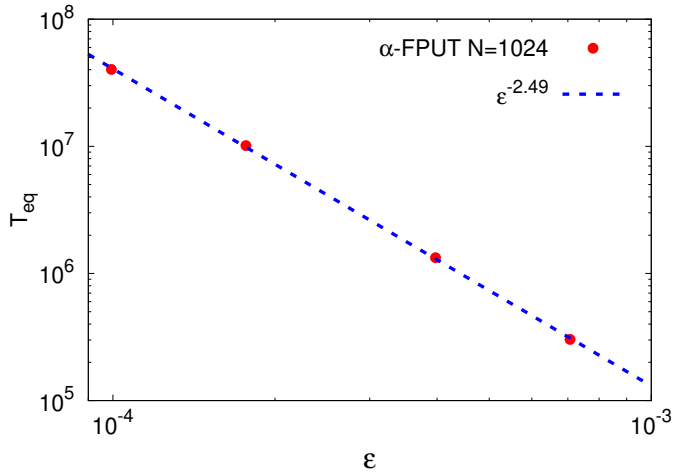
$$T_{\text{eq}} \propto \beta^{-2} \propto \epsilon_{\beta,KG}^{-2} \tag{6.6}$$

for β-FPUT and DNKG systems. The number of particles and the strength of the nonlinearity necessary to observe the continuous resonances are not something predicted by our treatment. However, estimates of the frequency broadening are available in [36,38], and they can be compared to the typical frequency spacing for a fixed number of particles. We point out that in general a higher value for m in Eq (2.8) makes the frequency gaps smaller among modes, hence the DNKG model in general requires a small number of particles to observe the continuous resonances.

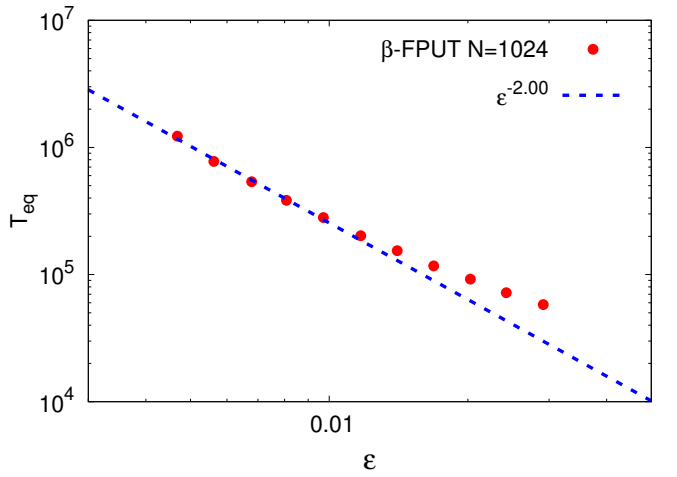
In Figure 5 we show the results for simulations with $N = 1024$, again with $m = 1$ for the DNKG system. The figure highlight a reasonable agreement between numerics and the theoretical predictions. A discussion of the numerical results is reported hereafter.

7. Discussion and relation to other results in the literature

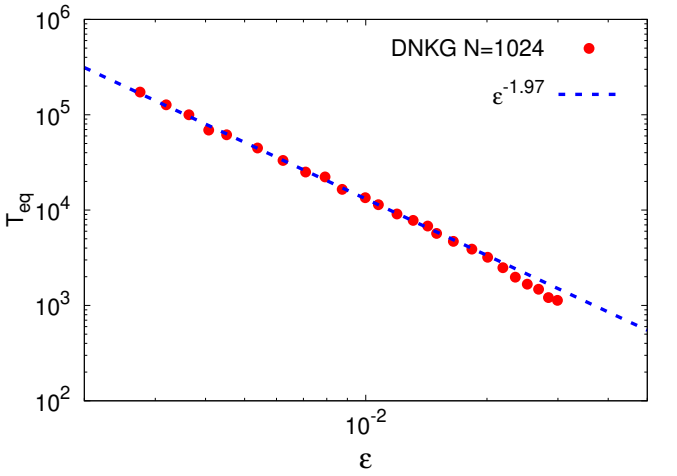
Before entering in the discussion, we summarise our numerical results here below:



(a) α-FPUT, $E = 16$, ensemble size 4096.



(b) β-FPUT, $E = 16$, ensemble size 1024.



(c) DNKG, $m = 1$, $E = 3.2$, ensemble size 4096.

Figure 5. T_{eq} as a function of ϵ for the α, β-FPUT and DNKG ($m = 1$) systems with $N = 1024$.

Thermodynamic limit:	**Discrete regime limit**:
WT prediction: $\nu = -2$	WT prediction $\nu = -4$
DNKG: $\nu = -1.97 \pm 0.04$	DNKG: $\nu = -3.79 \pm 0.02$
β-FPUT: $\nu = -2.00 \pm 0.08$	β-FPUT: $\nu = -3.72 \pm 0.07$
α-FPUT: $\nu = -2.49 \pm 0.02$	α-FPUT: $\nu = -3.84 \pm 0.19$

7.1. Discussion on the Thermodynamic limit

We mention that the formal derivation of the Wave Kinetic equation is made by taking the large box limit ($N \to \infty$) and then the small nonlinearity assumption is made. By looking at the above fitting numbers, for the DNKG and β-FPUT systems the agreement with the prediction from WT theory is remarkable, despite the fact that the simulations are done for large but finite N. Any numerical simulation, by definition, is performed with finite N. In this respect we recall that, besides transferring energy, two are the main effects of nonlinearity: (i) nonlinear shift of the frequencies and (ii) broadening of the frequencies [36]. If the broadening is large enough, then two adjacent frequencies can be connected and the system can be considered as continuous (even if the simulation is performed with a finite number of masses). The continuous limit assumed in the Wave Kinetic equation is therefore achieved numerically, with finite N, with a subtle balance between the Fourier spacing (which is proportional to $1/N$) and the broadening.

The correct recipe to be used in numerical simulations is still not fully understood theoretically; however, results show clearly that, for a given (large) N, if the nonlinearity is too small, then the discrete effects (lack of four-wave resonant interactions) may start taking place and consequently slopes steeper than predicted are observed. In our DNKG and β-FPUT simulations discussed in the manuscript, a proper balance between Fourier spacing and broadening is achieved. For the α-FPUT system, unfortunately the thermodynamic limit was never reached in our simulations. The reason is that the range in which simulations can be performed in the α-FPUT system corresponds to very weak nonlinearity, for which discrete effects start dominating. For larger nonlinearity (and large N), the probability of a blowup becomes very large [7], making it impossible to run simulations. In practice, the α-FPUT system would need a higher level of nonlinearity in order to show the expected scaling due to the continuous resonant manifold. More specifically, for $N = 1024$, the α-FPUT system starts showing very often blowups beyond $\epsilon \simeq 0.03$, making it difficult to obtain the ensemble averages that we use in our treatment. The thermodynamic limit is very difficult to reach in the α-FPUT and this is why a steeper slope is observed. One more thing should be considered: in order to derive the Wave Kinetic Equation, a canonical transformation that removes the three wave nonresonant interaction is performed. The convergence of such transformation is unknown in the thermodynamic limit. Note that for fixed N, if a simulation of the DNKG or β-FPUT would have been carried out for weaker nonlinearity, a steeper power law would have been observed because of discreteness and possible higher order effects.

In [3] the authors perform a large number of simulations mainly with the α+β model. For this model, our WT prediction for the equipartition time is $1/\epsilon^2$ when $\beta \neq \beta_T$; note that our ϵ is proportional to β. The result obtained in their simulation is $1/\epsilon^{2.25}$, which is definitely not far from the WT prediction. We may attribute this behaviour to an effect of discreteness in their simulations at low energies. As mentioned before, for a given value of N, as the nonlinearity becomes small, the effect of discreteness starts emerging and this leads to a deviation form the WT prediction. Interestingly, the WT theory

is able to predict the slopes shown in Figure 5 of [3]: For $\beta = \beta_T$, $\gamma = \gamma_T$ and $\delta \neq \delta_T$ (γ and δ are the coefficients of higher order interactions in the one dimensional chains and the subscript T corresponds to the Toda lattice coefficients) of the the prediction in the thermodynamic limit from WT theory would be $1/\epsilon^4$ (it is a simple and straightforward calculation to show this). The physical reason is that for those parameters the first nontrivial resonant interaction is the six-wave one (the coefficient in the collision integral calculated on the resonant manifold would be zero for lower-order resonances, namely five-wave and four-wave resonances). Numerical simulations in [3] show a slope of -3.95 and -3.97, which are very close to -4, the prediction from WT theory. For $\beta = \beta_T$ and $\gamma \neq \gamma_T$ then it is easy to show that the first nontrivial interactions are the five-wave interactions which, according to our theory, give a slope of -3. Numerical simulations show -2.98 and -3.05. We think that the agreement of our WT theory prediction with the numerical results provided in [3] is not a coincidence.

We also remark that in WT theory the time of thermalization is estimated from the kinetic equation using a dimensional argument; this implies that the collision integral should be of order one and the predicted time scale would then given by the power of the small parameter ϵ in front of the integral. When $\beta \simeq \beta_T$, the four-wave collision integral is not of order one (for $\beta = \beta_T$ it is actually zero). Interestingly, for $\beta \gg \beta_T$ the prediction of our theory is $1/\beta^2$, exactly what is observed in the simulations [3] (note that our small parameter ϵ is proportional to β).

At about the same time of the submission of the present manuscript, two paper have been submitted in the ArXiv, see [20, 21]. Those authors perform similar simulations to ours but including other models. Their results are consistent with ours. As in our case, some deviations from WT are observed for α-FPUT and for very low nonlinearity in the β-FPUT.

Concluding this part on the thermodynamic limit, we wish to stress that, despite not being fully rigorous, the Wave Turbulence theory offers a very good tool to predict the thermalization time scale in the broad regime consisting of the combination of weak nonlinearity and large-box limit, and its predictions apply to several results obtained from numerical simulations from different independent groups; the region of applicability of the theory (for example in terms of parameter space, discreteness, etc.) is still unclear and should be further investigated, both phenomenologically and rigorously. Further work should be performed for the α-FPUT which, apart from blowups, requires a canonical transformation before applying the WT theory. This could be problematic in the limit of small wave numbers.

7.2. Discussion on the discrete limit with finite N

While in the thermodynamic limit the boundary conditions should not matter, when N is not so large, boundary conditions may play a role. We must make clear from the beginning that in [3] and [4] the authors studied the FPUT system with fixed ends, i.e., in the case where all resonances are trivial, see [43]. In contrast, the FPUT system we study has nontrivial 4-wave resonances and 6-wave resonances: although in our case the resulting 4-wave resonant clusters of interacting modes are isolated and consequently lead to integrable systems, the resulting 6-wave resonant clusters are not isolated. They are typically connected via 4 common modes in the case of N even. This connectivity generically produces non-integrable systems. Therefore, from the point of view of energy transfers across Fourier modes in normal-form coordinates, the system we consider is much more active and thus *one expects to see differences* with respect to the system studied in [3, 4]. At the same time, it is difficult to assess quantitatively these differences, because of a lack of a rigorous theoretical

framework. Our scaling prediction for the time to equipartition relies on dimensional analysis of the kinetic equation, whose hypotheses do not necessarily hold in the discrete case.

Having made that clarification we now proceed to discuss our rationale. First, as mentioned previously, in the discrete case the hypotheses leading to the kinetic equation do not hold and therefore any prediction based on the kinetic equation is a conjecture. Second (not a conjecture), as mentioned above, in the discrete case 4-wave resonant interactions exist and are of umklapped type; they form isolated groups so these interactions are not efficient for thermalization. Therefore, for any specific value of N one applies formally a transformation of variables to normal-form coordinates, where the nonresonant terms are removed to obtain a discrete system containing 6-wave interactions (note that at this level, a thermodynamic limit would not make sense, because in such a limit, 4-wave interactions would dominate).

The normal-form coordinates satisfy a system of evolution equations where the only energy-transfer terms arise from 4-wave resonant interactions (forming isolated groups of 4 modes each) and 6-wave resonant interactions (which connect those isolated groups). With such a discrete system we may only conjecture that the time scale of interactions would be given by the square of the coefficient in front of the highest-order nonlinearity in the system (the prediction is made on spectrum and not amplitudes) This brings a scaling of -4. Now, numerical simulations show less steep values than predicted. In [3, 4] for the discrete case a stretched exponential is considered. Based on this, we tried to fit our data with a stretched exponential and the fitting parameter is somehow consistent with the one obtained in [4]. More specifically, we can fit a stretched exponential if we include in the fit large values of nonlinearity. In contrast, when we restricted the fit to small values of nonlinearity, a power law seemed to be more appropriate. We do not exclude that for lower nonlinearity, other effects, such as for example higher order resonances [20], may take place, changing the behaviour of the curve. Despite the fact that our predictions seem plausible and not far from numerics, at the moment it is very difficult to make a final statement on the scaling of these discrete weakly nonlinear systems and further work is definitely needed.

8. Conclusions

In this work, we have presented recent results on the dynamics of nonlinear chains, focusing on the thermalisation time. We have proposed to encompass the behaviour of all systems of this kind within the Wave Turbulence framework. In this context, the universal mechanism invoked to lead the systems to equipartition is the resonance among different modes, at least when generic initial conditions are considered. The scaling of the thermalization time T_{eq} as a function of the nonlinearity level of the system is obtained from the theory. The results show curves that can be fitted by power-laws for all the cases considered, with an exponent dependent on the order of the active resonances in the system. Although the point of view is rather different, present results are not in contradiction with the most recent and accurate estimates obtained also for particular low-wavenumber initial conditions [3]. Since the resonant manifold is significantly different in the discrete and thermodynamic $N \to \infty$ limit, we have considered the two cases separately, and they lead to two different scalings. Notably, large-size chains reach equipartition faster.

Our aim in this paper has been to present the results anticipated in works [36, 41, 42] in an unified manner to underline how our approach can be systematic, rather than dependent on the particular

features of the systems. In the thermodynamic limit, we have shown that there exists a resonant manifold of $2 \to 2$ waves. We have analysed numerically the thermodynamic limit $N \to \infty$ for the α and β-FPUT systems, as done previously for the DNKG chain. In the discrete case, for N a power of two (which excludes the existence of five-wave resonant interactions [5]), we have extensively verified that, for all the considered systems, the leading order interaction consists of six wave interactions. All the results seem to indicate a universal route to thermalization predicted. We have also shown that, besides the α and β-FPUT systems, also the DNKG equation is asymptotically integrable, if the expansion is truncated at four wave interactions.

Acknowledgments

M. O. has been funded by Progetto di Ricerca d'Ateneo CSTO160004. M.O. was supported by the "Departments of Excellence 2018-2022" Grant awarded by the Italian Ministry of Education, University and Research (MIUR) (L.232/2016). Simulations were run on GPU hardware provided at the OCCAM facility, University of Turin.

Conflict of interest

The authors declare no conflict of interest.

References

1. Benettin G, Christodoulidi H, Ponno A (2013) The Fermi-Pasta-Ulam problem and its underlying integrable dynamics. *J Stat Phys* 152: 195–212.
2. Benettin G, Livi R, Ponno A (2009) The Fermi-Pasta-Ulam problem: Scaling laws vs. initial conditions. *J Stat Phys* 135: 873–893.
3. Benettin G, Ponno A (2011) Time-scales to equipartition in the Fermi-Pasta-Ulam problem: Finite-size effects and thermodynamic limit. *J Stat Phys* 144: 793–812.
4. Berchialla L, Giorgilli A, Paleari S (2004) Exponentially long times to equipartition in the thermodynamic limit. *Phys lett A* 321: 167–172.
5. Bustamante MD, Hutchinson K, Lvov YV, et al. (2019) Exact discrete resonances in the Fermi-Pasta-Ulam-Tsingou system. *Commun Nonlinear Sci* 73: 437–471.
6. Bustamante MD, Kartashova E (2011) Resonance clustering in wave turbulent regimes: Integrable dynamics. *Commun Comput Phys* 10: 1211–1240.
7. Carati A, Ponno A (2018) Chopping time of the FPU α-model. *J Stat Phys* 170: 883–894.
8. Carati A, Galgani L, Giorgilli A, et al. (2007) Fermi-Pasta-Ulam phenomenon for generic initial data. *Phys Rev E* 76: 022104.
9. Chibbaro S, Dematteis G, Josserand C, et al. (2017) Wave-turbulence theory of four-wave nonlinear interactions. *Phys Rev E* 96: 021101.
10. Chibbaro S, Dematteis G, Rondoni L (2018) 4-wave dynamics in kinetic wave turbulence. *Phys D* 362: 24–59.

11. Chirikov BV (1979) A universal instability of many-dimensional oscillator systems. *Phys Rep* 52: 263–379.
12. Choi Y, Lvov YV, Nazarenko S (2004) Probability densities and preservation of randomness in wave turbulence. *Phys Lett A* 332: 230–238.
13. Choi Y, Lvov YV, Nazarenko S (2005) Joint statistics of amplitudes and phases in wave turbulence. *Phys D* 201: 121–149.
14. Düring G, Josserand C, Rica S (2017) Wave turbulence theory of elastic plates. *Phys D* 347: 42–73.
15. Dyachenko AI, Lvov YV, Zakharov VE (1995) Five-wave interaction on the surface of deep fluid. *Phys D* 87: 233–261.
16. Eyink GL, Shi YK (2012) Kinetic wave turbulence. *Phys D* 241: 1487–1511.
17. Falkovich G, Lvov VS, Zakharov VE (1992) *Kolmogorov Spectra of Turbulence*. Berlin: Springer.
18. Fermi E, Pasta J, Ulam S (1955) Studies of the nonlinear problems. *Technical report I, Los Alamos Scientific Lab Report* No. LA-1940.
19. Ford J (1961) Equipartition of energy for nonlinear systems. *J Math Phys* 2: 387–393.
20. Fu WC, Zhang Y, Zhao H (2018) Universality of energy equipartition in one-dimensional lattices. *arXiv preprint arXiv:1811.05697*.
21. Fu WC, Zhang Y, Zhao H (2019) Universal law of thermalization for one-dimensional perturbed toda lattices. *arXiv preprint arXiv:1901.04245*.
22. Fucito F, Marchesoni F, Marinari E, et al. (1982) Approach to equilibrium in a chain of nonlinear oscillators. *J Phys* 43: 707–713.
23. Gallavotti G (2008) *The Fermi-Pasta-Ulam Problem: A Status Report*. Springer.
24. Henrici A, Kappeler T (2008) Results on normal forms for FPU chains. *Commun Math Phys* 278: 145–177.
25. Arnold VI (1963) Small denominators and problems of stability of motion in classical and celestial mechanics. *Russ Math Surv* 18: 85–191.
26. Izrailev FM, Chirikov BV (1966) Statistical properties of a nonlinear string. *Sov Phys Dokl* 11: 30–32.
27. Janssen P (2004) *The Interaction of Ocean Waves and Wind*. Cambridge: Cambridge University Press.
28. Moser JK (1962) On invariant curves of area-preserving mappings of an annulus. *Nachr Akad Wiss Göttingen Math Phys kl* 166: 1–20.
29. Kartashova E (2007) Exact and quasiresonances in discrete water wave turbulence. *Phys Rev Lett* 98: 214502.
30. Khinchin A (1949) *Mathematical Foundations of Statistical Mechanics*. Courier Corporation.
31. Kolmogorov AN (1954) On the conservation of conditionally periodic motions under small perturbation of the Hamiltonian. *Dokl Akad Nauk SSR* 98: 527–530.

32. Krasitskii VP (1994) On reduced equations in the Hamiltonian theory of weakly nonlinear surface waves. *J Fluid Mech* 272: 1–20.
33. Landau LD, Lifshitz EM, Pitaevskii LP (1980) *Statistical Physics, Part I*. Oxford: Pergamon.
34. Laurie J, Bortolozzo U, Nazarenko S, et al. (2012) One-dimensional optical wave turbulence: Experiment and theory. *Phys Rep* 514: 121–175.
35. Lebowitz JL (1993) Boltzmann's entropy and time's arrow. *Phys Today* 46: 32–32.
36. Lvov YV, Onorato M (2018) Double scaling in the relaxation time in the β-fermi-pasta-ulam-tsingou model. *Phys Rev Lett* 120: 144301.
37. Matkowski J (2011) Subadditive periodic functions. *Opusc Math* 31: 75–96.
38. Nazarenko S (2011) *Wave Turbulence*. Springer Science Business Media.
39. Newell AC (1968) System of random gravity waves. *Rev Geophys* 6: 1–31.
40. Newell AC, Rumpf B (2011) Wave turbulence. *Annu Rev Fluid Mech* 43: 59–78.
41. Onorato M, Vozella L, Proment D, et al. (2015) Route to thermalization in the α-Fermi-Pasta-Ulam system. *Proc Natl Acad Sci* 112: 4208–4213.
42. Pistone L, Onorato M, Chibbaro S (2018) Thermalization in the discrete nonlinear klein-gordon chain in the wave-turbulence framework. *Europhys Lett* 121: 44003.
43. Rink B (2006) Proof of Nishida's conjecture on anharmonic lattices. *Commun Math Phys* 261: 613–627.
44. Spohn H (2006) The phonon boltzmann equation, properties and link to weakly anharmonic lattice dynamics. *J Stat Phys* 124: 1041–1104.
45. Yoshida H (1990) Construction of higher order symplectic integrators. *Phys Lett A* 150: 262–268.
46. Zabusky NJ, Kruskal MD (1965) Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Phys Rev Lett* 15: 240–243.
47. Zakharov VE, Schulman EI (1988) On additional motion invariants of classical Hamiltonian wave systems. *Phys D* 29: 283–320.
48. Zakharov VE, Schulman EI (1991) Integrability of nonlinear systems and perturbation theory. In: *What Is Integrability?* Springer Series in Nonlinear Dynamics. Berlin: Springer.
49. Zakharov VE, Odesskii AV, Onorato M, et al. (2012) Integrable equations and classical s-matrix. *arXiv preprint arXiv:1204.2793*.

Appendix

Let's define the function $f_k = 2|\sin(k/2)|$ as the linear dispersion relation for the α and β- FPUT models which corresponds to the general dispersion relation, ω_k, when $m = 0$. In [5] it has been shown that all processes $X \to 1$ are forbidden for the dispersion relation f_k.

$$f_{k_1} + f_{k_2} + ... + f_{k_X} \geq f_{k_1+k_2+...+k_X}, \tag{8.1}$$

where the equality holds only for wave numbers equal to $2l\pi$, with $l \in \mathbb{Z}$. Here we show that

$$\omega_{k_1} + \omega_{k_2} + ... + \omega_{k_X} > \omega_{k_1+k_2+...+k_X} \tag{8.2}$$

for any value of m >0. In order to show it, we square (8.2) and, after re-arranging, we get:

$$F(k_1, k_2, .., k_X; m) = (X-2)m + f_{k_1}^2 + f_{k_2}^2 + .. + f_{k_X}^2 - f_{k_1+k_2+..+k_X}^2 + \sum_{i=1}^{X} \omega_{k_i} \sum_{j=1}^{X} \omega_{k_j} - \sum_{j=1}^{X} \omega_{k_j}^2 > 0 \quad (8.3)$$

The function f_k is m-independent, while ω_k with $m \neq 0$ is proportional to m. The terms of the type $-\omega_{k_i}^2$ all cancel out. It is easy to observe that $F(k_1, k_2, .., k_X; m)$ is an increasing monotonic function of m. For $m = 0$, because of the inequality in (8.1), the $F(k_1, k_2, .., k_X; 0) \geq 0$; then, for its monotonicity, for any $m > 0$, $F(k_1, k_2, .., k_X; 0) > 0$. This proves that there are no resonances in the DNKG model of the type $X \to 1$.

AIMS Press

Chapter 4

nlchains: A fast and accurate time integration of 1-D nonlinear chains on GPUs

The simulations that we performed to verify our arguments were quite computationally intensive. For most of our datasets, an ensemble size in the thousands was used, and a large number of time steps was needed to observed thermalization (in some cases up to 10^9). The range of nonlinearities that we could explore was often bounded by the wall-clock duration of simulations.

A large part of my time during the PhD was devoted to developing efficient code that could handle large simulations of nonlinear lattices. We opted for using a parallel processing accelerator. Since I already had experience with Graphics Processing Units (GPU) programming, I implemented our simulations on GPU hardware. A GPU is a kind of processor that is usually employed for real-time graphics. Over the years, GPUs have become very complex machines with the capability of operating parallel operations on very large datasets. In the last decade, this power was recognized by the scientific community, and today GPUs (and other types of specialized hardware accelerators) are part of every HPC cluster.

A GPU computation is efficient only when the operations to be performed are uniform across a large dataset. This is because GPUs are mostly a SIMD (Single Instruction Multiple Data) architecture. Since GPUs are not general-purpose processors, writing a program for a GPU requires knowing and handling quite a lot of low-level details about the hardware. In traditional CPU programming, the programmer is often not concerned about the details of the hardware where the program will run. There are frameworks that can hide some of the complexities that come with GPU programming, but in general they add layers of abstraction that are very detrimental to the final performance.

The kind of simulations that we had to run are an obvious candidate for parallelization over GPU. In fact, ensemble averages must be taken, which means that

for most of the time the individual simulations do not interact. This scheme is very easy to implement on a SIMD architecture. The only time when they interact is at the time of an average, since it is useful to monitor the number of particles n_i, and the entropy to monitor equipartition, but the sampling can be quite infrequent.

We decided to publish the software as a proper peer-reviewed paper, because we believe my code can be reused by other groups. During the review process, we were also asked to verify the GPU code against a traditional CPU implementation, both for its correctness, and to prove the claimed performance gains. For this reason I implemented the same algorithms also as a traditional CPU code. In order to make the comparison fair I made use of the SIMD capabilities of recent CPUs (the AVX vector extensions), and made my program fully multithreaded. The performance comparison turned favourable for the GPU implementation compared to a 24 CPU threads implementation, for a minimum of a five-fold performance increase. Obviously, the end goal of optimizing performance is to spend less time in simulations and less money on renting HPC computing hours, a metric that in turn depends on which particular HPC cluster is available to a scientific group. In our case, we used the OCCAM facility at the University of Turin. Taking into account the relative cost of the resources used in the comparison, the GPU implementation provided a solution at least 6 folds more cost effective than the CPU implementation.

There is also a second reason besides reusability for the publication of this software. Very often, only the results are published in a scientific paper, but the code that is used to perform the simulations is often developed from the researchers themselves, and it is generally unavailable, or available only at request through informal channels and in a very non-redistributable form (it is bad code). In order for the scientific results to be verifiable, all logical steps and procedures that were used to obtain them should be known. There is really no reason to treat the code that provided the numerical result, which are often central to the scientific reasoning, as secondary material. The reproducibility of what is stated in a paper is of paramount importance, because repeatability is essential in science. By publishing the software that we used to support our claims, it is much easier for anyone else to verify them.

SoftwareX 10 (2019) 100255

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original software publication

nlchains: A fast and accurate time integration of 1-D nonlinear chains on GPUs

L. Pistone *, M. Onorato

Dipartimento di Fisica, Università di Torino, via Pietro Giuria 1, 10125, Torino, Italy

ARTICLE INFO

Article history:
Received 27 February 2019
Received in revised form 29 May 2019
Accepted 31 May 2019

Keywords:
FPU
Nonlinear chain
GPU
CUDA

ABSTRACT

We present nlchains, a software for simulating ensembles of one-dimensional Hamiltonian systems with nearest neighbor interactions. The implemented models are the α-β Fermi–Pasta–Ulam–Tsingou model, the discrete nonlinear Klein–Gordon model with equal or site-specific masses, the Toda lattice and the discrete nonlinear Schrödinger equation. The integration algorithm in all cases is a symplectic sixth order integrator, hence very accurate and suited for long time simulations. The implementation is focused on performance, and the software runs on graphical processing unit hardware (CUDA). We show some illustrative simulations, we estimate the runtime performance and the effective scaling of the cumulative error during integration. Finally, we give some basic pointers to extend the software to specific needs.

Code metadata

Current code version	v1.1.0
Permanent link to code/repository used for this code version	https://github.com/ElsevierSoftwareX/SOFTX_2019_56
Legal Code License	MIT
Code versioning system used	git
Software code languages, tools, and services used	C++14, CUDA, MPI, CMake
Compilation requirements, operating environments & dependencies	Linux, CUDA SDK, GCC, Boost, Armadillo (and one compatible LAPACK library)
Link to developer documentation/manual	https://bitbucket.org/pisto/nlchains/src/master/README.md
Support email for questions	lorenzo.pistone@unito.it

Software metadata

Current software version	v1.1.0
Permanent link to executables of this version	https://bitbucket.org/pisto/nlchains/downloads/nlchains-gite6c2001-opt64
Legal Software License	MIT
Computing platforms/Operating Systems	Linux
Installation requirements & dependencies	CUDA $\geq$ 10.0, NVIDIA card (compute capability $\geq$ 3.5, 6.0 or 7.0), Ubuntu 18.04 environment (see `Dockerfile-ubuntu1804-cuda100` in the source tree)
If available, link to user manual – if formally published include a reference to the publication in the reference list	https://bitbucket.org/pisto/nlchains/src/master/README.md
Support email for questions	lorenzo.pistone@unito.it

1. Introduction and motivation

Nonlinear chains, that is systems of points with nonlinear nearest-neighbor interactions, have an important role in the un-

* Corresponding author.
E-mail address: lorenzo.pistone@unito.it (L. Pistone).

https://doi.org/10.1016/j.softx.2019.100255

derstanding of the basic processes of nonlinear physics and statistical mechanics. Despite being in general toy models, or at best crude approximations of real physical systems, they show a rich phenomenology, and even basic questions such as the thermalization dynamics or heat transport law often do not have a definitive answer. Not by chance, one of the first numerical experiment in physics was the Fermi–Pasta–Ulam–Tsingou, that is the renowned FPUT experiment [1,2]. The aim of the experiment was to observe the thermalization of a linear chain of masses and springs, when a small nonlinearity is added. The apparent paradox of the FPUT experiment is that thermalization in the FPUT system cannot be attained in a short time, and so it could not be observed with the computers of the fifties.

Recently, we and collaborators have applied tools of Wave Turbulence [3,4], a statistical description of weakly interacting wave systems, to seek universal traits in the thermalization dynamics of nonlinear chains, [5–9] and also [10,11]. To this end, we needed a software to run large simulation of these systems. These numerical experiment are often time-consuming because thermalization oftentimes requires a long time compared to the wave periods of the corresponding linearized system, and also because large ensembles of realizations of the same chain are needed to extract meaningful statistics. For these reason, the code had to be written with performance but also accuracy in mind, because the Hamiltonian structure of the system needs to be preserved for long times. We could not find a properly published software that could fit our needs. Recently, an implementation of the FPUT system (possibly extendible to other model) has been published [12], however it was not designed with performance as a strong point, but rather it has a pedagogical value, stressing on fast prototyping and code readability. Our contribution here is the software that we used in our research, that is the implementation on graphical processing unit hardware (GPUs) of the time integration for several one-dimensional nonlinear chains. The principal value of this software is the effort that we have spent in tuning its performance.

2. The implemented algorithms

In Table 1 we list the models that are already implemented in *nlchains*. The q_j and p_j variables are the conjugate coordinates and momenta, while ψ_j is a complex variable, for j an index that runs from 0 to $N-1$, with N the length of the chain. We have included the original α-β Fermi–Pasta–Ulam–Tsingou model (FPUT), the discrete nonlinear Klein–Gordon model with equal masses (DNKG) and disordered, site-specific masses (dDNKG), the Toda lattice and the discrete nonlinear Schrödinger equation (DNLS). The specific version of the Toda lattice potential has been chosen so that it is tangent to the α-FPUT model, that is the potential energy between two adjacent masses $V(q_{j+1}-q_j) = V(\Delta q_j)$ is $V_{\text{Toda}}(\Delta q_j) = V_{\text{-FPUT}}\alpha(\Delta q_j) + O(\Delta q_j^4)$. Extending the software to other systems with similar potentials is expected to be easy. The parameters α, β, m and m_i can be set by the user.

2.1. The integration scheme

As mentioned in the introduction, the choice in the algorithm was driven by the need of a good performance but also accuracy in the conservation of the Hamiltonian structure of the nonlinear chains. We chose the 6th order symplectic Yoshida integrator [13]. This algorithm suited our needs for the following reasons. It allows for very long time simulation, as being symplectic it avoid secular growth of conserved quantities such as the Hamiltonian. It is explicit, which makes it direct and easy to implement. We chose the sixth order, as it resulted in an optimal trade-off between computational speed and accuracy in our field of research. Higher order schemes of the same type can be implemented with trivial modifications of the source code (details will be given in a later section).

For reference, we report here the principle of this integration scheme. Since all the systems considered are Hamiltonian, the formal solution to the time evolution of some initial state $z = (q_0, \dots, q_{N-1}, p_0, \dots, p_{N-1})$ (or $z = (\psi_0, \dots, \psi_{N-1})$ for the DNLS model) is given by the Poisson bracket

$$\dot{z} = \{z, H(z)\}. \tag{1}$$

Now the above equation can be formally solved by introducing the differential operator $D_H = \{\cdot, H(\cdot)\}$, to obtain

$$\dot{z} = D_H z,\ z(\delta t) = e^{\delta t D_H} z(0). \tag{2}$$

In general it is not possible to give an explicit form of $e^{\delta t D_H}$, as that coincides with integrating the system. However, since the Poisson brackets are bilinear, it is possible to take advantage of the fact that the Hamiltonians in consideration are the sum of different terms. If one splits the Hamiltonians in two contributions, $H = H_A + H_B$, then the differential operator also splits in two corresponding parts, $D_H = D_A + D_B$, to get

$$z(\delta t) = e^{\delta t D_H} z(0) = e^{\delta t (D_A + D_B)} z(0). \tag{3}$$

In general, D_A and D_B do not commute, so it is not possible to write $e^{\delta t D_H} = e^{\delta t D_A} e^{\delta t D_B}$, however it can be shown that it is possible to generate a scheme of the type

$$e^{\delta t D_H} = e^{c_1 \delta t D_A} e^{d_1 \delta t D_B} e^{c_2 \delta t D_A} e^{d_2 \delta t D_B} \dots e^{c_k \delta t D_A} e^{d_k \delta t D_B} + O(\delta t^x) \tag{4}$$

where the coefficients c_k and d_k are real and x is some positive integer, the order of the integration scheme. Given that, when the splitting $H = H_A + H_B$ is chosen carefully so that H_A and H_B are individually integrable, it is possible to generate a fully explicit integrations scheme.

For the real models, the most natural choice is to split the Hamiltonian into kinetic and potential energy, that is $H_A = \frac{1}{2} p_j^2$ and $H_B = V(q_{j+1} - q_j)$. For the DNLS model, we follow [14] and we split the Hamiltonian in the linear and nonlinear contributes, that is $H_A = \sum \frac{1}{2}\beta |\psi_j|^4$ and $H_B = \sum |\psi_{j+1} - \psi_j|^2$. The solution to the c_k and d_k of Eq. (4) that we hard-coded in the software is taken from [13], it is of the 6th order and the values of the coefficients c_k and d_k are

$$\begin{aligned}\{c_0, \dots, c_7\} = \{&0.392256805238780, 0.510043411918458,\\ &-0.471053385409757,\\ &0.068753168252518, 0.068753168252518,\\ &-0.471053385409757,\\ &0.510043411918458, 0.392256805238780\}\end{aligned} \tag{5}$$

and

$$\begin{aligned}\{d_0, \dots, d_7\} = \{&0.784513610477560, 0.235573213359357,\\ &-1.177679984178870,\\ &1.315186320683906, -1.177679984178870,\\ &0.235573213359357,\\ &0.784513610477560, 0\}.\end{aligned} \tag{6}$$

Note that the coefficient d_7 is 0, hence it is possible to merge the last integration of one step $e^{\delta t c_7 D_A}$ with the first integration of the next step $e^{\delta t c_0 D_A}$. We expect that if the user needs higher order integrators (e.g. 8th), modifying the software to this end should be quite trivial.

In concluding this section, we remark that for all the models except DNLS, the integration is performed in physical space. For the DNLS model however, because the linear and nonlinear

Table 1
The list of implemented models in *nlchains*. For real models, q_j and p_j are the conjugate coordinate and momentum at index $j \in [0, N)$ with N the length of the chain ($\dot{q}_j = p_j$). For DNLS, ψ_j is a complex variable, and i is the imaginary unit.

Subprogram	Equation of motion and Hamiltonian density
DNKG	$\dot{p}_j = q_{j-1} - 2q_j + q_{j+1} - mq_j + \beta q_j^3$ $H_j = \frac{1}{2}p_j^2 + \frac{1}{2}\left(q_{j+1} - q_j\right)^2 + \frac{1}{2}mq_j^2 + \frac{1}{4}\beta q_j^4$
dDNKG	$\dot{p}_j = q_{j-1} - 2q_j + q_{j+1} - m_j q_j + \beta q_j^3$ $H_j = \frac{1}{2}p_j^2 + \frac{1}{2}\left(q_{j+1} - q_j\right)^2 + \frac{1}{2}m_j q_j^2 + \frac{1}{4}\beta q_j^4$
FPUT	$\dot{p}_j = \left(q_{j-1} - 2q_j + q_{j+1}\right)\left(\alpha\left(q_{j+1} - q_{j-1}\right) + 1\right) +$ $+ \beta\left(\left(q_{j+1} - q_j\right)^3 - \left(q_j - q_{j-1}\right)^3\right)$ $H_j = \frac{1}{2}p_j^2 + \frac{1}{2}\left(q_{j+1} - q_j\right)^2 + \frac{1}{3}\alpha\left(q_{j+1} - q_j\right)^3 + \frac{1}{4}\beta\left(q_{j+1} - q_j\right)^4$
Toda	$\dot{p}_j = \frac{1}{2\alpha}\left(e^{2\alpha\left(q_{j+1} - q_j\right)} - e^{2\alpha\left(q_j - q_{j-1}\right)}\right)$ $H_j = \frac{1}{2}p_j^2 + \frac{1}{4\alpha^2}\left(e^{2\alpha\left(q_{j+1} - q_j\right)} - 2\alpha\left(q_{j+1} - q_j\right) - 1\right)$
DNLS	$i\dot{\psi}_j = -\left(\psi_{j-1} - 2\psi_j + \psi_{j+1}\right) + \beta\psi_j\left\lvert\psi_j\right\rvert^2$ $H_j = \lvert\psi_{j+1} - \psi_j\rvert^2 + \frac{1}{2}\beta\lvert\psi_j\rvert^4$

sub-Hamiltonians are diagonal in Fourier and physical space respectively, the integration scheme become essentially a refined split-step scheme [15], and the time evolution of the linear part is performed in Fourier space, as suggested (but not implemented) in [14]. This has the side effect that a large number of Fast Fourier Transforms (FFTs) are performed in integrating the DNLS model, that is 14 FFTs per step: the number of non-zero d_k coefficients, doubled to account for a direct and inverse FFT to go back and forth from physical space. As we will explain later, this has important consequences in the numerical accuracy and speed of the algorithm.

2.2. Other calculated quantities

In order to observe recurrence and equipartition, the software also calculates with a user-settable interval the average energy per eigenstate of the linearized system (that is $\alpha = 0$ and $\beta = 0$ in Table 1). For the DNKG, FPUT and Toda systems, it can be shown that the eigenstates of the linearized systems are the so called normal modes,

$$a_k = \frac{1}{\sqrt{2\omega_k}}\left(\widetilde{p}_k - i\omega_k\widetilde{q}_k\right) \tag{7}$$

where the $\widetilde{q}_k$ and $\widetilde{p}_k$ are the discrete Fourier transform of the q_j and p_j, and ω_k is the dispersion relation of the system, that is $\omega_k = \sqrt{m + 4\sin\left(\pi k/N\right)^2}$. For the DNLS model, the normal modes a_k are simply the discrete Fourier transform $\widetilde{\psi}_k$ of the physical space variables ψ_j, and the dispersion relation is $\omega_k = 4\sin\left(\pi k/N\right)^2$. In all these cases the energy per mode is defined as

$$e_k = \omega_k\langle\lvert a_k\rvert^2\rangle \tag{8}$$

with $\langle\cdot\rangle$ being the average over the ensemble. For the dDNKG model, the eigenstates v_k and corresponding eigenvalues ω_k^2 are calculated numerically from the matrix representation of the system of ordinary differential equations that correspond to the equations of motion of the chain. The energy per mode is then obtained from and explicit projection of the coordinates and momenta vectors $q = (q_0, \ldots, q_{N-1})$ and $p = (p_0, \ldots, p_{N-1})$ over the eigenstates,

$$e_k = \langle v_k \cdot \left(\omega_k^2 q_k + p_k\right)\rangle/2. \tag{9}$$

From the average energy per mode, the software calculates and outputs the associated information entropy,

$$S_{\text{inf}} = \sum_{k=0}^{N-1} e'_k \log(e'_k),\ e'_k = \frac{N}{E}e_k, \tag{10}$$

where N is the length of the chain and $E = \sum e_k$ the total linear energy, and the Wave Turbulence entropy

$$S_{\text{WT}} = -\sum_{k=0}^{N-1}\log(e'_k). \tag{11}$$

Traditionally, Eq. (10) has been used to monitor the route to thermalization of the FPUT problem, but in the framework of Wave Turbulence only Eq. (11) has the properties of an entropy function, that is it can be proven that statistically it is a monotonic decreasing function in time. However, it can be shown [8] that both these two entropies are greater than zero out of thermal equilibrium, and zero at perfect equipartition, and they are essentially equivalent for the purpose of monitoring the route to thermalization.

3. Software architecture

The software is packaged in a stand-alone Linux executable. Build requirements and compilation steps are documented in the readme (file `README.md`) that comes with the sources. An important detail of the building process is that all models except DNLS are optimized at compile-time for a specific chain length. A warning is generated if a build of *nlchains* is launched for a value of the chain length that does not correspond to the optimized value, but the computation is carried out anyway. For more details we refer the reader to Section 5.

The simulation is set up with a number of command line arguments. The invocation in the shell is in the following form:

```
[<MPI launcher>] nlchains <model> \\
                          <common options> \\
                          <model specific options>
```

The executable can be run as-is to run on a single GPU on the current host, or through MPI to split the ensemble of realizations on different GPUs. The user should assign to each compute node the same number of MPI processes as the number of GPU attached to the node itself. When running through MPI, the software expects the presence of a shared filesystem, because full-state dumps are written on disk with the MPI shared I/O facilities.

The GPU code is written in CUDA, and the host code is written in C++14, hence a NVIDIA GPU card (minimum compute capability 3.0) is required. Please note that the software is implemented in floating point double precision, and so a card of the Tesla line is suggested, because consumer-level cards are largely limited in the double precision performance.

Table 2
List of the output files. The value of *prefix* is set with the commmand line option `-p` (see Table 3)

Filename	Description
prefix-*step*	Full dump of the ensemble state
prefix-linenergies-*step*	Energy per eigenstate
prefix-entropy	List of entropies
dDNKG only:	
prefix-eigenvectors	eigenvectors
prefix-omegas	pulsation of the eigenvectors

Table 3
Command line options that are common to all models.

Argument	Description
`-i`	initial state of the ensemble
`-n`	length of the chain
`-c`	number of realizations in the ensemble
`--dt`	time step
`-b`	number of time steps per kernel invocation
`-s`	number of time steps to run
`-p`	prefix for output filenames

For the purpose of performance comparison, and as a fallback in the case that no GPU is available on the host, all the models also have a CPU-only implementation. To launch it, simply append to the model name `-cpu`, for example `DNLS-cpu`. The implementation is single-threaded but multiple cores and nodes can be used for parallelization in the same way that the GPU implementation can be launched on multiple GPUs. The implementation also makes use of advanced SIMD instructions present in modern processors (SSE, AVX and AVX-512), and as such it should be regarded competitive in performance. This implementation is used here to compare to the performance of the GPU implementation (see Section 5.2). However, it is not the main message of this paper to present a CPU implementation. The interested reader is referred to the implementation notes attached to this paper as supplementary material for further information.

3.1. File formats

The format of the inputs and outputs is raw binary. The list of output files is shown in Table 2.

The initial state of the ensemble, and the full state dumps are in the same format, that is as a C++ array `double[C][N][2]`, where `C` is the ensemble size, `N` is the chain length, and the last two-elements array contains the conjugate coordinate q_j and momentum p_j of the element in the chain. For the DNLS model, the format is `std::complex<double>[C][N]`, that is the real and imaginary part of the element in the chain take the places of the conjugate coordinate and momentum.

The energy per linear mode dumps is simply a plain array of doubles, `double[N]`. Note that when a discrete Fourier transform is involved in the calculation of the linear energies (all cases except dDNKG), we use the FFTW [16] convention (`FFTW_FORWARD`) for the sign of the exponent.

The list of entropies has the format `double[][3]`, where each triplet contains the absolute time (the step number times the timestep), the Wave Turbulence entropy of Eq. (11) and the information entropy of Eq. (10).

For the DNKG model, two additional files are created: the eigenvectors file in format `double[N][N]` (the first index being the index of the eigenvector), and the square root of the corresponding eigenvalues (corresponding to the pulsation of the eigenstates) as a `double[N]`.

3.2. Command line options

The first argument to *nlchains* is one of the subprogram listed in Table 1. If no other argument is printed, the list of arguments applicable to the model is printed, together with a short description. In Table 3 we describe the most important command line options that are common to all models and are necessary to run a simulation. Other common switches are available (such as for terminating when a given entropy threshold is reached), but for brevity refer to the readme available in the source tree.

The model parameters α, β, m and m_i (see Table 1) can be set respectively with `--alpha`, `--beta`, `-m`; for the DNKG model `-m` takes a value, for the dDNKG it takes a filename with a list of mass values in the format `double[N]`.

Note that the time granularity of the dumps and the entropy calculation is equal to the value of `-b`. A larger interval for the space-consuming full state and linear energy dumps can be set with the option `--dump_interval`.

4. Illustrative examples and accuracy

As an example of the use case of this software, we show in Fig. 1 the interesting case of the Toda lattice dynamics versus the α-FPUT dynamics, with the same initial state in terms of linear wave modes and the corresponding entropy curves given by Eq. (11). The initial state of the ensemble (empty circles in the top half of Fig. 1) has been initialized with random values for the energy per linear mode, the same set for all realizations in the ensemble and rescaled to have $E = 1$, but each realization had a different phase of the normal modes a_k. This scheme ensures that all realization have the same initial linear energy. The value of α is 0.5, so that the nonlinear terms of the Hamiltonians are small compared to the linear part, because in this regime it is easier to observe the equipartition of the linear modes. The timestep is set to $\delta t = 0.1$ and the total number of steps is $2 * 10^7$. We can see that for the Toda lattice the entropy (dashed line in the bottom half of Fig. 1) quickly settles to a value not far from the initial one, and that clearly signals that the system is not thermalized, while for the α-FPUT system (solid line) the system eventually reaches equipartition. This can be appreciated from direct inspection of the energy per mode at the final state: in the top half of Fig. 1, the filled circles are the energy per mode of the α-FPUT chain at equipartition (minus some statistical fluctuation due to the finite size of the ensemble), while empty squares are the final state for the Toda lattice. For more elaborate examples we refer the interested reader to the papers [7,8].

In Fig. 2 we show the scaling of the accuracy of the simulation as a function of the step size. In order to measure the accuracy, we control the value of a know exact integral of motion, that is the value of the Hamiltonian $H(t)$, or the total energy. The simulation is then run for a fixed total time $T = 100000$ (arbitrary units), but with a different time step, and hence a different total number of steps. The initial state of each simulation is initialized in a similar way to the data of Fig. 1. The accuracy is then calculated as the average relative deviation from the initial value of the energy,

$$\epsilon = \langle |H(T) - H(0)|/H(0) \rangle. \tag{12}$$

Since we use a sixth order symplectic integrator, it is expected that the scaling of the error is of the type $\epsilon = O(\delta t^6)$. We see in Fig. that we get the expected scaling of the error for the models FPUT, DNKG, Toda and dDNKG, up to a saturation around $\epsilon \sim 10^{-13}$. For the model DNLS, we get a saturation much earlier, at around $\epsilon \sim 10^{-9}$. This can be explained by the fact that the DNLS algorithm as we mentioned earlier requires 14 Fast Fourier Transforms (FFTs) for each single time step, hence

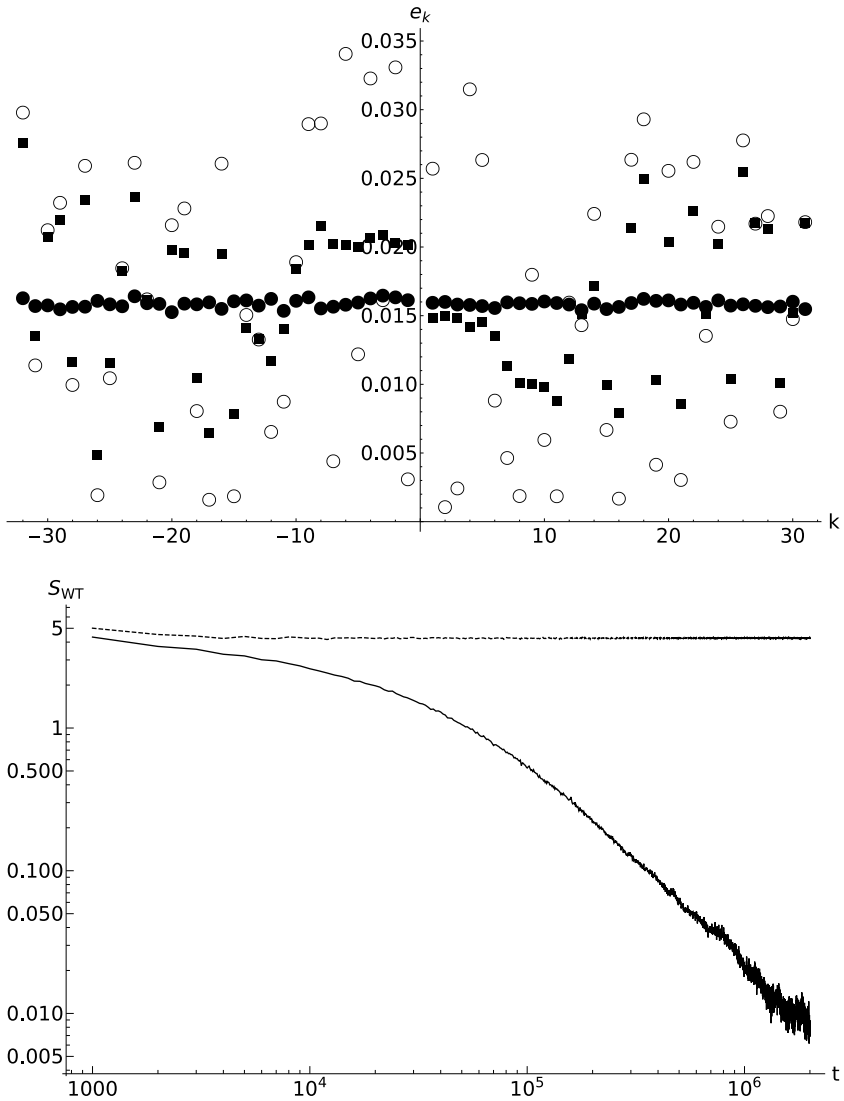


Fig. 1. A run from the same initial state of the Toda and α-FPUT model, for the same value of α, for a chain with length 64 and an ensemble size of 4096. Top figure, the energy per mode, Eq. (7): initial state ◯, thermalized α-FPUT chain •, final state for the Toda lattice ■. Bottom figure, the entropy curves, solid and dashed for the α-FPUT and Toda models respectively.

the data undergoes many more additions and multiplications compared to the other models, and numerical errors due to the finite accuracy of machine numbers accumulates faster. In fact, the behavior of ϵ when δt is small is of the type $\epsilon = O(\delta t^{-1})$, that is it is proportional to the number of FFTs. This is further corroborated by the fact that when the nonlinear parameter is set to zero, that is when the symplectic integration scheme is no longer a source of error as it becomes exact, the scaling of the error is still $\epsilon = O(\delta t^{-1})$ even for large values of δt.

All the data shown in this Section is attached as supplementary material.

5. Implementation and software extendibility

The software does not have an interface for adding new models, and essentially all the available settings can be accessed through command line options. This is a design choice because performance has been the top priority in developing this code: in the case of GPU programs good performance is in general attained with tight coupling between host and GPU code, and avoiding unnecessarily generalization (such as allowing to specify new models through virtual functions). In lieu of a runtime flexibility, *nlchains* is designed to be easily modifiable and extendible. In support of the source code comments, a manual with extensive implementation notes, a description of the internal utility interfaces and examples for their usage is provided in the supplementary material of this paper (also as a stand-alone document in the source code tree, `supporting-material/documentation/implementation-notes.pdf`). We refer the user who needs to adapt *nlchains* to his or her needs to this manual, to avoid cluttering this paper with implementation details.

5.1. Notes on the implementation for the end user

We mention here only some implementation details that are useful even for the user who does not intend to modify the functionality of *nlchains*.

All models except DNLS have three different implementations of the GPU kernels. One implementation is optimized for a chain length less than 32 (`move_chain_in_thread`), another one is generic for all chain lengths (`move_split`), and the last one (`move_chain_in_warp`) must be tuned at compile time for a specific chain length greater or equal to 32. The instruction on

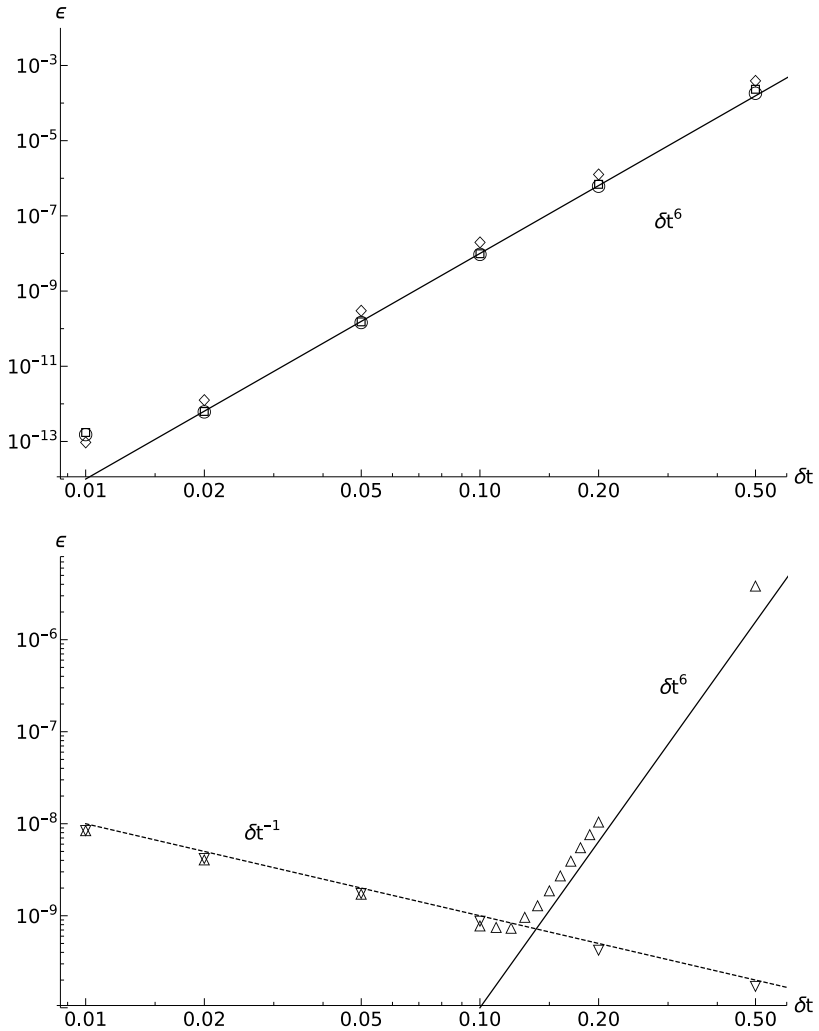


Fig. 2. The accuracy of the integration algorithms and their scalings. Top figure: FPUT ◯, Toda □, DNKG ♢, dDKNG +. Bottom figure: DNLS △, DNLS with zero nonlinearity ▽. Solid lines are the scaling δt^6, dashed line δt^{-1}.

how to select the target chain length optimization are detailed in the readme within the sources. This optimized implementation is much faster compared to the generic one (referred to as the "split" kernel for how the q_j and p_j variables are kept in separate buffers), because the ensemble state is kept in registers for all the duration of the kernel, rather than being read and written in global memory. Since register memory is limited, the maximum target chain length is around 1024, though it is not possible to give a precise upper bound as that depends on the GPU hardware and the compiler version. The user should always compare the runtime of the optimized version and the generic one by using the command line option `--split_kernel`.

We showed earlier how the integration of the DNLS model essentially turns into a refined split-step scheme with a large number of FFTs involved. We use the library cuFFT [17] for this purpose. In order to save some of the numerous round trips of the ensemble to and from registers and global memory, the linear and nonlinear operators have been implemented as cuFFT callbacks. In general, this leads to a large performance gain, however there might be circumstances where the non-callback version may be faster. The use of callbacks can be suppressed selectively with the command line options `--no_linear_callback` and `--no_nonlinear_callback`, and we encourage to benchmark the various combinations with and without callbacks for a defined ensemble size and hardware combination.

Table 4
Performance measurements on a single K40m card (clocked at 875MHz and 3004MHz for core and memory respectively) for a dataset with a chain length of 64 and 1024 copies, and a kernel batching size (option `-b`) of 100000 (1000 for the DNLS model). File dumping has been disabled in these runs.

Chain	Steps/second
DNKG (optimized kernel)	163904
dDNKG (optimized kernel)	150003
FPUT (optimized kernel)	83254
Toda (optimized kernel)	37431
FPUT (split kernel)	4012
DNLS (with cuFFT callbacks)	2303
DNLS (without cuFFT callbacks)	1149

5.2. Performance measures

Performance of the software depends on a large number of factors. For reference, in Table 4 we show some rough estimates of the number of steps per second for most of the implementations present in the software, for a fixed chain length size of 64 and ensemble size of 1024. These performance figures should roughly scale linearly with the number of GPUs, and linearly with the inverse of the chain length and ensemble sizes. As mentioned earlier, the big performance bottleneck for the FPUT model is the memory access in the case of the "split" kernel, hence it is crucial to recompile the software for a desired chain length. Such

Table 5
Performance measurements on a Intel Xeon Processor E5-2680, both single threaded and 24 threads, for a dataset with a chain length of 64 and 1024 copies (1008 for the multithreaded test), batching size (option `-b`) of 10000 (1000 for the DNLS model). File dumping has been disabled in these runs.

Chain	Steps/second (1 thread)	Steps/second (24 threads)
DNKG	1229	24510
dDNKG	1280	25623
FPUT	802	17123
Toda	197	4165
DNLS	90	1965

difference is expected also in the models DNKG, dDNKG and Toda. The DNLS model is the slowest of all, due to the large number of FFTs involved. However, it is possible to appreciate the speedup due to the use of cuFFT callbacks.

As previously mentioned, we have implemented the integrators as traditional CPU code, in order to compare the advantages of the GPU implementation. The code makes use of recent vectorization (SIMD) capabilities of recent CPUs, in particular it can use the AVX and AVX-512 instruction sets. The CPU implementation has been tested on an Intel Xeon Processor E5-2680 (which has a release date similar to the NVIDIA K40m card used in the GPU benchmarks), and as such it utilizes the AVX instruction set. It can be parallelized in the same way of the GPU implementation (see Section 3). The results are shown in Table 5, both for a single thread run and a 24-threads run (all the cores of the CPU in use), with the same parameters of the data shown in Table 4. We observe the following. The relative order in terms of performance of the models is essentially the same as the GPU implementation. The multithreaded performance is roughly linear in the number of threads, though it shows the typical signs of saturation: a 20-fold increase is observed when running the simulation with 24 threads. Finally, we see that the speedup of the (optimal) GPU implementation over the multithreaded implementation is significant. For the real models, we observer a speedup of 6.7x (DNKG), 5.9x (dDNKG), 4.9x (FPUT) and 9.0x (Toda). For the DNLS model, the speedup is much more modest (1.2x). This is due to the fact that the GPU implementation suffers from the high-latency global memory operations, while the CPU implementation operates on a single chain at once, and hence the CPU cache is utilized. Unfortunately it was not possible for us to test the code on more recent GPU hardware which should provide a much better memory technology.

6. Impact

The purpose of this software is to provide the scientific community a specialized tool for simulating nonlinear chains (and possibly other simple Hamiltonian, nearest-neighbor systems). While the simulation algorithms is not new, and trivial implementations of the Yoshida sixth order integration algorithms are not particularly difficult, to our knowledge there has been so far no effort to code a simulation with a strong focus on performance, modernity and extendibility of the code, nor software that makes use of heterogeneous hardware architectures such as GPUs.

This software has been essential in our research work [7,8]. Naive implementations of this kind of simulations can have a runtime of weeks. Exploiting the parallelization possibilities of GPU hardware, and a painstaking work of tuning and optimization of the code made the run of a simulation of a typical size a matter of at most hours.

7. Conclusions

In this paper we presented *nlchains*, a specialized software for simulating a number of one-dimensional Hamiltonian systems with nearest neighbor interactions on GPU hardware. The software has been coded during the study of the thermalization of these systems, but other applications are possible, as the speed and accuracy are very good. We have described briefly the usage of the software, and mentioned a few implementation details that can guide the interested user in adapting the software to his or her needs. We provided also a few simulation results, most importantly the scaling of the cumulative errors in the simulations, which matches very well the expected scaling $O(\delta t^6)$ with δt being the step size. To our knowledge, this is also the first implementation of the Yoshida 6th order symplectic integrator for the discrete nonlinear Schrödinger equation as described in [14] with the suggested optimization of leveraging the Fast Fourier Transform for the linear operator of the algorithm.

Declaration of competing interest

The authors wish to confirm that there are no known conflicts of interest associated with this publication.

Acknowledgments

We thank the OCCAM facility, Università di Torino, for providing the hardware necessary to the development of this software. M. O. has been funded by " Progetto di Ricerca d'Ateneo CSTO160004" and by the "Departments of Excellence 2018–2022" Grant awarded by the Italian Ministry of Education, University and Research (MIUR) (L.232/2016).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.softx.2019.100255.

References

[1] Fermi E, Pasta J, Ulam S. Studies of nonlinear problems. Tech. rep., I, Los Alamos Scientific Laboratory Report No. LA-1940; 1955.
[2] Gallavotti G. The Fermi-Pasta-Ulam problem: a status report, vol. 728. Springer; 2008.
[3] Falkovich G, Lvov V, Zakharov V. Kolmogorov spectra of turbulence. Berlin: Springer; 1992.
[4] Nazarenko S. Wave turbulence, vol. 825. Springer; 2011.
[5] Onorato M, Vozella L, Proment D, Lvov Y. Route to thermalization in the α-Fermi-Pasta-Ulam system. Proc Natl Acad Sci USA 2015;112(14). http://dx.doi.org/10.1073/pnas.1404397112.
[6] Lvov YV, Onorato M. Double scaling in the relaxation time in the β-Fermi-Pasta-Ulam-Tsingou model. Phys Rev Lett 2018;120(14):144301.
[7] Pistone L, Onorato M, Chibbaro S. Thermalization in the discrete nonlinear Klein-Gordon chain in the wave-turbulence framework. Europhys Lett 2018;121(4):44003.
[8] Pistone L, Chibbaro S, Bustamante M, L'vov Y, Onorato M. Universal route to thermalization in weakly-nonlinear one-dimensional chains. 2018, ArXiv preprint. arXiv:1812.08279.
[9] Bustamante MD, Hutchinson K, Lvov YV, Onorato M. Exact discrete resonances in the Fermi-Pasta-Ulam-Tsingou system. 2018, ArXiv preprint. arXiv:1810.06902.
[10] Spohn H. The phonon Boltzmann equation, properties and link to weakly anharmonic lattice dynamics. J Stat Phys 2006;124.
[11] Lukkarinen J. Kinetic theory of phonons in weakly anharmonic particle chains. In: Thermal transport in low dimensions. Springer; 2016, p. 159–214.
[12] Kashyap R, Sen S. PULSEDYN—A Dynamical simulation tool for studying strongly nonlinear chains. Comput Phys Comm 2019.

[13] Yoshida H. Construction of higher order symplectic integrators. Phys Lett A 1990;150(5):262–8.
[14] Boreux J, Carletti T, Hubaux C. High order explicit symplectic integrators for the Discrete Non Linear Schrödinger equation. 2010, ArXiv preprint. arXiv:1012.3242.
[15] Weideman J, Herbst B. Split-step methods for the solution of the nonlinear Schrödinger equation. SIAM J Numer Anal 1986;23(3):485–507.
[16] Frigo M, Johnson SG. The design and implementation of FFTW3. Proc IEEE 2005;93(2):216–31.
[17] NVIDIA Corporation. cuFFT. 2018, https://docs.nvidia.com/cuda/cufft/.

Chapter 5

Thermalization in the Klein-Gordon nonlinear chain in the presence of disorder

I present here the most recent results that we obtained concerning the disordered nonlinear Klein-Gordon (KG) lattice. This material is currently being edited for publication. We will consider the case of a fixed-in-time disorder, but random distribution of paramenters. The case of a random variation in time of the model parameters is also of interest in physics, see for example [41].

Disordered lattices are known to display the phenomenon of Anderson localization [42]: the modes of the lattices are no longer plain-waves that cover the entire space, but they are localized. Because of this, transmission of energy across the chain is not possible in the linearized dynamics. Travelling wave solutions are in fact only possible in ordered media with a translational invariance.

When nonlinearity is added, energy can be transmitted, so Anderson localization is destroyed [43, 44, 45, 46]. However, the disordered lattices still has a lower conductivity of an ordered lattices. This is because since the modes are localized in space, in general they interact much less with other modes compared to the spatially extended modes of ordered lattices. It can be shown in fact that the tails of the localized modes show an exponential decay [47]. Our intention is to observe this slow down, and to understand its relation with WT.

I have used the nonlinear disordered KG lattice as the master example in the previous Chapters. We reproduce here for convenience the Hamiltonian,

$$H = \sum_{j}^{N} \frac{1}{2}p_j^2 + \frac{1}{2}(q_{j+1} - q_j)^2 + \frac{1}{2}m_j q_j^2 + \frac{1}{4}\beta q_j^4, \tag{5.1}$$

where m_j is a random positive parameter, different for each site j. We choose the disordered version of the KG model rather than a FPUT model or others, because

extensive numerical simulations were run in [4] for the ordered version, and this data can be useful to understand the disordered case.

In a disordered lattice, the eigenvalue problem in general cannot be solved analytically, and one can only obtain a numerical solution. Different realizations of the disorder may have significantly different quantitative dynamics. In practice, as already mentioned in Chapter 3, the interaction parameter W_{1234} could have a significant dependence on the specific eigenvalue problem, and as such the dynamics of a single realization of the disorder of the chain may not be representative of any other realization. This is not too surprising, because we will consider lattices of a finite length N, well below the thermodynamic limit. We do not intend to characterize the eigenvalue problem and to build a statistic for the interaction coefficient W_{1234}, hence our approach on the linearized dynamics is mostly numerical.

Since the eigenvalue problem is random, also the frequencies of the modes are. The disorder that we will choose has no particular simmetries, hence the frequencies are essentially random number, and a resonant manifold do not exist. The role of quasiresonances is then investigated, and it is found to be quite determining in the route to thermalization of the lattice.

5.1 The model

As we shown in Chapter 1, to the Hamiltonian (5.1) corresponds an equation of motion in normal mode space,

$$i\dot{a}_1 = a_1\omega_1 - \beta \sum_{2,3,4} W_{1234} \left(a_2 a_3 a_4 - 3a_2 a_3 a_4^* + 3a_2 a_3^* a_4^* - a_2^* a_3^* a_4^* \right), \tag{5.2}$$

with the interaction coefficient

$$W_{1234} = \frac{1}{4\sqrt{\omega_1\omega_2\omega_3\omega_4}} \sum_i^N v_i^1 v_i^2 v_i^3 v_i^4. \tag{5.3}$$

We consider periodic boundary conditions. We choose to focus on one specific, yet quite generic type of disorder, that is an uniform distribution between a minimum and a maximum value $m_{\min}$ and $m_{\max}$, with no correlation across sites. We do not consider the presence of disorder in the nonlinear parameter β, as the nonlinear term is constrained to be small, and a random distribution of site-specific values of β would result in a higher-order perturbation. We measure the strength of the nonlinearity with the usual parameter,

$$\epsilon = \frac{< H_{\text{nlin}} >}{< H_{\text{lin}} >} \ll 1, \tag{5.4}$$

where H_{nlin} is the quartic part of the Hamiltonian, while H_{lin} is the quadratic part, and an appropriate time or ensemble average is taken. We note that $\epsilon \propto \beta$.

We now comment on the choice of the random disorder, and on an appropriate parameter range of $m_{\min}$ and $m_{\max}$. In the context of solid state physics, it would be more interesting to consider a random distribution of parameters m_i from a set of discrete values, to mimic the presence of impurities in an otherwise regular crystal. As we will see later however, the disordered system retains some similarities to the regular system in terms of the shape of the eigenvectors and eigenvalues or, in a terminology applicable only to ordered systems, to the original dispersion relation. It is known that biatomic lattices have different branches in the dispersion relation of ordered systems, and consequently band gaps [36]. This is intuitively explained by the fact that each different mass is associated to a different time scale of oscillations. The transmission of energy between these branches is weak, because of the large differences in frequencies that can hinder resonances. In this work we are interested in isolating as much as possible the effects of disorder alone in the equipartition process, rather than dealing with the combined effect with band gaps. Hence, the simplest model of disorder is an uniform distribution of the mass values.

We now turn onto delimiting the parameter space that we intend to study. To determine an appropriate range of the mass parameters $m_{\min}$ and $m_{\max}$, we refer to an argument first stated by [48], regarding the bounds of the eigenvalues of a disordered system, which essentially states that the bounds are the minimum and maximum eigenvalues obtained from the ordered system constructed with all the possible values of the random parameter. That is, we must consider the eigenvalue distribution of all the KG lattices with mass parameter m from $m_{\min}$ to $m_{\max}$. The dispersion relation of the ordered KG lattice is $\omega = \sqrt{m + 4\sin(\pi k/N)^2}$, hence the frequency bounds of the disordered lattice are $\omega_i \in [\sqrt{m_{\min}}, \sqrt{m_{\max} + 4}]$.

We note the following properties of the ordered KG lattice. When m is very large, the nearest neighbour interaction term becomes negligible in eq. (5.1), and all the frequencies tend to the same value $\sqrt{m}$. In this regime, resonance are expected to be abundant, but this is arguably a trivial case. On the other hand if the mass is close to zero, the system comes close to the FPUT problem, and in fact the linear dispersion relation becomes identical to the FPUT problem in the limit $m \to 0$. The FPUT dispersion relation poses some issues in our treatment. First, we note that $\omega(k) \to 0$ for $k \to 0$: this means that for some of the modes in the linearized dynamics the evolution timescale diverges. This blow up corresponds to the appearance of multiple linear timescales. Secondly, the modes with a very low frequency have a diverging number of particles at equipartition, because $e_j = \omega_j n_j$. This was also evidenced in the case of the ordered KG lattice in [5]. A diverging number of particles causes the nonlinearity associated to the low-frequency modes to diverge as well, breaking the assumption that the nonlinearity ϵ is small. We observe in fact a localized (in normal mode space) divergence from the expected equipartition energy. Since we want to first establish that WT arguments are applicable to the disordered system, we decide to avoid this complication by setting

$m_{\min} = 1$ and to control the level of disorder through $m_{\max}$, with $m_{\max} > m_{\min}$.

5.1.1 Numerical examination of the localized eigenstates

We now introduce methods to quantitatively measure the localization of a given disorder realization. A common measure of the localization of the eigenvectors is the inverse participation number [47],

$$P_k = 1 \Big/ \sum_i^N \left(v_i^k\right)^4 . \tag{5.5}$$

Note that since the eigenvectors satisfy the constraints $\sum_i^N (v_i^k)^2 = 1$, $(v_i^k)^2 \geq 0$, the measure P_k is minimized when the eigenvector is localized into a single site ($v_i^k = 1$, $v_{j\neq i}^k = 0$), and it is maximized for a flat eigenvector ($v_i^k = v_{i+1}^k$), and respectively $P_k = 1$ or $P_k = N$. We are interested in a measure of the localization of the whole set of eigenvectors of a system of a fixed size and disorder realization. The most immediate choice would be to average P_k over all the eigenvectors. However, this measure has some shortcomings. In fact, the inverse participation number is not equal to N for states that are evidently nonlocal, such as plane waves in an ordered system. It is easy to see in fact that in the ordered case, the eigenvectors are plane waves with wave number $k \in [0, N)$, and $P_k = N$ only for $k = 0$, else $P_k = 2/3N$, and in the case that N is even, then additionally $P_k = N$ for $k = N/2$ and $P_k = N/2$ for $k = N/4$ and $k = 3/4N$. To overcome this problem, we introduce the centered eigenvectors

$$w_i^k = v_{i+j(k)}^k, \tag{5.6}$$

where $j(k)$ is the index of the eigenvector k where $(v_i^k)^2$ is maximal. This shift operation is possible because we consider homogeneous disorder in a system with periodic boundary conditions. Since the centered eigenvectors are all peaked at the site $i = 0$, it is easier to estimate the average spatial decay of the eigenvectors. We can evaluate the average of the square amplitudes of the centered eigenvectors,

$$W_i = \frac{1}{N} \sum_k^N \left(w_i^k\right)^2 . \tag{5.7}$$

Note that the W_i have the properties $\sum_i^N W_i = 1$ and $W_i \geq 0$, which are the same properties of the squared eigenvectors, $\sum_i^N (v_i^k)^2 = 1$ and $(v_i^k)^2 \geq 0$. The average in eq. (5.7) may be performed additionally on an ensemble C of realizations of the disorder,

$$\overline{W}_i = \langle W_i \rangle_C, \tag{5.8}$$

if one needs to quantify the localization of multiple realizations of the disorder. In analogy to the inverse localization number, we introduce the localization length scale

$$L = 1 \Big/ \sum_i^N W_i^2 \tag{5.9}$$

The measure (5.9) is a generalization of eq. (5.5), and shares its desirables properties: if all the eigenvectors are localized in a single site then $L = 1$, while if $W_i = W_{i+1}$ then $L = N$. For an ordered system where the eigenvectors are plane waves, it can be easily checked that $L \simeq N$, and that is because the oscillations of the plane waves that result in a value of $P_k < N$ in eq. (5.5) are averaged out at each single site in eq. (5.3). When the localization length scale is calculated over an ensemble of disorder realizations (that is, using $\overline{W}_i$ rather than W_i) we use the symbol $\overline{L}$.

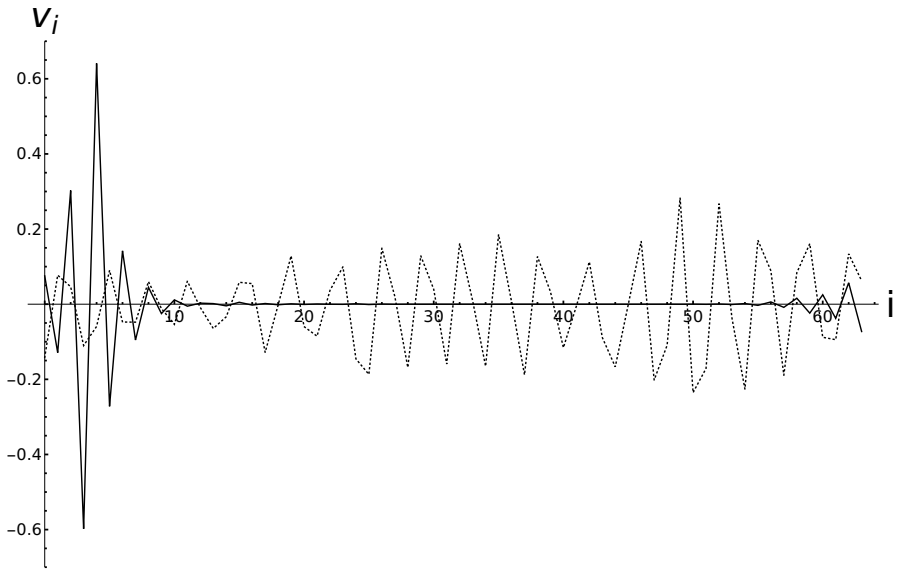


Figure 5.1: Two sample eigenstates of the KG lattice with $N = 64$, $m_{\max} = 3$. The eigenstate in solid and dotted lines have a participation number of 3.2 and 28.8 respectively.

We now examine the eigenstates of some realizations of the disordered lattice. In Figure 5.1 we show two eigenstates from a realization of the KG chain with $N = 64$ and $m_{\max} = 3$. We observe clear localization of the eigenstate in solid line, while the one in dotted line is closer to a plane wave. It is in fact to be expected that at finite $m_{\max}$ not all eigenstates are completely localized. We note however that for large system sizes, for the same disorder strength (in term of same value of $m_{\max}$) a larger portion of the eigenstates is localized, compared to the system size. This is a finite size effect and it is captured by the localization length scale.

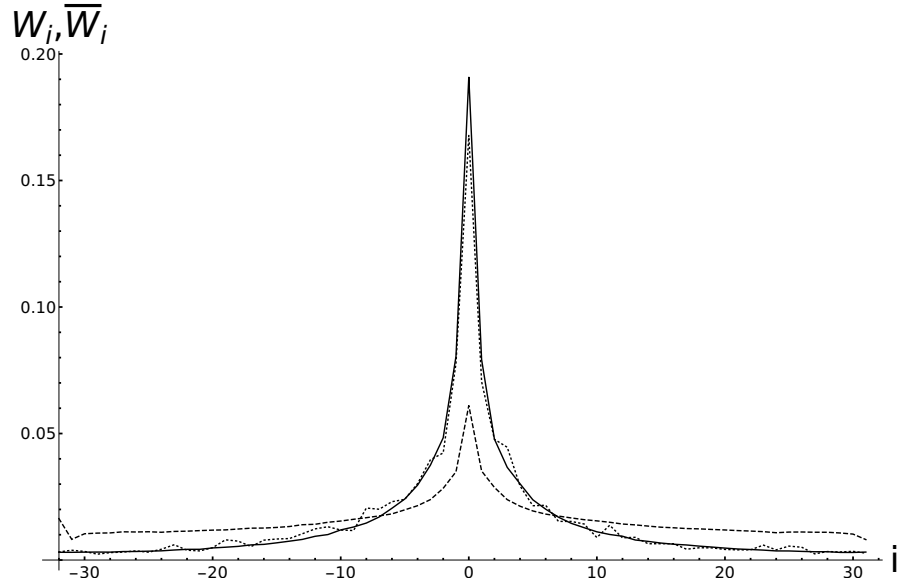


Figure 5.2: A plot of W_i (dotted line) and $\overline{W}_i$ (solid line, 100 realizations) for the KG lattice with $N = 64$ and $m_{\max} = 3$, the localization length scale is $L = 18.5$ and $\overline{L} = 15.8$. The dashed line is $\overline{W}_i$ for 100 realizations of the KG lattice with $N = 64$ and $m_{\max} = 1.5$, with localization length scale $\overline{L} = 50.7$.

In Figure 5.2 we show the plot of W_i for the same realization of Figure 5.1 (dotted line), together with the average $\overline{W}_i$ over 100 realizations (solid line). For comparison, the dashed line is a plot of $\overline{W}_i$ over 100 realizations with $N = 64$ and a lower disorder $m_{\max} = 1.5$. We first note that the effect of stronger disorder is immediately evident, as the tails of $\overline{W}_i$ for $m_{\max} = 1.5$ are much fatter than the case $m_{\max} = 3$. This is quantitatively described by the averaged localization length scale, which is 15.8 and 50.7 respectively. We also note that W_i and $\overline{W}_i$ are not too different. This is important, because in one dimension there is no guarantee that a disordered lattice is self-averaging, that is one realization of the disorder show similar dynamics compared to other realizations. However, with the restrictions we have made in the model of disorder, we do not expect to encounter corner cases (as for example extremely large masses that can effectively divide the chain in several, non-communicating halves), so the dynamics are effectively self-averaging. In comparing W_i and $\overline{W}_i$ in Figure 5.2 for $m_{\max} = 3$ we actually see that $W_i \simeq \overline{W}_i$ with only small fluctuations. The similarity is also confirmed by the fact that the localization length scales are similar: $L = 18.5$ and $\overline{L} = 15.8$ for the averaged case. We can safely assume then that in our case a single realization of the disorder represents sufficiently well all other possible disorder realizations.

It is interesting to see how the localization of individual eigenvectors correlate with the eigenvalue, or the distribution of masses. In Figure 5.3 we show a heat

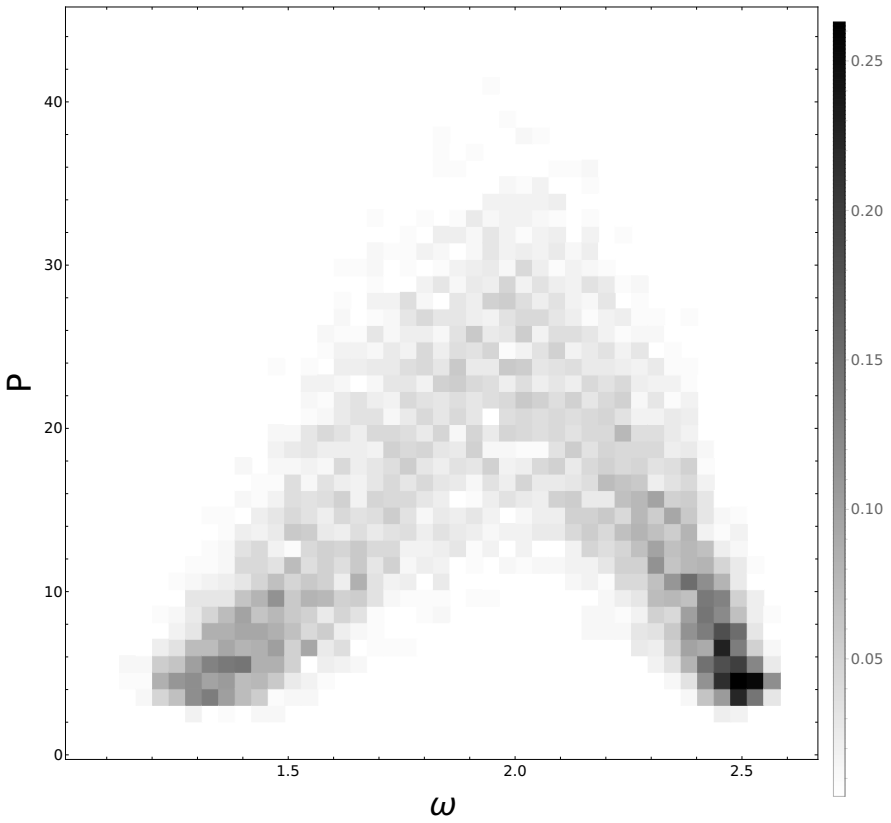


Figure 5.3: A heat map of the PDF of the frequencies ω and participation number P_k of 100 realizations of the KG lattice ($N = 64$, $m_{\max=3}$).

map of the PDF of the frequency ω versus the participation number P eq. (5.5), for 100 realizations of a moderately disordered chain ($N = 64$, $m_{\max} = 3$). We note that the eigenvectors can be localized either for the lower and upper boundaries of the frequency ($\omega \in [1, \sqrt{7}]$), and somewhere in between there is a large number of less localized states. The distribution of the frequencies is roughly uniform between the bounds (not shown). We can guess from this plot that the most localized eigenstates are centered either around areas where there is an accidental concentration of either large or small masses, and since larger masses are linked to larger frequencies in an ordered system, that explains the separation of frequencies in the localized eigenstates. This effect is shown in Figure 5.4, a heat map of the PDF of the of the frequency ω versus the average mass covered by the eigenvector,

$$\langle m \rangle_k = \sum_i^N \left(v_i^k\right)^2 m_i \tag{5.10}$$

(same parameters as in Figure 5.3). The correlation between light masses and low

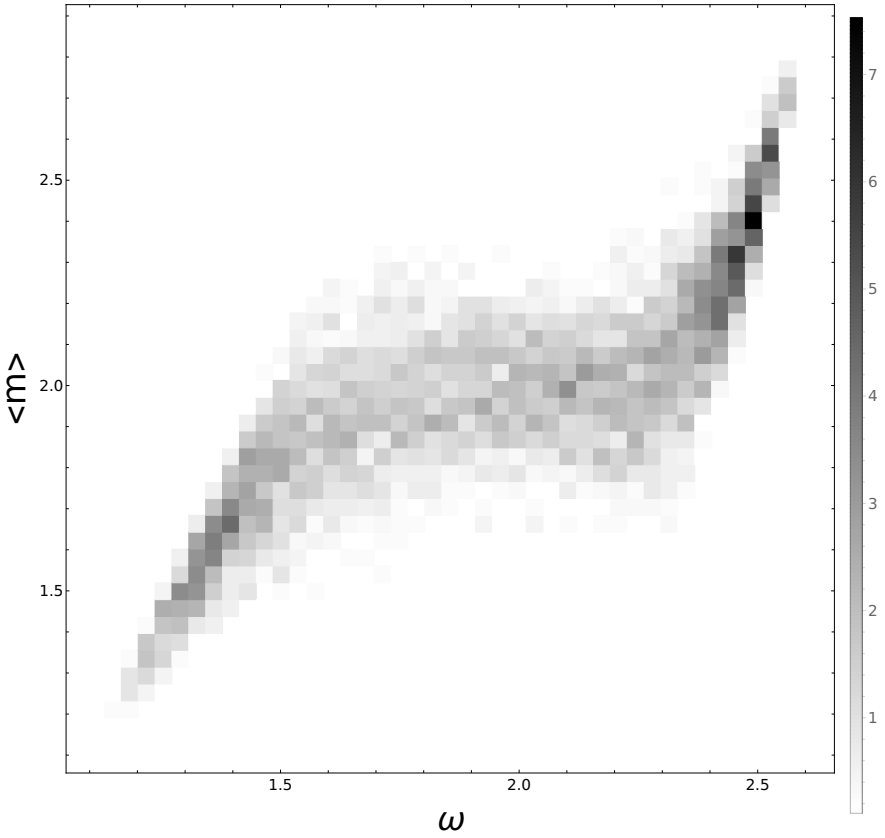


Figure 5.4: A heat map of the PDF of the frequency ω and average mass of the eigenvector $\langle m \rangle_k$ of 100 realizations of the KG lattice ($N = 64$, $m_{\max=3}$).

frequency localized eigenvectors, and larger masses and high frequency localized eigenvectors is evident. The more extended eigenvectors have obviously $\langle m \rangle_k \simeq (m_{\max} - m_{\min})/2$, and they show a rather uniform distribution of frequencies.

The localization length is also useful to quantify the the finite size effect for intermediate values of N. We already mentioned in fact that for a given disorder level, the eigenstates may not be fully localized as the system size is too small. This effect is shown in Figure 5.5. We can see that the localization length scale is the same for high values of $m_{\max}$ regardless of the system size, while for lower values it saturates around the maximum value possible for the system size $L \simeq N$. This means that for a large disorder, the dynamics are expected to be the same for any system size. This is different from what is found in the ordered models treated earlier in this these. It should be however not too surprising, because the thermodynamic limit for the Hamiltonian eq. (5.1) is not well defined: the random distribution of m_j creates a discontinuity in space. A continuous limit $N \to \infty$ is in fact only appropriate in weakly disordered lattices [49].

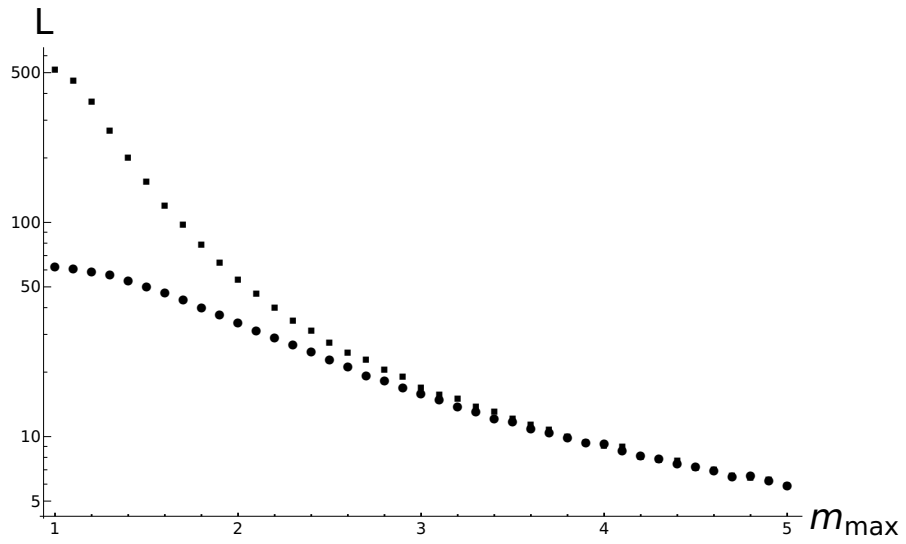


Figure 5.5: The localization length scale L as a function of the maximum mass parameter $m_{\max}$, for $N = 64$ (circles) and $N = 512$ (squares).

5.2 The route to equilibrium

The equations (1.12) and (5.2) largely resemble the ordered KG chain, and many well-studied nonlinear chains such as the FPUT and nonlinear Schrödinger chain can be cast in similar forms, regardless of the disorder strength. The effect of disorder essentially lies in the overlap integral, eq. (5.3). For ordered systems eq. (5.3) becomes a selection rule in the sum of wave numbers, that is the quadruplet of participating wave mode numbers k_1, k_2, k_3, k_4 must satisfy a constraint in the form of $k_1 \pm k_2 \pm k_3 \pm k_4 = 0 \mod N$. Not much else is known about the properties of the overlap integral in the disordered case (see [50]). Because of this similarity of the equations of motion, we will attempt to understand the thermalization dynamics in the framework of Wave Turbulence, and point out what is different from the ordered cases.

5.2.1 Quasiresonances

In the introduction we saw how resonances and quasiresonances enter the picture in the disordered lattices, see equation (1.16). The beat frequency $\Delta\omega$ of a quasiresonance quantifies its detuning,

$$\Delta\omega = \omega_1 \pm \omega_2 \pm \omega_3 \pm \omega_4. \tag{5.11}$$

The plus-minus signs account for the different possible resonances in eq. (1.16).

It is widely recognized that resonances need not to be exact in order to be effective at finite N. A quasiresonance can have a very low beat frequency, such

that its rotation is on a timescale several orders of magnitude larger than the linear timescale. It is evident that if for a given tuple of modes $|\Delta\omega| \ll 1$, then such tuple behaves essentially as a resonance. Unfortunately, it is not know analytically whether such a quasiresonance is just as "effective" as an exact resonance. In the WT community, quasiresonances are mostly studied for the cluster of interconnected modes that they create [40], but their actual effect on the dynamics of the system is surprisingly little studied. In [18], quasiresonances are believed to be effective with a modulation factor that is inversely proportional to the detuning $|\Delta\omega|$.

Quasiresonances are argued to be significant when exact resonances are scarce for some reason. We showed that in the ordered KG lattice, quasiresonances can become active above some nonlinearity ϵ threshold, and they can become the principal mechanism of transfer of energy across modes. This is because contrary to exact resonances, quasiresonances are believed to be sensitive to the nonlinearity of the system. In fact, a well-known effect of the nonlinearity is to introduce a stochasticization of the phases and frequencies of the normal modes, due to short-term interactions between modes mediated by the nonlinearity. This effect is called frequency broadening, because it appears as a dispersion of the effective normal mode frequencies around the theoretical value of the linearly truncated system, and we have already introduced it in Section 1.1.2. This broadening is essentially an uncertainty on the frequency of the linear modes, that carries over to the detuning parameter $\Delta\omega$ of a quasiresonance. When the broadening is strong enough, then it is impossible to distinguish between a resonance and a quasiresonance, because of the random jitter of the overall normal modes tuple phase.

The frequency fluctuations depend strongly on the lattice model details. This dependence is often overlooked, because the broadening is often measured by simply looking at the spectrum of the phase of the normal modes. In reality, the instantaneous speed (frequency) of the rotating phase of a normal mode can correlate with the value of the phase itself. It is easy to see why: a different value of the phase of the normal modes corresponds to a different split of the mode total energy between kinetic and potential energy of the eigenvector (see eq. (1.11)), and the nonlinearity may have a different effects along the orbit of the normal mode. In the KG lattice, the nonlinearity is present only in the potential term. During the evolution of the normal mode, the chain nodes that are part of the core of the eigenvector swing up and down, and when they reach the maximum distance from resting position $q_j = 0$, the effect of nonlinearity is maximal. For our parameter space ($\beta > 0$) this means that the chain node is pulled to the rest position faster than it would in a harmonic oscillator, with an short lived frequency greater than the normal frequency $\omega' > \omega$.

We can readily show how the broadening works at microscopic level. We measure the instantaneous frequency of the normal modes in an ensemble of realizations of the KG chain with $m_{\max} = 3$ and $N = 64$. The measurement procedure is as

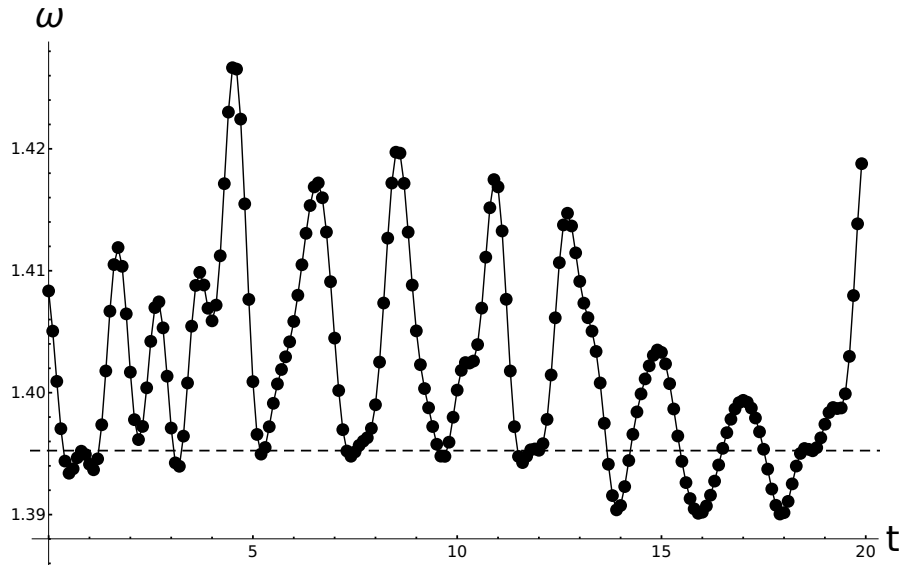


Figure 5.6: The fluctuations of the frequency of a mode for a chain with $m_{\max} = 3$, $N = 64$ and $\epsilon = 2.38 * 10^{-3}$. Black dots are the samples, the dashed line is the harmonic frequency $\omega = 1.395$ given by the solution to the eigensystem.

follows. We first wait for the ensemble to reach equipartition (see the results of the numerical simulations in Section 5.3 for details on the integration scheme). We then advance the simulations for a time $t = 1000$ in timesteps of $\delta t = 0.1$. At each time step we calculate the normal modes, and we calculate the instantaneous frequency

$$\omega_i(t) = \frac{\arg\left(\frac{a_i(t-\delta t)}{a_i(t)}\right)}{\delta t}. \tag{5.12}$$

Note that the time step δt should be chosen to be small enough in order to resolve the shortest timescale of the fluctuations. This can be easily tested by looking at the autocorrelation function of $\omega_i(t)$, which needs to be large at the sampling frequency.

The result data is plotted in Figures 5.6, 5.7 and 5.8. From Figure 5.6 we see that how the broadening and shift of frequencies happen: from the base linear frequency (dashed line) the actual frequency of the mode is "kicked" to slightly higher values. These kicks are due to the nonlinearity, as previously discussed, and it is highly correlated with the phase of the normal mode. This correlation is evident in Figure 5.7, where we plot a heatmap of the frequency versus the angle of the mode: when $\theta = k\pi$, $k \in \mathbb{Z}$ then the potential energy is minimum, hence in this configuration the effect of nonlinearity is negligible, and the most probable frequency is the normal linearized frequency of the eigenmode. In Figure 5.8 we show the PDF of the frequencies ω for three different levels of nonlinearity, $\epsilon \simeq$

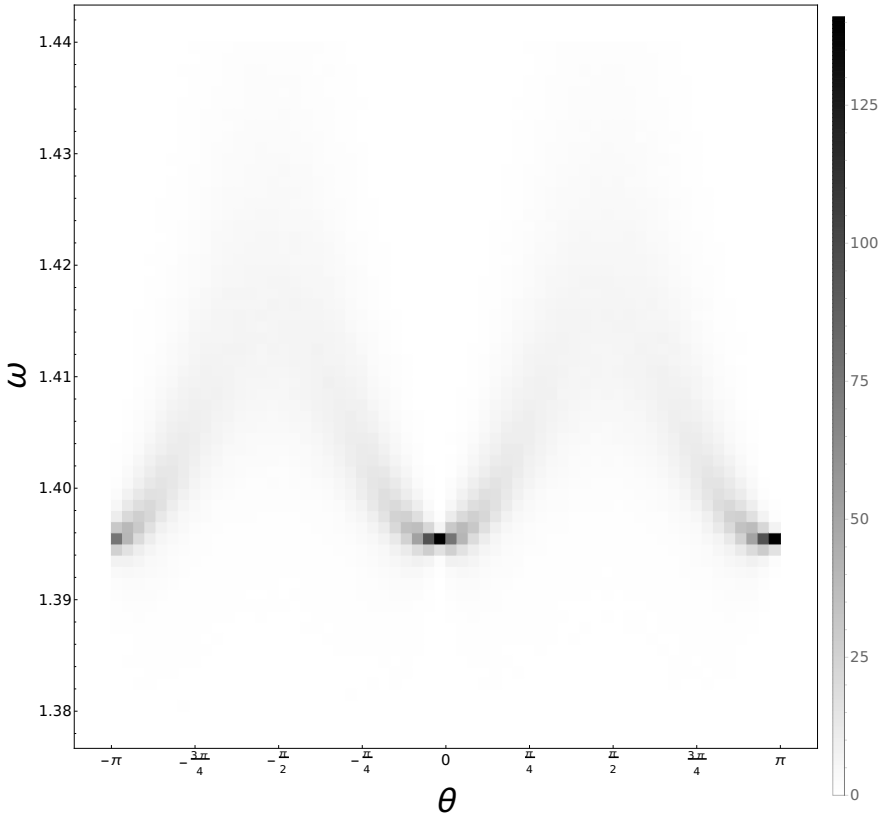


Figure 5.7: A heat map of the PDF of the instantaneous normal mode phase θ and frequency ω for the same model parameters as in Figure 5.6 ($m_{\max} = 3$, $N = 64$, $\epsilon = 2.38 * 10^{-3}$), averaged over an ensemble of 100 realizations for 10000 timesteps. Darker areas mean more probable.

$10^{-4}, 10^{-3}, 10^{-2}$ (solid, dashed, and dotted lines). In agreement with Figure 5.6, we observe that the mode of all three distributions of the instantaneous frequency ω_j is the same, regardless of the nonlinearity. The peak of the distribution lies at the linearized frequency (vertical dot-dashed line), while the distribution gets broader at higher nonlinearity, but in an asymmetric way. We see that the distribution is clearly non-gaussian: the evident shoulder at higher values of ω is the cause of both the frequency shift and broadening. We stress on the fact that this shoulder of the distributions has no universality across models or parameter space. For example, it is possible to create a negative frequency shift by using $\beta < 0$ (the so-called focusing nonlinear KG chain), and the shoulder would be observed at smaller on the left of the distribution.

The nontrivial distribution of the frequencies in the normal modes is conveniently hid when one considers the effect of frequency shift and broadening in quasiresonances. As we mentioned earlier, we expect that at some level of non-

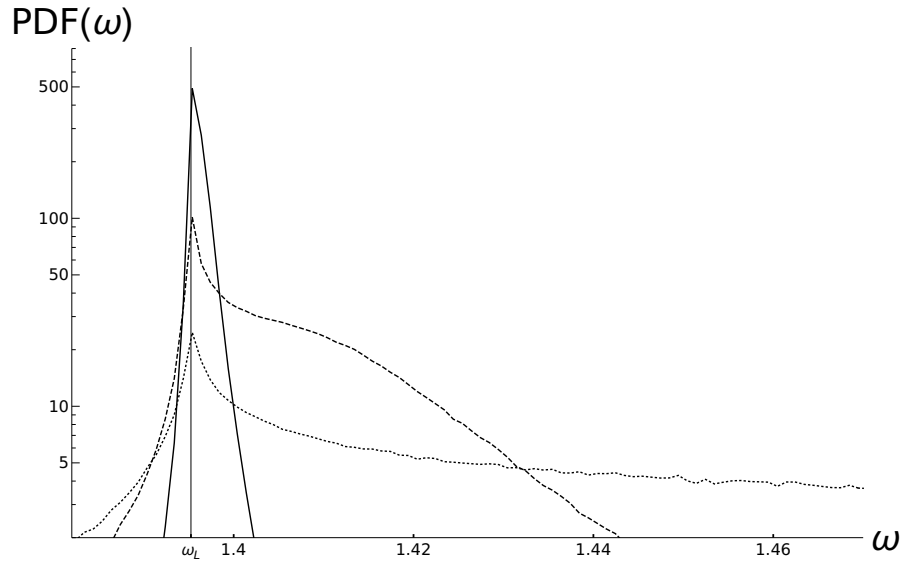


Figure 5.8: A plot of the PDF of ω for $m_{\max} = 3$, $N = 64$ and three different levels of nonlinearity, $\epsilon = 1.77 * 10^{-4}$ (solid line), $\epsilon = 2.38 * 10^{-3}$ (dashed line, same data as in Figure 5.6 and 5.7) and $\epsilon = 1.69 * 10^{-2}$ (dotted line). The vertical line is the expected linearized frequency $\omega_L = 1.395$, which coincides with the mode of all the three distributions.

linearity, the frequency broadening may render indistinguishable a quasiresonance from an exact resonance. The disordered KG model allows in theory all kinds of four-wave resonances. It is however simple to see that for all resonances not of the type $2 \to 2$ the detuning parameter $\Delta\omega$ is much different than zero. Resonances of the type $4 \to 0$ are evidently excluded, because the frequencies are all positive. Resonances of the type $3 \to 1$ have $\Delta\omega = \omega_1 + \omega_2 + \omega_3 - \omega_4$, and since $\omega \in [1, \sqrt{m_{\max} + 4}]$, unless $m_{\max} \gg 1$ then the positive terms in $\Delta\omega$ cannot be compensated by the negative term. For $3 \to 1$ quasiresonances to be active, we found that the $m_{\max}$ parameter must be so large that the chain essentially behaves as a number of nodes that oscillates with very little coupling to the neighbour nodes. We will consider $3 \to 1$ quasiresonances as inactive in the following, and we will verify numerically that this is indeed the case. We are left with $2 \to 2$ interactions, that is a sum of four random variables, with two positive signs and two negative signs. The random variables can be assumed to be uncorrelated, because of the weak nonlinearity, but also because (at least in the KG model) the nonlinearity is correlated with the phase of the normal mode, and the ratio of the frequency of two modes is in general irrational, leading to quick decorrelation between the phase of two distinct modes. The sum of these frequencies then is a sum of essentially independent random variables, and the standard deviation of $\Delta\omega$ can be estimated as the quadrature of the individual standard deviation of the

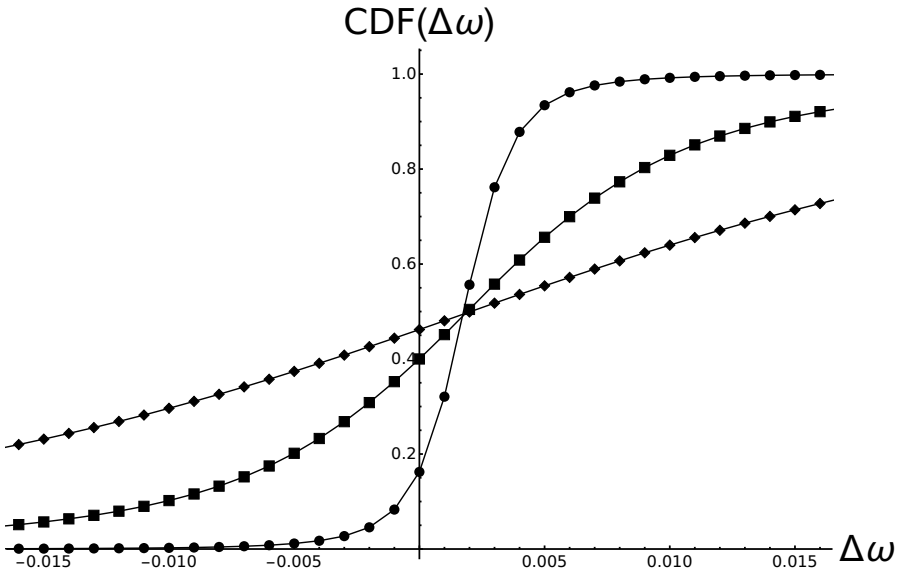


Figure 5.9: The CDF of the beat frequency $\Delta\omega$ for a $2 \to 2$ quasiresonance for a realization of the disordered KG lattice with $m_{\max} = 3$ and $N = 64$, with four distinct normal modes: $\omega_1 \simeq 1.443$, $\omega_2 \simeq 2.212$, $\omega_3 \simeq 1.45$, $\omega_4 \simeq 2.202$, $\Delta\omega \simeq 0.003$. The three datasets correspond to different levels of nonlinearity: $\epsilon \simeq 1.77 * 10^{-4}$ (circles), $\epsilon \simeq 7.95 * 10^{-4}$ (squares), $\epsilon \simeq 2.38 * 10^{-3}$ (diamonds).

frequencies,

$$\sigma_{\Delta\omega} = \sqrt{\sigma^2_{\omega_1} + \sigma^2_{\omega_2} + \sigma^2_{\omega_3} + \sigma^2_{\omega_4}} \simeq 2\langle\sigma_\omega\rangle. \tag{5.13}$$

The distribution of $\Delta\omega$ turns out to be quasi-gaussian around the mean. This is shown in Figure 5.9, where we plot the numerical CDF of $\Delta\omega$ of an arbitrary quasiresonance with $\Delta\omega \ll 1$. In the linear case, this quasiresonance would have a beat frequency $\Delta\omega = 0.003$. The three curves correspond to different and increasing nonlinearity, $\epsilon \simeq 1.77 * 10^{-4}$ (circles), $\epsilon \simeq 7.95 * 10^{-4}$ (squares), $\epsilon \simeq 2.38 * 10^{-3}$ (diamonds). Four random variables are definitely not enough to invoke the central limit theorem in a formal way, as done for example in [51, Chapter 1], yet the resulting distribution of $\Delta\omega$ is close to Gaussian [52]. In fact, the shoulders of the distributions that we found in Figure 5.8 tend to be cancelled out in $\Delta\omega$ because there is a balance between the two positive and two negative terms. This is because the broadening is quite homogeneous across different modes. For a quasiresonance to be essentially indistinguishable from a resonance, we require that

$$\mathrm{CDF}(0) \simeq 0.5, \tag{5.14}$$

that is the distribution is essentially centred around a zero beat frequency (an exact resonance). We see that this is not the case for the lowest nonlinearity ($\mathrm{CDF}(0) = 0.16$), but it is for the two curves that correspond for larger nonlinearity

(CDF(0) = 0.4 and CDF(0) = 0.46 respectively). The standard deviation of the distribution of $\Delta\omega$ is $\sigma_{\Delta\omega} = 2.6 * 10^{-3}, 2.6 * 10^{-3}, 1.5 * 10^{-2}, 5.2 * 10^{-2}$ respectively for the three nonlinearity levels. Note that we observe $\sigma_{\Delta\omega} \propto \sigma_\omega \propto \epsilon$, and the linear dependence of the frequency broadening over the nonlinearity was predicted in [33]. From this data we see that in practice our requirement that the CDF is around 0.5 at $\Delta\omega = 0$ is equivalent in practice to requiring that

$$\sigma_{\Delta\omega} \geq |\omega_1 + \omega_2 - \omega_3 - \omega_4|. \tag{5.15}$$

It is expected but interesting to note that all the CDF in Figure 5.9 cross roughly at the same the point $\Delta\omega = 0.003$, where the mean and centre of the distribution lies, around the linear beat frequency. This is because the average frequency shift $\widetilde{\omega}_k$ of a mode k in the weakly nonlinear regime has an expression of the type [33]

$$\widetilde{\omega}_1 \simeq \beta \left(2 \sum_2 W_{1122} I_2 - W_{1111} I_1 \right). \tag{5.16}$$

We see that even though in general the self-coupling W_{1111} is larger than the coupling to another mode W_{1122}, the frequency shift depends on a sum of the coupling to all the other modes, which evens the frequency shift among modes. Consequently, in a $2 \to 2$ quasiresonance the shift of the individual modes cancel out in the expression of the beat frequency $\Delta\omega$. This is the reason for which we can use the criteria (5.14) or (5.15) without considering the actual frequency shifts of the single modes.

5.2.2 Quasiresonances in Wave Turbulence

The kinetic equation (1.23) only includes exact resonances. In the previous section we have established that in practice the frequency of all normal modes becomes stochastic when the nonlinearity is active. Consequently, the beat frequency $\Delta\omega$ is stochastic as well, and the condition $\Delta\omega = 0$, which comes from the Dirac's delta in eq. (1.23), is relaxed, that is $\Delta\omega$ only needs to be close to zero. In other words, at a given level of nonlinearity some quasiresonances are indistinguishable from exact resonances, even in the case where exact-resonances do exist (such as in ordered lattices). In the previous section we have argued that a good criterion for considering a quasiresonance active is eq. (5.15). The main advantages of using $\sigma_{\Delta\omega}$ is that it is easy to compute numerically, and that it is easy to show that in the weakly nonlinear regime $\sigma_{\Delta\omega} \propto \epsilon$ [33].

The qualitative conclusion is that for larger nonlinearity levels more quasiresonances active in the system, and so the system can reach equipartition just as if exact resonances existed. But do these additional resonances have a quantifiable effect on the equipartition time T_{eq}? We can see that the kinetic equation (1.23)

is implicitly dependent on the number of active resonances through the parameter $|W_{1234}|^2$. The kinetic equation is in fact an integral over the active resonances, which means that the speed of the route to equipartition does not depend uniquely on the nonlinearity parameter β, but also on the number of resonances, which contribute to the overall integral value with their corresponding interaction coefficient $|W_{1234}|^2$. In ordered systems, exact resonances constitute a fixed core of resonances that are always active, regardless of the nonlinearity. In disordered systems, the number of active resonances (according to criterion (5.15)) is a function of β itself, and this dependence must be accounted for.

The number of active resonances at a certain broadening level is a difficult problem to solve. It depends on the fine details of the disordered lattice model. We do not attempt to find an analytic (probabilistic) solution to this problem. Here we show some numerical facts applicable to the disordered KG lattice, which allow us to quantify the correction to the rule (1.30) for the thermalization time in the disordered lattice.

First of all, the interaction coefficient appears under an integral in the kinetic equation (1.23), which means that the speed of the evolution of the averaged squared amplitude n_j for the normal mode j-th variable is roughly proportional to the sum of the interaction coefficients of all active resonances that include the j-th mode. Let us call this contribution $Q_j(\sigma_{\Delta\omega})$,

$$Q_j(\sigma_{\Delta\omega}) \simeq \sum_{234} \Big|_{\sigma_{\Delta\omega} \geq |\Delta\omega|} |W_{j234}|^2 \,, \tag{5.17}$$

where the sum is over all quasiresonances active (according to criterion (5.15)) that include the j-th mode. Then we can write

$$\dot{n}_j \sim \beta^2 Q_j(\sigma_{\Delta\omega}). \tag{5.18}$$

The equipartition timescale depends on the speed of evolution of all modes. In the ordered case, eq. (1.30) implies the assumption that there is not an extreme difference in the speed of evolution among the modes of the system. In the disordered case, in our parameter range (see the discussion in Section 5.1.1), we expect the same assumption to hold. Consequently, we can consider the average of the quantity $Q_i(\sigma_{\Delta\omega})$ across all modes, and so we can extend the timescale to equipartition for the ordered case (1.30) into

$$T_{\rm eq} \sim \beta^{-2} \langle Q_j \rangle_j^{-1}, \tag{5.19}$$

where $\langle Q_j \rangle_j$ is an average across all modes.

Eq. (5.19) can be further simplified. We have already assumed that all modes are all roughly equally likely to participate in a quasiresonance. In other words, the modes that participate in a quasiresonance that becomes active at some broadening

$\sigma_{\Delta\omega}$ are essentially random, because there are no symmetries in the disorder. This allows us to exchange the sum in eq. (5.17) with the sum that is implicit in the average, and we can finally write

$$T_{\rm eq} \sim \beta^{-2} \left(\frac{4}{N} \sum_{1234}\Big|_{\sigma_{\Delta\omega} \geq |\Delta\omega|} |W_{1234}|^2 \right)^{-1} . \tag{5.20}$$

Eq. (5.20) is easier to compute numerically than eq. (5.19) because it involves simply calculating the cumulative sum of all interaction coefficients of the quasiresonances that are active at some broadening threshold, regardless of which mode they connect. The factor of 4 appears because most quasiresonances include four distinct modes.

Numerical evaluation of the quasiresonances

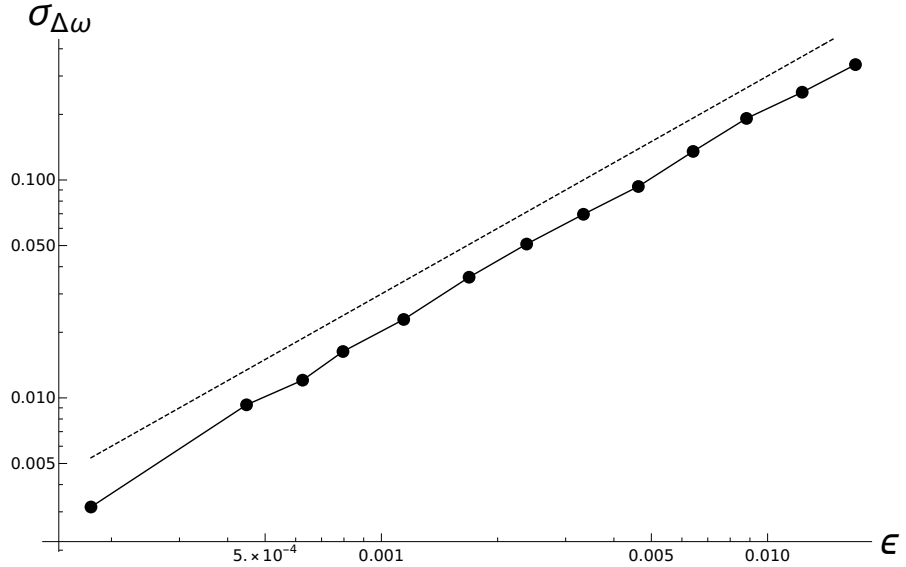


Figure 5.10: A plot of $\sigma_{\Delta\omega}$ (calculated with eq. (5.13)) as a function of the nonlinearity ϵ for a disordered KG lattice, with $N = 64$, $m_{\rm max} = 3$. The dots and solid line are the experimental data, while the the dashed line is a reference linear scaling $\sigma_{\Delta\omega} \propto \epsilon$.

Let us show some numerical result to support our arguments so far. We will show a number of results referred to a single realization of the disordered KG lattice with $N = 64$, $m_{\rm max} = 3$ ($\overline{L}/N \simeq 0.25$ same data as in Figure 5.2). We have checked that the conclusion that we draw here are applicable for multiple realizations of the lattice with the same parameters, and for a reasonable range of $m_{\rm max}$ ($2 \leq m_{\rm max} \leq 5$, roughly $0.5 \leq L/N \leq 0.1$).

In Figure 5.10 we show the dependence of the quasiresonance broadening $\sigma_{\Delta\omega}$ as a function of the nonlinearity ϵ, for a single realization of the disordered KG lattice ($N = 64$, $m_{\max} = 3$). The solid line and the dots are the experimental data, while the dashed line is a reference for the expected dependence $\sigma_{\Delta\omega} \propto \epsilon$. The broadening $\sigma_{\Delta\omega}$ was calculated with eq. (5.13), using the approximation that the broadening of all single modes σ_{ω_i} is comparable. We see that there is an excellent agreement with the expected linear dependence of the quasiresonance broadening.

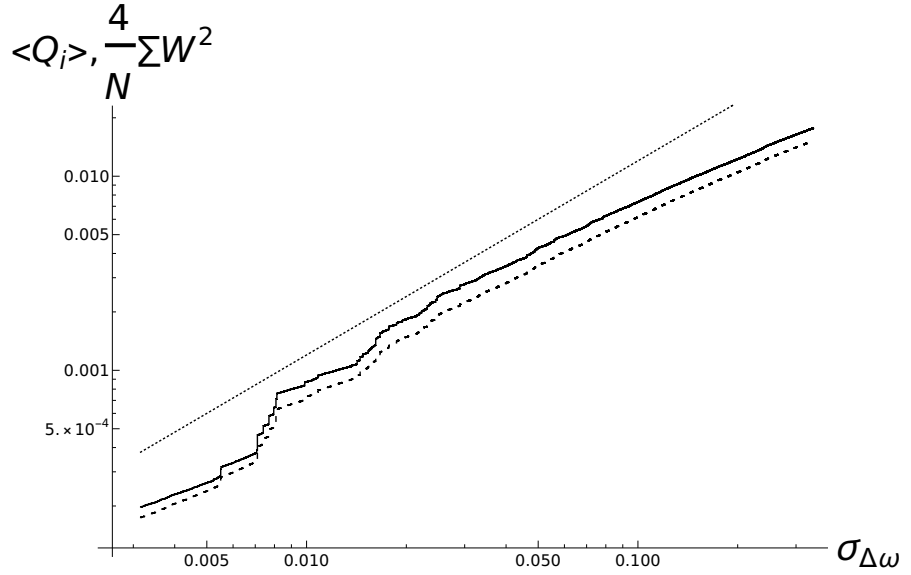


Figure 5.11: A plot of the cumulated squared interaction coefficients for the same model parameters of Figure 5.10: mean across mode (eq. (5.19), dashed line), total (eq. (5.20), solid line). The dotted line is a reference of a linear dependence.

In Figure 5.11 we show the contribution of the quasiresonances to the scaling of the equipartition time, calculated as in eq. (5.19) (dashed line) and eq. (5.20) (solid line). We see that the ways of calculating this contribution are essentially equal, but an almost constant multiplicative factor. This small difference is due to the fact that in eq. (5.20) we assumed that all resonances involve four distinct modes, while some may actually involve three, or even only two modes. Accounting for this correction would result in a factor lower than 4 in eq. (5.20), but we do not need this level of precision in our analysis. We observe that the behaviour is roughly linear, though there are evident jumps, corresponding of clumps of quasiresonances. We do not show the data here, but we have performed an average over ten realizations of the random lattices, and the resulting curves for $\langle Q_i(\sigma_{\Delta\omega})\rangle$ are essentially similar: we fitted the log-transformed data and extracted the coefficient, and the average across the realization gives 0.98 ± 0.05, which is consistent with a linear dependence.

From this numerical analysis, we conclude the following. We found numerically

that $\langle Q_i(\sigma_{\Delta\omega})\rangle \propto \sigma_{\Delta\omega}$. We do not claim any universality on this result (cp. Chapter 3), and we actually believe that different disorder models may give rise to different results. However, since in our case this proportionality holds, we can write

$$T_{\rm eq} \sim \beta^{-2}\langle Q_i(\sigma_{\Delta\omega})\rangle^{-1} \sim \beta^{-2}\sigma_{\Delta\omega}^{-1} \sim \beta^{-2}\epsilon^{-1} \sim \beta^{-3}, \tag{5.21}$$

where we have used (in order) eq. (5.19), the numerical results show in Figure 5.11, the proportionality between $\Delta\omega$ and ϵ (predicted in (cita onorato beta) and verified in the data show in Figure 5.10), and the definition (5.4). Obviously the nonlinearity also depends on the total energy in the system, but by rescaling time it is possible to map a change of the total energy in the lattice as a change of the controlling parameter β [34], and we prefer highlighting the dependence on β as we will show in the next section numerical results of $T_{\rm eq}$ as a function of β.

5.3 Simulation results

We now turn our attention to the simulation results. Our simulations consist in the following procedure. We use $N = 64$ and $1 \leq m_i \leq m_{\max}$, with $m_{\max} = 2$, 3, 10, corresponding to $L/N = 0.55$, 0.33, 0.038. Once the system parameters are chosen, we solve numerically the eigenvalue problem. The initial conditions are of the same type that we used in the ordered lattices, as explained in Section 1.3. We arbitrarily normalize the total sum of the normal mode energy (1.25) to $E = 1$. From this initial state, we integrate numerically the equation of motion, until equipartition is reached. We define equipartition by monitoring when the entropy (1.27) reaches an arbitrary threshold value.

The integration scheme is a symplectic integrator of the 6-th order, which conserves the Hamiltonian up to the sixth decimal digit, for the long simulation time that is needed to reach equipartition. We use the implementation provided by the software *nlchains* [6].

5.3.1 Conservation of the total number of particles

In the previous analysis, we claimed that the principal mechanism of energy transmission is four-wave dynamics. If this is the case, then we expect that the number of particles (1.26) is roughly constant in time, and that in the relaxation distribution (1.29) we can observe $\mu \neq 0$. Here we show some numerical simulations in order to support this argument.

We choose $m_{\max} = 3$. Fixed ω_j, it is possible to obtain a numerical relation between the quantities E and M versus T and μ,

$$\begin{cases} E = \sum_j \omega_j \frac{T}{\omega_j + \mu} \\ M = \sum_j \frac{T}{\omega_j + \mu} \end{cases}$$

We can then arbitrarily set $E = 1$, adjust μ for different experiments, generate the corresponding value of M, and arrange the generation of the random initial state so that it has the desired values for E and M. Then we let the state evolve, and when we observe a plateau of the entropy function 1.27, then the state should be thermalized according to the relaxation distribution (1.29), and the value of E and M should not have changed significantly from the initial state. To obtain the desired value of N in the initial state, the random initial energies are randomly reshuffled until the desired value of N is obtained (with a tolerance of 0.1%). This ensure that no additional unnecessary correlations are inserted in the initial conditions. We set $\epsilon \simeq 10^{-3}$ by calculating β numerically from the initial state.

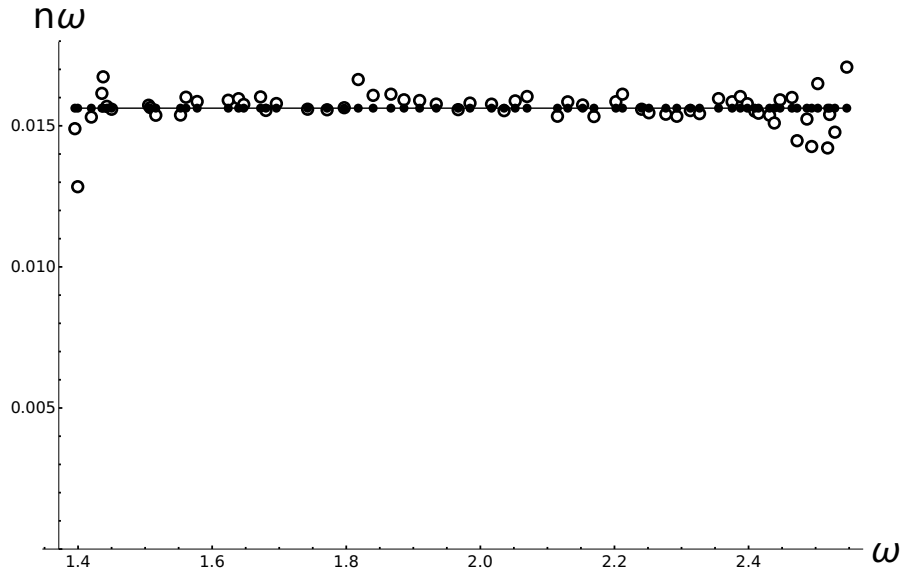


Figure 5.12: The relaxation distribution obtained for $m_{\max} = 3$, $E = 1$, $T = 0.016$, $N = 0.516$, $\mu = 0$. The solid dots (joined by a solid line for visibility) are the theoretical relaxation values of the energies, and the empty dots are the numerical data.

The state was evolved until a plateau was observed in the entropy function, the simulation run for at least $T = 10^5$ time units in physical time. In figure 5.12, 5.13 and 5.14 we compare the theoretical and numerical values of the relaxation state. The black dots are the theoretical values for $\mu = 0$, $\mu = 2$ and $\mu = -0.5$ respectively, and the empty dots the numerical relaxation values. We see that in all the three cases the relaxation of the energies is very close to the expected distribution. We can see that the edges of the curve, that is the eigenmodes with the largest or lowest frequencies are the one with the largest deviation from the equipartition value E/N. This is most evident in figure (5.12). This is however expected, because as we have argued in Section 5.1.1, in particular figure 5.3, the eigenmodes with the largest and lowest values of the system are generally the most

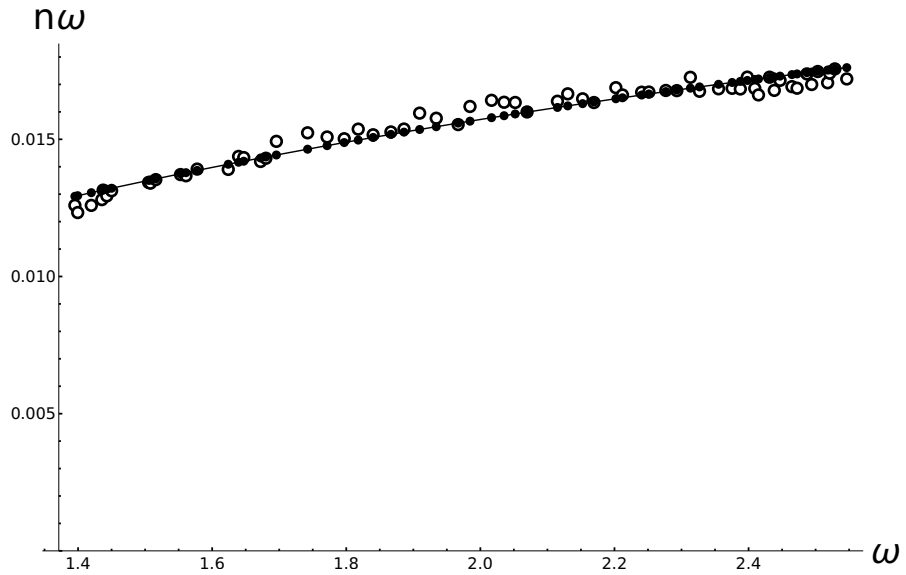


Figure 5.13: The relaxation distribution obtained for $m_{\max} = 3$, $E = 1$, $T = 0.031$, $N = 0.506$, $\mu = 2$. The solid dots (joined by a solid line for visibility) are the theoretical relaxation values of the energies, and the empty dots are the numerical data.

localized, so it takes a larger time for them to thermalize. The total linear energy E and number of particles M are conserved up to 0.006% of the initial value. No other statistical invariant quantities are expected. We propose this data to support the argument that the main process of exchange of energies is four-wave interactions of the type $2 \to 2$.

5.3.2 Scaling of the thermalization time as a function of the nonlinearity

We now turn our attention to the dependence of the thermalization time as a function of the nonlinearity. We recall that our simulations are run with parameters $N = 64$ and $1 \leq m_i \leq m_{\max}$, with $m_{\max} = 2, 3, 10$, corresponding to $L/N = 0.55, 0.33, 0.038$. We choose randomly a distribution of energies such that $\mu = 0$ (see the previous section for details). This is chosen so that we can better observe equipartition, that is the final distribution of energies is $e_i = E/N$.

We stress on the fact that the mechanism of four-wave interaction is not the same as a random noise, on the contrary it is based on a mechanism that has a meaning also in the deterministic mechanics of the single chain, that is the resonance. In fact, a resonance is a phase-locked contribution to $\dot{a}_i$ (cp. eq. (5.2)). In the terminology of [53], we are in the "strong chaos" regime. A comparison with the results of [53] is however difficult, because they give estimates for a related, but different quantity, that is the dependence of the second moment of the distribution

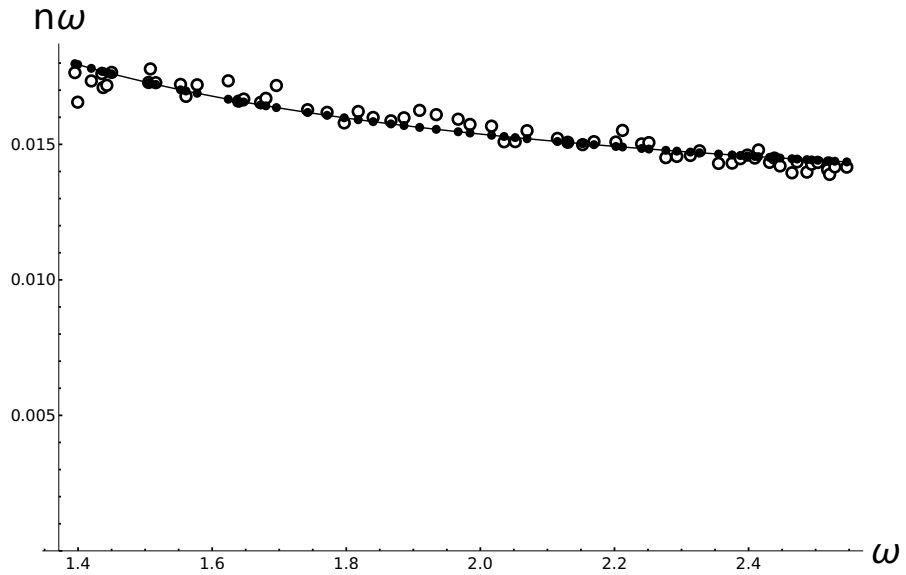


Figure 5.14: The relaxation distribution obtained for $m_{\max} = 3$, $E = 1$, $T = 0.012$, $N = 0.524$, $\mu = -0.5$. The solid dots (joined by a solid line for visibility) are the theoretical relaxation values of the energies, and the empty dots are the numerical data.

of energies as a function of time. It is not possible to calculate this quantity in our setting with $N = 64$, because in order to calculate meaningfully the second moment, one needs a much larger number of particles to avoid boundary effects. Additionally, by starting with a spatially localized distribution of energies and letting it spread out to neighbour modes, the states in the simulations of [54] will have at any moment in time a quite hard to track coexistence of different regimes, because the maximum energy density will decreases in time. By monitoring the second moment of the distribution of energies, one only obtains a measure of the *slowest* dynamics inside the wave packet. In our simulations the energy per site in the chain does not vary wildly even in the initial state (that is, there are no boundary layers in physical space, contrary to a diffusing packet), and such initial segregation of energy does not exist initially even in normal mode space, because the values for the individual energies are drawn at random.

We summarize our expectations. If the leading contribution to the equipartition of energy is four-wave interactions, then we expect the kinetic integral eq. (1.23) to hold approximately. From that, we can extract two elements that are function of the nonlinearity: β^2 (which is actually exactly proportional to ϵ), and the sum of active quasiresonances, defined according to eq. (5.15), which turns out to be essentially linear in the nonlinearity (cp. Section 5.2.2). We observe that the speed of evolution of the normal modes n_i is proportional to both, hence the equipartition time should be inversely proportional to both. We will plot the

modified equipartition time

$$T_q = \langle Q_j \rangle_j T_{\text{eq}} \tag{5.22}$$

which should display a power law of the form ϵ^{-2}, linked to four-wave interactions. In practice, we will calculate $\langle Q_i \rangle_i$ as the sum of the squared interaction coefficients of the quasiresonances that are active according to the criterion (5.15).

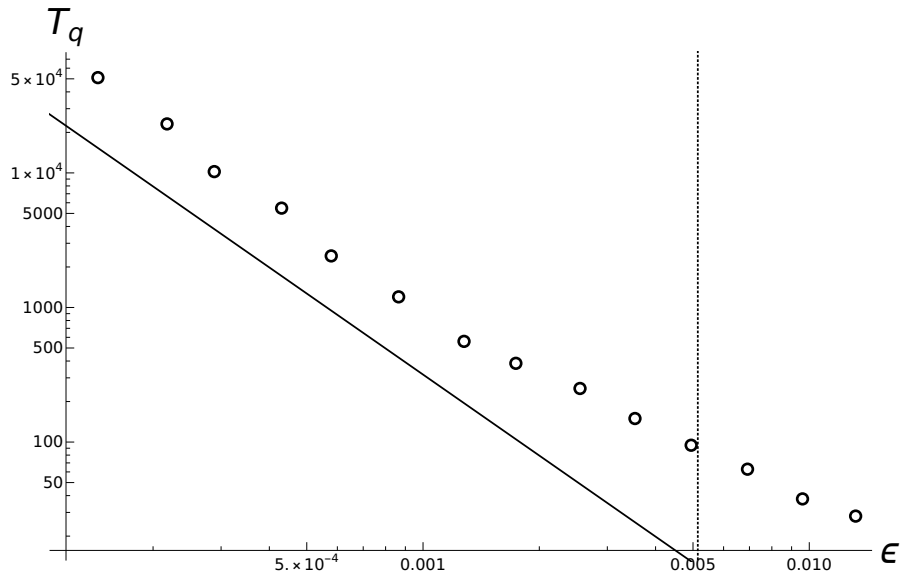


Figure 5.15: The dependence of the modified equipartition time T_q as a function of the nonlinearity ϵ in logarithmic scale, with $N = 64$, $m_{\max} = 2$. The empty dots are the experimental data, the black solid line is a power-law ϵ^{-2} to guide the eye, and the dotted vertical line is the level of nonlinearity ϵ when $\sigma_{\Delta\omega} \sim 0.1$.

In figures 5.15, 5.16, 5.17 we show the modified equipartition time (5.22), as a function of the nonlinearity ϵ (empty dots). We expect to see a power-law of the form ϵ^{-2}; to this scope we have plotted in a black line an arbitrarily scaled power-law of ϵ^{-2} for reference. We see that for the cases $m_{\max} = 2$ and $m_{\max} = 3$ there is a good agreement with the expected behaviour at the low levels of nonlinearity.

For larger nonlinearities, the data appears to deviate from a power-law with exponent -2. This is an effect that is observed during all the experiments of this kind, for all type of lattices, as we have shown in this thesis. A quantitative explanation of this deviation is elusive as of now. We can however argue that there is an upper bound for the validity of the argument put forth in the previous sections. It is evident that a quasiresonance will cease to exist when the dispersion of its beat frequency $\sigma_{\Delta\omega}$ is of the same order of magnitude of the typical frequency of the eigenmodes involved. In fact, when this happens, then the modes which participate in the resonance will decorrelate in a time comparable with the linear time scale $1/\omega$. In this regime, an effective noise theory is likely more appropriate,

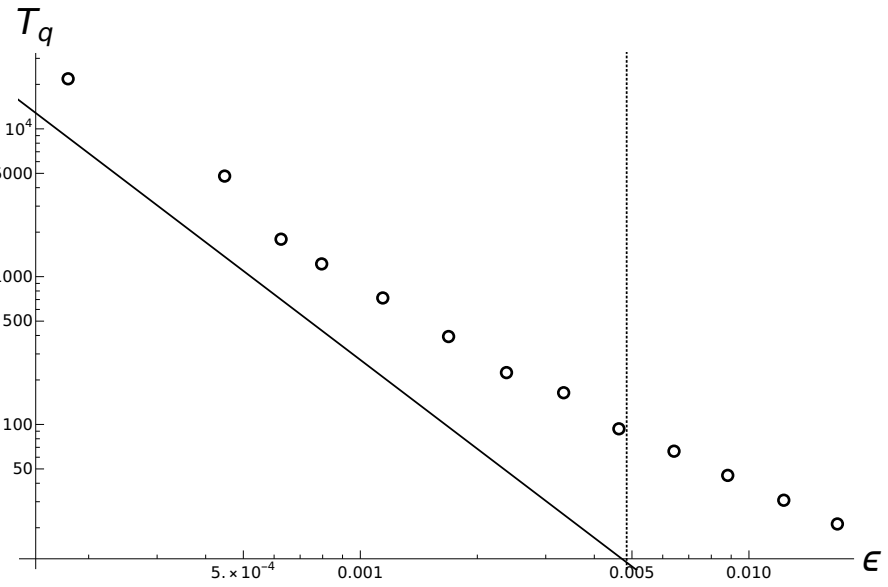


Figure 5.16: The dependence of the modified equipartition time T_q as a function of the nonlinearity ϵ in logarithmic scale, with $N = 64$, $m_{\max} = 3$. The empty dots are the experimental data, the black solid line is a power-law ϵ^{-2} to guide the eye, and the dotted vertical line is the level of nonlinearity ϵ when $\sigma_{\Delta\omega} \sim 0.1$.

as quasiresonances do not amount anymore to an almost phase-locked contribution to the equation of motion. We estimate this threshold in the nonlinearity in the following crude way. The frequencies ω_j in our realization are all of the order 1. We arbitrarily set a threshold $\sigma_{\Delta\omega} = 0.1$, that is the fluctuations of the beat frequency of the quasiresonances is one order of magnitude smaller than the typical linear frequency ω_j. We then mark with a dotted vertical line in the plots 5.15, 5.16, 5.17 the nonlinearity such that $\sigma_{\Delta\omega}$. This threshold value has been calculated numerically from the simulation data itself. We observe that this upper bound is consistent with the fact that we can fit a -2 power-law only with the leftmost datapoints of the cases $m_{\max} = 2$ and $m_{\max} = 3$.

The case $m_{\max} = 10$ shows that our arguments are clearly inapplicable when the system is extremely localized. We recall that in this case $L/N = 0.038$, which means that any eigenvalue will be essentially a single-site excitation ($L \sim 1$). The behaviour is definitely steeper than a ϵ^{-2} power-law. It is also difficult to provide more datapoints in a range comparable in nonlinearity to that of the previous datasets, because the simulations are obviously much longer due to the very low interaction coefficient between modes. Because of the discussed stochastic threshold $\sigma_{\Delta\omega} \ll \frac{1}{\omega_j}$, we also observe that the nonlinearity values for which we managed to observe thermalization within a reasonable wall-clock time are much larger than the regime where we expect four-wave dynamics to hold.

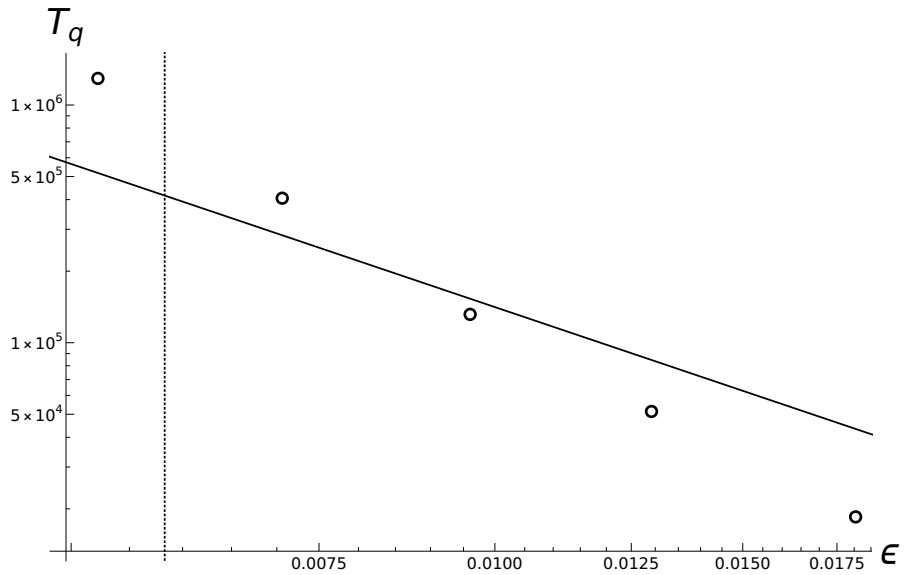


Figure 5.17: The dependence of the modified equipartition time T_q as a function of the nonlinearity ϵ in logarithmic scale, with $N = 64$, $m_{\max} = 10$. The empty dots are the experimental data, the black solid line is a power-law ϵ^{-2} to guide the eye, and the dotted vertical line is the level of nonlinearity ϵ when $\sigma_{\Delta\omega} \sim 0.1$.

We conclude that with the simulation data available to us at the moment, we observe a good agreement with the predictions of four-wave dynamics as the leading contribution to exchange of energy across modes, at least for a limited parameter space. This is the case for the experiments with $m_{\max} = 2, 3$ for a range of nonlinearity that spans roughly 10^{-4} to 10^{-3}. For larger nonlinearities, we foresee that already at levels of nonlinearity around $\epsilon = 10^{-2}$ the quick decorrelation of phases within a quasiresonance makes it difficult to claim that they are the leading contribution, and a effective noise approach is likely more applicable. We see that for extreme localization, our methodology does not give results consistent with experimental data.

Chapter 6

Conclusions and outlook

In this thesis, I have presented the results that my collaborators and I obtained on nonlinear lattice statistics. If one has to summarize our research in a single statement, it would be that we wanted to show that the exchange of energy between the modes that leads to equipartition can be explained by the presence of resonances and quasiresonances.

Our claims also provide insight into a class of initial conditions that are rarely studied in the context of nonlinear lattices. These initial conditions have many less symmetries than the ones usually studied (e.g. low or high-frequency modes) in this context. This should avoid the study of concurrent multiple leading dynamics, or different timescales that arise from inhomogeneous initial conditions. We also strived to understand "universal" traits across a number of classically studied nonlinear lattice.

Our analysis is primarily of qualitative nature. We understand the dynamics of the lattices in terms of WT arguments, but WT is a theory that only works in the large-box limit. The real dynamics are believed to be at most an approximation of the WT-predicted dynamics. We do not have a clear criterion for the minimum length of the lattice N so that our analysis to hold. A criterion for the maximum nonlinearity is proposed in Chapter 5 regarding the disordered KG lattice, but further theoretical and numerical validation is needed.

Our research highglighted that a large part of the parameter space of the lattices can be explained through the existence of quasiresonances. Quasiresonances are difficult to quantify, because it is difficult to construct an experiment in order to isolate the effect of resonances. The kinetic equation in fact holds only for large timescales, which means that the dynamics that can be observed by direct inspection in the short time are essentially irrelevant to the theory. And when one considers longer timescales, then the interactions among the modes cause the dynamics to be chaotic, and hence of difficult analysis.

In order to have more evidence that the resonances, and quasiresonances are active, several future experiments could be performed. One could look for example

at the time correlation of quasi-resonant terms, and try to observe directly the effect of the nonlinear frequency broadening. This could give definitive evidence from the microscopic level that the Chirikov criterion works.

It would be very interesting to be able to create arbitrary models with desired properties of the resonant manifold, as well as their interaction coefficient W_{1234}. From a computational point of view, this could be implemented by solving the equation of motion in normal modes space, where it would be possible to directly control and set W_{1234} to arbitrary values. It would be possible to observe the dependence of the thermalization dynamics as a function of the linear dynamics, which could give a definitive answer on how much the linearized dynamics matter in the route to thermalization.

It is important in further research to understand the role of the conservation of the number of particles. The fact that some equilibrium states can be generated with non-exact equipartition could be the sign that the dynamics that we are observing is actually a metastable state. We could not observe a second timescale where such equilibrium distribution would go to normal equipartition of energy. However, our time marching algorithm should be validated against such large timescales, assuming that they are even computationally accessible as of now.

Bibliography

[1] E. Fermi, J. Pasta, and S. Ulam, "Studies of nonlinear problems," tech. rep., I, Los Alamos Scientific Laboratory Report No. LA-1940, 1955.

[2] E. Fermi, "Dimostrazione che in generale un sistema meccanico normale è quasi ergodico," *Il Nuovo Cimento (1911-1923)*, vol. 25, no. 1, pp. 267–269, 1923.

[3] M. Onorato, L. Vozella, D. Proment, and Y. Lvov, "Route to thermalization in the α-Fermi-Pasta-Ulam system," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 14, 2015.

[4] L. Pistone, M. Onorato, and S. Chibbaro, "Thermalization in the discrete nonlinear klein-gordon chain in the wave-turbulence framework," *EPL (Europhysics Letters)*, vol. 121, no. 4, p. 44003, 2018.

[5] L. Pistone, S. Chibbaro, M. D. Bustamante, Y. V. Lvov, and M. Onorato, "Universal route to thermalization in weakly-nonlinear one-dimensional chains," *Mathematics in Engineering*, vol. 1, no. mine-01-04-672, p. 672, 2019.

[6] L. Pistone and M. Onorato, "nlchains: A fast and accurate time integration of 1-d nonlinear chains on gpus," *SoftwareX*, vol. 10, p. 100255, 2019.

[7] G. Gallavotti, *The Fermi-Pasta-Ulam problem: a status report*, vol. 728. Springer, 2008.

[8] F. M. Izrailev and B. V. Chirikov, "Statistical properties of a nonlinear string," in *Sov. Phys. Dokl*, vol. 11, pp. 30–32, 1966.

[9] M. Falcioni, U. M. B. Marconi, and A. Vulpiani, "Ergodic properties of high-dimensional symplectic maps," *Phys. Rev. A*, vol. 44, pp. 2263–2270, Aug 1991.

[10] N. J. Zabusky and M. D. Kruskal, "Interaction of "Solitons" in a Collisionless Plasma and the Recurrence of Initial States," *Phys. Rev. Lett.*, vol. 15, pp. 240–243, aug 1965.

[11] V. N. Serkin and A. Hasegawa, "Exactly integrable nonlinear schrodinger equation models with varying dispersion, nonlinearity and gain: application for soliton dispersion," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 8, pp. 418–431, May 2002.

[12] M. Toda, "Vibration of a chain with nonlinear interaction," in *Selected Papers of Morikazu Toda*, pp. 97–102, World Scientific, 1993.

[13] G. Benettin, H. Christodoulidi, and A. Ponno, "The Fermi-Pasta-Ulam Problem and Its Underlying Integrable Dynamics," *Journal of Statistical Physics*, pp. 1–18, 2013.

[14] B. Rink, "Symmetry and resonance in periodic fpu chains," *Communications in Mathematical Physics*, vol. 218, no. 3, pp. 665–685, 2001.

[15] B. Rink, "Proof of Nishida's conjecture on anharmonic lattices," *Communications in mathematical physics*, vol. 261, no. 3, pp. 613–627, 2006.

[16] M. D. Bustamante, K. Hutchinson, Y. V. Lvov, and M. Onorato, "Exact discrete resonances in the fermi-pasta-ulam–tsingou system," *Communications in Nonlinear Science and Numerical Simulation*, vol. 73, pp. 437–471, 2019.

[17] B. V. Chirikov, "Resonance processes in magnetic traps," *The Soviet Journal of Atomic Energy*, vol. 6, no. 6, pp. 464–470, 1960.

[18] C. Skokos, D. Krimer, S. Komineas, and S. Flach, "Delocalization of wave packets in disordered nonlinear chains," *Physical Review E*, vol. 79, no. 5, p. 056211, 2009.

[19] N. J. Zabusky and G. S. Deem, "Dynamics of nonlinear lattices i. localized optical excitations, acoustic radiation, and strong nonlinear behavior," *Journal of Computational Physics*, vol. 2, no. 2, pp. 126–153, 1967.

[20] H. Kantz, "Vanishing stability thresholds in the thermodynamic limit of non-integrable conservative systems," *Physica D: Nonlinear Phenomena*, vol. 39, no. 2-3, pp. 322–335, 1989.

[21] D. Bambusi and A. Ponno, "On metastability in fpu," *Communications in mathematical physics*, vol. 264, no. 2, pp. 539–561, 2006.

[22] A. Majda, D. McLaughlin, and E. Tabak, "A one-dimensional model for dispersive wave turbulence," *Journal of Nonlinear Science*, vol. 7, no. 1, pp. 9–44, 1997.

[23] D. Cai, A. J. Majda, D. W. McLaughlin, and E. G. Tabak, "Dispersive wave turbulence in one dimension," *Physica D: Nonlinear Phenomena*, vol. 152, pp. 551–572, 2001.

[24] K. Hasselmann, "On the non-linear energy transfer in a gravity-wave spectrum part 1. general theory," *Journal of Fluid Mechanics*, vol. 12, no. 4, pp. 481–500, 1962.

[25] A. Pushkarev, "On the kolmogorov and frozen turbulence in numerical simulation of capillary waves," *European Journal of Mechanics-B/Fluids*, vol. 18, no. 3, pp. 345–351, 1999.

[26] V. E. Zakharov, V. S. L'vov, and G. Falkovich, *Kolmogorov spectra of turbulence I: Wave turbulence*. Springer Science & Business Media, 2012.

[27] Y. V. Lvov, S. Nazarenko, and B. Pokorni, "Discreteness and its effect on water-wave turbulence," *Physica D*, vol. 218, pp. 24–35, 2006.

[28] E. Kartashova, "Nonlinear Resonance Analysis," *Nonlinear Resonance Analysis, by Elena Kartashova, Cambridge, UK: Cambridge University Press, 2010*, vol. 1, 2010.

[29] S. Nazarenko, *Wave turbulence*, vol. 825. Springer, 2011.

[30] A. C. Newell and B. Rumpf, "Wave turbulence," *Annual review of fluid mechanics*, vol. 43, pp. 59–78, 2011.

[31] C. Cercignani, U. Gerasimenko, and D. Y. Petrina, *Many-particle dynamics and kinetic equations*, vol. 420. Springer Science & Business Media, 2012.

[32] A. Picozzi, J. Garnier, T. Hansson, P. Suret, S. Randoux, G. Millot, and D. N. Christodoulides, "Optical wave turbulence: Towards a unified nonequilibrium thermodynamic formulation of statistical nonlinear optics," *Physics Reports*, vol. 542, no. 1, pp. 1–132, 2014.

[33] Y. V. Lvov and M. Onorato, "Double scaling in the relaxation time in the β-fermi-pasta-ulam-tsingou model," *Physical review letters*, vol. 120, no. 14, p. 144301, 2018.

[34] W. Fu, Y. Zhang, and H. Zhao, "Universal scaling of the thermalization time in one-dimensional lattices," *Physical Review E*, vol. 100, no. 1, p. 010101, 2019.

[35] W. Fu, Y. Zhang, and H. Zhao, "Universal law of thermalization for one-dimensional perturbed toda lattices," *New Journal of Physics*, vol. 21, no. 4, p. 043009, 2019.

[36] P. Martinsson and A. Movchan, "Vibrations of lattice structures and phononic band gaps," *Quarterly Journal of Mechanics and Applied Mathematics*, vol. 56, no. 1, pp. 45–64, 2003.

[37] S. Nazarenko, "Sandpile behaviour in discrete water-wave turbulence," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2006, no. 02, p. L02002, 2006.

[38] F. Fucito, F. Marchesoni, E. Marinari, G. Parisi, L. Peliti, S. Ruffo, and A. Vulpiani, "Approach to equilibrium in a chain of nonlinear oscillators," *Journal de Physique*, vol. 43, no. 5, pp. 707–713, 1982.

[39] J. De Luca and A. Lichtenberg, "Transitions and time scales to equipartition in oscillator chains: low-frequency initial conditions," *Physical Review E*, vol. 66, no. 2, p. 026206, 2002.

[40] E. Kartashova, "Exact and quasiresonances in discrete water wave turbulence," *Physical review letters*, vol. 98, no. 21, p. 214502, 2007.

[41] A. Fusaro, J. Garnier, K. Krupa, G. Millot, and A. Picozzi, "Dramatic acceleration of wave condensation mediated by disorder in multimode fibers," *Physical review letters*, vol. 122, no. 12, p. 123902, 2019.

[42] P. W. Anderson, "Absence of diffusion in certain random lattices," *Physical review*, vol. 109, no. 5, p. 1492, 1958.

[43] S. A. Gredeskul and Y. S. Kivshar, "Propagation and scattering of nonlinear waves in disordered systems," *Physics reports*, vol. 216, no. 1, pp. 1–61, 1992.

[44] T. Schwartz, G. Bartal, S. Fishman, and M. Segev, "Transport and anderson localization in disordered two-dimensional photonic lattices," *Nature*, vol. 446, no. 7131, pp. 52–55, 2007.

[45] A. Iomin and S. Fishman, "Localization length of stationary states in the nonlinear schrödinger equation," *Physical Review E*, vol. 76, no. 5, p. 056607, 2007.

[46] G. Kopidakis, S. Komineas, S. Flach, and S. Aubry, "Absence of wave packet diffusion in disordered nonlinear systems," *Physical Review Letters*, vol. 100, no. 8, p. 084103, 2008.

[47] B. Kramer and A. MacKinnon, "Localization: theory and experiment," *Reports on Progress in Physics*, vol. 56, no. 12, p. 1469, 1993.

[48] I. M. Lifshitz *et al.*, "Energy spectrum structure and quantum states of disordered condensed systems," *Soviet Physics Uspekhi*, vol. 7, no. 4, p. 549, 1965.

[49] S. Nazarenko, A. Soffer, and M.-B. Tran, "On the wave turbulence theory for the nonlinear schrödinger equation with random potentials," *Entropy*, vol. 21, no. 9, p. 823, 2019.

[50] D. O. Krimer and S. Flach, "Statistics of wave interactions in nonlinear disordered systems," *Phys. Rev. E*, vol. 82, p. 046221, Oct 2010.

[51] C. Besse and J.-C. Garreau, *Nonlinear Optical and Atomic Systems: At the Interface of Physics and Mathematics*, vol. 2146. 08 2015.

[52] X. Yu and S. Flach, "Enhancement of chaotic subdiffusion in disordered ladders with synthetic gauge fields," *Physical Review E*, vol. 90, no. 3, p. 032910, 2014.

[53] S. Flach, "Spreading of waves in nonlinear disordered media," *Chemical Physics*, vol. 375, no. 2-3, pp. 548–556, 2010.

[54] S. Flach, D. Krimer, and C. Skokos, "Universal spreading of wave packets in disordered nonlinear systems," *Physical Review Letters*, vol. 102, no. 2, p. 024101, 2009.