Perturbation theory at the thermodynamic limit

Alberto Mario Maiocchi

Contents

Introduction

Classical Hamiltonian perturbation theory deals with quasi-integrable systems and aims at controlling how much, during the dynamical evolution, some relevant quantities deviate with respect to the unperturbed case. Unfortunately, the classical methods do not apply in the thermodynamic limit (namely, when the number N of degrees of freedom tends to infinity and the energy divided by N, the specific energy, remains constant), and so, classical perturbation theory as it stands is useless in the frame of statistical mechanics. Essentially, the reason for the failure is that in the classical theory one pretends to control all initial data, even the ones that should be obviously considered exceptional (think, for instance, of the data for which all the energy is concentrated on a single degree of freedom). On the other hand, for the aims of statistical mechanics, only mean values with respect to a given probability measure are relevant, without the necessity of controlling all individual motions, so that a weaker approach is sufficient. Only a few years ago, in the work [1], an approach along these lines was proposed and an application to a particular model was made, by performing a few steps of perturbation theory, with estimates uniform in N at a finite specific energy. This opened the way to the construction of a "perturbation theory at the thermodynamic limit".

In the present thesis, we intend to expound such a theory as an organic whole. The principal aim is to find a notion which may be an analogue of the familiar adiabatic invariant of classical perturbation theory, although in a weaker sense, inasmuch as it should involve phase averages. The proposal we make is to consider as analogues of adiabatic invariants those quantities whose time autocorrelation function remains almost constant for long times (long, in the usual sense of perturbation theory, as, typically, in Nekhoroshev's sense). So, we decided to focus on the study of time autocorrelation functions in Hamiltonian dynamical systems endowed with an invariant probability measure. Indeed, time correlations are fundamental in ergodic theory, because, for instance, the property of mixing amounts to the requirement that the time correlations of all dynamical variables decay to zero. Thus, proving

the existence of an adiabatic invariant for long times amounts to proving that a system does not behave as a mixing one up to these times. In addition, time correlations are very well known to be fundamental ingredients in the whole field of linear response theory.

The main original result obtained along these lines is the construction of an adiabatic invariant for a Φ^4 model. Moreover, some general mathematical properties concerning time autocorrelation functions were discussed. Finally, an application of these ideas to plasma physics was made, giving a possible explanation of an instability concerning plasma magnetic confinement, which apparently was up to now unexplained.

The physical motivation for the interest of time autocorrelation functions, mainly in the context of linear response theory, is recalled in a heuristic way in Chapter 1, with a particular emphasis on the possible significance of results holding *for finite times*. The general features of the perturbation approach here developed are illustrated in Chapter 2, with which the proper mathematical part of the thesis begins. The application to the Φ^4 model is given in Chapter 3. This allows us to indicate on general grounds which are the problems arising in implementing the proposed perturbation scheme, in particular in connection with the estimates related to the marginal and conditional probabilities of the invariant measure. In Chapter 4 a study is presented concerning the general analytic properties of the time autocorrelation functions. Finally, in Chapter 5 a heuristic application to the problem of plasma magnetic confinement is given, together with some concluding remarks.

Chapter 1

The role of correlations in statistical physics

The aim of this chapter is to give a short discussion, with no pretension of rigour, about the role played by time correlations in statistical physics. We will give no justification for the formal passages performed, avoiding entering the interesting mathematical problems they raise.

More in general, one could ask whether the microscopic *dynamics* plays any role at all in determining the relevant macroscopic observables. Standard textbooks of statistical mechanics take as starting point the Gibbs measure, while dynamics is essentially avoided, being implicitly taken into account only through generic assumptions such as that integrals of motions should not exist, apart from energy. This suffices for the study of the means for macroscopic observables (for instance, density, pressure) in the commonly considered systems, as first shown by Khinchin (see [2]), but seems to be non completely satisfactory for dealing with response functions to perturbing fields, such as electric conductivity, magnetic susceptibility and compressibility.

In fact, in the literature one finds some different approaches to the definitions of such quantities. The first one is to use as thermodynamic definitions of response functions the derivatives of suitable effective thermodynamic potentials, in which no reference to the dynamics is made. Another common approach is that of linear response theory as formulated by Kubo and others (see [3]). Such quantities are indeed defined in this frame through time correlations of some observables, via the fluctuation–dissipation theorem. Thus, according to this approach, the response functions essentially depend on the dynamics, and, in particular, on the dynamics of the unperturbed system, to which no perturbing field is applied. Finally, a more rigorous way of dealing with these questions, which concern substantially nonequilibrium phenom-

ena, was presented by Gallavotti and Cohen (see [4]), whose theory relies on the so–called "chaotic hypothesis". Also in this case dynamics is taken into account, and time correlations can be viewed as indicators for establishing whether the chaotic hypothesis is satisfied or not.

In the case of small fields, all these approaches should lead to the same results if the decay of the time correlations involved is of a suitable kind: in particular, if the correlations go to zero sufficiently fast as time increases. It is clear, on the other hand, that this cannot happen in the case of completely integrable systems. Moreover, the differences could even become larger when one addresses the problem of response for finite times, where metastability phenomena could occur, as, for instance, in glasses.

We will first recall, in a quick way, Kubo's linear response theory. The significance of the results holding for finite times will be particularly stressed. The theory is introduced through the canonical example of electric conductivity in Section 1.1. In the subsequent Section 1.2 the case of magnetic susceptibility is expounded, for which the results for finite time are of particular interest, since they provide a possible microscopic explanation for diamagnetism within the frame of classical physics. In Section 1.3, we deal with the interesting case of specific heat, which does not properly fall within the frame of Kubo's theory. Here, too, the finite time value differs from the equilibrium value, if a certain time autocorrelation function does not vanish.

1.1 Linear response theory for finite times

Here, as in the whole thesis, we will consider Hamiltonian dynamical systems endowed with a measure on phase space $\mathcal{M}$ (a generic point of which will be denoted by z). We intend to study how a system with Hamiltonian $H = H(z)$, endowed with a measure μ invariant for that Hamiltonian, reacts to the introduction of a small external field as a perturbation. The latter is conceived as being of Hamiltonian nature and depending explicitly on time t, as described by the addition of a term of the form

$$\Delta H(z,t) \stackrel{\text{def}}{=} -h(t)A(z) \ .$$

Linear response theory shows that the perturbation induces an evolution of the measure μ such that, *at first order in the perturbing field*, the variation of the phase average of a given observable with respect to the unperturbed value depends on the *unperturbed (or equilibrium)* time correlation of two suitable observables. The response coefficients are then obtained from the variation of the mean values of some observables in response to suitable perturbing fields (see below).

In order to formulate the theory, we suppose that the measure μ has a density $\rho(z)$ and that for $t = -\infty$ the perturbation is absent (i.e., the limit of $h(t)$ as $t \to -\infty$ vanishes; one then says that the perturbation is adiabatically turned on). The density $\tilde{\rho}(z,t)$ of the perturbed system must then satisfy Liouville's equation

$$\begin{cases} \dfrac{\partial \tilde{\rho}}{\partial t} = [H, \tilde{\rho}] - h(t)\,[A, \tilde{\rho}] \\ \lim\limits_{t \to -\infty} \tilde{\rho}(t,z) = \rho(z) \end{cases} , \tag{1.1}$$

in which the symbol $[\cdot,\cdot]$ denotes canonical Poisson brackets.

Instead of solving the previous equation, we content ourselves to studying the linear approximation, namely, to consider the deviation $\Delta\rho$ from the unperturbed density ρ, i.e., to consider

$$\Delta\rho(z,t) \stackrel{\text{def}}{=} \tilde{\rho}(z,t) - \rho(z,t) \ ,$$

which, in turn, has to satisfy the equation

$$\begin{cases} \dfrac{\partial \Delta\rho}{\partial t} = [H, \Delta\rho] - h(t)\,[A, \rho] \\ \lim\limits_{t \to -\infty} \Delta\rho = 0 \end{cases} ,$$

from equation (1.1), using the fact that $[H,\rho] = 0$ (since μ is invariant for the flow induced by H), and by neglecting the term $-h(t)[A,\Delta\rho]$, which is formally of a higher perturbation order.

The solution of the previous equation can be expressed as

$$\Delta\rho(t) = -\int_{-\infty}^{t} \hat{U}_{t-s}\left([A,\rho]\right) h(s)\mathrm{d}s \ ,$$

where the time evolution operator $\hat{U}_t$ relative to the Hamiltonian H appears,[1] which acts on the dynamic variable $[A,\rho]$. This implies that the mean value $\langle \Delta B(t) \rangle$ of the deviation at time t of a given observable $B = B(z)$ from its equilibrium value can be written as

$$\begin{aligned} \langle \Delta B(t) \rangle &\stackrel{\text{def}}{=} \int_{\mathcal{M}} \Delta\rho(z,t)\, B(z)\mathrm{d}z \\ &= -\int_{\mathcal{M}} \mathrm{d}z \int_{-\infty}^{t} \mathrm{d}s\, h(s)\, \hat{U}_{t-s}\left([A,\rho]\right)(z)\, B(z) \ . \end{aligned} \tag{1.2}$$

[1] We will always adopt the convention that the time evolution operator $\hat{U}_t$ relative to an Hamiltonian H acts on a dynamic variable f by mapping $f(z)$ in $f(\Phi^{-t}z)$, where Φ^t is the flow at time t induced by H.

The relation of the previous formula with time correlation functions is quite simple if the initial measure is the Gibbs one, as is usually assumed. This means that we make the choice $\rho \stackrel{\text{def}}{=} \exp(-\beta H)/Z$, in which $\beta > 0$ and Z is the partition function, which normalizes to 1 the measure of the whole space. One has, then,

$$[A, \rho] = -\beta\, [A, H]\, \rho$$

and, by exchanging the order of integration,

$$\langle \Delta B(t) \rangle = \beta \int_{-\infty}^{t} \mathrm{d}s\, h(s) \int_{\mathcal{M}} \mathrm{d}z\, \hat{U}_{t-s} A'(z)\, B(z)\, \rho(z)\ , \tag{1.3}$$

where we have put $A' \stackrel{\text{def}}{=} [A, H]$. The second integral at the r.h.s. is the *equilibrium* time correlation function between A' and B at time $t - s$, which we presently denote by $\langle A'(t-s)B(0)\rangle$ (for a more precise mathematical setting for the correlations, see Chapter 2).

In order to elucidate the relation between the introduced quantities and the response coefficient we prefer to focus on one of them, because some small differences in dealing with different coefficients make a unitary treatment cumbersome. Maybe the simplest case is that of electric conductivity σ of a conducting body under a quasi–stationary electric field. To this end, we consider a system of N charged particles, each with charge e_j, governed by a Hamiltonian H, and introduce an electric field $\mathbf{E}$, which is assumed to be uniform in space and to depend on time only through the adiabatic switching. This means that we can write the perturbing term to the Hamiltonian as

$$\Delta H(t) = -h(t) \sum_{j=1}^{N} e_j x_j E_x\ , \quad \lim_{t \to -\infty} h(t) = 0\ , \quad h(t) = 1 \text{ for } t \geq 0\ ,$$

in which x_j denotes the coordinate along the x axis of a suitable reference frame of the j–th particle.

The electric conductivity[2] is defined as the ratio between the current induced along the x axis, whose microscopic counterpart is given by $J_x \stackrel{\text{def}}{=} \sum_{j=1}^{N} e_j \dot{x}_j$ (the dot denoting differentiation with respect to time), and the electric field. According to formula (1.3), the mean value of J_x at time t is, in the linear approximation,

$$\langle J_x(t) \rangle = \beta E_x \int_{-\infty}^{t} \mathrm{d}s\, h(s) \langle J_x(t-s) J_x(0) \rangle,$$

[2]One should more properly talk about the conductivity tensor, which, in the case of isotropy that we implicitly assume, reduces to a scalar.

where use is made of the relation $[\sum_j e_j x_j, H] = J_x$ and of the remark that the equilibrium mean value of J_x vanishes. Notice that on the r.h.s. there appears the time correlation of the current with itself, which is called the time autocorrelation of the current; indeed the response coefficients often involve the time autocorrelation of an observable. The conductivity in response to a stationary field is then given by

$$\sigma(t) \stackrel{\text{def}}{=} \frac{\langle J_x(t) \rangle}{E_x} = \beta \int_{-\infty}^{t} \mathrm{d}s \, h(s) \langle J_x(t-s) J_x(0) \rangle \ , \tag{1.4}$$

for $t > 0$. This expression, which is a fluctuation–dissipation relation, depends on time t. However, one usually supposes that, exception being made for a very short initial transient depending on the form of $h(t)$, the equilibrium time autocorrelation $\langle J_x(t) J_x(0) \rangle$ of the current decays so fast that the integral at the r.h.s. is practically constant. So, one can think of the conductivity as of a quantity independent of the instant at which the measurements are performed and of their duration. One writes, then, the following formula

$$\sigma = \beta \int_{-\infty}^{+\infty} \mathrm{d}s \, h(s) \langle J_x(t-s) J_x(0) \rangle \ . \tag{1.5}$$

However, the supplementary assumption on the decay of the correlation should be proved in terms of the dynamics and is certainly false, for instance, in the trivial integrable case in which the equilibrium Hamiltonian H is simply the sum of the kinetic energies of the particles[3]

This is a well known problem, and there exists an ample branch of ergodic theory which deals with asymptotic properties of the decay of correlations for different classes of dynamical systems (see [5]), determining under which hypotheses the integral in (1.5) converges. We want, instead, to omit for the moment the study of the asymptotic behaviour of the correlations (see, however, Chapter 4 for some remarks on this point), and focus on the behaviour for finite times. In fact, even if we admit that formula (1.5) provides a well–defined expression for the conductivity σ, nothing allows us to rule out that the actual value $\sigma(t)$ of the conductivity at time t, as defined by (1.4), could be essentially different from the infinite time value σ even for very large times, a fact that might be detected by experimental measurements. This is precisely what happens if the autocorrelation of the current does not completely decay within the measurement time. For this reason, our aim will be to show that there exist quasi–integrable systems in which the autocorrelation of some observables does not decay up to long times, which tend to infinity as the perturbation with respect to the integrable system goes to zero.

[3] In such a case, the integral in formula (1.5) would diverge.

1.2 Magnetic susceptibility and diamagnetism

A case in which the difference between the value of the response function at an infinite time and that at a finite time turns out to be particularly relevant is that of magnetic susceptibility. Indeed, in work [6] it has been proposed that such a difference provides a microscopic explanation of diamagnetism in the frame of classical physics, at least for what concerns magnetized plasmas.

Let us consider a system of N charged particles, each with charge e_j and unitary mass, subject to a Hamiltonian

$$H = \sum_{j=1}^{N} \frac{\mathbf{p}_j^2}{2} + V(\mathbf{q}_1, \ldots, \mathbf{q}_N) \ ,$$

where $\mathbf{p}_j$ and $\mathbf{q}_j$ denote the momentum and position vector of the j–th particle in a three dimensional region, the volume of which can be taken as unitary. The adiabatic switching of a uniform magnetic field $h(t)\,\mathbf{B}$ determines a modification to the Hamiltonian, which, in the linear approximation, takes the form

$$\Delta H(t) = -h(t) \sum_{j=1}^{N} e_j \mathbf{p}_j \cdot (\mathbf{B} \wedge \mathbf{q}_j) \ , \quad \lim_{t \to -\infty} h(t) = 0 \ , \quad h(t) = 1 \text{ for } t \geq 0 \ ,$$

for a suitable choice of units.

The magnetic susceptibility[4] is the ratio between the magnetization $\mathbf{M}$ induced along the axis of the magnetic field (which we take as z axis) and the field itself. One should use some caution here in defining the magnetization $\mathbf{M}$, because it depends explicitly on time if expressed as a function of p and q, as a consequence of the added perturbation. Indeed, we recall that for one particle the magnetic moment is just a multiple of the angular momentum, and that, in our case, one has

$$\mathbf{M}(p, q, t) \stackrel{\text{def}}{=} \sum_{j=1}^{N} e_j \mathbf{q}_j \wedge \dot{\mathbf{q}}_j = \sum_{j=1}^{N} e_j \mathbf{q}_j \wedge \mathbf{p}_j - h(t) \sum_{j=1}^{N} e_j^2 \mathbf{q}_j \wedge (\mathbf{B} \wedge \mathbf{q}_j) \ . \quad (1.6)$$

The magnetization of the unperturbed system, which is obtained from the previous one for $h = 0$, will be denoted by $\mathbf{M}^0$.

Due to formula (1.3) and to the fact that $\Delta H(t) = -h(t)\mathbf{B} \cdot \mathbf{M}^0$, one

[4] We should here, too, talk about a tensor (see footnote 2).

gets, in the linear approximation, the relation

$$\langle M_z(t)\rangle = B_z\Bigg(\beta\int_{-\infty}^{t}\mathrm{d}s\,h(s)\langle\hat{U}_{t-s}\left([M_z^0,H]\right)M_z^0\rangle \\ -h(t)\sum_{j=1}^{N}e_j^2\langle x_j^2+y_j^2\rangle\Bigg)\ . \tag{1.7}$$

Then, since $[M_z^0, H]$ is the total derivative with respect to time of M_z^0 under the unperturbed flow, we can write

$$\langle\hat{U}_{t-s}\left([M_z^0,H]\right)M_z^0\rangle = \frac{\mathrm{d}}{\mathrm{d}s}\langle M_z^0(t-s)M_z^0\rangle\ ,$$

in which the equilibrium autocorrelation of the unperturbed magnetization appears. After an integration by parts of the integral in the first row of (1.7), one gets a boundary term which cancels exactly the second row, so that the magnetic susceptibility χ turns out to be given by

$$\chi(t)\stackrel{\mathrm{def}}{=}\frac{\langle M_z(t)\rangle}{B_z} = -\beta\int_{-\infty}^{t}\mathrm{d}s\,\left(\frac{\mathrm{d}}{\mathrm{d}s}h(s)\right)\langle M_z^0(t-s)M_z^0\rangle\ . \tag{1.8}$$

The previous formula shows clearly that the vanishing value predicted by classical theory (see [7]) is obtained only when t exceeds the time at which the magnetic field becomes constant (i.e., after which the derivative of $h(s)$ vanishes) at least by the time needed for the autocorrelation of M_z^0 to vanish. If this does not happen, the susceptibility can take a nonvanishing value, which can also remain practically constant for long times, if the system lies in a metastable state. For the consequences of this fact for magnetized plasmas, see Chapter 5 later on.

1.3 The specific heat

Quite different is the case of specific heat, which cannot be reduced to the study of time correlations in an isolated system. We present here a treatment due to Carati and Galgani, in which the reversibility of the microscopic system plays a key role.

Let us consider a Hamiltonian system with Hamiltonian H_1, initially at equilibrium with respect to the Gibbs measure at inverse temperature β in its phase space $\mathcal{M}_1$, which is put in contact at time $t = 0$ with a system governed by a Hamiltonian H_2, at equilibrium with respect to the Gibbs measure on its phase space $\mathcal{M}_2$ at inverse temperature $\beta+\Delta\beta$. The latter system represents

a heat bath, so that we think of it as much larger than the first one, and we look at the way in which a global equilibrium at inverse temperature $\beta + \Delta\beta$ is attained. We aim at describing the occurrence of an exchange of heat with no macroscopic work. So, we impose that the phase spaces are not altered by the contact, which is described only by an interaction term H_{int} introduced in the global Hamiltonian, so small that it does not changes the probability measure in a sensible way, but only influences the dynamics. The global system is governed by

$$H_{tot}(p_1, q_1, p_2, q_2, t) = H_1(p_1, q_1) + H_2(p_2, q_2) + \theta(t) H_{int}(p_1, q_1, p_2, q_2) \ ,$$

in which $(p_1, q_1) \in \mathcal{M}_1$, $(p_2, q_2) \in \mathcal{M}_2$ and θ denotes the Heaviside step function.[5] In the following we will also denote by $z = (p, q)$ the canonical coordinates on the global phase space $\mathcal{M} \stackrel{\text{def}}{=} \mathcal{M}_1 \times \mathcal{M}_2$.

We suppose further that the Hamiltonians H_{tot}, H_1, H_2 are invariant under the parity transformation $\mathcal{P}$ which changes all momenta p in $-p$. As is well known, this implies that the global system and each of the subsystems are microscopically reversible, i.e., that $(\mathcal{P}\Phi^t)^{-1} = \mathcal{P}\Phi^t$ for each of the flows Φ^t at time t induced on phase space by each of the Hamiltonians.

In virtue of the first principle of thermodynamics, the heat Q absorbed by the first system in the absence of work is equal to the variation ΔU of its internal energy. If we identify ΔU with the mean value of the deviation of H_1 from its initial value, we have

$$Q = \Delta U \stackrel{\text{def}}{=} \int_{\mathcal{M}} \left(H_1(\Phi^t_{tot} z) - H_1(z) \right) \rho(z) \, \mathrm{d}z \ , \tag{1.9}$$

where Φ^t_{tot} denotes the flow induced at time t by H_{tot} and ρ is the probability density for the initial data, which corresponds to a nonequilibrium measure, given by[6]

$$\rho(z) = \frac{\exp(-\beta H_1(p_1, q_1) - (\beta + \Delta\beta) H_2(p_2, q_2))}{Z_1(\beta) Z_2(\beta + \Delta\beta)} \ ,$$

in which Z_i denotes partition function for system $i = 1, 2$.

Due to reversibility, we can write Q in an equivalent manner. We start from (1.9) and make the change of variables from z to $x \stackrel{\text{def}}{=} \mathcal{P}\Phi^t z$, getting that

$$Q = \int_{\mathcal{M}} \left(H_1(x) - H_1\left(\Phi^t_{tot} x\right) \right) \rho(z(x)) \mathrm{d}x \ ,$$

[5] For definiteness, we adopt the convention $h(0) = 0$.

[6] This approach is in some sense complementary to that of Kubo presented in the previous sections. Here only the initial measure is considered, instead of a measure depending explicitly on time, as a solution of Liouville's equation.

where use is made of the invariance of H_1 under the parity transformation and of the fact that the Jacobian of the coordinate transformation has determinant of modulus one, since it is a canonical transformation. In order to express the density probability in the new variables, we make use of the conservation of energy in the form

$$H_1(z) + H_2(z) = H_1(x) + H_2(x) \ ,$$

H_{int} having been neglected. This implies that

$$\rho(z) = \rho(x) \exp\left(-\Delta\beta \left(H_1(x) - H_1(z)\right)\right) \ ,$$

leading for Q to the expression

$$Q = \int_{\mathcal{M}} \left(H_1(x) - H_1\left(\Phi^t_{tot} x\right)\right) \exp\left[\Delta\beta \left(H_1(\Phi^t_{tot} x) - H_1(x)\right)\right] \rho(x) \, \mathrm{d}x \ . \quad (1.10)$$

Finally, one writes Q as one half the sum of the r.h.s. of formulas (1.9-1.10) and expands the exponential to first order in $\Delta\beta$, obtaining the equality (valid in the linear approximation)

$$Q = -\frac{1}{2}\Delta\beta \int_{\mathcal{M}} \left(H_1(\Phi^t_{tot} z) - H_1(z)\right)^2 \rho(z) \, \mathrm{d}z \ .$$

As a first remark, we stress that the previous formula proves the validity of the second principle of thermodynamics for this case, and thus proves macroscopic irreversibility in statistical terms, by making explicit use of microscopic *reversibility*. Indeed, due to the definite sign of all terms at the r.h.s. and to the fact that β is the inverse temperature, it is easy to check that heat always flows from the hot to the cold system.

Then, we point out that this expression enables one to get a theoretical formula for the specific heat C_V of the first system,[7] as is measured when it is put in a calorimeter (which is modeled by system 2)[8] for a time t. The specific heat can be defined by $C_V \stackrel{\text{def}}{=} Q/\Delta T$, with $T \stackrel{\text{def}}{=} 1/\beta$, so that one has

$$C_V(t) = \frac{1}{2T^2} \langle \left(\hat{U}_{-t} H_1 - H_1\right)^2 \rangle \ ,$$

in which $\langle \cdot \rangle$ denotes mean value with respect to the initial density ρ and $\hat{U}_t$ the time evolution operator for the global system.

[7] The index V emphasizes that we are dealing with the specific heat for transformations in which no macroscopic work is performed, as for the transformations at constant volume in a gas.

[8] This is the only point where use is made of the assumption that the second system is much larger than the first one.

We have thus obtained again a response function which varies with time, and it can actually be expressed in terms of the time autocorrelation of H_1. This is obtained by expanding the square in the previous expression and noting that, neglecting terms of a higher order in $\Delta\beta$, the following approximate equalities hold: $\langle(\hat{U}_t H_1)^2\rangle = \langle H_1^2\rangle$ and $\langle\hat{U}_{-t} H_1 H_1\rangle = \langle\hat{U}_t H_1 H_1\rangle$. The same expansion in $\Delta\beta$ shows that all mean values with respect to ρ can be replaced by mean values with respect to the *equilibrium* measure at inverse temperature β, which will be denoted by $\langle\cdot\rangle_{eq}$. Thus, one gets

$$C_V(t) = \frac{1}{T^2}\left(\langle H_1^2\rangle_{eq} - \langle H_1\rangle_{eq}^2\right) - \frac{1}{T^2}\left(\langle H_1(t) H_1\rangle_{eq} - \langle H_1\rangle_{eq}^2\right) \ ,$$

in which the first term in brackets is the specific heat as predicted by equilibrium statistical mechanics, i.e., the variance of the internal energy (divided by T^2) of the first system, while the second is, apart from a factor, the time autocorrelation of H_1.[9] Therefore, we must conclude that the specific heat attains its equilibrium value only when the time autocorrelation of the energy has decayed to zero. If this decay is slow, one should detect different values of the specific heat, depending on the observation time, as, on the other hand, is in some cases experimentally observed (see, for instance, [8, 9]).

[9]For the presence of the term $\langle H_1^2\rangle$ see the rigorous definition of time autocorrelation in equation (2.1) later on and the subsequent remark.

Chapter 2

Time correlations and perturbation theory

As said in the introduction, we will focus on the study the time autocorrelations of dynamical variables in isolated Hamiltonian systems. We have to construct a suitable Hamiltonian perturbation theory, in order to estimate the difference of behaviour between an integrable system, in which some observables (for instance, the actions) have a constant autocorrelation, and its perturbation.

In the present chapter we introduce the basic, albeit simple, results, which enable us to deal with the time autocorrelation of a dynamical variable through perturbation methods. The most relevant estimate is contained in Theorem 1 in Section 2.1, which is a slight modification to Theorem 1 of paper [1], and is followed by a summary of the perturbation scheme we intend to adopt. In Section 2.2 the picture is completed by some considerations on the Hilbert space in which we chose to frame the treatment. This also enables one to establish that, if two dynamical variables are correlated as random variables, then their time autocorrelations have a similar behaviour (see Theorem 2 therein).

2.1 The bases of perturbation theory

Our main purpose is to find stability results for finite times, bounding from below the time decay of autocorrelations. This means that we try to produce results similar to those of Nekhoroshev theorem, and we will find inspiration from its guidelines in the way of proceeding, too. Thus, let us recall that classical perturbation theory proceeds along two steps:

- formal reduction of the Hamiltonian of the system to that of an inte-

grable system through a sequence of canonical transformations (this is the so–called Birkhoff normal form) or, alternatively, construction of first integrals (i.e., constants of motions) through a formal power series expansion;

- estimate of the contribution to the dynamics given by "the remainder", namely, by what is left when the sequence of canonical transformations or, respectively, the series expressing the formal first integrals are truncated at a finite order (this will be called "estimate of the remainder").

For what concerns the first step, no problem arises at the thermodynamic limit if one applies the method explained in [10], which enables one to explicitly construct the formal integrals of motions.[1] The classical methods for estimating the remainder, instead, have the drawback of always providing values which diverge in that limit.

The problem arises because there exist exceptional points in phase space at which the whole energy is concentrated in a small number of degrees of freedom and the remainder takes values which grow with N, so that the typical estimates, which usually involve the upper bound to the values of the remainder on a region of phase space, present the same growth. On the other hand, it seems quite natural to expect that the values attained at isolated points should not matter for the observable properties, which always involve some averaging procedure. This is the reason why in paper [1] the proposal was advanced to replace in the estimates the usual norms with the L^2 norm with respect to an invariant measure, in order that the contribution due to each region of phase space is weighted with the probability of being therein. The knowledge of such a norm is sufficient to determine relevant mean properties for the system, as shown by the fundamental Theorem 1 in [1], which is a functional theoretic analogue of Lagrange finite increments theorem in a space endowed with an L^2 norm. As regards the time autocorrelation, a little modification of Theorem 1 of [1], given also in [11], expresses directly the relation between the time autocorrelation of a dynamical variable and the L^2 norm of its derivative, according to the following Theorem 1.

Before coming to its statement, we need to introduce some definitions and more precise statements on the mathematical frame in which we set ourselves.

In a dynamical Hamiltonian system with Hamiltonian H, we consider the time evolution operator $\hat{U}_t$ which acts on the space $L^2(\mu, \mathcal{M})$ of the square

[1] We have chosen this purely algebraic method, instead of the more usual of Birkhoff normal form, in order to avoid possible difficulties due to the fact that canonical transformations can become ill–defined at the thermodynamic limit.

integrable functions (with respect to the invariant measure μ) mapping the phase space $\mathcal{M}$ in $\mathbb{R}$. In terms of the flow Φ^t generated by H, the operator $\hat{U}_t$ maps f in $\hat{U}_t f = f \circ \Phi^{-t}$. It is well known, since Koopman (see [12]), that $\hat{U}_t$ defines a one parameter group of unitary operators in $L^2(\mu, \mathcal{M})$, i.e., of operators which preserve the norm in this space, which we will denote with $\| \cdot \|$.

The time autocorrelation of a dynamical variable $f \in L^2(\mu, \mathcal{M})$ is defined by

$$\mathbf{C}_f(t) \stackrel{\text{def}}{=} \langle f_t \, f \rangle - \langle f \rangle^2 = \langle (f_t - \langle f \rangle) \, (f - \langle f \rangle) \rangle \ , \tag{2.1}$$

in which f_t is a shortcut for $\hat{U}_t f$, and $\langle \cdot \rangle$ denotes mean value with respect to the probability measure μ. Notice that here, at variance with the previous Chapter 1, the square of the mean value $\langle f \rangle$ is subtracted. Such an additional term, in fact, was there neglected for the sake of simplicity, as the mean value of the derivative of any dynamical variable with respect to the flow vanishes, due to the invariance of the measure (the only case in which there appeared a dynamical variable with nonvanishing mean was that of specific heat, for we which we had to explicitly introduce the additional term).

The standard deviation σ_f of f is then defined through

$$\sigma_f^2 \stackrel{\text{def}}{=} \| f \|^2 - \langle f \rangle^2 = \| f - \langle f \rangle \|^2 \ . \tag{2.2}$$

The quantity σ_f^2 is called the variance of f and is related to the time autocorrelation of f by the relations

$$\mathbf{C}_f(0) = \sigma_f^2 \quad \text{and} \quad |\mathbf{C}_f(t)| \leq \sigma_f^2 \ . \tag{2.3}$$

Therefore, the variance of a dynamical variable is the natural scale of its time autocorrelation.

We are now able to state

Theorem 1 *Suppose that, for $f \in L^2(\mu, \mathcal{M})$, there exists a constant $\eta > 0$ such that*

$$\| [f, H] \| \leq \eta \, \sigma_f \ ; \tag{2.4}$$

then one has

$$\mathbf{C}_f(t) \geq \sigma_f^2 \left(1 - \frac{1}{2} \eta^2 t^2 \right) \ . \tag{2.5}$$

Proof. One introduces the difference $\delta(t) \stackrel{\text{def}}{=} f_t - f$, for which one has $\partial_t \delta = \partial_t f_t = -[H, f_t]$, since f_t satisfies Liouville equation and f is independent of time. One can then write

$$\partial_t \delta = -[H, \delta] + g \ , \tag{2.6}$$

with $g \stackrel{\text{def}}{=} -[H, f]$. In view of the unitarity of the evolution operator $\hat{U}_t$ and because $\delta(0) = 0$, the solution of (2.6) is given by

$$\delta(t) = \int_0^t \hat{U}_{t-s}\, g \, \mathrm{d}s \ ,$$

so that one gets the estimate

$$\|\delta(t)\| \leq \int_0^t \left\| \hat{U}_{t-s}\, g \right\| \mathrm{d}s = t \, \|g\| \leq \eta \, t \sigma_f \ . \tag{2.7}$$

One obtains then the thesis by making use if the simple identity

$$\mathbf{C}_f(t) = \sigma_f^2 - \frac{\|f_t - f\|^2}{2} \ .$$

Q.E.D.

Remark. The previous theorem says that, up to times of the order of the ratio $\sigma_f / \|[f, H]\|$, the autocorrelation of f is sensibly different from 0, so that the behaviour of the system cannot be considered as fully chaotic up to such a time. Notice that the considered ratio involves the standard deviations of two dynamical variables.[2] If both are sum variables in the sense of Khinchin, one should expect that the ratio is essentially independent of the number of degrees of freedom, due to a sort of central limit theorem, so that the estimate is valid when one passes to the thermodynamic limit.

In virtue of Theorem 1 we have got an analytic tool to estimate the remainder and can trace the broad outline of the scheme which we propose in order to get perturbation results at the thermodynamic limit.

One begins by choosing a constant of motion for the unperturbed system, and takes it as the starting point to construct a formal power series expansion of a first integral of the perturbed system. Then one truncates the series expansion and estimates the ratio $\sigma_f / \|[f, H]\|$ for the dynamical variable f constructed in this way, choosing the optimal truncation order by taking that for which the ratio takes the maximal value. We will then say that f is an adiabatic invariant for the system.

To the problems that arise in implementing this seemingly simple program is devoted the following Chapter 3, where we focus on a relevant physical model in order to show how to overcome the obstacles of probabilistic nature, in particular in giving an estimate independent of N of the aforementioned ratio.

[2]Since $[f, H]$ is a total derivative, it has vanishing mean (see above), so that its L^2 norm coincides with its standard deviation.

2.2 Correlations in their "natural" space

Before we go on, it is better to dwell a little more on the space in which we have framed our treatment, namely $L^2(\mu, \mathcal{M})$, in order to show how this choice leads us to explain in a natural way some general properties of the functions of time correlations, among which the already quoted relations (2.3) between variance and autocorrelation. Moreover, Theorem 2 to be given below shows how the "closeness" in this space translates into similar behaviours of the correlations. This is very important, because in perturbation theory one has to recover the behaviour of the functions of interest once one knows the behaviour of a small perturbation of them. A typical case, indeed, is that in which one manages to prove stability results for long times for a slight modification of the unperturbed action and then infers from this that the action itself has a similar behaviour: this is precisely what Theorem 2 allows one to do.

To the definitions (2.1-2.2) we add that of time correlation between two dynamical variables f and g in $L^2(\mu, \mathcal{M})$, which takes the form[3]

$$\mathbf{C}_{f,g}(t) \stackrel{\text{def}}{=} \langle f_t g \rangle - \langle f \rangle \langle g \rangle = \langle (f_t - \langle f \rangle)\,(g - \langle g \rangle) \rangle \ .$$

The space $L^2(\mu, \mathcal{M})$ is seen to be the natural setting for correlations if one notices that they acquire here a geometrical meaning. Indeed, they are closely tied to the covariance of two dynamical variables f and g, defined by

$$\mathrm{cov}(f, g) \stackrel{\text{def}}{=} \langle (f - \langle f \rangle)\,(g - \langle g \rangle) \rangle \ ,$$

since $\mathbf{C}_{f,g}(t) = \mathrm{cov}(f_t, g)$. On the other hand, in probability theory it often turns out to be useful to consider the covariance of two random variables as a scalar product, and so we will make in our functional space. This leads us to take as a norm[4] of a dynamical variable f its standard deviation σ_f, defined by (2.2), for which one has $\sigma_f^2 = \mathrm{cov}(f, f)$.

Within this frame, the analogue of the direction cosine is the correlation coefficient $\mathrm{r}(f, g)$ between f and g, defined according to

$$\mathrm{r}(f, g) \stackrel{\text{def}}{=} \mathrm{cov}(f, g)/\sigma_f \sigma_g \ .$$

[3]For the difference with respect to Chapter 1, see the remark following (2.1) in Section 2.1.

[4]As a matter of fact, it is a semi-norm; one could talk about a norm in strict sense only if all dynamical variables which differ by a constant are identified. Furthermore, for variables with vanishing mean the covariance $\mathrm{cov}(f, f)$ coincides with the usual L^2. We do not insist further on this point, pointing out that the symbol $\| \cdot \|$ will always denote the norm in L^2.

Indeed, one easily sees (in our case, for instance, by using Schwarz inequality) that one has $|\mathrm{cov}(f,g)| \leq \sigma_f \sigma_g$, so that $|\mathrm{r}(f,g)| \leq 1$. Two variables are orthogonal if they are uncorrelated, i.e., if $\mathrm{r}(f,g) = 0$, and collinear if $\mathrm{r}(f,g) = \pm 1$. Since then $\mathbf{C}_f(t) = \mathrm{cov}(f_t, f)$, hence easily follows (2.3)

The choice of this scalar product clarifies also in which sense one can pass the information on the autocorrelation of a dynamical variable to all those sufficiently "close". The following theorem (already appeared in [13]) shows that, if two variables are strongly correlated (i.e., if they are collinear enough in the scalar product given by the covariance), their time autocorrelations have a similar behaviour.

Theorem 2 *Let μ be a probability measure on the $\mathcal{M}$, invariant for the flow generated by H, and let $f, g \in L^2(\mu, \mathcal{M})$ be such that $|\mathrm{r}(f,g)| \geq 1 - \varepsilon^2/2$ for some $\varepsilon > 0$. Then there exists a multiple $\tilde{f} = \alpha f$ of f, with $\alpha \stackrel{\mathrm{def}}{=} (\mathrm{sign}(\mathrm{r}(f,g))\sigma_g/\sigma_f)$, such that*

$$\left|\mathbf{C}_{\tilde{f},g}(t) - \mathbf{C}_{\tilde{f}}(t)\right| \leq \varepsilon \sigma_g^2 \tag{2.8}$$

and

$$\left|\mathbf{C}_g(t) - \mathbf{C}_{\tilde{f}}(t)\right| \leq \left(\varepsilon^2 + 2\varepsilon\right) \sigma_g^2 \ . \tag{2.9}$$

Proof. Both inequalities follow from the remark that

$$\sigma_{g-\tilde{f}} = \sigma_g^2 + \sigma_{\tilde{f}}^2 - 2\mathrm{cov}(\tilde{f}, g) = 2\sigma_g^2 - 2\sigma_g^2 \mathrm{r}(f,g) \leq \varepsilon^2 \sigma_g^2 \ ,$$

which is due to the identities $\sigma_{\tilde{f}} = \sigma_g$ and $\mathrm{r}(\tilde{f}, g) = \left|\mathrm{r}(f,g)\right|$. In fact, (2.8) hence follows by noting that

$$\left|\mathbf{C}_{\tilde{f},g}(t) - \mathbf{C}_{\tilde{f}}(t)\right| = \left|\mathrm{cov}(\tilde{f}_t, g) - \mathrm{cov}(\tilde{f}_t, \tilde{f})\right| = \left|\mathrm{cov}(\tilde{f}_t, g - \tilde{f})\right| \leq \sigma_g \sigma_{g-\tilde{f}} \ ,$$

while, in a similar way, (2.9) comes from the following relations

$$\begin{aligned} \left|\mathbf{C}_g(t) - \mathbf{C}_{\tilde{f}}(t)\right| &= \left|\mathbf{C}_{g-\tilde{f}}(t) + \mathbf{C}_{g,\tilde{f}}(t) + \mathbf{C}_{\tilde{f},g}(t) - 2\mathbf{C}_{\tilde{f}}(t)\right| \\ &\leq \sigma_{g-\tilde{f}}^2 + 2\sigma_g \sigma_{g-\tilde{f}} \ . \end{aligned}$$

Here, in the second line use is made of the previous inequality and of the identity $\mathbf{C}_{g,\tilde{f}}(t) = \mathbf{C}_{\tilde{f},g}(-t)$, due to the invariance of the measure.

Q.E.D.

Remark. This theorem shows, among other things, that dynamical variables strongly correlated with the Hamiltonian cannot decay to zero, since the autocorrelation of the latter stays obviously constant. This suggests not to use variables correlated with the Hamiltonian to determine whether a system is chaotic or not, and that is better, instead, to confine oneself to the linear space of the variables uncorrelated with it through an orthogonalization procedure such as that of Gram–Schmidt.

Chapter 3

Exponentially long stability times for a nonlinear lattice

The major problems that arise in applying the perturbation methods at the thermodynamic limit along the scheme sketched can be divided in two categories: the accumulation of small divisors and the probabilistic estimates needed to estimate the remainder. In paper [1] a technique was provided in order to evaluate the contribution due to small divisors. This was explained on a chain of weakly coupled rotators, for which the probabilistic estimates become very simple. In the work [11], which we reproduce here almost completely, the situation is opposite. We treated a model of a completely resonant nonlinear lattice, so that no small denominator arises, whereas the problems linked to the use of the probability measure show up at full strength. Fortunately, most of them were already faced, and solved, in the most relevant case, namely that of Gibbs measures, in other frames, i.e., in the Dobrushin theory of Gibbsian fields and in the estimates on marginal probabilities on lattices and interacting gases. Thus our contribution consisted mainly in adapting such arguments to the present case and to bring them into harmony with perturbation techniques.

The chapter is organized as follows. The main result on the considered model, namely the construction of an adiabatic invariant in the thermodynamic limit, is stated in Section 3.1 (Theorem 3), together with a corollary concerning a control on the time evolution of the adiabatic invariant. Then, in Section 3.2, we present the scheme of the proof of Theorem 3, whereas the fundamental ingredients of the proof are separately given in the subsequent three sections. The first one (Section 3.3) concerns perturbation techniques and deals with the formal construction of the adiabatic invariant. The other two sections have a probabilistic nature: the estimate of the marginal probability is given in Section 3.4 together with the estimate of the norm of the

time derivative of the adiabatic invariant. In Section 3.5, we state Theorem 4 in which the estimate of the spatial correlations is given, which enables us to give an estimate on the variance of the adiabatic invariant. Some concluding remarks follow in Section 3.6, while most of the proofs of a more technical character are deferred to Appendix A.

3.1 Stability estimate in the Klein Gordon lattice

In the literature, as a prototype of several models, the so called Klein Gordon lattice is studied (see [14, 15, 16, 17, 18, 19]). From a physical point of view, it mimics a chain of particles, each free to move about a site of a lattice, subjected both to an on–site restoring nonlinear force and to a linear coupling with the nearest neighbours. It can also be seen as a discretization of the one–dimensional Φ^4 model, which plays a major role in field theory.

The Hamiltonian of such a system, in suitably rescaled variables, can be written as $H = H_0 + H_1$, in which

$$H_0 \stackrel{\text{def}}{=} \sum_{i=1}^{N} \omega \left(\frac{p_i^2}{2} + \frac{q_i^2}{2} \right) \quad \text{and } H_1 \stackrel{\text{def}}{=} \varepsilon \sum_{i=1}^{N-1} \frac{q_i q_{i+1}}{\omega} + \sum_{i=1}^{N} \frac{q_i^4}{4\omega^2} \ , \tag{3.1}$$

where $p = (p_1, \dots, p_N)$ and $q = (q_1, \dots, q_N)$ are canonically conjugated variables in the phase space $\mathcal{M}$, and ε is a positive parameter, while ω is defined by $\omega \stackrel{\text{def}}{=} \sqrt{1 + 2\varepsilon}$. Since we don't want to face here the problem of small divisors, which typically arises in perturbation theory, we confine ourselves to the case of small ε, i.e, of small coupling between the sites.

We consider as invariant probability measure the Gibbs one, i.e., that defined by

$$\mu(\mathrm{d}p\,\mathrm{d}q) \stackrel{\text{def}}{=} \frac{\exp(-\beta H(p,q))}{\mathcal{Z}(\beta)} \mathrm{d}p\,\mathrm{d}q \ , \tag{3.2}$$

where $\mathcal{Z}(\beta) \stackrel{\text{def}}{=} \int_{\mathcal{M}} \exp(-\beta H(p,q)) \mathrm{d}p\,\mathrm{d}q$ is the partition function, and $\beta > 0$ the inverse temperature. We aim at showing that, for small enough ε and sufficiently large β, there exists an adiabatic invariant for H up to times which grow as a stretched exponential in the perturbation parameters. This is the statement of the following

Theorem 3 (Estimate on the adiabatic invariant) *There exist positive constants ε^*, κ, independent of N, such that if $\varepsilon < \varepsilon^*$ and $\beta > \varepsilon^{-1}$, then*

there exists a polynomial function $\bar{X}$ uncorrelated with H such that

$$\frac{\| [\bar{X}, H] \|}{\sigma_{\bar{X}}} \leq \exp\left[- \left(\frac{1}{\kappa (\varepsilon + \beta^{-1})} \right)^{1/4} \right] . \tag{3.3}$$

This implies that

$$\mathbf{C}_{\bar{X}} \leq \sigma_{\bar{X}}^2 \left(1 - \frac{1}{2} \left(\frac{t}{\bar{t}} \right)^2 \right) , \tag{3.4}$$

where the time $\bar{t}$ is defined by

$$\bar{t} \stackrel{\text{def}}{=} \exp\left[\left(\frac{1}{\kappa (\varepsilon + \beta^{-1})} \right)^{1/4} \right] . \tag{3.5}$$

Remark. We require $\bar{X}$ to be uncorrelated with H in order that our adiabatic invariant be sufficiently different from the Hamiltonian, which is obviously a constant of motion (see the remark after Theorem 2 in Chapter 2).

Before the proof, we point out immediately that the result of this theorem can be viewed in another way, which sounds perhaps more familiar to the readers coming from perturbation theory. This will help clarifying in which sense $\bar{t}$ defined by (3.5) can be seen as a stability time. It concerns the probability $\mathbf{P}$ that the value of the variable $\bar{X}$ changes significantly from its original value. Indeed, it entails that the probability of such a change is practically negligible if $t < \bar{t}$ and is proved by combining equations (3.3) and (2.7) together with Chebyshev theorem (for a refinement of this kind of estimates, see also [20]),a sis seen from

Corollary 1 *In the hypotheses of Theorem 3, for any $\lambda > 0$ one has*

$$\mathbf{P}\left(\left| \bar{X}_t - \bar{X} \right| \geq \lambda \, \sigma_{\bar{X}} \right) \leq \frac{1}{\lambda^2} \left(\frac{t}{\bar{t}} \right)^2 .$$

3.2 Scheme of the proof of Theorem 3

First we use a variant of the classical construction scheme of approximate integrals of motion (see [21]) in order to perform the construction of the adiabatic invariant as a formal power series. Precisely, we use the scheme developed by Giorgilli and Galgani for a direct construction of integrals of motion (see [10] and Section 3.3 for the actual implementation). It is well known that the series thus obtained are, in general, divergent, so that the standard procedure consists in using as approximate integral of motion a

truncation of the series. Denoting by Y_n the series truncated at order $2n+2$, it turns out that it has the form

$$Y_n \stackrel{\text{def}}{=} H_0 + \sum_{j=1}^{n} P_j(p,q) \ , \tag{3.6}$$

where P_j are suitable polynomials. In order to make such a quantity uncorrelated with H, it is convenient to consider $X_n \stackrel{\text{def}}{=} Y_n - H$ instead of Y_n itself.

In order to make the construction rigorous, one has to add rigorous estimates of the variance $\sigma^2_{X_n}$ of X_n, and of the L^2 norm of $[X_n, H]$. The first step to get such estimates consists in controlling the structure of the polynomials P_j (which, in particular, contain only finite range couplings) and the size of their coefficients. This is done recursively, by a variant of the technique of the paper [22], which is implemented in Section 3.3 (see Lemma 2). We emphasize that, at variance with the original paper, we obtain here estimates independent of the number of degrees of freedom.

Then, due to the structure of the polynomials P_j, to get the needed L^2 estimates one has to compute the L^2 norm with respect to the Gibbs measure of the monomials appearing in P_j. The key step for this computation consists in giving an upper bound independent of N to the marginal probabilities of the Gibbs measure. Such an estimate is obtained by adapting techniques developed by Bogolyubov and Ruelle (see [23] and [24]) and is reported in Lemma 4 of Section 3.4. One thus obtains the following bound

$$\left\| \dot{X}_n \right\| \leq \sqrt{N} \left(\sqrt{2}\beta \right)^{-1} (n!)^4 \left(\beta^{-1} + \varepsilon \right)^n \kappa_1^n \ , \tag{3.7}$$

which is valid for a suitable constant $\kappa_1 > 0$, provided ε is small enough and β large enough (see Lemma 3 of Section 3.4).

We emphasize the presence of the factor $\sqrt{N}$ and that κ_1 is independent of N. It will be shown that actually the l.h.s. of (3.7) is of order $\sqrt{N}$ even if it is the square root of a sum of $O(N^2)$ terms. This is due to the fact that most of the terms have zero mean because the measure is even in p and furthermore the p's are independent variables.

To get the Theorem, one also needs an estimate of σ_{X_n} from below. This is obtained in two steps, which are based on the remark that $\sigma_{X_n} \geq \sigma_{X_1} - \sigma_{\mathcal{R}}$, where $\mathcal{R} \stackrel{\text{def}}{=} X_n - X_1$ is a remainder.

First we compute explicitly X_1 and estimate from below σ_{X_1}, obtaining a bound proportional to $\sqrt{N}$. Then, we estimate from above $\sigma_{\mathcal{R}}$. Precisely, we use techniques introduced by Dobrushin in papers [25, 26] to show that $\sigma_{\mathcal{R}}$ behaves as $\sqrt{N}$ (see Lemma 8 of Section 3.5). We remark that this is the

analogue of the law of large numbers. We recall that Dobrushin's techniques enable us to show that spatial correlations between variables pertaining to different lattice sites decrease exponentially with the distance between the sites, so that the monomials appearing in P_j are essentially independent, and the variance of P_n is essentially the sum of the variances of each monomial. This leads to Lemma 7 of Section 3.5, which shows that, for small enough ε and large enough β, for $n < \kappa_2^{-1/4}(\varepsilon + \beta^{-1})^{-1/4}$ there holds

$$\sigma_{X_n} \geq \sqrt{N}(\varepsilon + \beta^{-1})/(8\beta) \ , \tag{3.8}$$

where again κ_2 is a positive constant.

Then one finds the optimal n, call it $\bar{n}$, such that the ratio $\|[X_{\bar{n}}, H]\|/\sigma_{X_{\bar{n}}}$ takes the minimal value. Notice that, as n belongs to a bounded domain, the minimum can be attained at the boundary. The optimization is immediately done, once the estimates are given both for the L^2 norm $\|[X_n, H]\|$ of the time–derivative of the quasi integral of motion X_n, and for its variance $\sigma^2_{X_n}$. Then, the function $\bar{X}$ satisfying (3.3) of Theorem 3 is simply given by $\bar{X} \stackrel{\text{def}}{=} X_{\bar{n}} - H\rho_{X_{\bar{n}},H}\,\sigma_{X_{\bar{n}}}/\sigma_H$. The identity $\sigma^2_{\bar{X}} = (1 - \rho^2_{X_{\bar{n}},H})\sigma^2_{X_{\bar{n}}}$, together with the upper bound to $\rho_{X_{\bar{n}},H}$ given by Lemma 7, enables us to extend all conclusions from $X_{\bar{n}}$ to $\bar{X}$.

3.3 Construction of the adiabatic invariant

Following [10], we look for the formal integral of motion by looking for a sequence of polynomials $\chi = \{\chi_s\}_{s\geq 1}$ such that

$$[H, T_\chi H_0] = 0 \quad \text{at any order,} \tag{3.9}$$

where T_χ is a linear operator, whose action on a polynomial function f is formally defined by[1]

$$T_\chi f \stackrel{\text{def}}{=} \sum_{s\geq 0} (T_\chi f)_s \ , \quad \text{with } (T_\chi f)_0 \stackrel{\text{def}}{=} f \ , \quad (T_\chi f)_s \stackrel{\text{def}}{=} \sum_{j=1}^{s} \frac{j}{s}[\chi_j, (T_\chi f)_{s-j}] \ . \tag{3.10}$$

Inserting the expansion of $T_\chi H_0$ and H in (3.9) and equating terms of equal order one gets the system

$$\Theta_0 = H_0 \ , \quad \Theta_s - L_0\chi_s = \Psi_s \quad \text{for } s > 0 \ , \tag{3.11}$$

[1]Notice that in paper [10] the χ_s were required to be homogeneous polynomials of degree $s + 2$.However, there is no problem in considering the present more general case.

where

$$\begin{aligned}
&\Psi_1 \stackrel{\text{def}}{=} H_1 \ , \\
&\Psi_s \stackrel{\text{def}}{=} -\sum_{l=1}^{s-1} \frac{l}{s} \left[\chi_l, (T_\chi H_0)_{s-l}\right] - \sum_{l=1}^{s-1} (T_\chi \Theta_l)_{s-l} \quad \text{for } s \geq 2 \ ,
\end{aligned} \tag{3.12}$$

$L_0 \stackrel{\text{def}}{=} [H_0, \cdot]$ is the homological operator and (3.11) has to be read as an equation for the unknowns χ_s, Θ_s, which have to belong, respectively, to the range and to the kernel of the operator L_0. By defining the projections $\Pi_{\mathcal{N}}$, $\Pi_{\mathcal{R}}$, respectively on the kernel $\mathcal{N}$ and on the range $\mathcal{R}$ of L_0, one thus determines recursively

$$\chi_s = -L_0^{-1} \Pi_{\mathcal{R}} \Psi_s \ , \quad \Theta_s = \Pi_{\mathcal{N}} \Psi_s \quad \text{for } s \geq 1 \ . \tag{3.13}$$

The approximate integral of motion is then obtained by truncating the sequence $T_\chi H_0$ at a suitable order.

We have to estimate the action of the operator T_χ on the class of functions $f(p,q)$ we are interested in, in a norm which is well suited for our problem. Such a norm is defined as follows. Let $\mathcal{H}_s^{r,i}$ denote the class of monomials[2] $p^k q^l$ of degree s, i.e., with $|k| + |l| = s$, which furthermore depend on sites that are at most r lattice steps away from i, namely such that $k_j = l_j = 0$ if $|i - j| \geq r$. We denote by $\mathcal{P}_{s,r}$ the set of all homogeneous polynomials of degree s that can be decomposed as

$$f = \sum_{i=1}^{N} \sum_{j=1}^{|\mathcal{H}_s^{r,i}|} c_{ij} f_{ij} \ , \tag{3.14}$$

with $f_{ij} \in \mathcal{H}_s^{r,i}$, where $|\mathcal{H}_s^{r,i}|$ is the cardinality of $\mathcal{H}_s^{r,i}$. To $f \in \mathcal{P}_{s,r}$ we associate a norm,[3] defined by

$$\|f\|_+ \stackrel{\text{def}}{=} \min \left\{ \max_{i \in \{1,\ldots,N\}} \sum_{j=1}^{|\mathcal{H}_s^{r,i}|} |c_{ij}| \right\} \ , \tag{3.15}$$

where the minimum is taken over all possible decompositions of f.

Now, we can estimate the action of T_χ on any function $f \in \mathcal{P}_{s,r}$ according to the following Lemma, which is proved in Appendix A.1.

[2] We adopt here the multi–index notation: $k = k_1, \ldots, k_N$ and $l = l_1, \ldots, l_N$ are vectors of integers, with $|k| = |k_1| + \ldots + |k_2|$. So, $p^k q^l = p_1^{k_1} \cdot \ldots \cdot p_N^{k_N} q_1^{l_1} \cdot \ldots \cdot q_N^{l_N}$.

[3] One can check that this is indeed a norm.

Lemma 1 *Let T_χ be the operator defined by (3.10), relative to the sequence $\chi = \{\chi_s\}_{s\geq 0}$ which solves the system of equations (3.13–3.12) for the Hamiltonian (3.1). Then, for any $f(p,q) \in \mathcal{P}_{2s+2,r}$, one has $(T_\chi f)_n = \sum_{l=0}^n f_n^{(s+l)}$, where $f_n^{(s+l)} \in \mathcal{P}_{2s+2l+2,r+n-l}$ and*

$$\left\| f_n^{(s+l)} \right\|_+ \leq 2^{6n} 2^{5(n-1)} 2^{2s+l+2} n! \frac{(n+r)!}{r!} \frac{(n+s)!}{s!} \frac{n!}{l!(n-l)!} \varepsilon^{n-l} \left\| f \right\|_+ \ . \quad (3.16)$$

Lemma 2 below will give bounds to the adiabatic invariant obtained by truncating at a finite order the formal power series which defines $T_\chi H_0$. In particular, the adiabatic invariant will simply be $Y_n = \sum_{s=0}^n \left(T_\chi H_0\right)_s$, so that the polynomials P_j appearing at the r.h.s. of (3.6) of Theorem 3 are

$$P_j \stackrel{\text{def}}{=} \left(T_\chi H_0\right)_j \ , \quad (3.17)$$

while the quantity we will focus on will be

$$X_n \stackrel{\text{def}}{=} Y_n - H = -\Theta_1 + \sum_{j=2}^n P_j \ . \quad (3.18)$$

The time derivative of X_n is then given by

$$\dot{X}_n \stackrel{\text{def}}{=} [X_n, H] = [P_n, H_1] \ , \quad (3.19)$$

which is a polynomial of order $2n+4$. In order to obtain the estimates of the L^2–norm, eventually, it is of interest to take into account the parity properties of the operator T_χ, with respect to the canonical coordinate p. So we define as $\mathcal{P}^+$ the space of polynomials of even order in p, and $\mathcal{P}^-$ the space of those of odd order in p.

Finally, we can state

Lemma 2 *For the adiabatic invariant constructed through $T_\chi H_0$ (see (3.17)) one can write*

$$P_n = \sum_{l=0}^n \frac{n!}{l!(n-l)!} \varepsilon^{n-l} P_n^{(l)} \ , \quad (3.20)$$

where $P_n^{(l)} \in \mathcal{P}^+ \cap \mathcal{P}_{2l+2,n-l}$ and

$$\left\| P_n^{(l)} \right\|_+ \leq \mathcal{D}_n \ , \quad \text{with } \mathcal{D}_n \stackrel{\text{def}}{=} 2^{12n} \left(n!\right)^3 \ . \quad (3.21)$$

Furthermore, one has

$$[X_n, H] = \sum_{l=0}^{n+1} \frac{(n+1)!}{l!(n+1-l)!} \varepsilon^{n+1-l} \dot{X}_n^{(l)} \ ,$$

with $\dot{X}_n^{(l)} \in \mathcal{P}^- \cap \mathcal{P}_{2l+2,n+1-l}$ *and*

$$\left\| \dot{X}_n^{(l)} \right\|_+ \leq \mathcal{C}_n \ , \quad \textit{with } \mathcal{C}_n \stackrel{\text{def}}{=} 48 \cdot 2^{12n} n! \left((n+1)!\right)^2 \ . \tag{3.22}$$

Proof. The proof of the upper bounds is mainly based on the application of Lemma 1 to the function $H_0 \in \mathcal{P}_{2,0}$, together with the simple bound $\|H_0\|_+ = \omega \leq 2$, which holds for small enough ε. This proves equations (3.20), (3.21). Then, we use the fact that $[X_n, H] = [P_n, H_1]$ and the upper bound to the norm of the Poisson brackets of two variables provided by Lemma 10 of Appendix A.1. This gives equation (3.22).

The parity properties are obtained by observing that $[\mathcal{P}^\pm, \mathcal{P}^\pm] \subset \mathcal{P}^+$ and $[\mathcal{P}^\pm, \mathcal{P}^\mp] \subset \mathcal{P}^-$, as well as $\Pi_{\mathcal{N}}(\mathcal{P}^+) \subset \mathcal{P}^+$ and $\Pi_{\mathcal{N}}(\mathcal{P}^-) \subset \mathcal{P}^-$ and that the similar inclusions regarding $\Pi_{\mathcal{R}}$hold, and then working recursively.

Q.E.D.

3.4 Marginal probability estimates

The aim of this section is to prove the bound on the norm of $\dot{X}_n$ given by the following

Lemma 3 *There exist constants* $\bar{\beta} > 0$, $\bar{\varepsilon} > 0$, $\kappa_1 > 0$ *such that, for any* $\beta > \bar{\beta}$ *and for any* $\varepsilon < \bar{\varepsilon}$, *for* $\dot{X}_n$ *defined by (3.19) of Section 3.3 one has*

$$\left\| \dot{X}_n \right\| \leq \sqrt{N} \left(\sqrt{2}\beta \right)^{-1} (n!)^4 \left(\beta^{-1} + \varepsilon \right)^n \kappa_1^n \ . \tag{3.23}$$

The key tool of the proof is an estimate of the probability that the coordinates of a finite number s of sites are near some fixed values. Such an estimate is given in the following Subsection 3.4.1, whereas the proof of Lemma 3 is given in Subjection 3.4.2.

3.4.1 Estimates on the marginal probability

Everything is trivial for the p coordinates, for which the measure can be decomposed as a product: from a probabilistic point of view, this means that every p_j is independent of the q and of any p_i, for $i \neq j$. We focus, instead, on the q coordinates, which are independent of the p, but depend on each other. Then, we must study the relevant part of the density, which is given by

$$D_N(q_1, \ldots, q_N) \stackrel{\text{def}}{=} \frac{1}{Z_N} \exp\left[-\beta U_N(q_1, \ldots, q_N) \right] \ , \tag{3.24}$$

where Z_N is the "spatial" partition function

$$Z_N \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_N \, \exp\left[-\beta U_N(q_1, \ldots, q_N)\right] \tag{3.25}$$

and U_N the part of Hamiltonian (3.1) which depends on q, namely, the potential

$$U_N\left(q_1 \ldots, q_N\right) \stackrel{\text{def}}{=} \sum_{i=1}^{N} \left(\omega \frac{q_i^2}{2} + \frac{q_i^4}{4\omega^2}\right) + \varepsilon \sum_{i=1}^{N-1} \frac{q_i q_{i+1}}{\omega} \ .$$

The main point is then to estimate the marginal probability $F_{s,\mathfrak{x}}^{(N)}(q_{i_1}, \ldots, q_{i_s})$ that we are going to define. Given a set of indices $i_1 < i_2 < \ldots < i_s$ we say that they form a connected block if $i_{j+1} = i_j + 1$, i.e., if they label a "connected" chain. We say that a sequence of indices $i_1 < i_2 < \ldots < i_s$ form $\mathfrak{x}$ blocks if the set $\{i_j\}_{j=1}^s$ can be decomposed into $\mathfrak{x}$ connected blocks, which furthermore are not connected to each other. Given a set of indices $i_1 < i_2 < \ldots < i_s$ we define

$$F_{s,\mathfrak{x}}^{(N)}(q_{i_1}, \ldots, q_{i_s}) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \mathrm{d}q_{i_{s+1}} \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_{i_N} \, D_N(q_1, \ldots, q_N) \ , \tag{3.26}$$

where $\mathfrak{x}$ is the number of blocks in the set $\{i_j\}_{j=1}^s$. We remark here that such a quantity depends on the number of particles, N, but we will find for it an upper bound independent of N. In fact, the estimate will depend only on s and $\mathfrak{x}$, but not on the precise choice of the sites.

Define the two functions

$$\begin{aligned} n_{s,\mathfrak{x}}(q_{i_1}, \ldots, q_{i_s}) &\stackrel{\text{def}}{=} \exp\left[-\beta\left(\sum_{k=1}^{s}\left(\frac{q_{i_k}^2}{2\omega} + \frac{q_{i_k}^4}{4\omega^2}\right) + \varepsilon \sum_{k,l=1}^{s} \delta_{i_l, i_k+1} \frac{(q_{i_k} - q_{i_l})^2}{2\omega}\right)\right] \\ &\leq \exp\left(-\beta \sum_{k=1}^{s} \frac{q_{i_k}^2}{2\omega}\right) \ , \end{aligned} \tag{3.27}$$

$$\tilde{n}_{s,\mathfrak{x}}(q_{i_1}, \ldots, q_{i_s}) \stackrel{\text{def}}{=} \exp\left[-\beta\left(\sum_{k=1}^{s}\left(\frac{\omega q_{i_k}^2}{2} + \frac{q_{i_k}^4}{4\omega^2}\right) + \varepsilon \sum_{k,l=1}^{s} \delta_{i_l, i_k+1} \frac{q_{i_k} q_{i_l}}{\omega}\right)\right] \ , \tag{3.28}$$

where $\delta_{i,j}$ is the Krönecker delta.

Remark. Notice that $n_{s,\mathfrak{x}}$ is the configurational part of the Gibbs measure of the system with variables $q_{i_1}, \ldots, q_{i_s}$ and free boundary conditions, apart from the absence of the normalization factor (i.e., the partition function),

whereas $\tilde{n}_{s,\mathfrak{x}}$ is the analogous quantity for the same system, but with fixed boundary conditions. Thus, they differ only because of the different dependence on the coordinates at the sites lying on the boundary of the blocks, the number of which, γ, satisfies $\mathfrak{x} \leq \gamma \leq 2\mathfrak{x}$. If we denote by $m_1, \ldots, m_\gamma$ the indices of these sites, we can write the identity

$$\frac{n_{s,\mathfrak{x}}(q_{i_1}, \ldots, q_{i_s})}{\tilde{n}_{s,\mathfrak{x}}(q_{i_1}, \ldots, q_{i_s})} = \prod_{j=1}^{\gamma} \exp\left(\frac{\beta\varepsilon}{\omega}\alpha_{m_j} q_{m_j}^2\right) \leq \prod_{j=1}^{\gamma} \exp\left(\frac{\beta\varepsilon}{\omega} q_{m_j}^2\right) \ , \tag{3.29}$$

where the factor α_{m_j} is equal to 1 or $1/2$ according to whether the site m_j is isolated (i.e., the block is composed of only that site) or not.

Then the following lemma, which is the main result of the present subsection, holds

Lemma 4 *There exist constants $\bar{\beta} > 0$, $\bar{\varepsilon} > 0$, $K > 0$ and a sequence $\mathfrak{C}_\mathfrak{x} > 0$ such that, for any $\beta > \bar{\beta}$ and for any $\varepsilon < \bar{\varepsilon}$, one has the inequalities*

$$F_{s,\mathfrak{x}}^{(N)}(q_{i_1}, \ldots, q_{i_s}) \leq \mathfrak{C}_\mathfrak{x} K^s \left(\frac{\beta}{2\pi\omega}\right)^{s/2} n_{s,\mathfrak{x}}(q_{i_1}, \ldots, q_{i_s}) \tag{3.30}$$

and

$$F_{s,\mathfrak{x}}^{(N)}(q_{i_1}, \ldots, q_{i_s}) \geq \frac{1}{\mathfrak{C}_\mathfrak{x}} \left(\frac{\beta}{2\pi\omega}\right)^{s/2} \tilde{n}_{s,\mathfrak{x}}(q_{i_1}, \ldots, q_{i_s}) \exp\left(-8\varepsilon\mathfrak{x}\sqrt{\frac{\beta}{2\omega}} \sum_{j=1}^{\gamma} |q_{m_j}|\right) \ . \tag{3.31}$$

The proof of such a lemma is based on the techniques of paper [23], which apply quite simply to the case of periodic boundary conditions (see Lemma 5 below), on account of the translational invariance. Thus, it is also useful to introduce the density $\tilde{D}_N$ relative to the periodic system, defined by

$$\tilde{D}_N(q_1, \ldots, q_N) \stackrel{\text{def}}{=} \frac{1}{Q_N} \exp\left[-\beta U_N(q_1, \ldots, q_N) + \beta\varepsilon q_1 q_N\right] \ . \tag{3.32}$$

In this definition there appears the partition function for the periodic case

$$Q_N \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_N \ \exp\left[-\beta U_N(q_1, \ldots, q_N) + \beta\varepsilon q_1 q_N\right] \ . \tag{3.33}$$

For the periodic system it is simple to estimate two relevant quantities. The former is the ratio between the partition function for $N-1$ particles and that for N particles, i.e., the ratio Q_{N-1}/Q_N. The relation between Q_N and Z_N is then obtained as a particular case of Lemma 6, which will be stated

later on. The latter is the probability, evaluated with respect to the density for N particles, that the coordinates of r particles have an absolute value smaller than $\Theta\sqrt{2\omega/\beta}$, for a given Θ. In other terms, we need an estimate of the following quantity

$$\mathbf{P}_N\Big(|q_1| < \Theta\sqrt{\frac{2\omega}{\beta}} \wedge \ldots \wedge |q_r| < \Theta\sqrt{\frac{2\omega}{\beta}}\Big) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_N \mathbf{1}_{|q_1|<\Theta\sqrt{2\omega/\beta}} \times \ldots \times \mathbf{1}_{|q_r|<\Theta\sqrt{2\omega/\beta}}\, \tilde{D}_N(q_1, \ldots, q_N) \;, \tag{3.34}$$

in which $\mathbf{1}_A$ is the indicator function of the set A. We can now give the mentioned estimates by the following lemma, whose proof can be found in Appendix A.2.1.

Lemma 5 *There exist constants $\beta_0 > 0$, $\varepsilon_0 > 0$, $K_0 > 2$ such that, for any $\beta > \beta_0$ and $\varepsilon < \varepsilon_0$, one has*

$$\frac{Q_{N-1}}{Q_N} \leq K_0\sqrt{\frac{\beta}{2\pi\omega}} \;. \tag{3.35}$$

Furthermore, if $\Theta \geq 2\sqrt{r\log(4rK_0)}$, one has

$$\mathbf{P}_N\left(|q_1| < \Theta\sqrt{\frac{2\omega}{\beta}} \wedge \ldots \wedge |q_r| < \Theta\sqrt{\frac{2\omega}{\beta}}\right) \geq \frac{1}{2} \;. \tag{3.36}$$

This result enables us to give the proof of Lemma 4.

Proof of Lemma 4 First write

$$D_N\left(q_1 \ldots, q_N\right) = n_{N-s,\mathfrak{x}'}\left(q_{i_{s+1}}, \ldots, q_{i_N}\right)\big)\, n_{s,\mathfrak{x}}\left(q_{i_1}, \ldots, q_{i_s}\right) I\left(q_1, \ldots, q_N\right) \;,$$

for a suitable $\mathfrak{x}'$, with $\mathfrak{x} - 1 \leq \mathfrak{x}' \leq \mathfrak{x} + 1$ (the lower and the upper bound are attained, respectively, if both 1 and N are contained in $i_1, \ldots, i_s$ or none of them), where I contains the terms of interaction between the "internal" and the external part of the system. Remarking that $I \leq 1$, one gets

$$F^{(N)}_{s,\mathfrak{x}}(q_{i_1}, \ldots, q_{i_s}) \;\leq\; \frac{1}{Z_N}\left(\int_{-\infty}^{+\infty} \mathrm{d}q_{i_{s+1}} \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_{i_N} \times \times\, n_{N-s,\mathfrak{x}'}(q_{i_{s+1}}, \ldots, q_{i_N})\right) n_{s,\mathfrak{x}}(q_{i_1}, \ldots, q_{i_s}) \;. \tag{3.37}$$

We now have to estimate the integral appearing in (3.37). More in general, in the course of the proof we need to estimate integrals of a similar type.

This will be done in Lemma 6, which will be given in a while. Introduce the quantities:

$$\begin{aligned}\bar{\mathcal{Q}}_M^{\mathfrak{r}} &\stackrel{\text{def}}{=} \inf_{B(M,\mathfrak{r})} \int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_M n_{M,\mathfrak{r}}(q_1,\ldots,q_M) \ , \\ \mathcal{Q}_M^{\mathfrak{r}} &\stackrel{\text{def}}{=} \sup_{B(M,\mathfrak{r})} \int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_M n_{M,\mathfrak{r}}(q_1,\ldots,q_M) \ ,\end{aligned} \tag{3.38}$$

where $B(M,\mathfrak{r})$ denotes the collection of all possible partitions of M indices in $\mathfrak{r}$ blocks. It is also convenient to consider the quantities defined in a similar way, by integrating $\tilde{n}_{M,\mathfrak{r}}$ in place of $n_{M,\mathfrak{r}}$, namely

$$\begin{aligned}\bar{Z}_M^{\mathfrak{r}} &\stackrel{\text{def}}{=} \inf_{B(M,\mathfrak{r})} \int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_M \tilde{n}_{M,\mathfrak{r}}(q_1,\ldots,q_M) \ , \\ Z_M^{\mathfrak{r}} &\stackrel{\text{def}}{=} \sup_{B(M,\mathfrak{r})} \int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_M \tilde{n}_{M,\mathfrak{r}}(q_1,\ldots,q_M) \ .\end{aligned} \tag{3.39}$$

It is easily shown that for $\mathfrak{r}=1$ one has $\bar{Z}_M^1 = Z_M^1 = Z_N$. In order to link them to Q_N and to each other, we use the following lemma, the proof of which appears in Appendix A.2.2.

Lemma 6 *Let $\beta_0 > 0$, $\varepsilon_0 > 0$ and $K_0 > 2$ be constants such that Lemma 5 holds. Then, for any $\beta > \beta_0$ and any $\varepsilon < \varepsilon_0$, the inequalities*

$$\frac{\bar{\mathcal{Q}}_M^{\mathfrak{r}}}{Q_M} \geq 1 \ , \quad \frac{\bar{Z}_M^{\mathfrak{r}}}{Q_M} \geq \frac{1}{2} \left(8\mathfrak{r}K_0\right)^{-32\varepsilon_0\mathfrak{r}^2} \tag{3.40}$$

hold. Furthermore, the chain of inequalities

$$\frac{Z_M^{\mathfrak{r}}}{Q_M} \leq \frac{\mathcal{Q}_M^{\mathfrak{r}}}{Q_M} \leq 2K_0^{\mathfrak{r}} \exp\left(4\mathfrak{r}\varepsilon_0 \bar{\kappa}(\mathfrak{r},K_0)\right) \ , \tag{3.41}$$

holds, where $\bar{\kappa}(\mathfrak{r},K_0)$ is the solution of the equation

$$K_0^{2\mathfrak{r}} \Gamma(\mathfrak{r},\bar{\kappa}) = \frac{1}{2} \ , \tag{3.42}$$

$\Gamma(s,x)$ being the upper regularized Gamma function

$$\Gamma(s,x) \stackrel{\text{def}}{=} \frac{1}{(s-1)!} \int_x^{+\infty} t^{s-1} e^{-t} \mathrm{d}t \ . \tag{3.43}$$

The previous lemma enables one to see that $Z_N^{-1} \leq 2(8K_0)^{32\varepsilon_0}/Q_N$, while the integral appearing in (3.37) is estimated by

$$\mathcal{Q}_{N-s}^{\mathfrak{r}'} \leq 2K_0^{\mathfrak{r}+1} \exp\left[4(\mathfrak{r}+1)\varepsilon_0 \bar{\kappa}(\mathfrak{r}+1, K_0)\right] Q_{N-s} \ .$$

Thus, due to relation (3.35) of Lemma 5, one easily sees that

$$\frac{Q_{N-s}}{Q_N} = \prod_{i=1}^{s} \frac{Q_{N-i}}{Q_{N-i+1}} \leq K_0^s \left(\frac{\beta}{2\pi\omega}\right)^{s/2} \ , \tag{3.44}$$

so that (3.30) is proved, taking

$$\mathfrak{C}_{\mathfrak{r}} \geq 2^{96\varepsilon_0+2} K_0^{\mathfrak{r}+1+32\varepsilon_0} \exp\left[4(\mathfrak{r}+1)\varepsilon_0 \bar{\kappa}(\mathfrak{r}+1, K_0)\right] \ . \tag{3.45}$$

We come now to the proof of (3.31). To this end, we write

$$\begin{aligned} D_N(q_1, \ldots, q_N) = \frac{Q_{N-s}}{Z_N} \tilde{D}_{N-s}(q_{i_{s+1}}, \ldots, q_{i_N}) \tilde{n}_{s,\mathfrak{r}}(q_{i_1}, \ldots, q_{i_s}) \\ \times G(q_{m_1}, \ldots, q_{m_\gamma}, q_{l_1}, \ldots, q_{l_{\gamma'}}) \ , \end{aligned}$$

where the sites l_i are the ones which are contiguous to the blocks, but not contained in them, taken by keeping the relative order. Furthermore, due to the periodicity there appear factors depending on q_1 and q_N, if the sites 1 and N are not contained in $i_1, \ldots, i_s$. In this case, we put $q_{l_1} = q_1$ and $q_{l_{\gamma'}} = q_N$. We denote by γ', with $\gamma' \leq 2\mathfrak{r}+2$, the number of such indices. The explicit expression of the function G is complicated. Plainly, it represents the product of the factors $\exp(-\sum_j \beta\varepsilon q_{m_j} q_{l_i}/\omega)$ among all sites l_i contiguous to m_j, and just the factor $\exp(\beta\varepsilon q_{l_i} q_{l_{i+1}}/\omega)$, when l_i and l_{i+1} belong to different blocks.[4] In any case, a lower bound to G in the region

$$\mathcal{A} \stackrel{\text{def}}{=} \left\{ |q_{l_1}| < 2\mathfrak{r}\sqrt{\log(4K_0)}\sqrt{2\omega/\beta} \ \wedge \ \ldots \ \wedge \ |q_{l_{\gamma'}}| < 2\mathfrak{r}\sqrt{\log(4K_0)}\sqrt{2\omega/\beta} \right\}$$

is given by

$$G \geq \exp\left(-4\varepsilon\mathfrak{r}\sqrt{\frac{\beta}{2\omega}}\sqrt{\log(4K_0)} \sum_{j=1}^{\gamma} |q_{m_j}| \right) (4K_0)^{-8\varepsilon(\mathfrak{r}^2+\mathfrak{r})} \ .$$

So, we can write

$$\begin{aligned} F_{s,\mathfrak{r}}^{(N)}(q_{i_1}, \ldots, q_{i_s}) \geq \tilde{n}_s(q_{i_1}, \ldots, q_{i_s})(4K_0)^{-8\varepsilon(\mathfrak{r}^2+\mathfrak{r})} \frac{Q_{N-s}}{Z_N} \mathbf{P}_{N-s}(\mathcal{A}) \\ \times \exp\left(-4\varepsilon\mathfrak{r}\sqrt{\frac{\beta}{2\omega}}\sqrt{\log(4K_0)} \sum_{j=1}^{\gamma} |q_{m_j}| \right) \ , \end{aligned} \tag{3.46}$$

[4]Notice that the index i lies in $\{1, \ldots, \gamma'\}$. But if in the expression $\exp(\beta\varepsilon q_{l_i} q_{l_{i+1}}/\omega)$ there appears $q_{l_{\gamma'+1}}$ then one has to intend simply $l_{\gamma'+1}$ as l_1.

where the (positive) contribution of the integral over $\mathcal{A}^c$ was neglected.

The term with $\mathbf{P}_{N-s}(\mathcal{A})$ in (3.46) is bounded from below by relation (3.36) of Lemma 5. As for the fraction, by Lemma 6 we obtain

$$Q_{N-s} \geq \frac{1}{2K_0 \exp(8\varepsilon_0 \bar{\kappa}(1, K_0))} \mathcal{Q}^1_{N-s} \ .$$

Now, operating as in the deduction of formula (3.44), it is sufficient to observe that $\mathcal{Q}^1_{N-1} \geq \sqrt{\beta/(2\pi\omega)}\mathcal{Q}^1_N$ to obtain $\mathcal{Q}^1_{N-s} \geq (\beta/2(\pi\omega))^{s/2}\, \mathcal{Q}^1_N$. Then, choosing $\bar{\varepsilon}$, $\bar{\beta}$ such that $K_0 \leq e^4/4$ and observing that $\mathcal{Q}^1_N \geq Z_N$, one gets (3.31) with

$$\mathfrak{C}_{\mathfrak{r}} \geq (4K_0)^{8\varepsilon_0(\mathfrak{r}^2+\mathfrak{r})+1} \exp(8\varepsilon_0 \bar{\kappa}(1, K_0) \ . \tag{3.47}$$

Finally, $\mathfrak{C}_{\mathfrak{r}}$ can be chosen as the maximum of the r.h.s. of (3.45) and of (3.47) . This concludes the proof.

Q.E.D.

3.4.2 Estimate of the L^2 norm of $\dot{X}_n$

We apply directly inequality (3.30) to get the proof of Lemma 3, using the fact that such a quantity is a sum of polynomials depending at most on $2n+3$ sites, as can be seen by Lemma 2 of Section 3.3.

Proof of Lemma 3 The key ingredient of the proof is, as stated in Section 3.3, that the polynomials P_n are even in the p coordinates, so that the $\dot{X}_n$'s are odd in the p. On account of that, $\dot{X}_n^2$ is a sum in which the terms coming from the product of two monomials depending on separated groups of sites contain at least one p_i to an odd power. Since the measure is even with respect to any p, these terms have a vanishing integral.

We formalize this way of reasoning by decomposing $\dot{X}_n$ as $\dot{X}_n = \sum_{i=1}^N f_i$, where the f_i's are polynomials depending at most on the sites between $i-n-1$ and $i+n+1$. Then, the L^2–norm of $\dot{X}_n$ is expressed according to $\left\|\dot{X}_n\right\|^2 = \sum_{i,j=1}^N \langle f_i f_j \rangle$. In this sum, all the terms with $|i-j| > 2n+2$ vanish, while the other ones are estimated in terms of $\left\|\dot{X}_n\right\|_+$ in the following way.

On account of Lemma 2, we can write

$$f_i = \sum_{l=0}^{n+1} \frac{(n+1)!}{l!(n+1-l)!} \varepsilon^{n+1-l} \sum_{s=1}^{|\mathcal{H}^{n+1-l,i}_{2l+2}|} c_{is,l} f^{(l)}_{is} \ ,$$

in which $f_{is}^{(l)}$ is a monomial in $\mathcal{H}_{2l+2}^{n+1-l,i}$ and the decomposition in these monomials is performed in such a way that $\sup_{i,l} \sum_s |c_{is,l}| \leq \mathcal{C}_n$. Then, we sum on j and obtain that

$$\left\| \dot{X}_n \right\|^2 \leq (4n+5)\, \mathcal{C}_n^2 \sum_{i=1}^{N} \sum_{l=0}^{n+1} \frac{(2n+2)!}{l!(2n+2-l)!} \varepsilon^{2n+2-l} \sup_{g \in \mathcal{H}_{2l+4}^{2n+2-l,i}} \langle g \rangle \ ,$$

where we used the fact that the only nonvanishing contributions to the integral come from the product of $f_{js}^{(l-r)} \in \mathcal{H}_{2l-2r+2}^{n+1-l+r,j}$ and $f_{km}^{(r)} \in \mathcal{H}_{2r+2}^{n+1-r,k}$, for $|j-k| \leq 2n+2-l$, so that $g \stackrel{\text{def}}{=} f_{js}^{(l-r)} f_{km}^{(r)} \in \mathcal{H}_{2l+4}^{2n+2-l,i}$, for a suitable i between j and k.

Then, we make use of (3.30) together with the estimate (3.27) for $n_{s,\mathfrak{x}}$ to bound the mean value of any function in $\mathcal{H}_{2l+4}^{2n+2-l,i}$. In fact, one has

$$\begin{aligned} \sup_{g \in \mathcal{H}_{2l+4}^{2n+2-l,i}} \langle g \rangle &\leq \mathfrak{C}_1 K^{4n-2l+5} \sqrt{\frac{\beta}{2\pi\omega}} \int_{-\infty}^{\infty} x^{2l+4} \exp\left(-\frac{\beta}{2\omega} x^2\right) \mathrm{d}x \\ &= \mathfrak{C}_1 K^{4n-2l+5} \left(\frac{2\omega}{\beta}\right)^{l+2} \frac{(2l+3)!!}{2^{l+2}} \ . \end{aligned}$$

So, the inequality

$$\left\| \dot{X}_n \right\|^2 \leq \mathfrak{C}_1 K^{4n+5} (4n+5)(2n+4)!\, (2\omega)^{2n+4} \beta^{-2} \left(\varepsilon + \frac{1}{\beta}\right)^{2n+2} N \mathcal{C}_n^2$$

holds. Thus, choosing a suitable $\kappa_1 > 0$ and using the value of $\mathcal{C}_n$ given by (3.22) of Lemma 2, inequality (3.23) is satisfied.

Q.E.D.

3.5 Estimate of the variance of the adiabatic invariant

In the present Section we prove the following Lemma 7, which was used in the proof of Theorem 3. The lemma concerns estimates on the variance $\sigma_{X_n}^2$ and on the correlation $\rho_{X_n,H}$ of the adiabatic invariant and reads

Lemma 7 *There exist positive constants $\tilde{\varepsilon} > 0$, $\kappa_2 > 0$, $\kappa_3 > 1$, such that, for any $\varepsilon < \tilde{\varepsilon}$, for any $\beta > \varepsilon^{-1}$ and for $n < \kappa_2^{-1/4} (\varepsilon + \beta^{-1})^{-1/4}$, with X_n defined by (3.18), the following inequalities hold:*

$$\sigma_{X_n} \geq \sqrt{N}\, \frac{\varepsilon + \beta^{-1}}{8\beta} \tag{3.48}$$

and

$$|\rho_{X_n,H}| \leq \left(1 + \frac{1}{\kappa_3} \frac{\varepsilon^2}{(\varepsilon + \beta^{-1})^2}\right)^{-1/2} . \tag{3.49}$$

The proof of this lemma requires the study the spatial correlations between quantities depending on two separate blocks. The study of these properties has to be performed within the general frame of Gibbsian fields and conditional probabilities. In the present Section we provide the necessary notions and give a proposition of a general character concerning the decay of spatial correlations for lattices with finite range of interaction, i.e, Theorem 4 of Subsection 3.5.2, from which it will be possible to finally come to the proof of Lemma 7, which will be given in Subsection 3.5.4.

Our treatment of conditional probabilities is inspired in particular by the work of Dobrushin (see [25]). More precisely, we will make reference to paper [25] for the main ideas, and to the subsequent beautiful but underestimated paper [26], by Dobrushin and Pechersky, for a more direct relation to our problem. As a matter of fact, most of the ideas of this section are already contained in works [25] and [26], but the explicit result on the spatial correlations given here required some additional work. We recall that, since Gibbsian fields and the related techniques were introduced in order to deal with infinite lattices, our result holds even if the number of sites tends to infinity.

The present section is structured as follows: in Subsection 3.5.1 the link between spatial correlations and conditional probabilities is shown, and in Subsection 3.5.2 we state Theorem 4, whose proof is deferred to Appendix A.2.3. Such a result is used in order to obtain an upper bound to σ_{P_n}, stated in Lemma 8 of Subsection 3.5.3, whence the proof of Lemma 7 easily follows, as shown in Subsection 3.5.4.

3.5.1 Link between spatial correlations and conditional probability

In order to prove Lemma 7 we have to estimate quantities such as $\langle fg \rangle - \langle f \rangle \langle g \rangle$, relative to the Gibbs measure μ, where f is a function which depends on sites belonging to a set $\tilde{V}$, while g depends only on sites in V, with $V \cap \tilde{V} = \emptyset$. Our aim is to show that such correlations decrease as the distance between $\tilde{V}$ and V increases, where the distance $d(V, \tilde{V})$ is defined for example as $d(V, \tilde{V}) \stackrel{\text{def}}{=} \inf_{i \in V, j \in \tilde{V}} |i - j|$.

We start showing the relation between the spatial correlations and the conditional probability in a setting more general than ours. We consider as

given a measure μ on $\mathbb{R}^{|T|}$, with $T \subset \mathbb{Z}^\nu$, which induces on the measurable set $A \subset \mathbb{R}^{|\tilde{V}|}$ the probability

$$P_{\tilde{V}}(A) \stackrel{\text{def}}{=} \int_{\mathbb{R}^{|T|}} \mathrm{d}\mu(x) \mathbf{1}_{A \times T \setminus \tilde{V}}(x) \ ,$$

where $\mathbf{1}_A$ is the indicator function of the set A. One can express the quantity we are interested in as

$$\langle fg \rangle - \langle f \rangle \langle g \rangle = \int_{\mathbb{R}^{|\tilde{V}|}} f(\mathbf{x}) P_{\tilde{V}}(\mathrm{d}\mathbf{x}) \left(\int_{\mathbb{R}^{|V|}} g(\mathbf{y}) P_V(\mathrm{d}\mathbf{y}|\mathrm{d}\mathbf{x}) - \int_{\mathbb{R}^{|V|}} g(\mathbf{y}) P_V(\mathrm{d}\mathbf{y}) \right) \ , \tag{3.50}$$

where $P_V(B|A)$ represents the conditional probability of the measurable set $B \subset \mathbb{R}^{|V|}$, once A is given. So, in order to estimate the correlation between two functions, it is sufficient to estimate the difference enclosed in brackets at the r.h.s. of (3.50). Now, we notice that for any pair of probabilities P and $\tilde{P}$ on $\mathbb{R}^{|V|}$, one has

$$\begin{aligned} \left| \int_{\mathbb{R}^{|V|}} g(\mathbf{x}) P(\mathrm{d}\mathbf{x}) - \int_{\mathbb{R}^{|V|}} g(\mathbf{y}) \tilde{P}(\mathrm{d}\mathbf{y}) \right| \leq \int_{\mathbb{R}^{|V|} \times \mathbb{R}^{|V|}} & (|g(\mathbf{x})| + |g(\mathbf{y})|) \\ & \times \mathbf{1}_{\mathbf{x} \neq \mathbf{y}} \, Q(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y}) \ , \end{aligned} \tag{3.51}$$

in which Q is *any* probability on $\mathbb{R}^{|V|} \times \mathbb{R}^{|V|}$ such that P and $\tilde{P}$ are its marginal probabilities. In other terms Q is a joint probability of P and $\tilde{P}$, i.e., for any measurable $B \subset \mathbb{R}^{|V|}$ one has

$$P(B) = Q\left(B \times \mathbb{R}^{|V|}\right) \quad \text{and } \tilde{P}(B) = Q\left(\mathbb{R}^{|V|} \times B\right) \ . \tag{3.52}$$

Remark here that Q is not unique: indeed, such a probability provides also a way to define a distance between two probabilities defined on the same set V of indices, by

$$D(P, \tilde{P}) \stackrel{\text{def}}{=} \inf_Q \int_{\mathbb{R}^{|V|} \times \mathbb{R}^{|V|}} \mathbf{1}_{\mathbf{x} \neq \mathbf{y}} \, Q(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y}) \ . \tag{3.53}$$

We stress that the infimum is attained, i.e., there exists a probability measure $\bar{Q}(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y})$ such that $D(P, \tilde{P}) = \int \mathbf{1}_{\mathbf{x} \neq \mathbf{y}} \bar{Q}(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y})$ (see Lemma 1 of paper [26]). For the following, we suppose that it is possible to find a compact function h,[5] with domain in $\mathbb{R}$, such that

$$|g(\mathbf{x})| \leq \sum_{i \in V} h(x_i) \ , \tag{3.54}$$

[5]See later for the definition of a compact function, according to the convention of paper [26].

as is the case for the monomials we are dealing with. The bounds will then be given in terms of h. Now, observing that $\mathbf{1}_{\mathbf{x}\neq\mathbf{y}} \leq \sum_{i\in V} \mathbf{1}_{x_i\neq y_i}$, we can rewrite (3.51) as

$$\left|\int_{\mathbb{R}^{|V|}} g(\mathbf{x})P(\mathrm{d}\mathbf{x}) - \int_{\mathbb{R}^{|V|}} g(\mathbf{y})\tilde{P}(\mathrm{d}\mathbf{y})\right| \leq \sum_{i,j\in V}\int_{\mathbb{R}^2} (h(x_j)+h(y_j)) \times \mathbf{1}_{x_i\neq y_i} Q(\mathrm{d}\mathbf{x},\mathrm{d}\mathbf{y}) \ . \tag{3.55}$$

This way, we can make a direct connection with paper [26], in which the problem of estimating the r.h.s. of the above expression is dealt with. We summarize here the results and the methods we need.

3.5.2 Main argument and theorem on correlations

In the quoted work [26], the framework is more general than ours, because it deals with the problem of defining a "probability" for the configuration of an actually infinite ν-dimensional lattice of particle, in terms of the set of conditional probabilities on each site, which is called the *specification* Γ. Now, our case is in principle different, because our lattice is finite, and the probability is defined through the Gibbs measure. In particular, the specification too is assigned by such a measure.

However, as proved in [26], under suitable assumptions assigning the specification uniquely determines the probability, i.e. the Gibbsian field, which, in our case, turns out to be precisely that of Gibbs. So, in our case, it is equivalent to speak in terms of specification or in terms of measure. Indeed, in this subsection we will speak in terms of specifications, and in the following one we will show that the specification determined by Gibbs measure (3.2) with the Hamiltonian (3.1) satisfies the assumptions of [26] (i.e. Conditions 1 and 2 below).

We notice that the r.h.s. of (3.55) can be bounded from above if one estimates the quantity

$$\lambda(\mathbf{j},\mathbf{i}) \stackrel{\text{def}}{=} \max\left\{\mathbf{E}\left[\mathbf{1}_{\xi^1_{\mathbf{j}}\neq\xi^2_{\mathbf{j}}} h\left(\xi^1_{\mathbf{i}}\right)\right], \mathbf{E}\left[\mathbf{1}_{\xi^1_{\mathbf{j}}\neq\xi^2_{\mathbf{j}}} h\left(\xi^2_{\mathbf{i}}\right)\right]\right\} \ , \tag{3.56}$$

where ξ^1 and ξ^2 are two Gibbsian fields which assign, respectively, the probabilities P and $\tilde{P}$ appearing in (3.55) and the expectations are obtained by integrating over a joint probability Q of the two fields. Indeed, in [26] an upper bound just to $\lambda(\mathbf{j},\mathbf{i})$ is given, by requiring that two suitable conditions are satisfied.

So, by adopting the same techniques of [26] we bound from above the r.h.s. (3.55) and, thus, the l.h.s. of (3.50). Such a bound is given in Theorem 4 below. In order to state it, we recall the main notations of [26].

First, we consider a lattice of sites contained in $T \subset \mathbb{Z}^\nu$, and a finite–range specification with a radius of interaction r (this means that the conditional probability at site $\mathbf{i}$ does not depend on the conditioning at sites $\mathbf{j}$ for $|\mathbf{i} - \mathbf{j}| > r$). Then, for a vector $\mathbf{x} \in \mathbb{R}^{|T|}$ we denote by $P_{\mathbf{i},\mathbf{x}}(\mathrm{d}x)$ the probability distribution conditioned to $\mathbf{x}$ everywhere but at site $\mathbf{i}$. The specification Γ is defined by

$$\Gamma \stackrel{\text{def}}{=} \left\{ P_{\mathbf{i},\mathbf{x}} : \mathbf{i} \in T, \mathbf{x} \in \mathbb{R}^{|T|} \right\} \ .$$

Furthermore, we will say that a continuous positive function h on a metric space $\mathfrak{X}$ is compact if, for any $k \geq 0$, the set $\{x \in \mathfrak{X} : h(x) \leq k\}$ is compact.

For a fixed integer r, let $\partial_r V \stackrel{\text{def}}{=} \{\mathbf{j} \in T : \mathbf{j} \notin V, \min_{\mathbf{k} \in V} |\mathbf{j} - \mathbf{k}| \leq r\}$ be the boundary of thickness r of a set $V \subset T$. We call a the number of indices such that $|\mathbf{i}| \leq r$, $\mathbf{i} \neq 0$, where r is the range of interaction.

If Z_0 is a maximal subgroup of $\mathbb{Z}^\nu$ satisfying the condition $|\mathbf{j} - \mathbf{k}| > r$, for $\mathbf{j}, \mathbf{k} \in Z_0$, we denote by b the number of elements in the factor group $\mathbb{Z}^\nu \backslash Z_0$.

The conditions of paper [26] (which are hypotheses on the specification Γ, once the compact function h is given) are the following

Condition 1 (Compactness) *Let h be a compact function on $\mathbb{R}$ and let $C \geq 0$ and $c_{\mathbf{j}} \geq 0$, for $|\mathbf{j}| \leq r$, $\mathbf{j} \neq 0$, be some constants. We suppose that*

1. $\delta \stackrel{\text{def}}{=} \sum_{|\mathbf{j}| \leq r \, , \mathbf{j} \neq 0} c_{\mathbf{j}} < \frac{1}{a^b}$;

2. *for any $\mathbf{i} \in T$ and any $\mathbf{x} \in \mathbb{R}^{|T|}$ one has*

$$\int_{\mathbb{R}} h(x) P_{\mathbf{i},\mathbf{x}}(\mathrm{d}x) \leq C + \sum_{\mathbf{j} \in \partial_r \{\mathbf{i}\}} c_{\mathbf{j}-\mathbf{i}} h\left(x_{\mathbf{j}}\right) \ .$$

Condition 2 (Contractivity) *Let $\bar{K} \geq 0$ and $k_{\mathbf{j}} = k_{\mathbf{j}}(\bar{K}) \geq 0$, for $|\mathbf{j}| \leq r$, $\mathbf{j} \neq 0$, be constants and h be a compact function. We suppose that*

1. $\alpha \stackrel{\text{def}}{=} \sum_{|\mathbf{j}| \leq r \, , \mathbf{j} \neq 0} k_{\mathbf{j}} < 1$;

2. *for any $\mathbf{i} \in T$ and any pair of configurations $\mathbf{x}^1, \mathbf{x}^2 \in \mathbb{R}^{|T|}$ such that*

$$\max_{\mathbf{j} \in \partial_r \{\mathbf{i}\}} \max \left\{ h\left(x_{\mathbf{j}}^1\right), h\left(x_{\mathbf{j}}^2\right) \right\} \leq \bar{K} \ ,$$

one has the inequality

$$D\left(P_{\mathbf{i},\mathbf{x}^1}, P_{\mathbf{i},\mathbf{x}^2}\right) \leq \sum_{\mathbf{j} \in \partial_r \{\mathbf{i}\}} k_{\mathbf{j}-\mathbf{i}} \mathbf{1}_{x_{\mathbf{j}}^1 \neq x_{\mathbf{j}}^2} \ ,$$

where $D(\cdot, \cdot)$ is the distance defined by (3.53).

The set of specifications (i.e., the sets of conditional probabilities on every site) which satisfy Condition 1 for the constants C, δ and the compact function h will be denoted by $\Theta(h, C, \delta)$. We will instead denote by $\Delta(h, \bar{K}, \alpha)$ the set of specifications which satisfy Condition 2 for the constants $\bar{K}, \alpha$ and the compact function h.

If the specification satisfies Conditions 1 and 2, Dobrushin and Pechersky show that the specification uniquely determines the probability (see Theorem 1 of [26]). In particular, also the marginal probability P_V on V, and the probability $P_V(\cdot|\mathrm{d}\mathbf{x})$ conditioned to a vector $\mathbf{x}$ in $\tilde{V}$ are determined by Γ, and the maximum value taken by $\lambda(\mathbf{j}, \mathbf{i})$ of (3.56) is bounded from above.

As a matter of fact, in [26] the authors do not investigate the explicit dependence of $\lambda(\mathbf{j}, \mathbf{i})$ on $|\mathbf{i}-\mathbf{j}|$, which is what we need, but such a dependence can be obtained by a slight modification of their way of reasoning, which leads to the proof of Theorem 4 that appears in Appendix A.2.3. We need also to introduce an upper bound to the mean value of h with respect to the probability P_V on V and to the probability $P_V(\cdot, \mathrm{d}\mathbf{x})$, i.e.,

$$\langle h \rangle_{\mathbf{x}} \stackrel{\text{def}}{=} \sup_{\mathbf{i} \in V} \max \left\{ \int_{\mathbb{R}} h(y) P_{\mathbf{i}}(\mathrm{d}y|\mathrm{d}\mathbf{x}), \int_{\mathbb{R}} h(y) P_{\mathbf{i}}(\mathrm{d}y), C \right\} , \tag{3.57}$$

where $\mathbf{x} \in \tilde{V}$, while C is the constant entering Condition 1. So we can state

Theorem 4 (Decay of Correlations) *Let f be a measurable function from $\mathbb{R}^{|\tilde{V}|}$ to $\mathbb{R}$, depending on the sites lying in the set $\tilde{V}$. Let g be a measurable function from $\mathbb{R}^{|V|}$ to $\mathbb{R}$, depending on the sites contained in the set V, with $V \cap \tilde{V} = \emptyset$, and h be a compact function such that inequality (3.54) is satisfied.*

Let $\Gamma \in \Theta(h, C, \delta)$. Then, there exists a constant $\bar{K}_0$, depending on h, C, a, b only, such that, if $\Gamma \in \Delta(h, \bar{K}_0 C, \alpha) \cap \Theta(h, C, \delta)$ with any α, one can find constants $D, c > 0$, for which one has

$$|\langle fg \rangle - \langle f \rangle \langle g \rangle| \leq D\,|V|^2 \left| \int_{\mathbb{R}^{|V|}} f(\mathbf{x}) \langle h \rangle_{\mathbf{x}} P_{\tilde{V}}(\mathrm{d}\mathbf{x}) \right| \exp\left(-c d(V, \tilde{V}) \right) . \tag{3.58}$$

The constant D depends on a, b, α, δ only, while one has

$$c \stackrel{\text{def}}{=} -\frac{1}{br} \log \left[\frac{1}{2} \left(\max\{\alpha, \delta a^b\} + 1 \right) \right] . \tag{3.59}$$

3.5.3 Estimate of the variance of P_n

The previous way of proceeding can be fitted to our case by choosing as specification that given by the Gibbs measure relative to H. As f and g, we choose polynomials in p and q, depending on two disjoint sets of sites. In

fact, on account of Lemma 2, we know that the P_n are constituted by a sum of such terms. This way we can study σ_{P_n} and state that it is bounded from above as in the following

Lemma 8 *There exist constants $\bar\beta > 0$, $\bar\varepsilon > 0$, $k' > 0$ such that, for any $\beta > \bar\beta$ and any $\varepsilon < \bar\varepsilon$, one has, for $n < 1/\varepsilon$,*

$$\sigma_{P_n} \le \sqrt{N}\, n!\, (k')^n\, \mathcal{D}_n \left(\sqrt{2}\beta\right)^{-1} \left(\varepsilon + \beta^{-1}\right)^n \ , \tag{3.60}$$

where the polynomials P_n are defined by (3.17) and $\mathcal{D}_n$ are given in Lemma 2.

Proof. Lemma 2 provides the necessary estimates for the coefficients which appear in the sum defining P_n (see (3.20–3.21)): we can write $P_n = \sum_{i=1}^N f_i$, where f_i are polynomials depending at most on the sites between $i-n$ and $i+n$. The variance can be expressed as $\sigma_{P_n}^2 = \sum_{i,j=1}^N \left(\langle f_i f_j\rangle - \langle f_i\rangle\langle f_j\rangle\right)$. We then consider the set $\mathcal{S}_1 \stackrel{\text{def}}{=} \{(i,j) : |j-i| \le 2n\}$ and $\mathcal{S}_2 = \mathcal{S}_1^c$. The proof goes on by finding an upper bound separately for the contributions coming from these two sets: for the latter, we use the methods developed in the present section, while the terms of the former group are estimated in a way similar to that of Lemma 3.

We start from $\mathcal{S}_1$. Firstly, we observe that, in general, one has

$$|\langle f_i f_j\rangle - \langle f_i\rangle\langle f_j\rangle| \le \sigma_{f_i}\sigma_{f_j} \le \max\left\{\langle f_i^2\rangle, \langle f_j^2\rangle\right\} \ ;$$

so that it suffices to evaluate $\sup_i\langle f_i^2\rangle$ in a way similar to Lemma 3 and to sum over the j with $|j-i| \le 2n$ to see that the contribution due to the terms in this set is smaller than

$$\sum_{i,j\in\mathcal{S}_1} |\langle f_i f_j\rangle - \langle f_i\rangle\langle f_j\rangle| \le \mathfrak{C}_1 K^{4n+1}(4n+1)(2n+1)!(2\omega)^{2n}\beta^{-2}\left(\varepsilon + \frac{1}{\beta}\right)^{2n} N\mathcal{D}_n^2 \, .$$

We come now to $\mathcal{S}_2$. We will show below that the specification coming from the Gibbs measure satisfies the hypotheses of Theorem 4, and we make use of it in estimating the terms in the set $\mathcal{S}_2$ in the following way. We separate the terms of a definite degree by writing

$$f_i = \sum_{l=0}^{n} \frac{n!}{l!(n-l)!}\varepsilon^{n-l} \sum_{s=1}^{|\mathcal{H}_{2l+2}^{n-l,i}|} c_{is,l} f_{is}^{(l)} \ ,$$

in which $f_{is}^{(l)}$ is a monomial in $\mathcal{H}_{2l+2}^{n-l,i}$ and $\sup_{i,l}\sum_s |c_{is,l}| \le \mathcal{D}_n$. We fix an index i and use Theorem 4 with $f = f_i$ and $g = f_{js}^{(l)}$, for every $j \ne i$; then,

we sum over l, s, j and i, subsequently. For each l we choose the compact function $h_l(x)$ as $|x|^{2l+2}$, which satisfies (3.54) for any $f_{js}^{(l)}$. We will show that, for β large enough and ε small enough, one has

$$\sum_{l=0}^{n} \langle h_l \rangle_{\mathbf{x}} \leq k^n n! \frac{1}{\beta} \left(\varepsilon + \frac{1}{\beta} \right)^n \exp \left(\sum_{j=1}^{|\tilde{V}|} \left(\frac{\beta\varepsilon}{\omega} x_j^2 + 8\varepsilon \sqrt{\frac{\beta}{2\omega}} \, |x_j| \right) \right) \ , \quad (3.61)$$

for a suitable constant k. Then the sum on s brings in a factor $\mathcal{D}_n$. As regards the integration in $\tilde{V}$, we observe that we can write

$$\beta \frac{1 - 2\varepsilon}{2} x_j^2 - 8\varepsilon \sqrt{\frac{\beta}{2\omega}} \, |x_j| \geq \frac{\beta}{4} x_j^2 - 1 \ ,$$

provided ε is small enough. Thus, there exists $\bar{k} > 0$ such that

$$|\langle f_i f_j \rangle - \langle f_i \rangle \langle f_j \rangle| \leq \bar{k}^n (n!)^2 \mathcal{D}_n^2 \frac{1}{\beta^2} \left(\varepsilon + \frac{1}{\beta} \right)^{2n} \exp \left[-c(|i - j| - 2n - 1) \right] \ ,$$

where the constant c is defined by (3.59), and is, in the present case, equal to $\log(4/3)/2$. Since the sum over j of such terms converges as $N \to \infty$, the proof will be concluded if we show that (3.61) and the hypotheses of Theorem 4 are satisfied.

We proceed as in the proof of Theorem 2 of paper [26], starting from the explicit form of the conditional probability distribution given by the Gibbs measure: one has

$$P_{i,\mathbf{x}}(x) = \frac{1}{Z_{\mathbf{x}}} \exp \left[-\beta \left(\frac{\omega x^2}{2} + \frac{x^4}{4\omega^2} + \frac{\varepsilon}{\omega} x \cdot x_{i-1} + \frac{\varepsilon}{\omega} x \cdot x_{i+1} \right) \right] \ ,$$

in which there appears the conditional partition function

$$Z_{\mathbf{x}} \stackrel{\text{def}}{=} \int_{\mathbb{R}} \exp \left[-\beta \left(\frac{\omega x^2}{2} + \frac{x^4}{4\omega^2} + \frac{\varepsilon}{\omega} x \cdot x_{i-1} + \frac{\varepsilon}{\omega} x \cdot x_{i+1} \right) \right] \mathrm{d}x \ .$$

As regards Condition 1, we consider $\varepsilon < 2^{-5}$ fixed and define

$$\hat{y}(\mathbf{x}) \stackrel{\text{def}}{=} \max \left\{ |x_{i-1}| \, , |x_{i+1}| \right\} \quad \text{and } y(\mathbf{x}) \stackrel{\text{def}}{=} \min \left\{ 1/\hat{y}(\mathbf{x}), \sqrt{\beta} \right\} .$$

So, it is easily proved that, for $\beta > 1$, inequality $Z_{\mathbf{x}} \geq \bar{c} y(\mathbf{x})/\beta$, holds for some $\bar{c} > 0$ independent of $\mathbf{x}$, β and ε. To prove it, it is sufficient to observe that the integrand of $Z_{\mathbf{x}}$ is bounded away from zero if $|x| \leq y(\mathbf{x})/\beta$, and

then to integrate over such an interval. We now show item 2 of Condition 1 for $h(x) = |x|^{2l+2}$. We note that

$$\begin{aligned}\int_{|x|\geq\hat{y}(\mathbf{x})/4} h_l(x)P_{i,\mathbf{x}}(x)\mathrm{d}x &\leq \frac{2}{Z_{\mathbf{x}}}\int_{1/(4y(\mathbf{x}))}^{+\infty} x^{2l+2}\exp\left(-\beta\frac{x^2}{4}\right)\mathrm{d}x \\ &\leq \frac{2\sqrt{\beta}}{\bar{c}y(\mathbf{x})}\frac{1}{\beta^{l+1}}\int_{\sqrt{\beta}/(4y(\mathbf{x}))}^{+\infty} x^{2l+2}e^{-x^2/4}\mathrm{d}x\ ,\end{aligned}$$

which, in turn, is smaller than $(l+1)!(B/\beta)^{l+1}$ for a suitable constant B, independent of ε, β and l, since the integral decreases as an exponential function of $\sqrt{\beta}/y(\mathbf{x})$. Here, we have chosen $(l+1)!$ so that the previous relation is satisfied independently of l. Furthermore, one has

$$\int_{|x|\leq\hat{y}(\mathbf{x})/4} h_l(x)P_{i,\mathbf{x}}(x)\mathrm{d}x \leq 2^{-4l-4}\left(\hat{y}(\mathbf{x})\right)^{2l+2}$$

and $(h_l(x_{i-1}) + h_l(x_{i+1}))/h_l(\hat{y}(x)) \geq 1$, for any $\mathbf{x}$: this implies that

$$\int_{\mathbb{R}} h_l(x)P_{i,\mathbf{x}}(\mathrm{d}x) \leq (l+1)!\left(\frac{B}{\beta}\right)^{l+1} + \frac{1}{16}h_l\left(x_{i-1}\right) + \frac{1}{16}h_l\left(x_{i+1}\right)\ .$$

So, item 2 of Condition 1 holds with $C = (l+1)!\,(B/\beta)^{l+1}$ and $c_1 = c_{-1} = 1/16$. Since we have $a = 2$, $b = 2$, $r = 1$, item 1 of Condition 1 holds with $\delta a^b = 1/2$.

Condition 2 is proved by computing two limits, first letting β tend to infinity and then letting $\varepsilon \to 0$. In fact, let $\mathbf{x}^{(m)}$, for $m = 1, 2$, be two different configuration such that $|x_{i-1}^{(m)}| \leq \tilde{K}/e\sqrt{Bl/(e\beta)}$ and $|x_{i+1}^{(m)}| \leq \tilde{K}/e\sqrt{Bl/(e\beta)}$. Then it is easily checked that

$$\lim_{\beta\to\infty}\int_{\mathbb{R}}\left|P_{i,\mathbf{x}^{(1)}}(x) - P_{i,\mathbf{x}^{(2)}}(x)\right|\mathrm{d}x = \int_{\mathbb{R}}\mathrm{d}z\,e^{-z^2}\left|f(\varepsilon,z,\mathbf{z}^{(1)}) - f(\varepsilon,z,\mathbf{z}^{(2)})\right|\ ,$$

where

$$f(\varepsilon,z,\mathbf{z}^{(m)}) \stackrel{\mathrm{def}}{=} \exp\left(-\frac{\varepsilon}{\omega}z\left(z_{i-1}^{(m)} + z_{i+1}^{(m)}\right) + \frac{\varepsilon^2}{2\omega^2}\left(z_{i-1}^{(m)} + z_{i+1}^{(m)}\right)^2\right)$$

and $z = x\sqrt{\beta}$, $z_j^{(m)} = x_j^{(m)}\sqrt{\beta}$. Now, by the dominated convergence theorem, one has that the limit for $\varepsilon \to 0$ of $f(\varepsilon, z, \mathbf{z}^{(m)})$ is equal to 1. Here, use is made of the fact that $\varepsilon|z_j^{(m)}|$ can be bounded from above by $\varepsilon\tilde{K}/e\sqrt{Bl/e} \leq \tilde{K}/e\sqrt{B\varepsilon/e}$, because $l \leq n$, and n is smaller than $1/\varepsilon$, by hypothesis. So, for β sufficiently large and ε small enough one has

$$\int_{\mathbb{R}}\left|P_{i,\mathbf{x}^1}(x) - P_{i,\mathbf{x}^2}(x)\right|\mathrm{d}x \leq \frac{1}{4}\ .$$

We have chosen a bound to $\mathbf{x}^m$ of this particular form, because the constant $\bar{K}$ in Condition 2 turns out to be smaller than $\tilde{K}^{2l+2}$, so that it is independent of β. So Condition 2 holds with $\bar{K} = \tilde{K}^{2l+2}$, $k_{-1} = k_1 = 1/2$ and $\alpha = 1/2$. Thus, Theorem 4 holds.

There still remains to estimate $\langle h_l \rangle_{\mathbf{x}}$ in our case. By looking at its definition (3.57), we notice that we have to estimate the integrals $\int h(y) P_{\mathbf{i}}(\mathrm{d}y|\mathrm{d}\mathbf{x})$ and $\int h(y) P_{\mathbf{i}}(\mathrm{d}y)$. Now, on account of Lemma 4 and relation (3.29), the distribution functions of $P_{\mathbf{i}}(\mathrm{d}y|\mathrm{d}\mathbf{x})$ and $P_{\mathbf{i}}(\mathrm{d}y)$ can be bounded by

$$K^{2n+2}\mathfrak{C}_2\mathfrak{C}_1\sqrt{\frac{\beta}{2\pi\omega}}\exp\left(\sum_{j=1}^{|\tilde{V}|}\left(\frac{\beta\varepsilon}{\omega}x_j^2 + 8\varepsilon\sqrt{\frac{\beta}{2\omega}}\,|x_j|\right)\right)e^{-\beta y^2/(2\omega)} \ .$$

Then, we use the bound to C previously found, together with the fact that

$$\sqrt{\frac{\beta}{2\pi\omega}}\int_{\mathbb{R}} y^{2l+2}e^{-\beta y^2/(2\omega)}\mathrm{d}y = \left(\frac{2\omega}{\beta}\right)^{l+1}\frac{(2l+1)!!}{2^{l+1}},$$

and we get (3.61). This concludes the proof.

Q.E.D.

3.5.4 Estimate of the variance of X_n

Lemma 8 of Section 3.5.3 enables us to bound from below the variance of X_n defined by (3.18) and to estimate the correlation coefficient $\rho_{X_n,H}$, according to Lemma 7, which we prove here.

Proof of Lemma 7 We start by recalling that, on account of (3.18), one has $X_n = -\Theta_1 + \sum_{j=2}^n P_j$, with Θ_1 defined by equations (3.13–3.12) of Section 3.3. It is easily seen that $\Theta_1 = F + G + \mathcal{R}_1$, in which

$$F \stackrel{\text{def}}{=} -\frac{\varepsilon}{2\omega}\sum_{i=1}^{N-1} p_i p_{i+1} \quad \text{and } G \stackrel{\text{def}}{=} \frac{3}{32\omega^2}\sum_{i=1}^{N} p_i^4 \ ,$$

and $\mathcal{R}_1$ is the remainder. Then, we study the properties of F and G, for which the mean value, the variance and the correlation with H can be computed almost exactly, and we extend such properties to Θ_1, and to the whole X_n, by observing that, in some sense, Θ_1 is the term of first order in $\varepsilon + \beta^{-1}$.

As regards formula (3.48), we notice that, since F is odd in the momenta, while G, $\mathcal{R}_1$ and the measure are even, then F is uncorrelated both with G

and with $\mathcal{R}_1$. Furthermore, one can observe that

$$\langle G\, R_1\rangle - \langle G\rangle\, \langle R_1\rangle = \frac{9}{2^{10}\omega^4} \sum_{i=1}^{N} \langle q_i^2\rangle \left(\langle p_i^6\rangle - \langle p_i^4\rangle \langle p_i^2\rangle\right) \ ,$$

and use the estimates of Lemma 4 to bound from above $\langle q_i^2\rangle$, in order to prove that $\sigma^2_{\Theta_1} \geq \sigma_F^2 + \sigma_G^2 + 2C_{G,\mathcal{R}_1} \geq N(\varepsilon^2 + \beta^{-2})/(8\beta^2)$, where the second inequality holds for ε and β^{-1} small enough. On the other hand, making use of (3.60) of Lemma 8 together with the estimate for $\mathcal{D}_n$ given by (3.21) of Lemma 2, one has

$$\sigma_{X_n} \geq \sigma_{\Theta_1} - \sum_{j=2}^{n} \sigma_{P_j} \geq \sqrt{N}\, \frac{\varepsilon + \beta^{-1}}{4\beta} \left(1 - \sum_{j=2}^{n} (j!)^4 \left(\varepsilon + \beta^{-1}\right)^{j-1} \kappa_2^j\right) \ ,$$

for a suitable constant κ_2, if β^{-1} and ε are sufficiently small. Now, for $n < \kappa_2^{-1/4}(\varepsilon + \beta^{-1})^{-1/4}$, the sum is smaller than a constant multiplied by $\varepsilon + \beta^{-1}$ and this proves (3.48).

As for (3.49), we observe that, since H is even in the momenta, F and H are uncorrelated, so that, using $\rho_X, Y < 1$, one gets

$$|\rho_{X_n,H}| \leq \frac{1}{\sigma_{X_n}\sigma_H} \left(|C_{\Theta_1-F,H}| + \sum_{j=2}^{n} \left|C_{P_j,H}\right|\right) \leq \frac{\sigma_{\Theta_1-F}}{\sigma_{\Theta_1}} \frac{\sigma_{\Theta_1}}{\sigma_{X_n}} + \frac{\sum_{j=2}^{n} \sigma_{P_j}}{\sigma_{X_n}} \ .$$

As we have just shown, for $n < \kappa_2^{-1/4}(\varepsilon + \beta^{-1})^{-1/4}$ the last term at the r.h.s. tends to zero as $\varepsilon + \beta^{-1}$, and in the same way behaves $\sigma_{\Theta_1}/\sigma_{X_n} - 1$. So, we limit ourselves to study $\sigma_{\Theta_1-F}/\sigma_{\Theta_1} = 1/\sqrt{1 + \sigma_F^2/\sigma^2_{\Theta_1-F}}$. By computing explicitly σ_F^2 and applying the upper bound (3.60) to $\sigma^2_{\Theta_1} \geq \sigma^2_{\Theta_1-F}$, we get that there exists a constant $\bar{\kappa} \geq 1$ such that

$$\frac{\sigma_{\Theta_1-F}}{\sigma_{\Theta_1}} \leq \left(1 + \frac{1}{\kappa} \frac{\varepsilon^2}{\left(\varepsilon + \beta^{-1}\right)^2}\right)^{-1/2} \ .$$

Since the r.h.s. differs from 1 by a quantity larger than $\varepsilon^2\beta^2$, the corrections given by the other terms can be neglected if $\beta \geq \varepsilon^{-1}$. This completes the proof.

Q.E.D.

3.6 Concluding remarks

We now add some comments. The first one concerns the fact that in our model we have two perturbation parameters, ε and $1/\beta$. We believe however that the only really relevant parameter is $1/\beta$. Indeed, at least formally the parameter ε can be arbitrarily decreased by performing a suitable normal form change of coordinates (see [20]), and it seems to us that such a normal form does not alter in any fundamental feature the perturbation H_1 (i.e., its local character). At the moment, however, we are unable to say anything definite on this point.

In any case, while the estimate of Section 3.5 could presumably be applied to models more general than that studied in this paper, the estimates on the marginal probabilities of Section 3.4 are especially adapted to our model. It would be important to improve our method, making it more flexible in order to cover more general situations.

In our opinion, the real big problem that remains is the construction of adiabatic invariants for problems in which small denominators appear. This problem could be overcome in particular cases (see, for example, paper [1]), but no precise strategy exists yet for the general case.

Chapter 4

A series expansion for the time autocorrelation

In the present chapter we focus on the study of the power series expansion with respect to time of the time autocorrelation of a given dynamical variable, and show how such an expansion provides some information on the asymptotic behaviour of autocorrelation for large times. In particular, we expound a criterion which enables one to establish whether the correlations decay exponentially fast to zero. It is well known, indeed, that an exponential decay to zero is a relevant chaoticity indicator. Our criterion admits a partial numerical check, as we will try to show here through an example of application to the celebrated Fermi–Pasta–Ulam problem (FPU, see [27]).

We will reproduce here most of the content of paper [13], starting in Section 4.1, where Theorem 5 for the series expansion is given. Section 4.2 is devoted to a study of some general properties of the series expansion. Here an interesting link with the Stieltjes moment problem is pointed out (see Theorem 6), and a necessary and sufficient criterion for the time-decay of the autocorrelation to be exponential is given as Corollary 2. Finally, the results of some numerical computations on the time autocorrelations of a significant variable in a FPU chain are reported in Section 4.3 and in Section 4.4 some comments on the usefulness of the present method are also given.

4.1 Main theorem on the power series expansion

The series expansion of $\mathbf{C}_f(t)$ with respect to time, for a sufficiently regular dynamical variable f, is provided by the following theorem. Here, use is made of the definition of the k–th order Lie derivative of f, for $k \geq 1$,

namely $f^{(k)} \stackrel{\text{def}}{=} [f^{(k-1)}, H]$, where $f^{(0)} \stackrel{\text{def}}{=} f$.

Theorem 5 *Let μ be a probability measure on the phase space $\mathcal{M}$, invariant for the flow generated by H, and let $n > 0$. Then, for any dynamical variable $f \in L^2(\mu, \mathcal{M})$ such that $f^{(k)} \in L^2(\mu, \mathcal{M})$ for all $k \leq n$, one has*

$$\begin{aligned} \mathbf{C}_f(t) =& \sigma_f^2 + \sum_{k=1}^{n} (-)^k \left\| f^{(k)} \right\|^2 \frac{t^{2k}}{(2k)!} \\ &+ (-1)^{n+1} \int_0^t \mathrm{d}t_1 \ldots \int_0^{t_{2n-1}} \mathrm{d}t_{2n} \left\| f^{(n)}_{t_{2n}} - f^{(n)} \right\|^2 . \end{aligned} \tag{4.1}$$

Proof. The thesis is proved by iterating Theorem 1 of Chapter 2. We use the simple chain of identities

$$\|f_t - f\|^2 = 2\sigma_f^2 - 2\mathbf{C}_f(t) = 2\|f\|^2 - 2\langle f_t\, f\rangle \ , \tag{4.2}$$

which hold true for any invariant measure, and the remark that

$$f_t - f = \int_0^t \hat{U}_{t-s} \dot{f} \mathrm{d}s, \tag{4.3}$$

where $\dot{f} = [f, H]$. One can check equations (4.2) by writing down the square at the l.h.s., while (4.3) comes[1] from the variation of constants formula applied to $f_t - f$.

We go on by writing the l.h.s. of (4.2) as

$$\begin{aligned} \|f_t - f\|^2 &= \int \mathrm{d}\mu \int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \big(\hat{U}_{t-t'} \dot{f}\big)\big(\hat{U}_{t-t''} \dot{f}\big) \\ &= \int \mathrm{d}\mu \int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t'' \big(\hat{U}_{t'-t''} \dot{f}\big) \dot{f} \qquad (4.4) \\ &= 2 \int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t'' \left(\left\| \dot{f} \right\|^2 - \frac{1}{2} \left\| \dot{f}_{t''} - \dot{f} \right\|^2 \right), \qquad (4.5) \end{aligned}$$

where, in the second line, the invariance of the measure with respect to the time evolution is used, so that the equality proceeds from a simple change of coordinates, and in the last line the same argument is used, as well as relation (4.2) applied to the variable $\dot{f}$. The order of the integrals can be exchanged if $\|\dot{f}\|$ is finite, because the integral (4.4) is absolutely convergent

[1] See the proof of Theorem 1.

in this case, in virtue of Schwarz inequality. In (4.5) we repeat the same calculation for $\|\dot{f}_{t''} - \dot{f}\|^2$, using $\dot{f} = [f, H]$. This gives us the relation

$$\begin{aligned} \|f_t - f\|^2 &= 2\left(\left\|\dot{f}\right\|^2 \frac{t^2}{2!} - \left\|f^{(2)}\right\|^2 \frac{t^4}{4!}\right) \\ &\quad + \int_0^t \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_3 \int_0^{t_3} \mathrm{d}t_4 \left\|f^{(2)}_{t_4} - f^{(2)}\right\|^2 , \end{aligned}$$

where $f^{(2)} = [\dot{f}, H]$. The same procedure can be iterated up to $k = n$ by recursively defining $f^{(k)} \stackrel{\text{def}}{=} [f^{(k-1)}, H]$, starting with $f^{(0)} \stackrel{\text{def}}{=} f$, and using the hypothesis that $f^{(k)} \in L^2(\mu, \mathcal{M})$. This completes the proof.

Q.E.D.

Remark. Theorem 5 suggests that, at least in a formal way, one can express the time autocorrelation of a dynamical variable f as a power series with respect to time

$$\mathbf{C}_f(t) = \sigma_f^2 + \sum_{k=1}^{+\infty} (-)^k \left\|f^{(k)}\right\|^2 \frac{t^{2k}}{(2k)!} . \tag{4.6}$$

To simplify the notations, we define

$$c_0 \stackrel{\text{def}}{=} \sigma_f^2 , \quad c_n \stackrel{\text{def}}{=} \left\|f^{(n)}\right\|^2 \text{ for } n > 0 . \tag{4.7}$$

We point out that all these coefficients are positive. We could have got to formula (4.6) also by expressing f_t as a time power series through its Lie derivatives and by integrating $\langle f_t\, f\rangle$ by parts, taking into account the observation that the mean value of any function which can be written as $[g, H]$ vanishes[2], since $\big\langle [g, H] \big\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle g\rangle$.

Remark. We emphasise that the term in the second line in equation (4.1), i.e., the remainder, turns out to be positive or negative according to n being odd or even, respectively. Therefore, the truncations of series (4.6) provide upper bounds for the time autocorrelation of f if truncated at an odd n, and lower bounds if truncated at an even n. Such bounds provide some information on the behaviour of the time autocorrelation of f for finite times. The simplest example is that of the first order truncation, which is precisely that of Theorem 1.

[2]For the same reason, $\sigma_{f^{(k)}}$ can replace $\|f^{(k)}\|$ in all the previous formulae. We will keep the given notation, because this way it is more straightforward to understand how to compute the coefficients of the series.

4.2 Asymptotic behaviour of the correlations

In many cases one is interested in determining whether the time correlation tends to zero as $t \to \infty$ and, if this is the case, which is the law controlling the decay. Indeed, it is commonly stated that if the correlations decay exponentially fast (for a suitable wide class of functions) then the system is chaotic. The decay is linked to the analyticity property of the Fourier transform of the correlations (see for example [28]): in particular it is well known that the Fourier transform is analytic in a strip of width $1/\tau$ if and only if the correlation decays at least as $\exp(-t/\tau)$. We recall that the correlations can always be expressed as the Fourier transform of an even positive measure, the so called "spectral measure", namely as

$$\mathbf{C}_f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega t} \mathrm{d}\alpha(\omega) \ , \tag{4.8}$$

where α is a positive Borel measure on $\mathbb{R}^+$ (see, for example, [29]).

On the other hand, it is not so easy to get information about the spectrum: we devise to get it using the Laplace transform $F(s)$ of the time autocorrelation, namely,

$$F(s) \stackrel{\text{def}}{=} \int_0^{+\infty} e^{-st} \mathbf{C}_f(t) \quad \text{for } s \in \mathbb{C} \ . \tag{4.9}$$

It is well known that the Laplace transform of a function which decays faster than $\exp(-t/\tau)$ is analytic in the half-plane $\mathrm{Re}\, s > -1/\tau$, so that, again, the control on the decay of correlations in linked to the analyticity properties of $F(s)$, as in the case of the Fourier transform. The great advantage of the latter approach lays in the fact that the Laplace transform of the correlation is essentially the Stieltjes transform of the spectral measure. This is stated as Theorem 6 below.

The point is that, as first shown by Stieltjes (see [30]), if one knows the asymptotic expansion in powers of $1/s$ of a functions $F(s)$ which is the Stieltjes transform of a positive measure, then one can recover $F(s)$ itself through a convergent scheme of continued fractions expansions, defined in terms of the coefficients of the asymptotic expansion. In other words, one can construct a rational approximation

$$F_n(s) = \frac{P_n(s)}{Q_n(s)} \ ,$$

which converges uniformly in $\mathrm{Re}\, s > 0$, as $n \to +\infty$, to $F(s)$. For the case of the time autocorrelation $\mathbf{C}_f(t)$, the asymptotic expansion is obtained just by

integrating the expansion (4.6) term by term. In fact, the Laplace transform (4.9) can be formally written as

$$F(s) = \int_0^{+\infty} \left(\sum_{n=0}^{+\infty} (-)^n \frac{c_n}{(2n)!} t^{2n} \right) e^{-st} \, \mathrm{d}t = \sum_{n=0}^{+\infty} (-)^n \frac{c_n}{s^{2n+1}} \ . \tag{4.10}$$

This expansion would give a convergent expansion, provided one can exchange the sum of the series with the integration over time. This is in general not permissible, in view of the fact that the expansion in power series of time converges (in general) only in a circle of finite radius and not up to infinity. However, it is easy to show that series (4.10) nevertheless provides the asymptotic expansion of $F(s)$ we are looking for. One has to use formula (4.1) and check that the remainder grows at most as a power of time.

We are now able to state the main result of this section, namely, the following theorem, together with a relevant corollary on the decay of the correlations.

Theorem 6 *Let $\mathbf{C}_f(t)$ be analytic in t about the origin and continuous for any $t \in \mathbb{R}$. Then the following statements hold:*

1. *in the half-plane $\mathrm{Re}\, s > 0$ the Laplace transform $F(s)$ of $\mathbf{C}_f(t)$ is analytic and one has*
$$F(s) = \frac{s}{\pi} \int_0^{+\infty} \frac{\mathrm{d}\alpha(\omega)}{s^2 + \omega^2} \ , \tag{4.11}$$
where $\alpha(\omega)$ is the positive Borel measure such that (4.8) holds;

2. *the continued fraction which approximates the asymptotic series (4.10) converges to $F(s)$ for $\mathrm{Re}\, s > 0$ and the Borel positive measure $\Phi(u)$ defined on $\mathbb{R}^+$ by $\Phi(u) \stackrel{\mathrm{def}}{=} (1/\pi)\alpha(\sqrt{u})$ solves the Stieltjes moment problem for the coefficients c_n defined by (4.7). Moreover, $\alpha(\omega)/\pi$ solves the symmetric Hamburger moment problem with moments $c'_{2n} = c_n$, $c'_{2n+1} = 0$.*

Corollary 2 *$\mathbf{C}_f(t)$ decays exponentially fast as t goes to infinity if and only if the coefficients c_n are such that $\alpha'(\omega)$ exists and $\sqrt{u}\Phi'(u)$ is analytic.*

Remark. We will state our results in terms of the regularity properties of α or Φ according to convenience. We prefer the study of the properties of Φ, when the convergence properties of Stieltjes continued fraction approximation are needed.

Proof of Theorem 6. First of all, we observe that (4.11) follows by expressing $\mathbf{C}_f(t)$ via (4.8) in the definition (4.9) of $F(s)$ and exchanging the order

of integration. This can be done for any s for which $\mathrm{Re}\, s > 0$, on account of the Tonelli-Fubini theorem for the Stieltjes integral (see [31]), and it is well known that the so defined $F(s)$ is analytic in the same region.

We come now to the proof of statement 2. We notice that the form of series (4.10) recalls the Stieltjes work [30] on the link between the asymptotic series, their expansion in continued fractions and the Stieltjes transform. In dealing with this subject he met with the moment problem that now bears his name. We recall that, given a sequence of positive numbers c_n, the Stieltjes moment problem consists in checking whether there exists a measure Φ on $\mathbb{R}^+$ such that

$$\int_0^{+\infty} u^n \mathrm{d}\Phi(u) = c_n \ . \tag{4.12}$$

The solution exists if and only if

$$\det(\Delta_n) \geq 0 \quad \text{and} \quad \det(\tilde{\Delta}_n) \geq 0 \quad \forall n \ , \tag{4.13}$$

where the matrices Δ_n, $\tilde{\Delta}_n$ are defined by

$$\Delta_n \stackrel{\text{def}}{=} \begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n} \end{bmatrix} \ , \ \tilde{\Delta}_n \stackrel{\text{def}}{=} \begin{bmatrix} c_1 & c_2 & \cdots & c_{n+1} \\ c_2 & c_3 & \cdots & c_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n+1} & c_{n+2} & \cdots & c_{2n+1} \end{bmatrix} \ , \tag{4.14}$$

and if the equality in (4.13) is obtained for Δ_k, then it is obtained for Δ_n, with $n \geq k$ and for $\tilde{\Delta}_n$, with $n \geq k-1$. Furthermore, the solution is unique if there exists $D > 0$ such that

$$c_n \leq D^n (2n)! \ , \tag{4.15}$$

holds (for a proof with more recent methods, see [32]). We point out that, in our case, since $C_f(t)$ is analytic about the origin, there should exist $D > 0$ such that the previous condition is satisfied, in view of the Cauchy-Hadamard criterion.

Coming back to our problem, we collect the results of paper [30] of interest to us in the following statement: *if the function $F(s)$ admits the representation (4.11), then the continued fraction which approximates the asymptotic expansion (4.10) converges to $F(s)$ for* $\mathrm{Re}\, s > 0$*; moreover, the Stieltjes moment problem for the sequence c_n is soluble.*

The previous statement is proved this way. In equation (4.11) we put $z \stackrel{\text{def}}{=} s^2$, $u \stackrel{\text{def}}{=} \omega^2$, $\Phi(u) \stackrel{\text{def}}{=} \alpha(\sqrt{u})$. Then we have the formal chain of equalities

$$s \int_0^{+\infty} \frac{\mathrm{d}\Phi(u)}{z+u} = F(s) = s \left(\frac{c_0}{z} - \frac{c_1}{z^2} + \frac{c_2}{z^3} + \ldots \right) \ , \tag{4.16}$$

where again Φ is a Borel positive measure on $\mathbb{R}^+$. Stieltjes showed that the continued fraction which approximates the term in brackets at the r.h.s. converges to the integral at the l.h.s, and that the Stieltjes moment problem for the sequence c_n has a solution for the measure Φ. Then it is straightforward to check that α/π solves the associated symmetric Hamburger moment problem, which means that

$$\frac{1}{\pi}\int_{-\infty}^{+\infty}\omega^{2n}\mathrm{d}\alpha(\omega) = c_n \ , \quad \frac{1}{\pi}\int_{-\infty}^{+\infty}\omega^{2n+1}\mathrm{d}\alpha(\omega) = 0 \ .$$

Q.E.D.

In view of Corollary 2, the question of the asymptotic behaviour of the time autocorrelations can be restated as a regularity problem for the measure Φ, which solves Stieltjes moment problem for the c_n, or, equivalently, for α which solves the associated Hamburger moment problem. An answer to this question can be given only if the whole sequence of moments c_n is known. However, it is also of interest to understand which is the qualitative difference between the approximation (at a given order) of an analytic measure and that of a measure with at least a singular point. We try here to give an answer to the latter question, deferring to Appendix B.1 the discussion of which is the link between the behaviour of the moments and the regularity of the solution (see Proposition 1, 2, 3 therein).

We have already pointed out that a rational approximation $F_{2n}(s)$ of the Laplace transform $F(s)$ can be constructed from (4.16) by standard methods (see Chapter 2 of Stieltjes memoir [30]), which give

$$F_{2n}(s) = s\frac{P_{2n}(s)}{Q_{2n}(s)} = s\sum_{k=0}^{n}\frac{\rho_k}{s^2+\omega_k^2} \ .$$

Such a function, in turn, can be rewritten as the following Stieltjes integral

$$F_{2n}(s) = s\int_0^{+\infty}\frac{\mathrm{d}\alpha_n(u)}{s^2+u} \ ,$$

in which the measure α_n is piecewise constant and has a jump ρ_k at the points $u=\omega_k^2$. Therefore, in order that the limiting measure $\alpha = \lim\alpha_n$ be continuous, it is necessary that, as the approximation order grows, the poles become dense and their residues tend to zero. As a consequence, *we will take as a qualitative indication of a sub-exponential decay the property that the rational approximations have at least a pole which remains isolated, with a stable residue, as the order grows.*

4.3 Numerical study of the FPU chain

Checking whether the spectral measure is analytic for a concrete problem is a major task, because it requires information on an infinite number of coefficients. Thus, as a first attempt to implement the ideas of Section 4.2, we computed numerically some coefficients c_n for a model which is widely studied in the literature and is of great relevance for the foundations of statistical mechanics, namely the Fermi–Pasta–Ulam model. It describes a one-dimensional chain of $N+1$ particles interacting through nonlinear springs. The Hamiltonian of such a system, after a suitable rescaling, can be written as

$$H = \sum_{j=0}^{N} \frac{p_j^2}{2} + \sum_{j=0}^{N-1} \left[\frac{(q_j - q_{j+1})^2}{2} + \frac{\alpha\,(q_j - q_{j+1})^3}{3} + \frac{\beta\,(q_j - q_{j+1})^4}{4} \right] ,$$

where $p = (p_0, \ldots, p_N)$ and $q = (q_0, \ldots, q_N)$ are canonically conjugated variables, while α and β are parameters that control the size of the nonlinearity. We choose as invariant probability measure the Gibbs one, i.e., $\mathrm{d}\mu(p,q) \stackrel{\text{def}}{=} Z^{-1}(T) \exp(-H(p,q)/T)\mathrm{d}p\,\mathrm{d}q$, where $Z(T)$ is the partition function and $T > 0$ the temperature, and for the purpose of speeding up the numerical evaluation of the integrals involving the measure, we consider here a chain with one end point fixed and the other one free, i.e, we impose only the boundary condition $q_0 = 0$, $p_0 = 0$ (see footnote 3 below). We note however that a few computations made on the model in which both end-points are fixed have shown no significant difference. Concerning the parameters, we consider the often studied case in which $\alpha = \beta = 1/4$. The relevant fact is that as T approaches 0 the contribution due to the nonlinear terms becomes statistically negligible.

It is well known that if the nonlinear terms are neglected then the Hamiltonian is integrable, admitting N normal modes of oscillation, and that the motion is quasi periodic so that, for every $k = 0, \ldots, N-1$, the energy E_k of the k–th mode is a constant of motion (and thus, as obvious, the time autocorrelation of such a dynamical variable is constant). Our aim is to investigate what happens when the nonlinearity is introduced as a perturbation, namely, for T positive but close to 0. We are particularly interested in the exchange of energy among the normal modes. To this end, a good choice could be to consider the fraction of energy localized on the low frequency modes, defined, for example, as

$$\mathcal{E} = \frac{1}{N} \sum_{k=0}^{N/2} E_k \ ,$$

since its time variation is a convenient indicator of the flow of energy from low to high frequency modes, and vice versa. As a numerical tool we want to compute numerically the coefficients of the series (4.6) for this function, since, in view of Theorem 6, they give indications on the asymptotic behaviour of the autocorrelation.

As a matter of fact, we notice that the function $\mathcal{E}$ may not represent the best choice. For, such a function is strongly correlated with the Hamiltonian, so that, in virtue of Theorem 2, its autocorrelation remains close to that of the Hamiltonian, thus remaining far from zero. So, we rather took a suitable modification of $\mathcal{E}$, namely,t he projection of $\mathcal{E}$, via the Gram-Schmidt orthogonalization process, on the space of the dynamical variables uncorrelated with H. Thus we consider the dynamical variable $\tilde{\mathcal{E}} \stackrel{\text{def}}{=} \mathcal{E} - \text{cov}(\mathcal{E}, H)H/\sigma_H^2$.

Our calculation proceeds as follows. We extract a sample in phase space according to the Gibbs measure[3] and we estimate the L^2-norm $\|\tilde{\mathcal{E}}^{(n)}\|^2$ of the functions $\tilde{\mathcal{E}}^{(n)} = [\tilde{\mathcal{E}}^{(n-1)}, H]$ generated by $\tilde{\mathcal{E}}^{(0)} = \tilde{\mathcal{E}}$ taking the mean value on our sample. For this purpose, we observe that $\tilde{\mathcal{E}}^{(n)}$ is a certain function of p and q, which can be expressed via Faà di Bruno's formula for the derivatives of any order of a composed function. This formula requires the computation of some combinatorial coefficients, which needs a smart implementation in order to reduce the computation time (see Appendix B.2 for the scheme we have followed).

In order to study the Laplace transform of the time autocorrelation of $\tilde{\mathcal{E}}$, following the procedure of Stieltjes, we approximate it up to order n with rational functions, and look for the poles of such an approximating function in the complex plane (see chapter 2 of the memoir [30]). Our aim is to compare the behaviour of the time autocorrelation of $\tilde{\mathcal{E}}$ with some functions which behave in a somehow known way, by comparing the poles of the corresponding rational approximations of the Laplace transforms. The first comparison function is the time autocorrelation of a variable which should lose correlation in a very short time, i.e., (the orthogonal projection of) the kinetic energy $\tilde{K}$ of the first half of the chain's particles. The second one will be a given function of time which has an exponential decay (see below for its actual form).

Let us describe the results in some detail. We know that the rational approximation of the Laplace transform at order $2n$ will have $2n$ complex conjugated poles, $s_k = i\omega_k$, $\bar{s}_k = -i\omega_k$ say. The inverses of the ω's so found for the time autocorrelation of $\tilde{\mathcal{E}}$ are plotted in fig. 4.1 versus temperature

[3]We extract p_j and $r_j \stackrel{\text{def}}{=} q_j - q_{j-1}$, for $j = 1, \ldots, N$, which are stochastically independent variables for the Gibbs measure with respect to the present Hamiltonian. Here lies the great advantage of studying the chain with one end-point free.

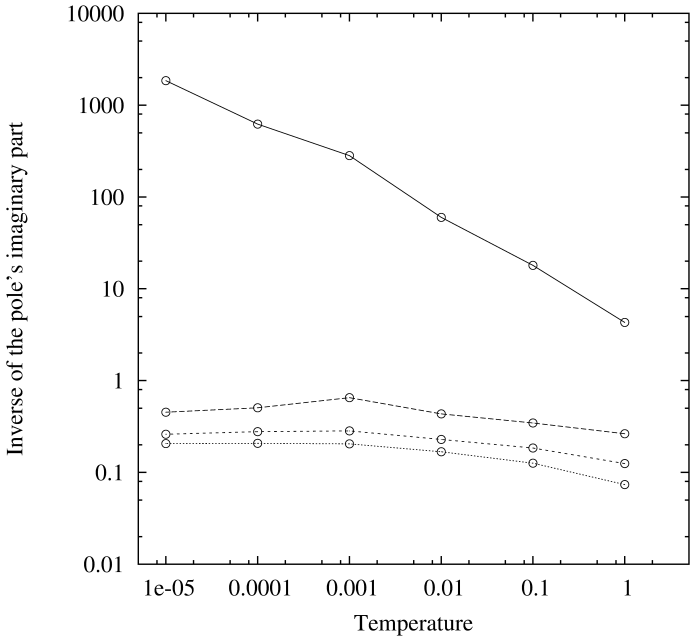


Figure 4.1: Inverse of the imaginary part of the four poles of the Laplace transform of the time autocorrelation of $\tilde{\mathcal{E}}$ versus temperature, in logarithmic scale. Here the number of particles is $N = 40$.

T. The pole that shows up at order 1 produces the upper curve in the figure. It corresponds to oscillation of a quite large period, growing as $1/T^{1/2}$ as T decreases to zero. The actual value changes a little when the approximation increases, keeping however the same order of magnitude. The plotted data correspond to the approximation order $n = 4$. We have also noted that the values found do not significantly change with the number N of particles, for N up to 364. The frequencies represented by the lower curves show up at successive orders $n = 2, 3, 4$, respectively, and correspond to oscillations of a short period, which, at variance with the first period, are of the order of magnitude of the typical periods of the normal modes. They appear to remain almost constant with T.

A further information is provided by the residues of the poles, which represent the amplitudes of the oscillations. The quantities for the four poles of fig. 4.1 are reported in fig. 4.2, after a normalization defined by setting the sum of the residues equal to one. The remarkable fact is that the amplitude corresponding to the longest period is the largest one. Even for $T = 1$ the amplitudes of the shorter periods do not exceed one percent of the total.

It is now of interest to look at the corresponding graphs (fig. 4.3a-4.3b)

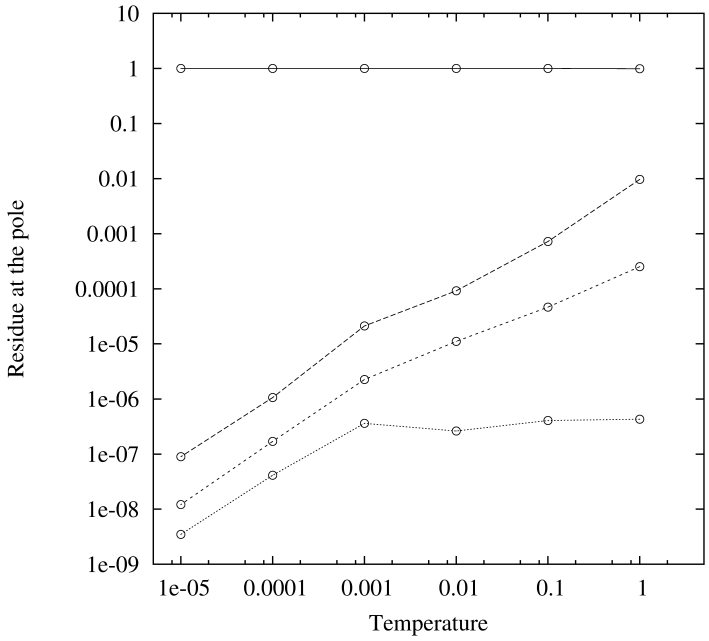
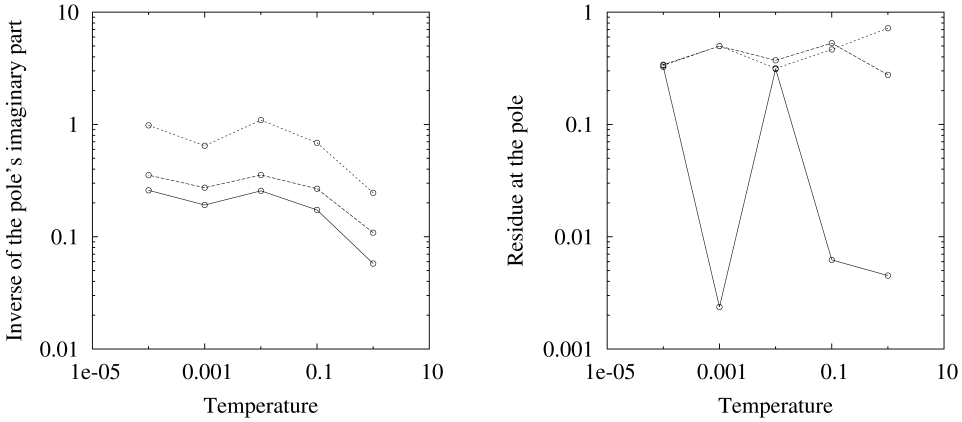


Figure 4.2: Residues at the four poles, whose imaginary parts are reported in fig. 4.1 versus temperature, in logarithmic scale. Here the number of particles is $N = 40$.

for the time autocorrelation of the kinetic energy of half of the chain, $\tilde{K}$. As expected, the time-scales involved are much shorter, as is seen by inspection of fig. 4.3a. In any case, an indication of a bigger chaoticity is given by the fact that all the residues are of the same order of magnitude and are closer to each other, as can be seen in fig. 4.3b. This means that in the spectrum of the time autocorrelation of $\tilde{K}$ at least three frequencies are excited, at variance with the spectrum of $\tilde{\mathcal{E}}$, for which just a frequency is actually excited.

Then, we compare the behaviour of $\mathbf{C}_{\tilde{\mathcal{E}}}(t)$ with that of $g(t) \stackrel{\text{def}}{=} \cosh^{-1}(bt)$, where $b > 0$, in order to look for possible differences with respect to an exponential decay. Here the constant b is so chosen that the approximation of the Laplace transform of such a function at order $n = 1$ has a pole at the same place as the Laplace transform of $\mathbf{C}_{\tilde{\mathcal{E}}}(t)$, for a fixed temperature T (as we have seen, $b \sim 1/T^{1/2}$). In fig. 4.4 the poles and the corresponding residues of the Laplace transforms of $\cosh^{-1}(bt)$ and $\mathbf{C}_{\tilde{\mathcal{E}}}(t)$ are plotted at the successive orders of approximation $n = 3, 4$.

In both cases a pole of order 10^{-3} has the largest residue, but it is apparent that, while for $g(t)$ the other poles are close to that one, in the case of $\mathbf{C}_{\tilde{\mathcal{E}}}(t)$ there seems to remain a jump between the first pole and the others. This fact



(a) Inverse of the imaginary part of the poles versus temperature, in logarithmic scale.

(b) Residues at the poles versus temperature, in logarithmic scale. Here the number of particles is $N = 40$.

Figure 4.3: Location and residues of the first three poles of the Laplace transform of the time autocorrelation of kinetic energy $\tilde{K}$ versus temperature, in logarithmic scale. Here the number of particles is $N = 40$.

is in agreement with the remark, already made by Stieltjes, that the poles of the Laplace transform of $g(t)$ become dense on the imaginary axis and corresponds with the fact that the spectral measure is in this case analytic. On the other hand, the same fact seems to suggest a qualitatively different behaviour of $\mathbf{C}_{\tilde{\mathcal{E}}}(t)$, and in particular that $\alpha(\omega)$ is not analytic in this case, so that the decay of the time autocorrelation of $\tilde{\mathcal{E}}$ is not exponential.

As a final remark, we add a few considerations concerning the actual difficulty occurring in the calculations. In view of Theorem 6, all poles lie on the imaginary axis and all the determinants (4.14) are positive. So, we are certain that the outcome of our computation ceases to be reliable at the order at which we encounter a negative determinant. Since the coefficients are determined via a Monte Carlo approximation of an integral, they are subject to a numerical error. The difficulty is that increasing the order requires also a corresponding increase of the precision, and so a longer and longer calculation. This makes the whole procedure unpractical even at not too high orders, and for this reason we performed most calculations up to order 8 only, thus finding 4 frequencies.

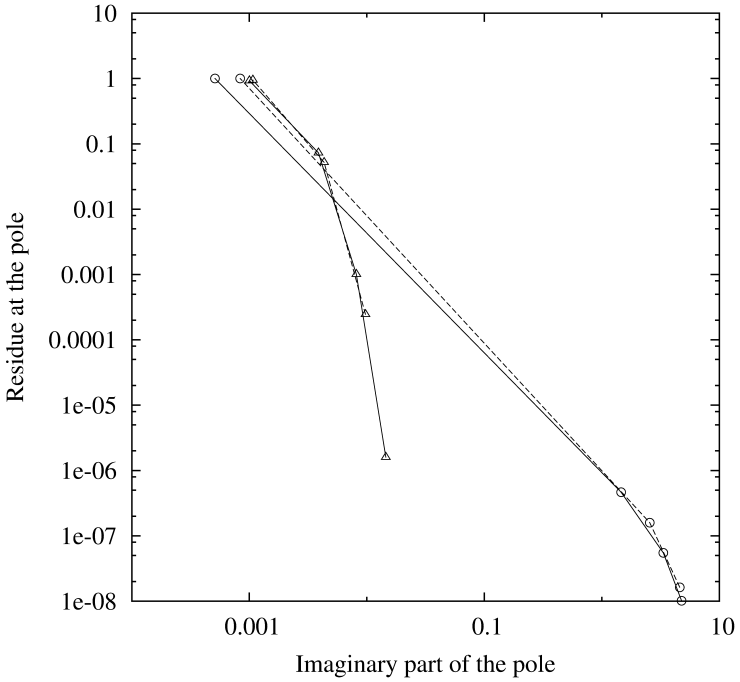


Figure 4.4: Residues of the poles of the Laplace transforms of $\mathbf{C}_{\tilde{\mathcal{E}}}(t)$ and of $g(t)$ versus their imaginary parts, in logarithmic scale. The circles refer to $\mathbf{C}_{\tilde{\mathcal{E}}}(t)$ and the triangles to $g(t)$, while the dashed lines indicate that the order of approximation is 3, and the continuous lines indicate order 4. Here the number of particles is $N = 40$ and $T = 10^{-5}$

4.4 Concluding remarks

We have shown that there is a convergent algorithm which, in principle, could be implemented to get the Laplace transform of the time autocorrelation of a dynamical variable, and so to get conclusions about the rate of its decay as $t \to \infty$. In our opinion, the most relevant aspect is the connection we found between the spectral measure and the solution of the Stieltjes moment problem, which implies a relation between the smoothness property of the measure and the fulfilment of an infinite set of constraints (see, as an example, Proposition 2 in Appendix B.1). This leads us to the question, whether smooth measures are in some sense exceptional, namely, to ask whether these constraints can still be satisfied when small changes in the moments are performed (i.e., by generic small perturbations of the Hamiltonian flow).

Implementing such an algorithm on a computer raises major difficulties with respect to both the computational time and the reliability of the results. As a matter of fact, the main problem is connected with the calculation of large determinants. However, we tried to apply the method to two functions of interest for the FPU problem, namely, the energy of a packet of low frequency modes and the kinetic energy of half of a FPU chain. For the first function, our results seem to support the conjecture that the time autocorrelations exhibit a sub-exponential decay, while no definite conclusions can be drawn for the second quantity. We remark that this is a known characteristic for chains of FPU type, namely that quantities related to the particles typically exhibit a behaviour coherent with the predictions of Statistical Mechanics, while this often does not happen with quantities related to the modes.

Chapter 5

A physical application to magnetized plasmas

The idea itself that perturbation theory can apply at the thermodynamic limit, and that there exists some continuity in the behaviour of correlations under perturbations of an ordered system, makes us ponder upon some assumptions which are common in statistical mechanics. If one refers to linear response theory, one should admit that response functions can take on values which differ from the equilibrium ones even up to experimentally observable times. More generally, if some time correlation is known not to decay for very large times, this appears as indicating that the dynamics is not completely chaotic, and preserves some part of its ordered character.

To better elucidate some physical consequences implied by this fact, we expound here some heuristic arguments for the case of magnetized plasmas. Such plasmas can be modeled as a completely resonant system (as all electrons turn with the same frequency), and one could think that also in this case an adiabatic invariant can be constructed, at least for small values of some perturbation parameter, in a way similar to Chapter 3. Actually, we are not able, at the moment, to implement our perturbation techniques for such a system, but we intend to investigate how the presence of an adiabatic invariant would affect the macroscopic behaviour of the system. Indeed, if one supposes that some relevant quantity has an essentially constant time autocorrelation up to times which grows as the perturbation parameter goes to zero, one can say that the system presents a transition from a chaotic to an ordered regime.[1] Furthermore, if the dependence of the stability time on

[1]The concepts of chaos and order are actually not free from ambiguities, but we will consider the fast or slow decay of the autocorrelation of the adiabatic invariant we study (namely, the magnetization) as an indicator of a chaotic or ordered regime, respectively. This may seem an arbitrary definition of chaos and order, but the relevance of magneti-

the perturbation parameter is exponential (or just steep enough, as in the case of Theorem 3 of Chapter 3), the transition between the two regimes occurs suddenly, and one can think that there exists a threshold value for the perturbation parameter, below which the system is ordered, and above which it is chaotic.

In the case of magnetized plasmas, the arguments just given lead us (see Section 5.1) to conjecture that the perturbation parameter depends only on the density n of electrons and the intensity B of a uniform magnetic field which is imposed to the plasma from outside. This entails the existence of a limit density (depending only on B) above which the system is chaotic and cannot thus be confined by the magnetic field. The conjectured limit density can be estimated in an extremely idealized case, but this value however turns out to fit surprisingly well the data on the maximum densities at which plasma fusion devices can operate.

The conjecture is provided, together with a comparison with experimental data, in the following Section 5.1, which reports part of paper [33]. Some numerical computations seem to confirm such a conjecture, as shown in Section 5.2, while some concluding comments follow in Section 5.3.

5.1 Density limit in magnetized plasmas

The model chosen for the magnetized plasma is the simplest one we could conceive. We consider the extremely idealized case of a plasma subject to an uniform magnetic field, say $\mathbf{B} = B\mathbf{e}_z$, where $\mathbf{e}_z$ is the unit vector along the z axis. Concerning the plasma itself, it is conceived as a dynamical system of point charges, in which any charge will be subject, in addition to the Lorentz confining force due to $\mathbf{B}$, also to the force due to the Coulomb electric field $\mathbf{E}$ generated by all other particles, which is usually called the microfield. Mutual magnetic forces and retardation effects are neglected.

So we have a system of several kinds of charges, and the Newton equation for the j–th charge (in the nonrelativistic approximation) is then

$$m_j \ddot{\mathbf{x}}_j = e_j \mathbf{v}_j \wedge \mathbf{B} + e_j \mathbf{E}_j \tag{5.1}$$

where m_j and e_j are the mass and the charge of the particle, $\mathbf{x}_j$ and $\mathbf{v}_j = \dot{\mathbf{x}}_j$ its position vector and velocity, and $\mathbf{E}_j = \mathbf{E}(\mathbf{x}_j)$ the microfield evaluated at $\mathbf{x}_j$. Its value is obviously given by

$$\mathbf{E}_j \stackrel{\text{def}}{=} \frac{1}{4\pi\varepsilon_0} \sum_{k \neq j} \frac{e_k(\mathbf{x}_k - \mathbf{x}_j)}{|\mathbf{x}_k - \mathbf{x}_j|^3} \ .$$

zation (see below) suggests such qualifications are adequate.

Finally, in order that the dynamical system be defined within the standard approach of ergodic theory, we consider as given also an invariant measure, a few minimal properties of which will be mentioned later.

If the microfield is neglected, the transverse motion of each particle is a uniform gyration about a field line with its characteristic cyclotron frequency $\omega_{cj} = |e_j| B/m_j$. So the system is integrable, the z component of the angular momentum of each particle being a constant of motion. Moreover, one can check that the considered system with N particles, in the absence of microfield, can be transformed through a time dependent change of variables into a Hamiltonian system of $2N$ independent harmonic oscillators and N free particles on a line, which, if all masses and all charges are equal, is completely resonant (apart from the free particle terms). The microfield, acting as a perturbation, makes the system no more integrable.

5.1.1 Conjecture for the transition from order to chaos

As already explained, in order to establish whether a transition from order to chaos occurs, we investigate the behaviour of a sensible observable. Our purpose is to look at the magnetization of the system, to which each charge contributes through the z component of its angular momentum, in view of its relevance for diamagnetism. In fact, magnetization is a constant of motion if the electric field is neglected. Moreover, as said in Section 1.2 of Chapter 1, in the frame of linear response theory the magnetic susceptibility vanishes in classical physics if the system is chaotic enough, i.e., if the time autocorrelation of magnetization decays sufficiently fast. Instead, susceptibility can remain bounded away from zero for very long times if the time autocorrelation of magnetization stays almost constant. Therefore, if we make a further idealization in our extremely simplified treatment, the transition from order to chaos we are looking at corresponds to a passage from a diamagnetic to a nonmagnetic plasma. This seems quite consistent, since it is known that the existence of diamagnetism plays a key role in magnetic confinement of plasmas, due to the presence of the diamagnetic pressure, which pushes particles from regions of lower magnetic field (think of the border of the fusion machine) to regions of larger magnetic field (see [6]).

On the other hand, it is well known that only the electrons are relevant, the ions contribution to magnetization being negligible. So, in order to make the treatment as simple as possible, we will consider the contribution to magnetization due to any single electron, i.e., the z component of its angular momentum. Thus, as zeroth order approximation for the adiabatic invariant

we take the quantity[2]

$$L = \frac{m^2}{eB}\, v_\perp^2 \tag{5.2}$$

($\mathbf{v}_\perp$ denoting transverse velocity of the chosen electron), which is proportional to the transverse kinetic energy of the electron. One immediately checks (see, for instance, page 16 of the book of Alfvén [34]) that L is the z component of the angular momentum of the chosen electron, referred to its instantaneous gyration center (or guiding center), the latter being calculated in the approximation in which the perturbing force is neglected.

We also add here the formula for the time derivative $\dot{L}$ of L, as we will need it in a moment. As L is a multiple of $\mathbf{v}_\perp \cdot \mathbf{v}_\perp$, $\dot{L}$ is immediately obtained through dot multiplication of Newton's equation (5.1) by $2\mathbf{v}_\perp$, which gives

$$\dot{L} = \frac{2m}{B}\, \mathbf{v}_\perp \cdot \mathbf{E}_\perp \ . \tag{5.3}$$

We come now to the main point, i.e., to find the dimensionless perturbation parameter ϵ which determines the threshold of the transition from order to chaos corresponding to the destruction of the chosen adiabatic invariant (and of all the adiabatic invariants corresponding to each electron). This would require performing the corresponding perturbation estimates at all orders, which at the moment we are unable to do. What we can easily do is to estimate the time autocorrelation function of L, by using the fundamental result of Theorem 1 of Chapter 2. This leads to

$$\frac{C_L(t)}{\sigma_L^2} \geq 1 - \frac{1}{2}\frac{\sigma_{\dot{L}}^2}{\sigma_L^2} t^2 \ . \tag{5.4}$$

Thinking of this as the first step of a perturbation scheme, we expect that this is the zeroth order approximation of an adiabatic invariant $I^{(s)}$ at order s, for which one should have something similar to the estimates in Section 3.2 of Chapter 3, namely,

$$\frac{C_{I^{(n)}}(t)}{\sigma_{I^{(n)}}^2} \geq 1 - \frac{1}{2}(n!)^\alpha \epsilon^{n+1} \left(\frac{t}{\tau}\right)^2 \ ,$$

for some constants α and τ. The conjecture that one has $\tau = 1/\omega_c$ seems then quite natural, because this is the microscopic time of the unperturbed electron motion and thus, from (5.4) we find that the relevant dimensionless parameter of the problem is

$$\epsilon = \frac{\sigma_{\dot{L}}}{\omega_c \sigma_L} \ , \tag{5.5}$$

[2]Since now on, the index j referring to a chosen electron will be left understood.

($\sigma_X = \sqrt{\sigma_X^2}$ denoting standard deviation), and this leads to a chaoticity threshold given by

$$\frac{\sigma_{\dot{L}}}{\omega_c \sigma_L} = 1 \ . \tag{5.6}$$

Through formula (5.8) of the next section it will be seen that the condition for the threshold can also be expressed in the form $(\sigma_{E_\perp}/B\sigma_v) = 1$. Thus one sees that the threshold occurs when the typical value of the perturbing force due to the microfield equals the typical value of the Lorentz force which characterizes the unperturbed motions. So the condition $\epsilon = 1$, which we have assumed as a definition of the threshold within a rather abstract point of view, is just what one would immediately guess, as the naivest implementation of the idea that a threshold occurs when the perturbing force equals the unperturbed one.

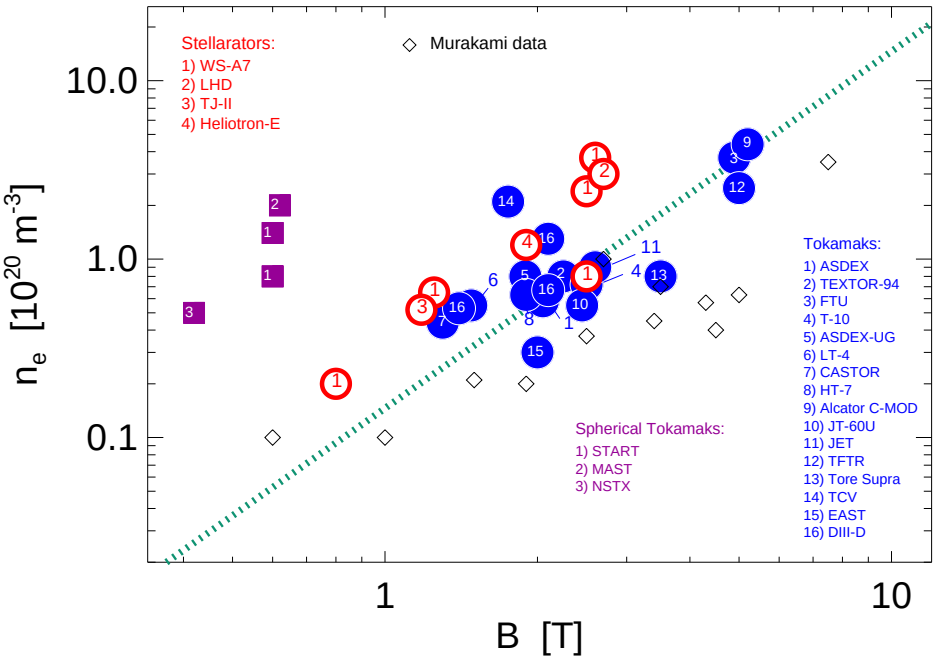


Figure 5.1: Density limit values *vs* B for various machines: conventional tokamaks, for which recent data are shown (see references [38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53]) along with the original ones of Murakami (see [37]), stellarator machines [55, 56, 57, 58], and spherical tokamaks [59, 60, 61]. Dotted line is the theoretical density limit (5.10).

5.1.2 The chaoticity threshold in terms of macroscopic parameters. The theoretical density limit

Our aim is now to express the chaoticity threshold (5.6) in terms of the macroscopic parameters T, n, B, temperature, electron number density and field strength. Recalling the expressions (5.2), (5.3) of L and $\dot{L}$, and the definition $\omega_c = |e|\, B/m$, the threshold (5.6) takes the form

$$\frac{2}{B}\,\frac{\sigma_{\mathbf{v}_\perp\cdot\mathbf{E}_\perp}}{\sigma_{v_\perp^2}} = 1\ . \tag{5.7}$$

It is clear that the standard deviations appearing in (5.7) depend on the model of plasma adopted, which determines the microfield, as well as on the chosen invariant measure. The choice of the invariant measure is a quite delicate problem, particularly in a nonequilibriun situation as the one we are discussing here. A general introduction may be found in the book [4]. For example, it is obvious that $\sigma_{\mathbf{v}_\perp}$ and $\sigma_{v_\perp^2}$ should be expressed in terms of temperature, albeit with coefficients which depend on the assumptions made for the velocity distribution. Analogously, the statistical properties of the microfield may be different for a system composed by electrons plus a neutralizing background, rather than for a system of electrons and ions.

Quite natural assumptions on the measure are:

i) that velocities and positions are independent variables;

ii) that the distribution of the transverse velocities is Maxwellian at a temperature T;

iii) that the distribution of positions is isotropic.

Under these natural assumptions the equation (5.7) for the threshold is seen to take the form

$$\sqrt{\frac{2}{3}}\,\frac{\sigma_{\mathbf{E}}}{B\,\sqrt{k_B T/m}} = 1 \tag{5.8}$$

Indeed, from i) and ii) one gets

$$\sigma^2_{\mathbf{v}_\perp\cdot\mathbf{E}_\perp} = \frac{1}{2}\,\sigma^2_{\mathbf{v}_\perp}\,\sigma^2_{\mathbf{E}_\perp}\ ,$$

the variance of a vector $\mathbf{F}$ being defined by $\sigma^2_{\mathbf{F}} = \sigma^2_{F_x} + \sigma^2_{F_y} + \sigma^2_{F_z}$. One also gets

$$\frac{1}{2}\,\sigma^2_{\mathbf{v}_\perp} = \frac{k_B T}{m}$$

(k_B being the Boltzmann constant) and furthermore, as one easily checks,

$$\sigma^2_{v^2_\perp} = 4\left(\frac{k_B T}{m}\right)^2 \ .$$

Finally, from iii) one gets $\sigma^2_{\mathbf{E}_\perp} = (2/3)\,\sigma^2_{\mathbf{E}}$

The form (5.8) of the equation for the threshold already constitutes in our opinion a significant result. Indeed, the fluctuation $\sigma^2_{\mathbf{E}}$ of the microfield should in principle be itself a measurable quantity, which depends on the macroscopic state of the plasma, namely, electron number density n and the temperatures of the several constituents. So the previous relation provides in principle the density limit as a function of the macroscopic state of the plasma.

However, we were unable to find in the literature sufficient experimental information on the fluctuation $\sigma^2_{\mathbf{E}}$ of the microfield. So, in order to have a definite theoretical formula to be compared with the experimental data, we limit ourselves to the consideration of a particular model for which an estimate of $\sigma^2_{\mathbf{E}}$ is available. In fact a formula for $\sigma^2_{\mathbf{E}}$ was given by Iglesias, Lebowitz and MacGowan [35] for the model of a one component plasma with neutralizing background, at equilibrium with respect to the Gibbs distribution, for which they found

$$\sigma^2_{\mathbf{E}} = \frac{n\,k_B T}{\varepsilon_0} \ , \tag{5.9}$$

where ε_0 is the vacuum dielectric constant and n the electron number density (see [35], formula (2.5), substituting n for ρ and $1/\varepsilon_0$ for 4π).

So, for a one component plasma with neutralizing background at Gibbs equilibrium at temperature T, the chaoticity threshold (5.8) takes the form

$$n = \frac{3}{2}\,\frac{\varepsilon_0}{m}\,B^2 \ , \tag{5.10}$$

in which temperature disappeared, so that the threshold only involves density and field strength. Notice however that this might not be true with a more realistic model of a plasma, in which the temperature appearing in (5.9), which refers to the plasma as a whole, may be different from the electron transverse temperature which enters the previous formulas.

Formula (5.10) for the limit density (holding for a one component plasma with neutralizing background, at Gibbs equilibrium at temperature T) is the type of result we were looking for, inasmuch as it provides a definite theoretical formula for the density limit that can be compared to the available empirical data for collapses in fusion machines, as will be done in the next section.

5.1.3 Comparison with the empirical data for plasma collapses in fusion machines

We now check whether the transition from order to chaos discussed here has anything to do with the empirical data for collapses in fusion machines. We recall that a proportionality of the density limit to the square of the magnetic field in tokamaks was suggested by Granetz [36] on the basis of empirical data, but apparently was not confirmed by later observations [53, 54]. It is well known that, while at first a proportionality to the magnetic field (through B/R, where R is the major radius of the torus) had been proposed on an empirical basis for tokamaks by Murakami [37], in the plasma physics community the common opinion is rather that the density limit for tokamaks should be proportional to the Greenwald parameter I_p/r_a^2, where I_p is the plasma current and r_a the minor radius of the torus (see [54]).

We do not enter here a discussion of this point, and only content ourselves with plotting in Figure 5.1 a collection of available data of the density limit for several fusion machines versus their operating magnetic field B in log–log scale, comparing the data to the theoretical formula (5.10). The first thing that comes out from the figure is that the order of magnitude of the theoretical threshold is correct, and this without having introduced any phenomenological parameter. There is no adjustable parameter in the theory, and no fitting at all. One is thus tempted to say that the essence of the phenomenon has perhaps been captured, especially in consideration of the extreme simplicity of the model, with respect to the variety of machines and of operational conditions to which the experimental data refer.

Entering now in some more details, one sees that the theoretical law appears to correspond not so badly to the data for the high field machines (tokamak and stellarators), whereas a sensible discrepancy is met for the low field machines (spherical tokamaks), for which the experimental data are larger by even an order of magnitude. Perhaps this discrepancy might be attributed to the fact that we are discussing here a model describing an isolated, non sustained, system (i.e., with no input heating power), whereas the low field machines considered in the figure are just the ones characterized, in general, by lower confinement time and thus by larger sustainment. Indeed, (see the empirical Sudo limit for stellarators [58]) larger densities are expected to be accessible as the input power is increased (although this is not so clear for tokamaks [54]). This is illustrated, in the figure, by the three points reported for the same device (the stellarator WS-A7 [55]) at essentially the same applied field, which however correspond to three different (increasing) input heatings.

5.2 Time autocorrelation of magnetization: numerical computations

Although the density limit we have found seems to agree rather well with experimental data, this doesn't help too much in clarifying whether our conjecture on the behaviour of the idealized system we are considering is well grounded. Indeed, real plasmas are too complex to be well approximated by our model, and one can still ask whether the dynamical system of point particles subject to equations of motion (5.1) actually displays a transition from order to chaos. Moreover, it is also interesting to study the behaviour of the magnetization itself, instead of that of the transverse kinetic energy of a single electron.

We tried to answer these questions by performing some numerical computations of the time autocorrelation of magnetization for a pure electron plasma. The results thus obtained seem to confirm the theoretical estimate (5.10) for the transition from order to chaos, by exhibiting below threshold an ordered behaviour of correlations, which do not decay up to long times, and a chaotic behaviour above it.

The model we consider is just that of a system of N point electrons with mutual Coulomb interactions subject to a uniform magnetic field, enclosed in a box of side L, so that, obviously, for the electron number density one has $n = N/L^3$. So, the equations of motion are simply given by (5.1), in which $m_j = m$ and $e_j = e$. Caution, however, should be paid in the definition of the microfield. Indeed, we chose to introduce periodic boundary conditions, so that we are actually dealing with a system of infinitely many electrons. In such a case, the electric field can be computed (see [62]) by the Ewald summation of the field due to an infinite cubic lattice of charges of the form $\mathbf{x}_i + L\mathbf{n}$. Here, $\mathbf{n}$ is a vector with integer coordinates, i.e., $\mathbf{n} \stackrel{\text{def}}{=} (n_x\mathbf{e}_x + n_y\mathbf{e}_y + n_x\mathbf{e}_z)$, with n_x, n_y and $n_z \in \mathbb{Z}$, while $\mathbf{e}_x, \ldots$ are the unit vectors along their axes.

Rescaling time by the electron cyclotron frequency ω_c (see paragraph 5.1.1 for its definition) and position vectors by the mean interparticle distance $a \stackrel{\text{def}}{=} n^{-1/3}$, i.e., introducing $\tau \stackrel{\text{def}}{=} \omega_c t$ and $\mathbf{y}_j \stackrel{\text{def}}{=} \mathbf{x}_j/a$, the equations of motion (5.1) take the form

$$\ddot{\mathbf{y}}_j = \mathbf{e}_z \wedge \dot{\mathbf{y}}_j + \left(\frac{\omega_p}{\omega_c}\right)^2 \tilde{\mathbf{E}}(\mathbf{y}_j) \ , \tag{5.11}$$

where the plasma frequency $\omega_p \stackrel{\text{def}}{=} \sqrt{e^2 n/(4\pi\varepsilon_0 m)}$ and the dimensionless field $\tilde{\mathbf{E}}$ appear, while the dots now denote derivatives with respect to τ. In the

simulations, the explicit form of the Ewald resummed field $\tilde{\mathbf{E}}$ acting on the j–th particle is given by

$$\tilde{\mathbf{E}}(\mathbf{y}_j) = \sum_{\mathbf{n}} \sum_{l} \frac{\mathbf{r}_{l,\mathbf{n}}}{|r_{l,\mathbf{n}}|^3} \Big[\operatorname{erfc}(\alpha r_{l,\mathbf{n}}) + \frac{\alpha r_{l,\mathbf{n}}}{\sqrt{\pi}} \exp(-\alpha^2 r_{l,\mathbf{n}}^2) \Big] + \frac{4\pi}{L^3} \sum_{\mathbf{k}\neq 0} \sum_{l} \frac{\mathbf{k}}{|\mathbf{k}|^2} \exp(-\frac{\mathbf{k}^2}{4\alpha}) \sin(\mathbf{k}\cdot\mathbf{r}_l) \ .$$

Here $\mathbf{r}_l \stackrel{\text{def}}{=} \mathbf{y}_j - \mathbf{y}_l$, while $\mathbf{r}_{l,\mathbf{n}} \stackrel{\text{def}}{=} \mathbf{y}_j - \mathbf{y}_l + L\mathbf{n}/a$; the function $\operatorname{erfc}(x)$ is the usual error function, and α is the Ewald convergence parameter which we chose as $\alpha \stackrel{\text{def}}{=} \pi^{1/2} N^{-1/6}$. Notice also that, in the first sum, the term corresponding to the self-force on the j–th particle is excluded.

It is important to notice that equation (5.11) contains only one (dimensionless) parameter, namely, ω_p/ω_c, while the rescaled density is obviously equal to 1. On the other hand, it is easy to check that the threshold (5.10) takes, in virtue of the definitions, the approximate form $\omega_p/\omega_c \approx 1$, which will be used later on.

These are the equations of motion that were actually integrated numerically, using a symplectic splitting method. The conservation of energy in every run was better than a part over 10^3. The integration time was chosen proportional to ω_c in order that all different cases be integrated for the same physical time. In any case, the time was always some hundreds cyclotron periods.

The initial data were chosen in the following way: the electron positions $\mathbf{y}_j$ were taken uniformly distributed in the box of side $N^{1/3}$, while the velocities were extracted from a Maxwellian with a given temperature T. This introduces in the model (in addition to ω_p/ω_c) the further parameter T, or equivalently the dimensionless Coulomb coupling parameter Γ defined by

$$\Gamma \stackrel{\text{def}}{=} e^2/(4\pi\epsilon_0 a k_B T) \ ,$$

to be used in the Maxwell distribution for the velocities (here, k_B denotes the Boltzmann constant).

For what concerns the number N of electrons in the box, our computational power allows us to go up to $N = 512$. This induces a lower bound on Γ, namely, $\Gamma \geq N^{-2/3}$. Indeed, in order to correctly simulate the Coulomb cumulative force acting on an electron in a plasma, the side of the box has to be at least equal to the Debye length $\lambda_D \stackrel{\text{def}}{=} \sqrt{k_B T/ne^2}$, which, in our rescaled units, takes the value $\lambda_D = \Gamma^{-1/2}$. We took $\Gamma = N^{-2/3}$. Computations were performed both for $N = 128$ and $N = 512$, which correspond to $\Gamma = 128^{-2/3} \simeq 0.04$ and $\Gamma = 512^{-2/3} = 1/64 \simeq 0.016$ respectively.

We now come to the main issue, i.e., whether the motions are ordered or chaotic. Obviously what plays the role of the unperturbed system with completely ordered motions is the limit case with $\omega_p/\omega_c = 0$, for which the Coulomb interaction disappears and one has pure Larmor gyrations. The problem then is to determine whether a threshold for chaotic motions takes place as the parameter ω_p/ω_c is increased and Γ is varied. To this end we looked at the z component of the magnetization of a box, $M_z \stackrel{\text{def}}{=} (e/2m)\sum(\mathbf{x}_j \wedge \dot{\mathbf{x}}_j)_z$, looking at its autocorrelation function (normalized by Nk_BT)

$$\mathcal{C}_{M_z}(t) \stackrel{\text{def}}{=} \frac{\mathbf{C}_{M_z}(t)}{Nk_BT} \ ,$$

Remark that here we are considering the actual magnetization, which is different from the equilibrium magnetization M_z^0, the latter involving the momenta p instead of the velocities $\dot{x}$ (see formula (1.6) of Section 1.2). Moreover, the autocorrelation is computed along the flow of the global Hamiltonian. Thus $\mathcal{C}_{M_z}(t)$ is different from the correlation considered in Section 1.2, but it is related to the magnetic susceptibility by the same formula, as is proved in [6] (actually, the differences between them disappear in the linear approximation). We will also consider the Fourier transform $\hat{\mathcal{C}}_{M_z}(\omega)$ of the autocorrelation. The latter is a physically very relevant quantity because, according to linear response theory (just apply the Fourier transform to equation (1.8) with $h(t) = e^{i\omega t}$, for instance) the quantity $i\omega\hat{\mathcal{C}}_{M_z}(\omega)$ gives the susceptibility $\chi(\omega)$ at frequency ω.

As regards the actual computation of such quantities, we stress that, whereas in the formula for the time–autocorrelation $\mathcal{C}_{M_z}(t)$, the average is meant, as usual, as a phase–average with respect to the global Gibbs measure, in our computations we estimated it by the time–average along an orbit (with initial data extracted as previously explained), as often done in numerical works. We did not investigate the relations between the two averages. Moreover, the Fourier transform $\hat{\mathcal{C}}_{M_z}(\omega)$ was estimated by the amplitude of the discrete Fourier transform of $\mathcal{C}_{M_z}(t)$, which will be simply called the spectrum. So we report figures of the time–autocorrelation $\mathcal{C}_{M_z}$ versus t, and of the corresponding spectrum versus angular frequency ω/ω_c.

Having fixed $\Gamma = 1/64$, by increasing ω_p/ω_c we found that a threshold occurs for ω_p/ω_c between 0.25 and 2. This is exhibited in Fig. 5.2, where the results are reported for such two values of ω_p/ω_c, 0.25 on the left and 2 on the right. The autocorrelations are reported in the upper part of the figure, and the spectra in the lower part.

For $\omega_p/\omega_c = 0.25$ the autocorrelation is seen to display regular oscillations with a decreasing amplitude: we were unable to follow this relaxation process

up to the end. The oscillations are apparently peaked about the cyclotron frequency and its harmonics of low order (as should be, due to the nonlinearities in the equations of motions). This is clearly exhibited by the spectrum, with its large peak at $\omega/\omega_c = 1$, and the smaller ones about the harmonics $\omega/\omega_c = 2, 3, ...$ Of special relevance is the peak at $\omega = 0$, which corresponds to the existence of a nonvanishing static susceptibility, i.e., to the existence of diamagnetism, which could be also noted in the corresponding autocorrelation graph, as they remain always different from zero. In the spectrum also appears a continuous component, which accounts for the extremely slow drift towards equilibrium. This case clearly corresponds to prevalently ordered motions with a corresponding nonvanishing diamagnetism, and should be interpreted as an indication that the perturbation due to the Coulomb interactions is not yet sufficiently large to produce prevalent chaotic motions. The passage to chaos, however, already occurred at $\omega_p/\omega_c = 2$. Indeed in this case the autocorrelation is seen to go to zero in an extremely short lapse of time (even shorter than one cyclotron period $2\pi/\omega_c$), so that the peaks disappear from the spectrum and one only remains with the continuous part. This means that for $\Gamma = 1/64$ the threshold in ω_p/ω_c lies between 0.25 and 2. For $\Gamma = 128^{-2/3}$ the corresponding figures at those same values of ω_p/ω_c are qualitatively similar to the above ones, and are not reported here.

So, the numerical results obtained for $\Gamma = 128^{-2/3} \simeq 0.04$ and $\Gamma = 1/64 \simeq 0.016$ are in rather good agreement with the theoretical prediction of formula (5.10), namely: at $\omega_p/\omega_c = 1$ the interactions become strong enough as to make the motions chaotic.

5.3 Conclusions

The case of magnetized plasmas shows how the perturbation theory here presented can be used in a heuristic way to get some information on the physics of a macroscopic system, thus complementing the rather abstract example of Chapter 3. In doing this, we made use of extremely simplified arguments, neglecting some relevant physical and mathematical features, but this allowed us to provide a possible explanation of the relevant phenomenon of the existence of a density limit in fusion devices, for which there seems to lack any first principles rationale. In fact, the comparison between theory and experiments exhibited in Figure 5.1 appears quite encouraging.

This fact stimulates us to insist in our research line and to try to broaden the application field of the mathematical methods, which are at the very beginning and need to be consolidated in order to deal with different physical situations. We hope, however, that the present result may be a good

starting point for a rigorous study of macroscopic systems within the frame of perturbation theory.

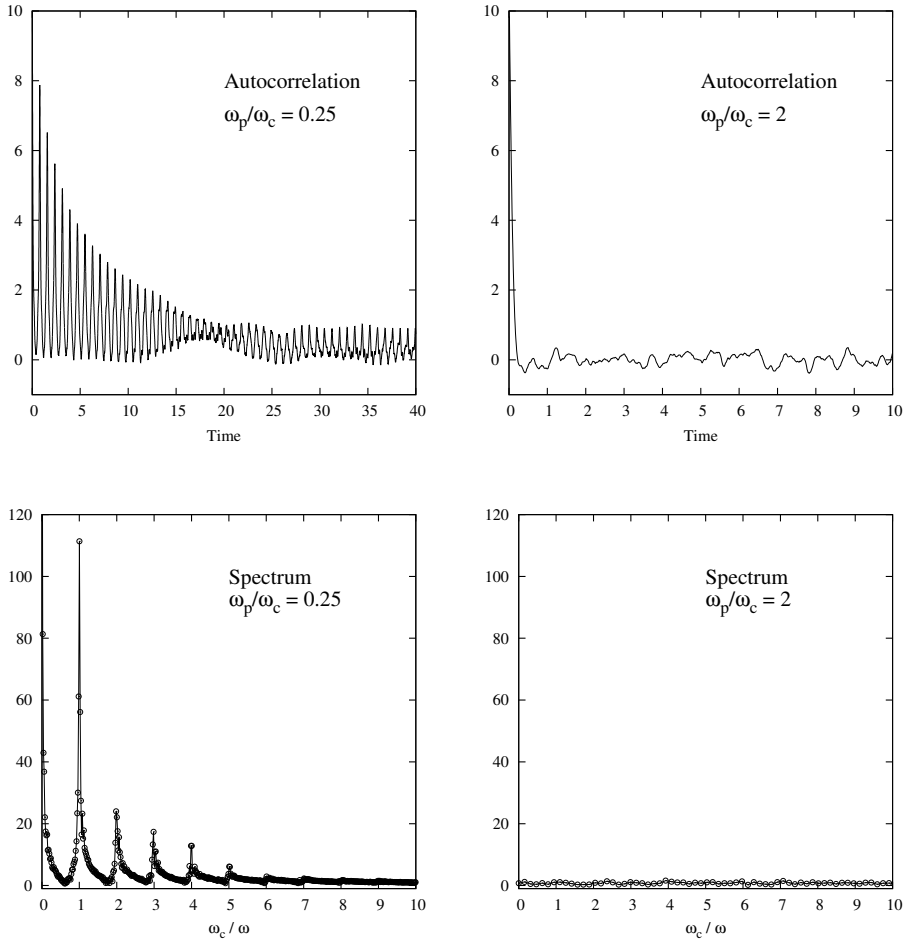


Figure 5.2: Top: Autocorrelation $\mathcal{C}_M(t)$ of magnetization versus time for $\omega_p/\omega_c = 0.25$ (left) and $\omega_p/\omega_c = 2$ (right). The time scale is the same in both figures, and one has $\omega_c = 1$ at the right, $\omega_c = 8$ at the left. Notice the fast decay to zero at the right. Bottom: Discrete Fourier transform (absolute value) of $\mathcal{C}_M(t)$ versus ω/ω_c for $\omega_p/\omega_c = 0.25$ (left), and $\omega_p/\omega_c = 2$ (right). Peaks (and thus also magnetization) have disappeared at the right. Here $\Gamma = 1/64$.

Appendix A

Technical tools for Chapter 3

A.1 Estimates for the construction of the adiabatic invariant

Here we intend to prove Lemma 1 of Section 3.3. In order to do that, we need to recollect the usual algebraic properties used in perturbation theory (see [22]), adapted to our norm $\| \cdot \|_+$, defined by (3.15). Such properties are stated in Lemmas 9–11 later on, then the proof of Lemma 1 is briefly sketched.

In order to develop the perturbation theory, a primary role is played by the action of the operator L_0 and by the projections on its kernel and its range (see Section 3.3), and these are more easily discussed in terms of the complex variables which diagonalize L_0. These are implicitly defined by

$$q_l = \frac{1}{\sqrt{2}}(\xi_l + i\eta_l) \ , \quad p_l = \frac{1}{\sqrt{2}}(\xi_l - i\eta_l) \ , \quad 1 \leq l \leq N \tag{A.1}$$

and in such variables one has

$$L_0 \xi^j \eta^k = i\omega(|k| - |j|)\xi^j \eta^k \ . \tag{A.2}$$

We must, however, take into account the fact that the norm $\| \cdot \|_+$ is not invariant under such a change of coordinates. In fact, such a norm is formally well defined also for polynomials depending on the variables (ξ, η) if, in the definition of $\mathcal{H}^{r,i}_s$ and $\mathcal{P}_{s,r}$, we simply substitute for (p, q) the pair (ξ, η). In that case, denoting by f' the transform of f via (A.1), one will have, in general, $\|f\|_+ \neq \|f'\|_+$. On the other hand, the following lemma, whose proof is identical to that of Lemma A.1 of paper [22], enables one to estimate the difference between the norms of the two functions.

Lemma 9 *Let $f(q,p)$ be in $\mathcal{P}_{s,r}$ and let $f'(\xi,\eta)$ be the transform of f via (A.1). Then, one has $f' \in \mathcal{P}_{s,r}$ and $\|f'\|_+ \leq 2^{\frac{s}{2}} \|f\|_+$. Moreover, let $g'(\xi,\eta)$ be in $\mathcal{P}_{s,r}$ and let $g(q,p)$ be the transform of g via the inverse of (A.1). Then, one has $g \in \mathcal{P}_{s,r}$ and $\|g\|_+ \leq 2^{\frac{s}{2}} \|g'\|_+$.*

We need also the following lemmas

Lemma 10 *Let f be in $\mathcal{P}_{s,r}$ and g in $\mathcal{P}_{s',r'}$. Then, $[f,g] \in \mathcal{P}_{s+s'-2,r+r'}$ and one has, both in real and in complex variables, the inequality*

$$\|[f,g]\|_+ \leq (2r+2r'+1)ss' \|f\|_+ \|g\|_+ \ .$$

Proof. See Lemma A.2 of [22], noticing that, for any fixed i, each term of f contained in $\mathcal{H}_s^{r,i}$ has Poisson bracket different from 0 only with the monomials of $\mathcal{H}_{s'}^{r',k}$ such that $|i-k| \leq r+r'$. The number of such monomials appearing in the decomposition of g is smaller than $2r+2r'+1$.

Q.E.D.

Lemma 11 *Let $f \in \mathcal{P}_{s,r}$ be a polynomial in complex variables. Then $\Pi_{\mathcal{N}} f$, $\Pi_{\mathcal{R}} f$ and $L_0^{-1}\Pi_{\mathcal{R}} f$ belong to $\mathcal{P}_{s,r}$ and the following inequalities hold:*

$$\|\Pi_{\mathcal{N}} f\|_+ \leq \|f\|_+ \ , \quad \|\Pi_{\mathcal{R}} f\|_+ \leq \|f\|_+ \ , \quad \left\|L_0^{-1}\Pi_{\mathcal{R}} f\right\|_+ \leq \|f\|_+ \ .$$

Proof. The fact that $L_0^{-1} f$ belongs to $\mathcal{P}_{s,r}$ comes directly from Lemma 10, as H_0 is in $\mathcal{P}_{2,0}$. The remaining statements are a consequence of the fact that L_0 is diagonal in complex coordinates and that the smallest eigenvalue of L_0 on $\mathcal{R}$ has modulus $\omega \geq 1$, in virtue of (A.2).

Q.E.D.

Proof of Lemma 1 We pass to complex variables via Lemma 9 and proceed by induction on n, checking at each step even two supplementary inductive hypotheses:

i) Ψ_n can be decomposed as $\Psi_n = \sum_{l=0}^{n} \Psi_n^{(l)}$, where $\Psi_n^{(l)} \in \mathcal{P}_{2l+2,n-l}$;

ii) the following bound holds

$$\left\|\Psi_n^{(l)}\right\|_+ \leq 2^n 2^{10(n-1)} (n!)^2 (n-1)! \frac{n!}{l!(n-l)!} \varepsilon^{n-l} \ .$$

As a matter of facts, on account of Lemma 11, such an estimate enables one to control the contributions due to χ_n and Θ_n, which appear in the recurrent procedure that determines χ_s, for $s \geq n$. Then, we come back to real variables via lemma 9 again.

Q.E.D.

A.2 Technical proofs

A.2.1 Proof of Lemma 5

We start by proving formula (3.35). On account of the symmetry of the periodic system, one can pass from a system with $N-1$ particles to one with N by inserting one more particle after the i–th site, for $i = 1, \ldots, N-1$. The potential energy of the corresponding system is given by

$$U_N(q_1, \ldots, q_N, q) = U_{N-1} - \frac{\varepsilon}{2\omega}\,(q_{i+1} - q_i)^2 + \frac{\varepsilon}{2\omega}\,(q - q_i)^2 + \frac{\varepsilon}{2\omega}\,(q - q_{i+1})^2 + \frac{q^2}{2\omega} + \frac{q^4}{4\omega^2}\,.$$

Neglecting the second term at the r.h.s. (which gives a contribution to the partition function which can be bounded from below by 1, and averaging over i in order to get a translational invariant system, one gets

$$\begin{aligned}\frac{Q_N}{Q_{N-1}} \geq & \frac{1}{N-1}\sum_{i=1}^{N-1}\int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_{N-1}\, \tilde{D}_{N-1}(q_1, \ldots, q_{N-1}) \times \\ & \times \int_{-\infty}^{+\infty} \mathrm{d}q \exp\left[-\frac{\beta}{2\omega}\left(q^2 + \frac{q^4}{2\omega} + \varepsilon(q - q_i)^2 + \varepsilon(q - q_{i+1})^2\right)\right]\,.\end{aligned} \tag{A.3}$$

Here we have put $q_N = q_1$. Then, we introduce the function $\varphi_{q_i}(q) \stackrel{\text{def}}{=} 1 - \exp\left[-\beta\varepsilon(q - q_i)^2/(2\omega)\right]$, for which the inequality

$$\exp\left[-\frac{\beta\varepsilon}{2\omega}(q - q_i)^2 + \frac{\beta\varepsilon}{2\omega}(q - q_{i+1})^2\right] \geq 1 - \varphi_{q_i}(q) - \varphi_{q_{i+1}}(q)$$

holds. We will show now that $\varphi_{q_i}(q)$ is small except for a set of small measure. Making use of the previous inequality, relation (A.3) becomes

$$\begin{aligned}\frac{Q_N}{Q_{N-1}} \geq & \, a(\beta, \varepsilon) - \int_{-\infty}^{+\infty} \mathrm{d}q_1 \ldots \int_{-\infty}^{+\infty} \mathrm{d}q_{N-1}\, \tilde{D}_{N-1}(q_1, \ldots, q_{N-1}) \times \\ & \times \int_{-\infty}^{+\infty} \mathrm{d}q\, \frac{2}{N-1}\sum_{i=1}^{N-1} \varphi_{q_i}(q) \exp\left[-\frac{\beta}{2\omega}\left(q^2 + \frac{q^4}{2\omega}\right)\right]\,,\end{aligned} \tag{A.4}$$

in which the function $a(\beta,\varepsilon)$ is defined by[1]

$$a(\beta,\varepsilon) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \mathrm{d}q \exp\left[-\frac{\beta}{2\omega}\left(q^2+\frac{q^4}{2\omega}\right)\right] = \frac{\sqrt{2\omega}\, e^{\frac{\beta}{8}}}{2} K_{\frac{1}{4}}\left(\frac{\beta}{8}\right) \ ,$$

where $K_\alpha(x)$ is the Bessel modified function of second kind. The well known properties of $K_\alpha(x)$ imply that $a(\beta,\varepsilon)$ can be written as $a(\beta,\varepsilon) = G(\beta,\varepsilon)\sqrt{2\pi\omega/\beta}$, where G is a function always smaller than 1, approaching 1, at fixed ε, as $\beta \to +\infty$. We go on by dealing with the integral in (A.4), first giving an upper bound for the innermost integral over q. We estimate it by splitting the phase space of the $N-1$ particles periodic system in two sets: we will fix $\kappa > 0$ and consider $\Omega(N-1,\kappa)$, which is defined by

$$\Omega(N-1,\kappa) \stackrel{\text{def}}{=} \left\{(q_1,\ldots,q_{N-1}) \text{ such that } \sum_{i=1}^{N-1} q_i^2 < \frac{2\omega}{\beta}\kappa(N-1)\right\} \ , \quad \text{(A.5)}$$

and its complement. In the latter set, the integral is simply bounded from above by $2a(\beta,\varepsilon)$. On the other hand, in order to estimate the integral in the set $\Omega(N-1,\kappa)$, we observe that, for any κ_1, the number of particles for which $|q_i| \geq \sqrt{\kappa_1\kappa 2\omega/\beta}$ holds cannot exceed $(N-1)/\kappa_1$. For these particles the integral is estimated again by $2a(\beta,\varepsilon)$. For the purpose of estimating the contribution of the remaining particles, we introduce the function

$$I(\beta,\varepsilon,\kappa,\kappa_1) \stackrel{\text{def}}{=} \frac{1}{a(\beta,\varepsilon)} \sup_{|y|<\sqrt{\kappa_1\kappa 2\omega/\beta}} \int_{-\infty}^{+\infty} \varphi_y(q) \exp\left(-\frac{\beta}{2\omega}q^2\right) \mathrm{d}q \ .$$

We point out that $I(\beta,\varepsilon,\kappa,\kappa_1)$ tends to 0 as ε tends to 0, for β,κ,κ_1 fixed. Then, in the region $\Omega(N-1,\kappa)$, for any $\kappa_1 > 1$, one has the bound

$$\int_{-\infty}^{+\infty} \mathrm{d}q \, \frac{2}{N-1} \sum_{i=1}^{N-1} \varphi_{q_i}(q) \exp\left[-\frac{\beta}{2\omega}\left(q^2+\frac{q^4}{2\omega}\right)\right] \leq \left[\frac{2}{\kappa_1} + 2I(\beta,\varepsilon,\kappa,\kappa_1)\right] a(\beta,\varepsilon) \ .$$

We notice that we have provided estimates independent of q_i, so the integrals over $q_1,\ldots,q_{N-1}$ appearing in (A.4) can simply be estimated as the product of these upper bounds times the measures of the sets in which the bounds hold. Now, we observe that the measure of $\Omega^c(N-1,\kappa)$ is estimated by

$$\int_{\Omega^c(N-1,\kappa)} \mathrm{d}q_1 \ldots \mathrm{d}q_{N-1} \, \tilde{D}_{N-1}(q_1,\ldots,q_{N-1}) \leq \frac{R_{N-1}(\beta,\kappa)}{Q_{N-1}}$$

[1]Remark that the function $a(\beta,\varepsilon)$ depends on ε only via the term $\omega = \sqrt{1+2\varepsilon}$.

where the function $R_{N-1}(\beta,\kappa)$ is defined by

$$\begin{aligned} R_{N-1}(\beta,\kappa) &\stackrel{\text{def}}{=} \int_{\Omega^c(N-1,\kappa)} \mathrm{d}q_1 \dots \mathrm{d}q_{N-1} \, \exp\left(-\frac{\beta}{2\omega}\sum_{i=1}^{N-1} q_i^2\right) \\ &= \left(\frac{2\pi\omega}{\beta}\right)^{\frac{N-1}{2}} \Gamma\left(\frac{N-1}{2}, \kappa(N-1)\right) \end{aligned} \tag{A.6}$$

and $\Gamma(s,x)$ is defined by (3.43). This way one obtains, finally,

$$\frac{Q_N}{Q_{N-1}} \geq \left(1 - \frac{2}{\kappa_1} - 2I(\beta,\varepsilon,\kappa,\kappa_1) - 2\frac{R_{N-1}(\beta,\kappa)}{Q_{N-1}}\right) a(\beta,\varepsilon) \ . \tag{A.7}$$

¿From this expression one can prove (3.35) by induction on N.

We now come to the proof of (3.36). We make use of the trivial inequality $\mathbf{P}(A_1 \cap \ldots \cap A_r) \geq 1 - \sum_{i=1}^r \mathbf{P}(A_i^c)$, which holds for any probability and any collection of sets $A_1, \ldots, A_r$. Consequently, we obtain

$$\mathbf{P}_N\left(|q_1| < \Theta\sqrt{\frac{2\omega}{\beta}} \ \wedge \ \ldots \ \wedge \ |q_r| < \Theta\sqrt{\frac{2\omega}{\beta}}\right) \geq 1 - r \cdot \mathbf{P}_N\left(|q_1| \geq \Theta\sqrt{\frac{2\omega}{\beta}}\right) \ .$$

because, due to the translation invariance of the periodic system, every set has the same measure. Recall that $\mathbf{P}_N\Big(|q_1| \geq \Theta\sqrt{2\omega/\beta}\Big)$ is just the integral of $\tilde{D}_N$ times $\mathbf{1}_{|q_i|\geq\Theta\sqrt{2\omega/\beta}}$. A bound to this integral can be found proceeding as above, i.e., by symmetrizing on q_i, fixing $\kappa > 0$ and integrating separately over $\Omega(N,\kappa)$ and its complement (recall that $\Omega(N,\kappa)$ is defined by (A.5)). This way we get

$$\begin{aligned} \mathbf{P}_N\left(|q_1| \geq \Theta\sqrt{\frac{2\omega}{\beta}}\right) \ \leq \ & \frac{1}{NQ_N}\sum_{i=1}^N \int_{\Omega(N,\kappa)} \mathbf{1}_{|q_i|\geq\Theta\sqrt{2\omega/\beta}} \tilde{D}_N(q_1,\ldots,q_N) \\ & + \frac{1}{Q_N} R_N(\beta,\kappa) \ , \end{aligned}$$

where R_N is defined by (A.6), and we bound $\mathbf{1}_{|q_i|\geq\Theta\sqrt{2\omega/\beta}}$ by 1 in $\Omega^c(N,\kappa)$. It is straightforward to notice that the number of sites for which $|q_i| \geq \Theta\sqrt{2\omega/\beta}$, in the interior of $\Omega(N,\kappa)$, cannot exceed $N\kappa/\Theta^2$. Therefore, the former term at the r.h.s. of the previous formula is smaller than $1/4r$ if $\Theta \geq 2\sqrt{\kappa r}$. As far as the latter is concerned, we can choose κ such that $R_N(\beta,\kappa)/Q_N \leq 1/4r$, as we have shown above. For example, we can fix $\kappa = \log(4rK_0)$. This suffices to infer that, for $\Theta \geq 2\sqrt{r\log(4rK_0)}$, (3.36) is valid.

A.2.2 Proof of Lemma 6

The first inequality in (3.40) comes directly from the fact that the integrand appearing in the definition of Q_M is smaller than the function $n_{M,\mathfrak{x}}$, i.e., the integrand in the definition of $\bar{\mathcal{Q}}^{\mathfrak{x}}_M$.

As regards the second inequality in (3.40), we note that the integrand of Q_M is equal to the one of $\bar{Z}^{\mathfrak{x}}_M$ multiplied by $\mathfrak{x}$ terms of the form $\exp(-\beta\varepsilon q_{m_i} q_{m_{i+1}}/\omega)$ at the sites, in number $2\mathfrak{x}$, on the boundary of the blocks, which we denote by $m_1, \dots, m_{2\mathfrak{x}}$, with the convention that $m_{2\mathfrak{x}+1} = m_1$. Then, we integrate only in the region in which the q coordinate of each of these sites is smaller than $\Theta\sqrt{2\omega/\beta}$, with $\Theta = 2\sqrt{2\mathfrak{x}\log(8\mathfrak{x}K_0)}$, and we observe that

$$\begin{aligned}\bar{Z}^{\mathfrak{x}}_M &\geq \exp\left(-4\mathfrak{x}\varepsilon\Theta^2\right) Q_M \mathbf{P}_M \left(|q_{m_1}| < \Theta\sqrt{2\omega/\beta} \wedge \ldots \wedge |q_{m_{2\mathfrak{x}}}| < \Theta\sqrt{2\omega/\beta}\right) \\ &\geq \frac{Q_M}{2}\left(8\mathfrak{x}K_0\right)^{-32\varepsilon_0\mathfrak{x}^2} \ .\end{aligned}$$

Here, Q_M comes from the normalization of the probability, and in the second line use is made of Lemma 5.

We come now to inequalities (3.41) and observe at once that the first one is trivial, because, on account of the identity in (3.29), one has $\tilde{n}_{M,\mathfrak{x}} \leq n_{M,\mathfrak{x}}$. The second one is more complicated: we begin by proving it in the case in which each block is constituted by an even number of elements.

In order to estimate $\mathcal{Q}^{\mathfrak{x}}_M$, we divide again the phase space of the system in the region $\tilde{\Omega}$ in which $|q_{m_1}| < \sqrt{2\omega\kappa/\beta}, \dots, |q_{m_{2\mathfrak{x}}}| < \sqrt{2\omega\kappa/\beta}$, where $\kappa > 0$ is a constant to be determined, and in its complement $\tilde{\Omega}^c$. The integral over $\tilde{\Omega}$ is smaller than $Q_M \cdot \exp\left(4\mathfrak{x}\varepsilon\kappa\right)$, while, as regards the complement, we notice that it is contained in the set in which $\sum_{i=1}^{2\mathfrak{x}} q_i^2 \geq 2\kappa\omega/\beta$. Thus, the integral over such a region is bounded from above by $\mathcal{Q}^{\mathfrak{x}_1}_{M-2\mathfrak{x}}\left(2\pi\omega/\beta\right)^{\mathfrak{x}}\Gamma(\mathfrak{x},\kappa)$, with $\mathfrak{x}_1 \leq \mathfrak{x}$, where we have dropped some positive term in the potentials, then we have integrated first over $q_{m_1}, \dots, q_{m_{2\mathfrak{x}}}$ (which gives the term $(2\pi\omega/\beta)^{\mathfrak{x}}\Gamma(\mathfrak{x},\kappa)$, with $\Gamma(\mathfrak{x},\kappa)$ defined by (3.43)); then the blocks made of just 2 particles disappear, so that the integration over the remaining positions gives the term $\mathcal{Q}^{\mathfrak{x}_1}_{M-2\mathfrak{x}}$. This way we get

$$\mathcal{Q}^{\mathfrak{x}}_M \leq Q_M \cdot \exp\left(4\mathfrak{x}\varepsilon\kappa\right) + \left(\frac{2\pi\omega}{\beta}\right)^{\mathfrak{x}} \Gamma(\mathfrak{x},\kappa) \cdot \mathcal{Q}^{\mathfrak{x}_1}_{M-2\mathfrak{x}} \ , \quad \text{with} \quad \mathfrak{x}_1 \leq \mathfrak{x} \ .$$

Now, we apply the previous inequality to the function $\mathcal{Q}^{\mathfrak{x}_1}_{M-2\mathfrak{x}}$ at the r.h.s, and we end up with a relation similar to the previous one, in which however there appears the function $\mathcal{Q}^{\mathfrak{x}_2}_{M-2\mathfrak{x}-2\mathfrak{x}_1}$, with $\mathfrak{x}_2 \leq \mathfrak{x}_1$. So, we can iterate this

procedure, observing that $\Gamma(\mathfrak{x},\kappa)$ is an increasing function of $\mathfrak{x}$, and we get

$$\mathcal{Q}_M^{\mathfrak{r}} \leq \exp\left(4\mathfrak{x}\varepsilon\kappa\right) \sum_{j=0}^{J} \left(\frac{2\pi\omega}{\beta}\right)^{\sigma_j} Q_{M-2\sigma_j} \left(\Gamma(\mathfrak{x},\kappa)\right)^j \ , \quad \text{with} \quad Q_0 \stackrel{\text{def}}{=} 1 \ ,$$

where we define $\sigma_j = \sum_{k=0}^{j} \mathfrak{x}_k$, with $\mathfrak{x}_0 = \mathfrak{x}$, and J represents the integer such that $\sigma_J = M/2$. We make use of inequality (3.35) of Lemma 5 and finally get, if the series converges, $\mathcal{Q}_M^{\mathfrak{r}} \leq \exp\left(4\mathfrak{x}\varepsilon\kappa\right) Q_M \sum_{j=0}^{\infty} (K_0^{2\mathfrak{x}} \Gamma(\mathfrak{x},\kappa))^j$. We point out that the common ratio of this geometric series is a decreasing function of κ, which tends to 0 as $\kappa \to +\infty$: thus, we choose $\bar{\kappa} = \bar{\kappa}(\mathfrak{x}, K_0)$ so as to satisfy (3.42), and obtain the relation

$$\mathcal{Q}_M^{\mathfrak{r}} \leq 2 \exp\left(4\mathfrak{x}\varepsilon\bar{\kappa}(\mathfrak{x},K_0)\right) Q_M \ . \tag{A.8}$$

If a number $\lambda \leq \mathfrak{x}$ of blocks is constituted by an odd number of elements, we integrate on one of the sites on the boundary of each of these blocks, in order that each of the blocks, in number $\mathfrak{x}'$, of the resulting lattice contains an even number of elements. By dropping some suitably chosen interaction terms in the potential, one gets $\mathcal{Q}_M^{\mathfrak{r}} \leq (2\pi\omega/\beta)^{\lambda/2} \mathcal{Q}_{M-\lambda}^{\mathfrak{r}'}$, with $\mathfrak{x}' \leq \mathfrak{x}$, where the blocks made of just one particles disappear. Now we can use (A.8) with $Q_{M-\lambda}$ instead of Q_M. Then, making use of Lemma 5 to express $Q_{M-\lambda}$ in terms of Q_M, we get (3.41).

A.2.3 Proof of Theorem 4

As already said, the proof is performed by bounding from above every term at the r.h.s. of (3.55), i.e., $\lambda(\mathbf{j},\mathbf{i})$, for $\mathbf{j},\mathbf{i} \in V$.

We point out that the expectations in the definition of $\lambda(\mathbf{j},\mathbf{i})$ depend on the choice of Q, which is not completely fixed by its marginal probabilities (see comments on relations (3.52)). In fact, the main part of paper [26] consists in introducing a suitable reconstruction operator (on the space of the joint probabilities) which enables one to find a joint probability distribution that minimizes λ, starting from an initially chosen one. We adopt the same technique, with the only difference that we apply it not at all sites of the lattice, but only at those lying on the complement of a fixed set $\bar{V}$ (we will call it $\bar{T} \stackrel{\text{def}}{=} T \backslash \bar{V}$).

We also need to control, together with $\lambda(\mathbf{j},\mathbf{i})$, the auxiliary quantity

$$\gamma(\mathbf{i}) \stackrel{\text{def}}{=} \mathbf{E}\left[\mathbf{1}_{\xi_{\mathbf{i}}^1 \neq \xi_{\mathbf{i}}^2}\right] \ ,$$

where ξ^1 and ξ^2 are the same Gibbsian fields entering in (3.56).

We introduce, then, the main tool of the proof, i.e., the reconstruction operator $U_{\mathbf{i}}$, with $\mathbf{i} \in T$, which will enable us to construct the joint probabilities $Q(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y})$ on V (see formula (3.55)). This operator is defined on a couple of fields (ξ^1, ξ^2) having the same conditional probability at $\mathbf{i}$, as follows. For each pair of configurations $\mathbf{x}^1, \mathbf{x}^2 \in \mathbb{R}^{|T|}$, we denote by $P^{\mathbf{i}}_{\mathbf{x}^1,\mathbf{x}^2}$ the measure on $\mathbb{R}^2$ for which the minimum of the distance between $P_{\mathbf{i},\mathbf{x}^1}$ and $P_{\mathbf{i},\mathbf{x}^2}$ is attained, i.e., such that, for any measurable $B \subset \mathbb{R}$, one has

$$P^{\mathbf{i}}_{\mathbf{x}^1,\mathbf{x}^2}(\mathbb{R} \times B) = P_{\mathbf{i},\mathbf{x}^1}(B) \ , \quad P^{\mathbf{i}}_{\mathbf{x}^1,\mathbf{x}^2}(B \times \mathbb{R}) = P_{\mathbf{i},\mathbf{x}^2}(B) \ ,$$
$$\text{and} \int_{\mathbb{R}^2} \mathbf{1}_{x \neq y} P^{\mathbf{i}}_{\mathbf{x}^1,\mathbf{x}^2}(\mathrm{d}x, \mathrm{d}y) = D\left(P_{\mathbf{i},\mathbf{x}^1}, P_{\mathbf{i},\mathbf{x}^2}\right) \ .$$

Such a definition enables us to describe the action of $U_{\mathbf{i}}$, because this operator maps the couple (ξ^1, ξ^2) into $(\hat{\xi}^1, \hat{\xi}^2)$ such that, for any measurable $C \subset \mathbb{R}^2$,

$$P\left(\left(\hat{\xi}^1_{\mathbf{i}}, \hat{\xi}^2_{\mathbf{i}}\right) \in C \,|\, \hat{\xi}^1_{T\setminus\{\mathbf{i}\}} = \mathbf{x}^1_{T\setminus\{\mathbf{i}\}}, \hat{\xi}^2_{T\setminus\{\mathbf{i}\}} = \mathbf{x}^2_{T\setminus\{\mathbf{i}\}}\right) = P^{\mathbf{i}}_{\mathbf{x}^1,\mathbf{x}^2}(C) \ ,$$

and, for any finite $V \subset T$ not containing $\mathbf{i}$, the joint probability distribution of $(\hat{\xi}^1_V, \hat{\xi}^2_V)$ coincides with that of (ξ^1_V, ξ^2_V).

The effect of $U_{\mathbf{i}}$ on $\gamma(\mathbf{i})$ and $\lambda(u, \mathbf{i})$ is described in detail in Lemmas 2, 3 and 4 of the work [26]. Following such a paper, we adopt the convention that the quantities relative to the reconstructed couple $(\hat{\xi}^1, \hat{\xi}^2)$ are distinguished from the corresponding ones relative to (ξ^1, ξ^2) by adding the symbol $\hat{}$. For every set $S \subset T$ we define the operator

$$U_S \stackrel{\text{def}}{=} U_{\mathbf{i}_1} \circ U_{\mathbf{i}_2} \circ \ldots \circ U_{\mathbf{i}_m} \ , \tag{A.9}$$

where the order of the points $\mathbf{i}_1, \ldots, \mathbf{i}_m$, contained in $S \cup \partial_{br} S$ is chosen in a suitable way. This is described in full detail in the proof of the following Lemma 12. If we define $\gamma_S \stackrel{\text{def}}{=} \sup_{\mathbf{i} \in S} \gamma(\mathbf{i})$ and $\lambda_S \stackrel{\text{def}}{=} \sup_{\mathbf{j},\mathbf{i} \in S} \lambda(\mathbf{j}, \mathbf{i})$, we can describe the action of U_S on a couple of fields having the same conditional probability on $S \cup \partial_{br} S$ accordingly to the following lemma, which is proved in Appendix A.2.4.

Lemma 12 *Let (ξ^1, ξ^2) be a couple of fields having the same conditional probability on $S \cup \partial_{br} S$, given by a specification $\Gamma \in \Theta(h, C, \delta) \cap \Delta(h, \bar{K}C, \alpha)$, and $(\hat{\xi}^1, \hat{\xi}^2) = U_S(\xi^1, \xi^2)$. Then, one has*

$$\begin{pmatrix} \hat{\gamma}_S \\ \hat{\lambda}_S \end{pmatrix} \leq A \begin{pmatrix} \gamma_{S \cup \partial_{br} S} \\ \lambda_{S \cup \partial_{br} S} \end{pmatrix} \ ,$$

in which the matrix A is defined by

$$A \stackrel{\text{def}}{=} \begin{pmatrix} \alpha + N\bar{K}^{-1} & C^{-1} M \bar{K}^{-1} \\ C\left(R + N\bar{K}^{-1}\right) & \delta a^b + M\bar{K}^{-1} \end{pmatrix} \ ,$$

where N, M and R are constants depending on a and b only.

We remark that, if the eigenvalues of A are smaller than 1, the reconstructed quantities are smaller than the initial ones. So, we want to iterate the reconstruction procedure as much as possible. It turns out that we can iterate the procedure at most a number of times proportional to the distance between V and $\tilde{V}$. The reason is the following.

In our case, ξ^1 is the field relative to the equilibrium Gibbs measure and ξ^2 that relative to the probability conditioned to the configuration $\mathbf{x}$ on $\tilde{V}$, which we consider as fixed. It is apparent that such fields have the same conditional probability on every set which does not intersect $\tilde{V}$, but not on the whole T; by hypothesis, this conditional probability is that given by $\Gamma \in \Theta(h, C, \delta) \cap \Delta(h, \bar{K}C, \alpha)$. Since the reconstuction procedure shrinks the set S on which we can control γ and λ, we can iterate it until $V \subset S$. So, the maximum number of iterations is attained if we start by reconstructing on $V \cup \partial_{nbr} V$, where n is the largest number such that $\partial_{(n+1)br} V \cap \tilde{V} = \emptyset$. We use Lemma 12 as the first step of a recurrent scheme, by applying each time U_{V_m}, where $V_{m+1} = V_m \cup \partial_{br} \tilde{V}_m$, $V_0 = V$. In virtue of Lemma 12, after the application of U_{V_m}, one has

$$\begin{pmatrix} \hat{\gamma}_{V_m} \\ \hat{\lambda}_{V_m} \end{pmatrix} \leq A \begin{pmatrix} \gamma_{V_{m+1}} \\ \lambda_{V_{m+1}} \end{pmatrix} .$$

Thus we get that the final values of γ_V and λ_V are smaller than the result of the application of the matrix A^n to the vector with components γ, λ. Moreover, we observe that we can write $A = J^{-1}\tilde{A}J$, where $\tilde{A}$ and J are defined by

$$\tilde{A} \stackrel{\text{def}}{=} \begin{pmatrix} \alpha + N\bar{K}^{-1} & M\bar{K}^{-1} \\ R + N\bar{K}^{-1} & \delta a^b + M\bar{K}^{-1} \end{pmatrix} \quad \text{and } J \stackrel{\text{def}}{=} \begin{pmatrix} C & 0 \\ 0 & 1 \end{pmatrix} .$$

This way we get $A^n = J^{-1}\tilde{A}^n J$. As the component λ_V, which is the one we are interested in, is not affected by the action of J^{-1}, we can write that it is the second component of the matrix product

$$\tilde{A}^n \begin{pmatrix} C\gamma \\ \lambda \end{pmatrix} .$$

Since the eigenvalues of $\tilde{A}$ are smaller than $G \stackrel{\text{def}}{=} \max\{(\alpha+1)/2, (\delta a^b+1)/2\} < 1$, if $\bar{K}$ is large enough, there exists $\bar{K}_0$ such that

$$\lambda_V \leq (D/2) \max\{\lambda, C\gamma\} G^n ,$$

where D is a constant depending on a, b, α, δ only. On the other hand, $n = d(V, \tilde{V})/(br)$ and $\gamma \leq 1$, from which there follows

$$\lambda_V \leq \frac{D}{2} \langle h \rangle_{\mathbf{x}} \exp\left(-c\, d(V, \tilde{V})\right) ,$$

where c is defined in (3.59).

In order to show (3.58), we need only the use of (3.55) in estimating the term in brackets of (3.50). We then observe that the r.h.s. of (3.55) is smaller than $2|V|^2\lambda_V$, for the joint probability we have just found, and this concludes the proof.

A.2.4 Proof of Lemma 12

Lemma 5 of work [26] shows the result of the application of $U_{\mathbf{i}}$, in a suitably chosen order, to every site of T in sequence: one obtains that, for the couple of fields (ξ^1, ξ^2), with the same specification $\Gamma \in \Theta(h, C, \delta) \cap \Delta(h, \bar{K}C, \alpha)$ and $\bar{K} \geq 1$, and for the reconstructed couple $(\hat{\xi}^1, \hat{\xi}^2)$ the following matrix relation holds[2]

$$\begin{pmatrix} \hat{\gamma}_T \\ \hat{\lambda}_T \end{pmatrix} \leq A \begin{pmatrix} \gamma_T \\ \lambda_T \end{pmatrix} . \tag{A.10}$$

In the proof of Lemma 5 of [26], the order of the $U_{\mathbf{i}}$'s is chosen in the following way: the lattice is partitioned in b disjoint sublattice, $Z_0, \ldots, Z_{b-1}$, which are the cosets in T of $\mathbb{Z}^\nu$ with respect to Z_0. Then the reconstruction is applied in sequence to each sublattice, and use is made of the fact that, if $\mathbf{i} \in Z_l$, there exists a bound to $\gamma(\mathbf{j})$, $\lambda(\mathbf{j}, \mathbf{k})$ and $\lambda(\mathbf{k}, \mathbf{j})$ for $\mathbf{j} \in Z_l \backslash \{\mathbf{i}\}$ and $\mathbf{k} \in T \backslash \{\mathbf{i}\}$ which does not change after the application of $U_{\mathbf{i}}$. In particular, this implies that the bounds do not change for the already reconstructed sites in Z_l. Neither does the reconstruction at site $\mathbf{i}$ change, on account of Lemma 4 of paper [26], the value of $\lambda(\mathbf{i}, \mathbf{j})$ and $\lambda(\mathbf{j}, \mathbf{i})$, for $|\mathbf{j} - \mathbf{i}| > r$. In this sense the reconstruction is local.

So, the values of $\hat{\gamma}_V$ and $\hat{\lambda}_V$, after one application of $U_{\mathbf{i}}$, depend at most on the values of $\gamma(\mathbf{i})$ and $\lambda(\mathbf{j}, \mathbf{i})$ in $V \cup \partial_r V$. It is thus apparent that we can control the values of the reconstructed quantities only in a set V smaller than the set V' on which we control γ and λ initially. In particular, V can be chosen so that $V' = V \cup \partial_r V$. Therefore, for any $V \subset S$, we define for

[2] As a matter of fact, A is not the same matrix which appears in [26], since we needed to make the dependence on C explicit. Our statement can be proved by checking, in proving the induction (11)-(14) of [26], that the constants $N(\cdot, \cdot)$ are proportional to C^2, the constants $N(\cdot)$ and $M(\cdot, \cdot)$ are proportional to C and the constants $M(\cdot)$ are independent of C.

$l = 0, \ldots, b-1$ a nested sequence of sets $V_{l+1} \stackrel{\text{def}}{=} V_l \cup \partial_r V_l$, with $V_0 \stackrel{\text{def}}{=} V$ and the operator U_V, as $U_V = U_{V_0 \cap Z_0} \circ \ldots \circ U_{V_{b-1} \cap Z_{b-1}}$, and notice that (see the above remarks) the order in which the sites in $V_l \cap Z_l$ are chosen does not matter. Then, after the application of U_V, one has that

$$\begin{pmatrix} \hat{\gamma}_V \\ \hat{\lambda}_V \end{pmatrix} \leq A \begin{pmatrix} \gamma_{V_b} \\ \lambda_{V_b} \end{pmatrix} ,$$

for the same matrix A appearing in (A.10). This concludes the proof.

Appendix B

Technical tools for chapter 4

B.1 Conditions on the regularity of the measure which solves the moment problem

The problem of determining a necessary and sufficient condition in order that the solution of an Hamburger moment problem admit a bounded positive derivative has been solved by Akhiezer and Krein (see [63]). Unfortunately, their criterion cannot be applied to study the derivative itself. Thus, we report the criterion of Akhiezer and Krein for checking the existence of a derivative for the Hamburger moment problem for the measure α, then we give an alternative procedure for studying the derivatives of $\sqrt{u}\Phi'(u)$,[1] at least in the case in which the time autocorrelation is analytic on the whole real axis. We convert the problem into an Hausdorff moment problem (i.e., a moment problem on a finite interval), in order to get condition on the regularity of the successive derivatives.

To state the result of [63] of interest to us, let us introduce the definition of a non-negative definite sequence of moments c_n, by requiring that

$$\sum_{i,j=1}^{n} c_{i+j} X_i \bar{X}_j \geq 0 \ ,$$

for any n and any sequence $(X_1, \ldots, X_n)$ of complex numbers.[2] Then one has

Proposition 1 *In order that $\alpha'(\omega)$ exist and be bounded, it is necessary and sufficient that there exist $L > 0$ such that the sequence $t_0(L), t_1(L), \ldots$ defined*

[1] Recall that $\alpha'(\omega)$ is analytic if and only if $\sqrt{u}\Phi'(u)$ is (see Corollary 2).

[2] Notice that this implies that the determinants of Δ_n defined in (4.14) are all non-negative, but the converse is not true.

by[3]

$$t_k(L) = \frac{1}{(k+1)!L^{k+1}} \begin{vmatrix} c'_0 & -L & 0 & \cdots & 0 & 0 \\ 2c'_1 & c'_0 & -2L & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ kc'_{k-1} & (k-1)c'_{k-2} & (k-2)c'_{k-3} & \dots & c'_0 & -kL \\ (k+1)c'_k & kc'_{k-1} & (k-1)c_{k-2} & \dots & 2c'_1 & c'_0 \end{vmatrix} ,$$

is non-negative definite.

For the study of the successive derivatives, we observe that in paper [64] Hausdorff found a solution method for the moment problem on the interval $[0,1]$, which also allows to give a criterion to check whether, at any order, the derivatives of the measure exist and are bounded. Such results can be also applied to our problem if the time autocorrelation is analytic on the whole real axis, because, through an analytic change of variable, one can map $[0,1]$ on $\mathbb{R}^+$ and obtain the next Proposition 2, which provides necessary and sufficient conditions in order that the n–th derivative of $\sqrt{u}\Phi'(u)$ exist and be bounded. For its statement, we need to define the auxiliary moments μ_k and $\tilde{\mu}_k$: we need to preliminarily define

$$\mathfrak{F}(z) \stackrel{\text{def}}{=} \int_0^{+\infty} \frac{\mathrm{d}\Phi(u)}{z+u} ,$$

then we set

$$\begin{aligned} \mu_0 \stackrel{\text{def}}{=} c_0 \quad \mu_k \stackrel{\text{def}}{=} c_0 - \frac{1}{(k-1)!}\frac{\mathrm{d}^{k-1}}{\mathrm{d}^{k-1}z}\left(z^k\mathfrak{F}(z)\right)_{z=1} , \text{ for } k>0 , \\ \tilde{\mu}_k \stackrel{\text{def}}{=} \sum_{n=0}^{\infty}\binom{k+1/2}{n}\sum_{l=0}^{n}\binom{n}{l}(-)^{n-l}2^{l-k-1/2}\mu_l . \end{aligned} \tag{B.1}$$

These quantities are well defined, since it is known that $\mathfrak{F}(z)$ is analytic in $z=1$ and the binomial expansion converges.[4] Then, we define the quantities

$$\begin{aligned} \lambda^0_{p,m} \stackrel{\text{def}}{=} \binom{p}{m}\sum_{k=0}^{p-m}(-1)^k\binom{p-m}{k}\tilde{\mu}_k , \quad \text{for } 0\le m\le p ; \\ \lambda^k_{p,m} \stackrel{\text{def}}{=} (p+1)\left(\lambda^{k-1}_{p,m} - \lambda^{k-1}_{p,m-1}\right) , \quad \text{for } k>0, k\le m\le p-k , \end{aligned} \tag{B.2}$$

and state the following

[3] Recall that we have defined $c'_{2n} = c_n$ and $c'_{2n+1} = 0$.

[4] See the discussion in the proof below.

Proposition 2 *Let $\mathbf{C}_f(t)$ be an analytic function of t, for $t \in \mathbb{R}$, and let Φ be the measure which solves the related Stieltjes moment problem. A bounded derivative of order n of $\sqrt{u}\Phi'(u)$ exists if and only there exist $L_k > 0$ such that, for any $p > 0$, $0 \leq k < n$ and $k \leq m \leq p - k$, one has*

$$(p+1)\left|\lambda_{p,m}^{k}\right| \leq L_k \ . \tag{B.3}$$

Proof. The proof is performed through the Euler transformation $w \stackrel{\text{def}}{=} \frac{u}{u+1}$. This enables one to define the positive Borel measure $\Psi(w)$ on $[0,1]$, by $\Psi(w) \stackrel{\text{def}}{=} \Phi(w/(1-w))$. The relation between the moments of Ψ on the interval $[0,1]$ and the moments c_n of Φ can be obtained by the following chain of equalities, for $k > 0$,

$$\int_0^1 w^k \mathrm{d}\Psi(w) = \int_0^{+\infty} \frac{u^k}{(u+1)^k} \mathrm{d}\Phi(u) = \sum_{l=0}^{k} (-)^k \binom{k}{l} \int_0^{+\infty} \frac{\mathrm{d}\Phi(u)}{(u+1)^l} \ , \tag{B.4}$$

as can be seen by expanding $u^k = (u+1-u)^k$ via the binomial formula. It is easy to see that this implies that the coefficients μ_k defined by (B.1) are the moments of Ψ.

In order to study the derivatives of $\alpha'(\omega) = 2\sqrt{u}\Phi'(u)$, we pass to the the moments given by the distribution function $\sqrt{w}\Psi'(w)$. This function exists if $C_f(t)$ is analytic on the whole real axis, because in such a case its spectral measure $\alpha(\omega)$ tends exponentially fast towards its limit for $\omega \to \infty$ and, in consequence, $\Psi(w)$ tends towards the same limit as $w \to 1$ at least as fast as $e^{-\kappa/\sqrt{1-w}}$. Moreover, $\sqrt{w}\Psi'(w)$ admits a bounded derivative of order n if and only if $\sqrt{u}\Phi'(u)$ does. We need, thus, to find the moments of $\sqrt{w}\Psi'(w)$, which are given by

$$\int_0^1 w^{k+1/2}\Phi'(w)\mathrm{d}w = \int_0^1 \sum_{n=0}^{\infty} \binom{k+1/2}{n} 2^{k+1/2-n} \left(w - \frac{1}{2}\right)^n \Phi'(w)\mathrm{d}w \ ,$$

in view of the binomial expansion of $\sqrt{1/2 + (w-1/2)}$. Since this expansion converges uniformly for $w \in [0,1]$, one can exchange the order of summation and integration and prove that these moments are indeed the $\tilde{\mu}_k$ defined by (B.1).

As Hausdorff proved, in order that the moment problem for the sequence $\tilde{\mu}_k$ have a solution ψ_0 on $[0,1]$, which is almost everywhere differentiable with a limited derivative (in our case, it is $\psi_1(w) = \sqrt{w}\Psi'(w)$), it is necessary and sufficient that $(p+1)|\lambda_{p,m}^0|$ be bounded, where $\lambda_{p,m}^0$ are defined by equation (B.2). This corresponds to

$$\lambda_{p,m}^0 = \binom{p}{m} \int_0^1 w^m (1-w)^{p-m} \mathrm{d}\psi_0(w) \ .$$

Then, one considers the Hausdorff moment problem in which the measure (not necessarily positive) is ψ_1, and expresses its moments $\tilde{\mu}_k^1$ through the original moments $\tilde{\mu}_k$, by integrating by parts. So one obtains that the coefficients $\lambda_{p,m}^1$ defined by (B.2) play for ψ_1 the same role which the $\lambda_{p,m}^0$ play for ψ_0, because, for $1 \leq m \leq p-1$,[5]

$$\lambda_{p,m}^1 = (p+1)(\lambda_{p,m}^0 - \lambda_{p,m-1}^0) = \binom{p+1}{m} \int_0^1 w^m (1-w)^{p+1-m} \mathrm{d}\psi_0(w) \ .$$

Thus, one can express the condition that ψ_1 admits a limited derivative by asking that $(p+1)|\lambda_{p,m}^1|$ have an upper bound. The proof is then concluded by iterating this way of proceeding.

Q.E.D.

Finally, we add the following Proposition, which can be proved by the easy remark that a power series with positive coefficients, which has a radius of convergence ρ, has a singularity for $z = \rho$ (this is known as Vivanti's theorem).

Proposition 3 *In order that the decay of the time autocorrelation of the dynamical variable f be not exponentially fast it is sufficient that there exists $\rho > 0$ such that* $\limsup_{n\to+\infty} \sqrt[n]{c_n} \leq \rho$.

Remark. The condition of Proposition 3 requires only to have an upper bound on the c_n, whereas those of Propositions 1-2 ask for a more detailed knowledge on the coefficients c_n. We believe that the first one is seldom fulfilled (an example in which it is fulfilled is that of a harmonic oscillator). For, the n–th derivative of a function usually grows as $n!$, so that

$$\limsup_{n\to+\infty} \sqrt[n]{c_n} = +\infty \ .$$

The other requirements enables us to deal with a larger class of variables, but, as just said, it has the drawback of needing a very detailed knowledge of all coefficients.

[5]We neglected here the terms for $m = 0$ and $m = p$, since they give indications only on the discontinuities of ψ_1, which can be disposed of, since ψ_1 can be defined at will on a set of measure zero.

B.2 Computation of the coefficients needed for the Faà di Bruno formula

Let us recall the Faà di Bruno's formula (see [65]). Let a function $f\big(g(x)\big)$ be given, where f and g possess a sufficient number of derivatives. Then one has

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} f\big(g(x)\big) = \sum \frac{n!}{k_1!\, k_2! \cdots k_n!} \cdot f^{(k_1+\cdots+k_n)}\big(g(x)\big) \cdot \prod_{j=1}^{n} \left(\frac{g^{(j)}(x)}{j!} \right)^{k_j} ,$$

where the sum is over all n-tuples of nonnegative integers $(k_1, \ldots, k_n)$ satisfying the constraint

$$1 \cdot k_1 + 2 \cdot k_2 + 3 \cdot k_3 + \cdots + n \cdot k_n = n \ .$$

In our case, we apply twice the formula, the first time to find the successive derivatives of the canonical coordinates q and $p = \dot{q}$ with respect to time, the second one to find the derivatives of $\tilde{\mathcal{E}}(p,q)$ and $\tilde{K}(p,q)$. Whereas the latter application is trivial, we remark that the former is obtained by explicitly expressing the force $\ddot{q} = \dot{p}$ as a function of q and p, then by writing

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} q = \frac{\mathrm{d}^{n-2}}{\mathrm{d}t^{n-2}} \ddot{q}(p,q) \ .$$

As the term at the r.h.s. involves only derivatives of q up to order $n-1$, the scheme can be applied iteratively to obtain all the derivatives of q (and p) with respect to time.

Since the derivatives of f and g are easily calculated in our case, the problem is just to have an effective algorithm that produces all the required n-tuples $(k_1, \ldots, k_n)$.

For positive integers n, s we introduce the sets

$$\mathcal{K}_{n,s} = \{k \in \mathbb{N}_0^n \ : \ k_1 + 2k_2 + \ldots + nk_n = s\}$$

where $\mathbb{N}_0$ is the set of non negative integers. For $n = 1$ we have $\mathcal{K}_{1,s} = \{(s)\}$, i.e., just one element. For $n > 1$ the following statement is obviously true: *every element* $(k_1, \ldots, k_n) \in \mathcal{K}_{n,s}$ *can be written as* (k', k_n) *where* $k' = (k_1, \ldots, k_{n-1}) \in \mathcal{K}_{n-1,s-nk_n}$. This is the key of our algorithm.

In $\mathcal{K}_{n,s}$ we consider the inverse lexicographic order, right to left. More precisely, the ordering is recursively defined as follows: for $n = 1$ the ordering is trivial, since there is just one element; for $j, k \in \mathcal{K}_{n,s}$ with $n > 1$ we say that $j < k$ if either case applies: (i) $j_n > k_n$, or (ii) $j_n = k_n$ and

$(j_1, \ldots, j_{n-1}) < (k_1, \ldots, k_{n-1})$. E.g., setting $n = 3$ and $s = 5$ we get the ordered set

$$\big\{(0,1,1)\,,\ (2,0,1)\,,\ (1,2,0)\,,\ (3,1,0)\,,\ (5,0,0)\big\}\ .$$

Remark that the last vector $k \in \mathcal{K}_{n,s}$ is $(s, 0, \ldots, 0)$.

We come now to stating the algorithm. We use two basic operations, namely: (a) find the first vector $k \in \mathcal{K}_{n,s}$ according to the ordering above; (b) for a given $k \in \mathcal{K}_{n,s}$ find the next vector.

The first vector is easily found by setting

$$k_n = \left\lfloor \frac{s}{n} \right\rfloor\ ,\ k_{n-1} = \left\lfloor \frac{s - nk_n}{n-1} \right\rfloor\ , \ldots,\ k_1 = s - nk_n - \ldots - 2k_2\ .$$

For a given k the next vector is found as follows: starting from k_2 find the first k_m which is not zero, so that $k_2 = \ldots = k_{m-1} = 0$ and $k_m \neq 0$. Then replace k_m with $k_m - 1$ and $(k_1, \ldots, k_{m-1})$ with the first vector of $\mathcal{K}_{m-1,k_1+m}$. The algorithm stops when $k = (s, 0, \ldots, 0)$, namely the last vector, because the index m can not be found.

Bibliography

[1] A. Carati, J. Stat. Phys. **128**, 1057-1077 (2007).

[2] A.I. Khinchin, *Mathematical foundations of statistical mechanics*, Dover (New York, 1949).

[3] R. Kubo, J. Phys. Soc. Japan **12**, 570-586 (1957).

[4] G. Gallavotti, *Statistical mechanics. A short treatise*, Springer Verlag (Berlin, 1999).

[5] C. Liverani, Ann. Math. **142**, 239-301 (1995).

[6] A. Carati, F. Benfenati, L. Galgani, Chaos **21**, 023134 (2011).

[7] J.H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities*, Oxford University Press (London,1932).

[8] M. Meissner and K. Spitzmann, Phys.Rev. Lett. **46**, (1981) 265-268.

[9] N. O. Birge, Phys. Rev. B **34**, (1986) 1631-1642. Phys.Rev. Lett. **46**, (1981) 265-268.

[10] A. Giorgilli, L. Galgani, Cel. Mech. **17**, (1978) 267-280.

[11] A. Carati, A.M. Maiocchi, Commun. Math. Phys. **314**, (2012) 129–161.

[12] B.O. Koopman, Proc. Natl. Acad. Sci. USA **17**, (1931) 315-318.

[13] A.M. Maiocchi, A. Carati, A. Giorgilli, J. Stat. Phys. **148**, (2012) 1054-1071.

[14] E. Fucito, F. Marchesoni, E. Marinari, G. Parisi, L. Peliti, S. Ruffo, A. Vulpiani, J. Phys.–Paris **43**, (1982) 707-714.

[15] G. Parisi, Europhys. Lett. **40**, (1997) 357-362.

[16] S. Flach, Phys. Rev. E **58**, (1998) R4116-R4119.

[17] D. Bambusi, D. Muraro, T. Penati, Phys. Lett. A **372**, (2008) 2039-2042.

[18] D. Bambusi, A. Carati, T. Penati, Nonlinearity **22**, (2009) 923-946.

[19] A. Carati, L. Galgani, F. Santolini, Chaos **19**, (2009) 023123.

[20] A. Giorgilli, S. Paleari, T. Penati, J. Stat. Phys **148**, (2012) 1106-1134.

[21] T.M. Cherry, Proc. Camb. Phil. Soc. **22**, (1924) 325349.

[22] A. Giorgilli, Ann. Inst. Henri Poincaré **48**, (1988) 423-439.

[23] N.N. Bogolyubov, B.I. Khatset, D.Ya. Petrina, Ukr. J. Phys. **53**, Special Issue, (2008) 168-184. Russian original in Teor. Mat. Fiz., **1:2**, (1969) 251-274.

[24] D. Ruelle, Commun. Math. Phys. **50**, (1976) 189-194.

[25] R.L. Dobrushin, Theory Probab. Appl. **13**, (1968) 197-224.

[26] R.L. Dobrushin, E.A. Pechersky, in J.V. Prokhorov, K. Ito (eds.), *Probability Theory and Mathematical Statistics*, (Springer, Berlin 1983), 97-110.

[27] E. Fermi, J. Pasta, S. Ulam, in *E. Fermi Collected Papers*, vol. 2, pp. 977–988, The University Chicago Press (Chicago, 1965).

[28] D. Ruelle, Phys. Rev. Lett. **56**, (1986) 405–407.

[29] W. Feller, *An introduction to probability theory and its applications*, vol. II., Wiley & Sons (New York, 1971).

[30] T. J. Stieltjes, Ann. Fac. Sci. Toulouse Sér. 1 **8**, (1894) J1–J122 and **9**, (1895) A5–A47.

[31] D.V. Widder, *The Laplace transform*, Princeton University Press (Princeton, 1946).

[32] N.I. Akhiezer, *The classical moment problem and some related questions in analysis*, Oliver & Boyd (Edinburgh, 1965).

[33] A. Carati, M. Zuin, A. Maiocchi, M. Marino, E. Martines, L. Galgani, Chaos **22**, 033124 (2012).

[34] H. Alfvén, *Cosmical electrodynamics*, Oxford U.P. (Oxford, 1950).

[35] C.A. Iglesias, J.L. Lebowitz, D. MacGowan, Phys. Rev. A **28**, 1667 (1983).

[36] R.S. Granetz, Phys. Rev. Lett. **49**, 658 (1982).

[37] M. Murakami, J.D. Callen, L.A. Berry, Nucl. Fusion **16**, 347 (1976).

[38] Ä. Stabler, et al., Nucl. Fusion **32**, 1557 (1992).

[39] P.C. de Vries, J. Rapp, F.C. Schüller, M.Z. Tokar, Phys. Rev. Lett. **80**, 3519 (1998).

[40] D. Frigione, et al., Nucl. Fusion **36**, 1489 (1999).

[41] V.G. Merezhkin, 33rd EPS Conference on Plasma Phys., Rome 19-23 June 2006 ECA Vol. 30I, P-4.085 (2006).

[42] V. Merthens, et al., Nucl. Fusion **37**, 1607 (1997).

[43] J. Howard, M. Person, Nucl. Fusion **32**, 361 (1992).

[44] K.S. Dyabilin, et al., Czech. J. Phys. **37**, 713 (1987).

[45] M. Asif, et al., Phys. Lett. A **336**, 61 (2005).

[46] B. LaBombard, et al., Phys. Plasmas **8**, 2107 (2001).

[47] H. Takenaga, et al., Nucl. Fusion **45**, 1618 (2005).

[48] G. Saibene, et al. Nucl. Fusion **39**, 1133 (1999).

[49] M.G. Bell, et al., Nucl. Fusion **32**, 1585 (1992).

[50] N. Commaux, et al., 33rd EPS Conference on Plasma Phys., Rome, 19 - 23 June 2006 ECA, Vol.30I, P-5.105 (2006).

[51] H. Weisen, et al. Plasma Phys. Control. Fusion **38**, 1137 (1996).

[52] H.Q. Liu, et al., Plasma Phys. Control. Fusion **49**, 995 (2007).

[53] T.W. Petrie, A.G. Kellman, M.Ali Mahdavi, Nucl. Fusion **33**, 929 (1993).

[54] M. Greenwald, Plasma Phys. Contr. Fusion **44**, R27 (2002).

[55] L. Giannone et al., Plasma Phys. Control. Fusion **42**, 603 (2000).

[56] J. Mijazawa, et al., Nucl. Fusion **48**, 015003 (2008).

[57] F.L. Tabarés, et al., Plasma Phys. Control. Fusion **50**, 124051 (2008).

[58] S. Sudo . et al., Nucl. Fusion **30**, 11 (1990).

[59] A. Sykes, Nucl. Fusion **39**, 1271 (1999).

[60] R. Martin et al., 27th EPS Conference on Contr. Fusion and Plasma Phys., Budapest, 12-16 June 2000 ECA, Vol. 24B, 221 (2000).

[61] S.M. Kaye et al., Phys. Plasmas **8**, 1977 (2001).

[62] P. Gibbon, G. Sutmann, in *Quantum Simulation of Complex Many–Body Systems: from Theory to Algorithms*, J. Grotendorst, D. Marx, A. Muramatsu eds., NIC Series **10**, 467–506 (2002).

[63] N.I. Akhiezer, M.G. Krein, *Some questions in the theory of moment*, AMS (Providence, 1962).

[64] F. Hausdorff, Math. Z. **16**, (1923) 220–248.

[65] F. Faà di Bruno, Ann. Sci. Mat. Fis. **6**, (1855) 479–480. French version in: Note sur une nouvelle formule de calcul différentiel, Quart. J. Pure Appl. Math. **1**, (1857) 359–360.