An averaging theorem for FPU in the thermodynamic limit *

A. Maiocchi[†]   D. Bambusi[‡]   A. Carati[§]

April 3, 2014

Abstract

Consider an FPU chain composed of $N \gg 1$ particles, and endow the phase space with the Gibbs measure corresponding to a small temperature β^{-1}. Given a fixed K, we construct K packets of normal modes whose energies are adiabatic invariants (i.e., are approximately constant for times of order β^{1-a}, $a > 0$) for initial data in a set of large measure. Furthermore, the time autocorrelation function of the energy of each packet does not decay significantly for times of order β. The restrictions on the shape of the packets are very mild. All estimates are uniform in the number N of particles and thus hold in the thermodynamic limit $N \to \infty$, $\beta > 0$.

KeywordsFPU — perturbation theory — thermodynamic limit — energy exchanges

---

†Università degli Studi di Milano, Via Saldini 50, Milano, I–20133, ITALY. E–mail: alberto.maiocchi@unimi.it.

‡Università degli Studi di Milano, Via Saldini 50, Milano, I–20133, ITALY. E–mail: dario.bambusi@unimi.it.

§Università degli Studi di Milano, Via Saldini 50, Milano, I–20133, ITALY. E–mail: andrea.carati@unimi.it.

1 Introduction

In 1954 Fermi, Pasta and Ulam, being interested in the problem of the foundations of statistical mechanics, started the study of the energy exchanges among the normal modes of a nonlinear chain of particles with nearest neighbor interaction. In the present paper we prove a result bounding the variation of the energy of packets of normal modes for times of the order β^{1-a}, with $a > 0$, where $\beta > 0$ is the inverse temperature of the chain. The bound holds for initial data in a set of large Gibbs measure. We also prove that the time autocorrelation function of each packet remains significantly away from zero at least for times of order β. As far as we know this is the first rigorous result on energy exchange among packets of modes of the FPU model in the thermodynamic limit.

The FPU model has been the object of a huge number of studies (see e.g. [13] for a report and [15, 5, 3] for some numerical works strictly related to the present one), and many techniques have been used in order to give significant analytical predictions about the dynamics of the chain. We recall in particular the averaging type results of [24, 2], the results on the dynamics of solitary waves of [9, 10, 11, 12, 17, 21, 22], and the results of [1] on the Toda chain (which should be usable for studying the FPU dynamics). However all known results cover only the case of small *total* energy, so that they are unable to deal with the thermodynamic limit (with finite *specific* energy) which is the relevant one for foundations of statistical mechanics.

A technique allowing one to deal with the thermodynamic limit was introduced in [4, 6] (see also [7, 16]); this is the technique that we extend here to deal with the FPU system. We recall that the idea of those papers was to consider a "resonant" linear combination $\Phi_0 := \sum_k g_k E_k$ of the energies E_k of the normal modes of the linearized system and to construct a modification $\Phi := \Phi_0 + \Phi_1$ whose Poisson bracket with the Hamiltonian has a zero of high order at the origin. Then one uses methods from statistical mechanics in order to estimate the ratio between the standard deviation of $\dot\Phi$ and that of Φ. Finally one can use standard probabilistic techniques to deduce the result on the variation in time of Φ (and of Φ_0), and of their time autocorrelation functions.

In order to apply such ideas to the FPU system we have to tackle two kinds of difficulties, which we think should appear also in typical models of crystal dynamics. The first one is related to the fact that low temperature FPU is a perturbation of a linear system presenting a *continuum of frequencies*, so that the problem of small denominators (which was absent in [6]) occurs here in a new way.[1] This problem is here overcome exploiting two properties, the first one is that, due to the translational invariance of the interactions, there occurs a selection of the coefficients actually appearing in the interaction, which in turn implies a selection rule on the small denominators. The second property is that if one stops the construction at order three, then the small denominators always appear with a numerator which depends on the coefficients g_k defining

---

[1]Small denominators appear also in [4, 7], where however the frequencies occur essentially as iid random variables.

Φ_0. Thus, with an appropriate choice of g, the numerators are made to vanish exactly when the denominators do. Surprisingly enough, such a procedure only imposes a constraint on the behavior of g_k as $k \to 0$ (see Theorem 1 below) and thus one has a great freedom in the choice of the adiabatic invariants. The fact that at order four more complicated small denominators appear constitutes an obstruction to a naive extension of the present result to longer time scales. We recall that numerical computations (see e.g. [3]) suggest that some stability properties should hold also for longer time scales, so we think that our result is not optimal. However we have no idea on how to get longer times, neither we know whether the problem of going to higher order is only technical or not.

The second difficulty tackled here is related to the fact that the normal modes of the unperturbed system (linearized FPU) are the Fourier modes, while the measure presents in a simple way if it is written in the space of the particles. So we have to work quite a lot in order to perform, in an efficient way, the averages of the quantities of interest.

The paper is organized as follows: in Sect. 2 we give a precise statement of our results; in Sect. 3 we prove the result on the adiabatic invariance of the energies of packets of normal modes; such a section is split into two subsections: in the first one we give the proof of the main theorem using the result of the main technical Lemma 1 which is proved in the subsequent subsection. In Sect. 4 we prove Lemma 3 which gives the characterization of the allowed functions g. Finally, in the Appendix A we give a the proof of a more or less standard auxiliary Lemma useful for the computation of averages.

2 Stability estimate for the FPU model

The Hamiltonian of the FPU–system with fixed end points can be written, in suitably rescaled variables, as

$$H = H_0 + H_1 + H_2 \tag{1}$$

where

$$H_0 \stackrel{\text{def}}{=} \sum_{j=0}^{N} \left(\frac{p_j^2}{2} + \frac{(q_{j+1} - q_j)^2}{2} \right) \ ,$$

$$H_1 \stackrel{\text{def}}{=} \frac{1}{3} \sum_{j=0}^{N} (q_{j+1} - q_j)^3$$

$$H_2 \stackrel{\text{def}}{=} \frac{A}{4} \sum_{j=0}^{N} (q_{j+1} - q_j)^4 \ ,$$

and $p = (p_1, \ldots, p_N)$, $q = (q_1, \ldots, q_N)$ are canonically conjugated variables in the phase space $\mathcal{M} \equiv \mathbb{R}^{2N}$, $p_0 = p_{N+1} = q_0 = q_{N+1} = 0$, and $A > 0$ is a positive parameter.

We endow the phase space by the Gibbs measure at inverse temperature β, namely

$$\mathrm{d}\mu(p,q) \overset{\mathrm{def}}{=} \frac{\mathrm{e}^{-\beta H(p,q)}}{Z(\beta)} \mathrm{d}^n p \, \mathrm{d}^n q \; ; \tag{2}$$

as usual $Z(\beta)$ is the partition function, i.e. the normalization constant such that the measure of $\mathcal{M}$ equals 1. Given a function F on the phase space, we will use this measure to compute its average $\langle F \rangle$, its L^2-norm $\|F\|$ and its variance σ_F^2 defined by

$$\langle F \rangle \overset{\mathrm{def}}{=} \int_{\mathcal{M}} F \mathrm{d}\mu \; , \tag{3}$$

$$\|F\|^2 \overset{\mathrm{def}}{=} \int_{\mathcal{M}} |F|^2 \mathrm{d}\mu \; , \tag{4}$$

$$\sigma_F^2 \overset{\mathrm{def}}{=} \|F - \langle F \rangle\|^2 \; . \tag{5}$$

We define also the correlation of two dynamical variables F, G by

$$C_{F,G} := \langle FG \rangle - \langle F \rangle \langle G \rangle$$

and the time autocorrelation of a dynamical variable by

$$C_F(t) := C_{F,F(t)} \; , \tag{6}$$

where $F(t) := F \circ g^t$ and g^t is the flow of the FPU system.

The unperturbed Hamiltonian H_0 can be put in diagonal form by passing to the normal modes of oscillation. The coordinates of the normal modes, denoted by $\hat{p} = (\hat{p}_1, \ldots, \hat{p}_N)$ and $\hat{q} = (\hat{q}_1, \ldots, \hat{q}_N)$ are obtained through the canonical change of variables

$$p_j = \sqrt{\frac{2}{N+1}} \sum_{k=1}^{N} \hat{p}_k \sin\left(\frac{\pi j k}{N+1}\right) \; ,$$

$$q_j = \sqrt{\frac{2}{N+1}} \sum_{k=1}^{N} \hat{q}_k \sin\left(\frac{\pi j k}{N+1}\right) \; .$$

In such variables, H_0 takes the form

$$H_0 = \sum_{k=1}^{N} \frac{\hat{p}_k^2 + \omega_k^2 \hat{q}_k^2}{2} = \sum_{k=1}^{N} E_k \; ,$$

where we have defined the energies of the normal modes

$$E_k \overset{\mathrm{def}}{=} \frac{\hat{p}_k^2 + \omega_k^2 \hat{q}_k^2}{2}$$

and the corresponding frequencies $\omega_k = 2\sin\left(\frac{\pi k}{2(N+1)}\right)$.

Remark 1. As one can expect from the equipartition theorem, it will turn out that $\|H_1\| \, / \, \|H_0\| \sim \beta^{-1/2}$. This will allow to consider the FPU system at low temperature as a small perturbation of H_0, the perturbation parameter being $\beta^{-1/2}$.

Let $g \in \mathcal{C}^2([0,1], \mathbb{R}^+)$ be a twice differentiable function; as anticipated above, we are interested in the time evolution of quantities of the form

$$\Phi_0 \overset{\text{def}}{=} \sum_{k=1}^{N} g\left(\frac{k}{N+1}\right) E_k \ .$$

In the following we will sometimes denote $g_k \overset{\text{def}}{=} g(k/(N+1))$; furthermore we define $\omega(x) := 2\sin(\pi x/2)$ so that $\omega(k/(N+1)) = \omega_k$.

Our main result is the following theorem, which will be proved in the rest of the paper.

Theorem 1 *Let* $g \in \mathcal{C}^2([0,1]; \mathbb{R}^+)$ *be a function fulfilling* $g'(0) = 0$. *Define* $\Phi_0 \overset{\text{def}}{=} \sum_k g(k/(N+1))E_k$, *then there exist constants* $\beta^* > 0$, $N^* > 0$ *and* $C > 0$ *s.t., for any* $\beta > \beta^*$ *and for any* $N > N^*$, *there exists a polynomial of third order* Φ_1 *with the property that* $\Phi \overset{\text{def}}{=} \Phi_0 + \Phi_1$ *fulfills*

$$\frac{\|\dot{\Phi}\|}{\sigma_\Phi} \le \frac{C}{\beta} h \ . \tag{7}$$

$$\sigma_{\Phi_1}/\sigma_{\Phi_0} < Ch/\sqrt{\beta} \ , \tag{8}$$

where

$$h \overset{\text{def}}{=} \frac{g(0) + \sup_{x \in [0,1]} |g''(x)| + 1}{\int_0^1 g^2(x)\mathrm{d}x} \ . \tag{9}$$

Remark 2. In the perturbative construction the assumptions on g are used to control the small denominators. Actually the small denominators always appear with a corresponding numerator. It turns out that such a numerator vanish exactly at the same points where the denominators vanish. Indeed the quantity that one actually meets is

$$h_1(\nu) \overset{\text{def}}{=} \max_{\tau_i = \pm 1} \sup_{x,y \in [0,1]} \left| \frac{\tau_1 \nu(x) + \tau_2 \nu(y) + \tau_3 \nu(z(x,y))}{\tau_1 \omega(x) + \tau_2 \omega(y) + \tau_3 \omega(z(x,y))} \right| \ , \tag{10}$$

$$\tag{11}$$

where

$$\nu(x) \overset{\text{def}}{=} \omega(x)g(x) \ , \quad z = z(x,y) \overset{\text{def}}{=} \begin{cases} x + y & \text{if } x + y \le 1 \\ 2 - x - y & \text{if } x + y > 1 \end{cases} \ .$$

In Lemma 3, it will be proved that under our assumptions $h_1(\nu)$ is bounded.

It is worth to point out some consequences of the main theorem:

Corollary 1 *Under the hypotheses of Theorem 1, there exists C_1 s.t.*

$$\frac{\mathbf{C}_{\Phi_0}(t)}{\sigma_{\Phi_0}^2} \geq \frac{1}{2} , \quad \forall |t| \leq \frac{\beta}{C_1}. \tag{12}$$

Proof. One starts by observing that, in virtue of Theorem 3 of [6], (7) implies that

$$\mathbf{C}_{\Phi}(t) \geq \sigma_{\Phi}^2 \left(1 - \frac{C^2 h^2}{2\beta^2} t^2 \right) , \tag{13}$$

whereas applying Schwartz inequality one gets

$$\left| \sigma_{\Phi}^2 - \sigma_{\Phi_0}^2 \right| = \left| \sigma_{\Phi_1}^2 + 2\langle (\Phi_1 - \langle \Phi_1 \rangle); (\Phi_0 - \langle \Phi_0 \rangle) \rangle \right| \leq \sigma_{\Phi_1}^2 + 2\sigma_{\Phi_1}\sigma_{\Phi_0} .$$

On the other hand (cf. also Theorem 1 of [20]) one also has

$$\left| \mathbf{C}_{\Phi}(t) - \mathbf{C}_{\Phi_0}(t) \right| = \left| \mathbf{C}_{\Phi_1}(t) \right| + 2 \left| \langle \Phi_0; \Phi_1 \circ g^t \rangle \right| \leq \sigma_{\Phi_1}^2 + 2\sigma_{\Phi_1}\sigma_{\Phi_0} .$$

Since (8) provides the upper bound

$$\sigma_{\Phi_1}^2 + 2\sigma_{\Phi_1}\sigma_{\Phi_0} \leq \sigma_{\Phi_0}^2 \left(\frac{C^2 h^2}{\beta} + 2\frac{Ch}{\sqrt{\beta}} \right) , \tag{14}$$

the thesis follows.   Q.E.D

We have also the following corollary concerning the probability $\mathbf{P}$ that the time change of Φ_0 is large:

Corollary 2 *Under the hypotheses of Theorem 1, there exists C_2 s.t. $\forall 0 \leq a \leq 1/2$ one has*

$$\mathbf{P}\left(|\Phi_0(t) - \Phi_0| \geq \frac{\sigma_{\Phi_0}}{\beta^{a/2}} \right) \leq \frac{C_2}{\beta^a} , \quad \forall |t| \leq \beta^{1-a} , \tag{15}$$

where, as above, $\Phi_0(t) = \Phi_0 \circ g^t$.

Proof. The proof is easily obtained by using the relations

$$\sigma_{\Phi_0(t) - \Phi_0}^2 = 2 \left(\sigma_{\Phi_0}^2 - \mathbf{C}_{\Phi_0}(t) \right) \leq 2\sigma_{\Phi_0}^2 \left(\frac{C_1 h^2}{\sqrt{\beta}} + \frac{C_1 h^2}{2\beta^2} t^2 \right) , \tag{16}$$

where in the upper bound use is made of (13), (14). Then one applies the Chebyshev inequality to $\Phi_0(t) - \Phi_0$, which gives, for any $\lambda > 0$:

$$\mathbf{P}\left(|\Phi_0(t) - \Phi_0| \geq \lambda \sigma_{\Phi_0} \right) = \mathbf{P}\left(|\Phi_0(t) - \Phi_0| \geq \frac{\lambda \sigma_{\Phi_0}}{\sigma_{\Phi_0(t) - \Phi_0}} \sigma_{\Phi_0(t) - \Phi_0} \right)$$

$$\leq \frac{\sigma_{\Phi_0(t) - \Phi_0}^2}{\lambda^2 \sigma_{\Phi_0}^2} .$$

By choosing $\lambda = \beta^{-a/2}$ and inserting relation (16) the thesis is proved. Q.E.D

Remark 3. Following [15] it is also possible to bound the probability that the time average and the time variance of $\Phi_0(t) - \Phi_0$ are not small. Here, for simplicity we choose to state just the previous Corollary.

Of course one can repeat the argument for different choices of the function g. Having in mind to partition the modes into packets, each one composed by adjacent modes, one partitions the interval $[0, 1]$ of the variable $k/(N + 1)$ in K sub-intervals. So one can define K different functions $g^{(1)}, g^{(2)}, ..., g^{(K)}$, for example with disjoint support, each one fulfilling the assumptions of Theorem 1, so that the quantities $\Phi_0^{(l)} \overset{\text{def}}{=} \sum_k g_k^{(l)} E_k$ are adiabatic invariants. Precisely one has

Corollary 3 *Assume that $\forall l = 1, ..., K$, the function $g^{(l)}$ fulfills the assumptions of Theorem 1. Then there exists C_3 (which depends on K, but not on N) s.t. $\forall 0 \le a \le 1/2$ one has*

$$\mathbf{P}\left(\exists l \ : \ |\Phi_0^{(l)}(t) - \Phi_0^{(l)}| \ge \frac{\sigma_{\Phi_0^{(l)}}}{\beta^{a/2}}\right) \le \frac{C_3}{\beta^a} \ , \quad \forall |t| \le \beta^{1-a} \ . \qquad (17)$$

Remark 4. The proof of Corollary 3 is straightforward, but one gets a constant C_3 which depends on K. We expect that the corollary should be true also with a constant independent of K, provided the functions $g^{(l)}$ are chosen in an appropriate way. However we did not investigate such a point.

We give now an explicit example of a possible choice of the form of the packets whose energy is approximately conserved according to Corollary 3: fix a positive function $\tilde{g} : \mathbb{R} \to \mathbb{R}^+$ fulfilling

$$\tilde{g}(0) = 1 \ , \ \tilde{g}(x) = \tilde{g}(-x) \ , \ \text{supp}\,\tilde{g} = \left[-\frac{1}{2}, \frac{1}{2}\right] \ , \qquad (18)$$

and define

$$g^{(1)}(x) := \tilde{g}\left(\frac{K}{2}x\right) \ , \ g^{(l)}(x) := \tilde{g}(Kx - l - 1) \ , \ l = 2, ..., K \qquad (19)$$

then the corresponding approximate integrals are $\Phi_0^{(l)} = \sum_{k=1}^N g^{(l)}\left(\frac{k}{N+1}\right) E_k$. We remark that as N increases the result holds for fixed number K of packets. On the contrary, the number of modes contained in each packet increases in a way proportional to N.

3 Proof of Theorem 1

3.1 The proof

In this section we give the proof of Theorem 1 using the results of the main technical Lemma 1, which will be proved in the subsequent subsection.

The proof consists in performing the first step of the formal construction of an integral of motion which is a perturbation of Φ_0, and in estimating its time derivative. Define $\Phi = \Phi_0 + \Phi_1$, with Φ_1 a polynomial of order three determined by the condition that $\{\Phi, H\}$ is of order four, where $\{.,.\}$ denotes the Poisson bracket. Then Φ_1 must fulfill the equation

$$\{H_0, \Phi_1\} = -\{H_1, \Phi_0\} \ . \tag{20}$$

The formal construction is standard (see, for instance [14]), but the estimate of the remainder requires a special care and is the main difficulty we have to address here.

To start with we pass to the complex coordinates

$$\xi_k = \frac{\hat{p}_k + i\omega_k \hat{q}_k}{\sqrt{2}} \ , \eta_k = \frac{\hat{p}_k - i\omega_k \hat{q}_k}{\sqrt{2}} \ ,$$

such that $\{\xi_k, \eta_k\} = i\omega_k$ and $H_0 = \sum_k \xi_k \eta_k$. Then the nonlinearity is a linear combination of monomials of the form

$$\Xi^s_{\tau,k} \overset{\text{def}}{=} \xi_{k_1}^{(1+\tau_1)/2} \eta_{k_1}^{(1-\tau_1)/2} \ldots \xi_{k_s}^{(1+\tau_s)/2} \eta_{k_s}^{(1-\tau_s)/2} \ , \quad s \geq 3$$

where

$$\tau = (\tau_1, ..., \tau_s) \ , \ \tau_l = \pm 1 \ , \quad k = (k_1, ..., k_s) \ , \ k_l = 1, ..., N \ ; \tag{21}$$

furthermore, the index k is such that

$$\lfloor \tilde{\tau} \cdot k \rceil = 0 \ , \quad \text{where } \lfloor n \rceil \overset{\text{def}}{=} n \bmod [2(N+1)] \ , \tag{22}$$

for some

$$\tilde{\tau} = (\tilde{\tau}_1, ..., \tilde{\tau}_s) \ , \quad \tilde{\tau}_l = \pm 1 \ . \tag{23}$$

In the following we will use denote by $\mathcal{I}_s$ the set of the indexes $(\tau, \tilde{\tau}, k)$ of the form (21), (23). Finally, for $i \in \mathbb{Z}$ we will denote

$$\delta_i = \begin{cases} 1 & \text{if } i = 0 \\ 0 & \text{otherwise} \end{cases} \ .$$

Definition 1. We say that $f \in \mathcal{P}_s$ if it has the form

$$f = \frac{1}{(N+1)^{(s-2)/2}} \sum_{(\tau, \tilde{\tau}, k) \in \mathcal{I}_s} f_{\tau, \tilde{\tau}} \left(\frac{k_1}{N+1}, \ldots, \frac{k_s}{N+1} \right) \Xi^s_{\tau,k} \delta_{\lfloor \tilde{\tau} \cdot k \rceil} \ , \tag{24}$$

where $f_\tau : [0,1]^s \to \mathbb{C}$ are continuous functions.

This is the class of polynomials which will enter the perturbative construction.

8

We define in $\mathcal{P}_s$ the norm

$$\|f\|_+ \stackrel{\text{def}}{=} \max_{(\tau,\tilde{\tau},k)\in\mathcal{I}_s} \left| f_{\tau,\tilde{\tau}}\left(\frac{k_1}{N+1},\ldots,\frac{k_s}{N+1}\right)\right| \delta_{\lfloor\tilde{\tau}\cdot k\rceil} \ . \tag{25}$$

The variance of a dynamical variable in $\mathcal{P}_s$ is related to the above defined norm by the following lemma which is the main technical lemma of the paper and whose proof is deferred to subsection 3.2.

Lemma 1 *For any integer[2] $s \geq 2$ there exist $N_0 > 0$ and C such that, for any $N > N_0$, and any $f \in \mathcal{P}_s$ one has*

$$\sigma_f^2 \leq N\frac{C}{\beta^s}\|f\|_+^2 \ .$$

The norm of the Poisson brackets of two variables is controlled by the following lemma whose simple proof is omitted.

Lemma 2 *If $f \in \mathcal{P}_s$, $g \in \mathcal{P}_r$, then $\{f,g\} \in \mathcal{P}_{r+s-2}$. Moreover, one has*

$$\|\{f,g\}\|_+ \leq 2^4 \max(s,r)\|f\|_+ \|g\|_+ \ .$$

In order to find a solution of equation (20), we express H_1 in complex coordinates, namely

$$H_1 = \frac{i}{6}\sqrt{\frac{1}{N+1}} \sum_{k_1,k_2,k_3=1}^{N} (\xi_{k_1} - \eta_{k_1})(\xi_{k_2} - \eta_{k_2})(\xi_{k_3} - \eta_{k_3})$$
$$\times \left(3\delta_{k_1+k_2-k_3} + \delta_{k_1+k_2+k_3-2(N+1)}\right)$$

so that $H_1 \in \mathcal{P}_3$ (one can similarly check that $H_2 \in \mathcal{P}_4$). Then, by using the properties of Poisson brackets and the fact that $\Phi_0 = \sum_k (\nu_k/\omega_k)\xi_k\eta_k$, where $\nu_k = \nu(k/N+1)$ $\nu(x) = g(x)\omega(x)$, one can check that a formal solution of (20) is given by the expression

$$\Phi_1 = \frac{i}{3}\sqrt{\frac{1}{N+1}} \sum_{\substack{\tau_i=\pm1 \\ k_i=1,\ldots,N}} \tau_1\tau_2\tau_3 \frac{\tau_1\nu_{k_1} + \tau_2\nu_{k_2} + \tau_3\nu_{k_3}}{\tau_1\omega_{k_1} + \tau_2\omega_{k_2} + \tau_3\omega_{k_3}} \Xi_{\tau,k}^3$$
$$\times \left(3\delta_{k_1+k_2-k_3} + \delta_{k_1+k_2+k_3-2(N+1)}\right) \ . \tag{26}$$

In order to show that Φ_1 is well defined we have to show that the fraction in (26) is bounded. This is the content of the following lemma. We recall that the supremum of the above fraction is denoted by $h_1(\nu)$ (cf. eq. (10)).

---

[2]In the present paper we use the lemma only with $s = 3, 4$, however we decided to prove the lemma in the general case because the consideration of the cases $s = 3, 4$ only would not lead to significant simplifications of the proof.

Lemma 3 *Let $g(x) \in \mathcal{C}^2([0,1], \mathbb{R})$ fulfill $g'(0) = 0$; set $\nu(x) \stackrel{\text{def}}{=} g(x)\omega(x)$, $c_0 \stackrel{\text{def}}{=} g(0)$, $c_2 \stackrel{\text{def}}{=} \sup_{x \in [0,1]} |g''(x)|$. Then there exists a constant $C > 0$, independent of g, such that one has*

$$h_1(\nu) \leq C(c_0 + c_2) .$$

Moreover, if $g'(0) \neq 0$, $h_1(\nu)$ is not bounded.

The proof of this lemma is postponed to Sect. 4.

Proof of Theorem 1. We bound the numerator of the fraction at the l.h.s. of (7) by using Lemma 2 and Lemma 1 (notice that $H_0, \Phi_0 \in \mathcal{P}_2$):

$$\|\dot{\Phi}\| = \|\{\Phi_1, H_1 + H_2\} + \{\Phi_0, H_2\}\| \leq \sqrt{N}\frac{C_0}{\beta^2}(h_1(\nu) + 1) ,$$

for some $C_0 > 0$. Concerning the denominator of (7), we write

$$\sigma_\Phi \geq \sigma_{\Phi_0} - \sigma_{\Phi_1} \tag{27}$$

and we estimate σ_{Φ_0} from below using $\sigma_{\Phi_0} \geq \sigma_F$ with $F \stackrel{\text{def}}{=} \sum_k (\nu_k/\omega_k)\hat{p}_k^2/2$, where the last inequality is due to the stochastic independence of $\hat{p}_k$ and $\hat{q}_k$. Thus one has

$$\sigma_F^2 = \frac{1}{2\beta} \sum_k \left(\frac{\nu_k}{\omega_k}\right)^2 \geq \frac{N}{4\beta} \int_0^1 \left(\frac{\nu(x)}{\omega(x)}\right)^2 \mathrm{d}x , \tag{28}$$

where the last estimate is obtained through Euler summation formula, which in turn can be applied in virtue of the regularity hypotheses on $\nu(x)/\omega(x)$. Moreover, notice that as a consequence one has $h = h(\nu) < \infty$. On the other hand, one can apply Lemma 1 and get

$$\sigma_{\Phi_1} \leq \sqrt{N}\frac{C_1}{\beta^{3/2}}h_1(\nu) ,$$

for some $C_1 > 0$. This, together with (28), proves formula (8). Furthermore, making use again of (28) and inserting it in (27), formula (7) is proved too. Q.E.D

3.2 Proof of Lemma 1

The proof consists in some steps, the first of which is the choice of suitable coordinates in which the integrals with respect to Gibbs measure become tractable. The rest of the proof consists of a careful analysis of the expression obtained through the integration.

Concerning the choice of coordinates, first we go back to the variables $\hat{p}$, $\hat{q}$, then the integration over the $\hat{p}$'s is easy (they are iid Gaussian variables with zero average). The integration with respect to the $\hat{q}$ variables is more complicated. In order to do it we use the fact that the Hamiltonian is a simple function of

$r_j \overset{\text{def}}{=} q_{j+1} - q_j$, for $j = 0, \ldots, N$. In fact, the potential part of the Hamiltonian can be written as

$$\sum_{j=0}^{N} V(r_j) \quad \text{with } V(r) \overset{\text{def}}{=} \frac{r^2}{2} + \frac{1}{3}r^3 + \frac{A}{4}r^4 \ ,$$

so that the configurational part of the probability measure is factorized in terms of the variables r_j, which are independently distributed, apart from the constraint $\sum_j r_j = q_{N+1} - q_0 = 0$ (this implies that they are exchangeable random variables as defined e.g. in [8]). The situation is similar to that of the micro-canonical ensemble for the perfect gas, in which the energies of the particles are independently distributed, except for the constraint that their sum is fixed. In such a case one can compute mean values and variances of sensible observables in the canonical ensemble, in which all energies are independent, and then estimate the error introduced. For this reason, we will use the mixed coordinates $\hat{p}$, r, and adopt the methods developed in the frame of statistical mechanics to deal with the integration over the r's (see [18]). The corresponding lemma[3] (see Lemma 4 below) is more or less standard, however, we were not able to find an adapted statement in literature, so we give its proof in Appendix A.

Lemma 4 *There exist $K, N_0 > 0$ such that, for any multi–index k, j with length n and m, respectively, and any $N > N_0$, one has*

$$\left| \langle r^k r^l \rangle - \langle r^k \rangle \langle r^l \rangle \right| \leq K^{n+m} \sqrt{n!m!} \beta^{-(n+m)/2} \ . \tag{29}$$

Moreover, if the sets of sites $\operatorname{supp} k$ *and* $\operatorname{supp} l$ *are disjoint, one has*

$$\left| \langle r^k r^l \rangle - \langle r^k \rangle \langle r^l \rangle \right| \leq \frac{1}{N} K^{n+m} \sqrt{n!m!} \beta^{-(n+m)/2} \ . \tag{30}$$

The variance of $f \in \mathcal{P}_s$, can be written as

$$
\sigma_f^2 = \frac{1}{(N+1)^{s-2}} \sum_{\substack{(\tau, \tilde{\tau}, k) \in \mathcal{I}_s \\ (\tau', \tilde{\tau}', k') \in \mathcal{I}_s}} f_{\tau, \tilde{\tau}} \left(\frac{k_1}{N+1}, \ldots, \frac{k_s}{N+1} \right) f_{\tau', \tilde{\tau}'} \left(\frac{k_1'}{N+1}, \ldots, \frac{k_s'}{N+1} \right)
$$
$$
\times \left(\langle \Xi_{\tau, k}^s \Xi_{\tau', k'}^s \rangle - \langle \Xi_{\tau, k}^s \rangle \langle \Xi_{\tau', k'}^s \rangle \right)
$$
$$
\times \delta_{\lfloor \tilde{\tau} \cdot k \rceil} \delta_{\lfloor \tilde{\tau}' \cdot k' \rceil}
\tag{31}
$$

Introducing the coordinates $\hat{p}$, $\hat{q}$, each term of the second line of (31) gives rise to at most 2^{2s} terms of the form

$$\langle \hat{r}_{k_1} \ldots \hat{r}_{k_{s_1}} \hat{r}_{k_1'} \ldots \hat{r}_{k_{s_1'}'} \rangle \langle \hat{p}_{k_{s_1+1}} \ldots \hat{p}_{k_s} \hat{p}_{k_{s_1'+1}'} \ldots \hat{p}_{k_s'} \rangle \tag{32}$$

$$- \langle \hat{r}_{k_1} \ldots \hat{r}_{k_{s_1}} \rangle \langle \hat{r}_{k_1'} \ldots \hat{r}_{k_{s_1'}'} \rangle \langle \hat{p}_{k_{s_1+1}} \ldots \hat{p}_{k_s} \rangle \langle \hat{p}_{k_{s_1'+1}'} \ldots \hat{p}_{k_s'} \rangle \ ,$$

---

[3] In its statement, we adopt the multi–index notation: $k = (k_0, \ldots, k_N)$ and $j = (j_0, \ldots, j_N)$ are vectors of nonnegative integers, with the norm defined by $|k| = k_0 + \ldots + k_N$. So, $r^k = r_0^{k_0} \cdot \ldots \cdot r_N^{k_N}$. Moreover, $\operatorname{supp} k$ denotes the set of sites i for which $k_i \neq 0$.

where $\hat{r}_k \overset{\text{def}}{=} \omega_k \hat{q}_k$.

The main step of the proof consists in computing a representation formula for the quantity

$$\hat{A}_{k,k'} \overset{\text{def}}{=} a\langle \hat{r}_{k_1} \cdots \hat{r}_{k_{s_1}} \hat{r}_{k'_1} \cdots \hat{r}_{k'_{s'_1}} \rangle - b\langle \hat{r}_{k_1} \cdots \hat{r}_{k_{s_1}} \rangle \langle \hat{r}_{k'_1} \cdots \hat{r}_{k'_{s'_1}} \rangle , \qquad (33)$$

where a, b are complex constants.

We start by establishing some notation. We will denote

$$S \overset{\text{def}}{=} s_1 + s'_1 , \quad L \equiv (L_1, ..., L_S) \overset{\text{def}}{=} (l_1, ..., l_{s_1}, l'_1, ..., l'_{s'_1})$$

$$K \equiv (K_1, ..., K_S) \overset{\text{def}}{=} (k_1, ..., k_{s_1}, k'_1, ..., k'_{s'_1}) .$$

Inserting the definition of the Fourier coefficients, one has

$$\hat{r}_k = \sqrt{\frac{2}{N+1}} \sum_{l=0}^{N} r_l \cos\left[\frac{\pi}{N+1}\left(l + \frac{1}{2}\right)k\right] , \qquad (34)$$

in (33) one gets

$$\hat{A}_{k,k'} \equiv \hat{A}_K = \frac{2^{S/2}}{(N+1)^{S/2}}$$

$$\times \sum_{L_1, ..., L_S} A_L \cos\left[\frac{\pi}{N+1}\left(L_1 + \frac{1}{2}\right)K_1\right] ... \cos\left[\frac{\pi}{N+1}\left(L_S + \frac{1}{2}\right)K_S\right] \quad (35)$$

where

$$A_L \overset{\text{def}}{=} a\langle r_{l_1} \cdots r_{l_{s_1}} r_{l'_1} \cdots r_{l'_{s'_1}} \rangle - b\langle r_{l_1} \cdots r_{l_{s_1}} \rangle \langle r_{l'_1} \cdots r_{l'_{s'_1}} \rangle . \qquad (36)$$

In order to compute $\hat{A}_K$ we proceed by reducing iteratively the number of variables to be summed. We will start by summing over L_S. At each step one gets that the quantity to be summed is the linear combination of quantities of the form (35) with coefficients enjoying a suitable property which is the same fulfilled by averages of exchangeable variables.

Precisely, let S be an integer and we consider the sequences $B_{L_1, ..., L_S}$ of complex numbers with the property that , if one fixes $S-1$ indexes, say $L_1, ..., L_{S-1}$, then $B_{L_1, ..., L_S}$ has the same value for all values of the remaining index, say L_S, s.t.

$$L_S \neq L_1 \text{ and } L_S \neq L_2 \text{ and } ... \text{ and } L_S \neq L_{S-1} .$$

Definition 2. We will denote by $\tilde{B}_{K_1, ..., K_S} \equiv \tilde{B}_K$ the rescaled Fourier transform of one of these sequences, precisely

$$\tilde{B}_K \overset{\text{def}}{=} \sum_{L_1, ..., L_{S-S_1}} B_{L_1, ..., L_{S-S_1}} \qquad (37)$$

$$\times \cos\left[\frac{\pi}{N+1}\left(L_1 + \frac{1}{2}\right)K_1\right] ... \cos\left[\frac{\pi}{N+1}\left(L_{S-S_1} + \frac{1}{2}\right)K_S\right]$$

12

where $B_{L_1,...,L_{S-S_1}}$ has the property just described.

The main remark needed in order to start the induction is contained in the following Lemma.

Lemma 5 *The following formula holds:*

$$
\begin{aligned}
\tilde{B}_K =&(N+1)\delta_{\lfloor K_S\rceil}\widetilde{B^0}_{K_1,K_2,...,K_{S-1}} \\
&+\widetilde{B^1}_{K_1+K_S,K_2,...,K_{S-1}} + \widetilde{B^1}_{K_1-K_S,K_2,...,K_{S-1}} \\
&+...+\widetilde{B^{S-1}}_{K_1,K_2,...,K_{S-1}+K_S} + \widetilde{B^{S-1}}_{K_1,K_2,...,K_{S-1}-K_S} \ ,
\end{aligned}
\tag{38}
$$

where the $\widetilde{B^j}$ are obtained through (37) from the sequences

$$
B^0_{L_1,...,L_{S-1}} \overset{\text{def}}{=} B_{L_1...L_S}\big|_{L_S\neq L_1,...,L_S\neq L_{S-1}} \ ,
\tag{39}
$$

$$
B^j_{L_1,...,L_{S-1}} \overset{\text{def}}{=} \frac{B_{L_1,...,L_{S-1},L_j} - B_{L_1...L_S}\big|_{L_S\neq L_1,...,L_S\neq L_{S-1}}}{2} \ .
\tag{40}
$$

Proof. It is a computation which exploits the formula

$$
\sum_{L=0}^{N}\cos\left[\frac{\pi}{N+1}\left(L+\frac{1}{2}\right)k\right] = (N+1)\delta_{\lceil k\rceil} \ .
\tag{41}
$$

In order to use it we rewrite $\tilde{B}_K$ by separating the sum over L_S, namely

$$
\tilde{B}_K = \sum_{L_1,...,L_{S-1}}\cos\left[\frac{\pi}{N+1}\left(L_1+\frac{1}{2}\right)K_1\right]...\cos\left[\frac{\pi}{N+1}\left(L_{S-1}+\frac{1}{2}\right)K_{S-1}\right]
$$

$$
\times\left\{B_{L_1...L_S}\big|_{L_S\neq L_1,...,L_S\neq L_{S-1}}\sum_{L_s:L_S\neq L_1,...,L_S\neq L_{S-1}}\cos\left[\frac{\pi}{N+1}\left(L_S+\frac{1}{2}\right)K_S\right]\right.
$$

$$
+B_{L_1,...,L_{S-1},L_1}\cos\left[\frac{\pi}{N+1}\left(L_1+\frac{1}{2}\right)K_S\right]+...
$$

$$
\left.+B_{L_1,...,L_{S-1},L_{S-1}}\cos\left[\frac{\pi}{N+1}\left(L_{S-1}+\frac{1}{2}\right)K_S\right]\right\}
$$

but the curly bracket is equal to

$$
B_L\big|_{L_S\neq L_1,...,L_S\neq L_{S-1}}\sum_{L_S=0}^{N}\cos\left[\frac{\pi}{N+1}\left(L_S+\frac{1}{2}\right)K_S\right]
$$

$$
+\left(B_{L_1,...,L_{S-1},L_1} - B_L\big|_{L_S\neq L_1,...,L_S\neq L_{S-1}}\right)\cos\left[\frac{\pi}{N+1}\left(L_1+\frac{1}{2}\right)K_S\right]+...
$$

$$
+\left(B_{L_1,...,L_{S-1},L_{S-1}} - B_L\big|_{L_S\neq L_1,...,L_S\neq L_{S-1}}\right)\cos\left[\frac{\pi}{N+1}\left(L_{S-1}+\frac{1}{2}\right)K_S\right]
$$

$$
= B_L\big|_{L_S\neq L_1,...,L_S\neq L_{S-1}}\delta_{\lfloor K_s\rceil}(N+1) + \sum_{j=1}^{S-1}B^j_{L_1,...,L_{S-1}}\cos\left[\frac{\pi}{N+1}\left(L_j+\frac{1}{2}\right)K_S\right]
$$

13

where

$$B^j_{L_1,...,L_{S-1}} \stackrel{\text{def}}{=} B_{L_1,...,L_{S-1},L_j} - B_L\big|_{L_S \neq L_1,...,L_S \neq L_{S-1}}. \tag{42}$$

In order to conclude the proof insert such a formula in the expression for $\tilde{B}_K$ and remark that except for the term containing the δ, all the other addenda contain the expression

$$\cos\left[\frac{\pi}{N+1}\left(L_j + \frac{1}{2}\right)K_S\right]\cos\left[\frac{\pi}{N+1}\left(L_j + \frac{1}{2}\right)K_j\right] =$$

$$\frac{1}{2}\left\{\cos\left[\frac{\pi}{N+1}\left(L_j + \frac{1}{2}\right)(K_S + K_j)\right] + \cos\left[\frac{\pi}{N+1}\left(L_j + \frac{1}{2}\right)(K_S - K_j)\right]\right\}$$

so that the thesis follows. Q.E.D

With formula (38) at hand we can iterate the construction in order to get the general structure of the terms involving $\hat{A}_K$.

Actually, in order to get the proof of Lemma 1, we need quite precise information on the structure of $\hat{A}_K$. To this end we still need some more preliminary definitions.

Having fixed a positive integer S, we consider vectors $\tau \equiv (\tau_1, ..., \tau_S)$, with $\tau_j \in \{-1, 0, 1\}$. The set of such vectors will often be denoted by $\mathbb{Z}_3^S$.

Definition 3. A collection $\tau^{(1)}, ..., \tau^{(S_1)}$, $S_1 \leq S$ of vectors of $\tau^{(i)} \in \mathbb{Z}_3^S$ will be said to be S-admissible, or simply admissible, if the following properties hold

1) the supports $\text{supp}(\tau^{(i)})$ are disjoint.

2) $\bigcup_{i=1}^{S_1}\text{supp}(\tau^{(i)}) = \{1, ..., S\}$.

We are now ready for the main lemma of this section. It gives the representation formula for the A's.

Lemma 6 $\hat{A}_K$ is the sum of a number independent of N of addenda, each one of the form

$$B_\tau\left[\prod_{i=1}^{S_1}\delta_{\lfloor\tau^{(i)}\cdot K\rceil}\right](N+1)^{S_1-S/2} , \quad S_1 \leq S \tag{43}$$

where $\tau = (\tau^{(1)}, ..., \tau^{(S_1)})$ is an S- admissible collection of vectors.

Furthermore, B_τ is a linear combination of the quantities A_L (cf. eq. (36)), such that the indexes L assume only those values s.t. the following property holds

$$\left[I \in \text{supp}(\tau^{(i)}) , \ J \in \text{supp}(\tau^{(j)}) , i \neq j\right] \Longrightarrow L_I \neq L_J . \tag{44}$$

The number of terms in the linear combination is bounded independently of N, the coefficients are bounded uniformly with respect to N.

Proof. The proof is obtained by applying iteratively Lemma 5. We claim that, after R steps of decomposition, $\tilde{B}_K$ turns out to be the sum of terms of the form

$$\tilde{B}_{K \cdot \tau^{(1)}, \ldots, K \cdot \tau^{(S-R)}} \left[(N+1)^{S_1} \prod_{i=1}^{S_1} \delta_{\lfloor K \cdot \tau^{(S-R+i)} \rceil} \right] , \qquad (45)$$

$$\text{with} \quad S_1 \leq R \leq S ,$$

where $\tau \equiv (\tau^{(1)}, \ldots, \tau^{(S+S_1-R)})$ is an admissible collection and the B's fulfill a variant of the selection property (44). Precisely, define

$$\bar{\tau} \overset{\text{def}}{=} (\tau^{(1)} + \ldots + \tau^{(S-R)}, \tau^{(S-R+1)}, \ldots, \tau^{(S-R+S_1)}),$$

then B fulfills (44) with respect to such a collection of vectors (which is S-admissible).

We prove (45) by induction on R. The formula is true for $R = 0$ with $S_1 = 0$. We assume it is true for R and we prove it for $R+1$.

Applying (38) to (45), such a quantity turns out to be the sum of

$$\widetilde{B^0}_{K \cdot \tau^{(1)}, \ldots, K \cdot \tau^{(S-R-1)}} \left[(N+1)^{S_1+1} \prod_{i=1}^{S_1+1} \delta_{\lfloor K \cdot \tau^{(S-R+i)} \rceil} \right] , \qquad (46)$$

and of the quantities

$$\widetilde{B^j}_{K \cdot \tau^{(1)}, \ldots, K \cdot \tau^{(j)} + K \cdot \tau^{S-R}, \ldots, K \cdot \tau^{(S-R-1)}} \left[(N+1)^{S_1} \prod_{i=1}^{S_1} \delta_{\lfloor K \cdot \tau^{(S-R+i)} \rceil} \right] , \qquad (47)$$

so also at step $R+1$ we have the wanted representation. Still we have to verify that the new τ's form an admissible collection and that the involved coefficients A_L fulfill the selection rule (44).

We start by the τ's in the term (46). In this term the collection of the τ's coincides with the previous one, so it is still an admissible collection.

We come to the selection rule on B^0. The new collection $\bar{\tau}$ is

$$(\tau^{(1)} + \ldots + \tau^{(S-R-1)}, \tau^{(S-R+1)}, \ldots, \tau^{(S-R-1+S_1)}) ,$$

so that one is adding a further restriction on the values of the indexes of the A_L's entering in B^0, namely that L_{S-R} (and thus also the indexes labeled by $\text{supp}(\tau^{(S-R)})$) must be different from the other indexes. But, by formula (39) one has

$$B^0_{L_1, \ldots, L_{S-R-1}} = B_L \big|_{L_{S-R} \neq L_1, \ldots, L_{S-R} \neq L_{S-R-1}} ,$$

and therefore the selection rule is fulfilled.

We come to (47). In this term the elements of the collection of the τ's are the same as before except for the fact that τ^{S-R} is missing and that $\tau^{(j)}$ is substituted by

$$\tau^{(j)\pm} \overset{\text{def}}{=} \tau^{(j)} \pm \tau^{(S-R)} .$$

By the properties of the supports one has thus

$$\text{supp}(\tau^{(j)\pm}) = \text{supp}(\tau^{(j)}) \cup \text{supp}(\tau^{(S-R)}) \ , \tag{48}$$

from which one immediately sees that properties 1) and 2) of definition 3.2 are fulfilled also by the new collection.

Concerning the selection property for B, we just notice that the new collection $\bar{\tau}$ coincides with the old one, and therefore the new B's automatically fulfill the needed property. Q.E.D

We have now to insert the averages of the $\hat{p}$'s. To get a useful formula we have to analyze quite in detail the corresponding terms.

First remark that a possible expression of $\langle \hat{p}_{k_1}...\hat{p}_{k_s} \rangle$ is constructed as follows: consider the distinct partitions of $1,...,s$ into subsets composed by an even number of elements. Let $\Sigma^s \equiv (\Sigma_1^s, ..., \Sigma_{s_1}^s)$ be one of these partitions (of course $s_1 \leq s/2$), denote $\ell_J \stackrel{\text{def}}{=} \#\Sigma_J^s$, and let $j_1^{(J)}, ..., j_{\ell_J}^{(J)} \in \Sigma_J^s$ be the elements of Σ_J^s, then to the partition Σ^s we associate the quantity

$$D_{\Sigma^s} \stackrel{\text{def}}{=} \left[\prod_{J=1}^{s_1} \langle \hat{p}_1^{\ell_J} \rangle \right] 2^{s/2} \beta^{s/2} = \prod_{J=1}^{s_1} (\ell_J - 1)!! \tag{49}$$

and one has

$$\langle \hat{p}_{k_1}...\hat{p}_{k_s} \rangle = \frac{1}{2^{s/2}\beta^{s/2}} \sum_{\Sigma^s} D_{\Sigma^s} \delta_{k_1,...,k_s}^{\Sigma^s} \ , \tag{50}$$

$$\delta_{k_1,...,k_s}^{\Sigma^s} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } k_{j_1^{(J)}} = ... = k_{j_{\ell_J}^{(J)}} \ \forall J \\ 0 & \text{otherwise} \end{cases} \ ,$$

where, of course the sum is over all the distinct partitions described above.

Remark 5. Defining $\mathbf{k}^{(J)} \equiv (k_1^{(J)}, ..., k_s^{(J)})$ with $k_{j_i^{(J)}}^{(J)} = 1$ for $i = 1, ..., \ell_J$ and zero otherwise, one has

$$\delta_{\mathbf{k}}^{\Sigma} \neq 0 \iff \mathbf{k} = \sum_{J=1}^{s_1} n_J \mathbf{k}^{(J)} \ , \tag{51}$$

for some integers n_J. This means that for every fixed partition Σ, the subspace of vectors in $\mathbb{Z}^s$ such that $\delta_{\mathbf{k}}^{\Sigma} \neq 0$ has dimension $s_1 \leq s/2$, where the equality is attained only if $\ell_J = 2, \forall J$.

For this reason the partitions for which $\ell_J = 2$ for all J's will play a special role. In such a case one can write

$$\delta_{\mathbf{k}}^{\Sigma^s} = \prod_{i=1}^{s/2} \delta_{\mathbf{k} \cdot \tau^{(i)}},$$

where $\tau = \{\tau^{(i)}\}_{i=1}^{s/2}$ is an s-admissible collections s.t. each of the $\tau^{(i)}$'s has only one component equal to 1 and one component equal to -1.

We will denote by $\mathcal{T}^s$ the set of the s-admissible collections with such a property.

We will denote by $\mathcal{S}_4^s$ the set of partitions Σ^s such that $\ell_J \geq 4$ for at least one J.

In order to obtain a useful expression for the covariance we consider $\mathcal{T}^{s+s'}$ and decompose it as

$$\mathcal{T}^{s+s'} = \mathcal{T}^s \oplus \mathcal{T}^{s'} = \mathcal{T}^s \cup \mathcal{T}^{s'} \cup \mathcal{T}^{s,s'} \ , \tag{52}$$

where $\mathcal{T}^{s,s'}$ is composed by the $(s + s')$-admissible collections s.t. at least one of the vectors $\tau^{(i)}$ has one non vanishing component in the $\mathcal{T}^s$ and one nonvanishing component in $\mathcal{T}^{s'}$.

Lemma 7 *The following formula holds*

$$a\langle \hat{p}_{k_1} \cdots \hat{p}_{k_s} \hat{p}_{k'_1} \cdots \hat{p}_{k'_{s'}} \rangle - b\langle \hat{p}_{k_1} \cdots \hat{p}_{k_s} \rangle \langle \hat{p}_{k'_1} \cdots \hat{p}_{k'_{s'}} \rangle$$

$$= \frac{1}{(2\beta)^{\frac{s+s'}{2}}} \left[\sum_{\substack{\tau \in \mathcal{T}^s \\ \tau' \in \mathcal{T}^{s'}}} (a - b) \left(\prod_{i=1}^{s/2} \delta_{k \cdot \tau^{(i)}} \right) \left(\prod_{i=1}^{s'/2} \delta_{k \cdot \tau'^{(i)}} \right) \right.$$

$$\left. + a \sum_{\tau \in \mathcal{T}^{s,s'}} \prod_{i=1}^{(s+s')/2} \delta_{K \cdot \tau^{(i)}} + \sum_{\Sigma^{s+s'} \in \mathcal{S}_4^{s+s'}} E_{\Sigma^{s+s'}} \delta_K^{\Sigma^{s+s'}} \right]$$

where $K = (k, k')$ and $E_{\Sigma^{s+s'}}$ is a (possibly vanishing) constant fulfilling

$$|E_{\Sigma^{s+s'}}| \leq C(|a| + |b|).$$

The proof is a simple computation which is omitted.

We have now at hand the tools that enable us to estimate σ_f. In the forthcoming formulas we will use the following notations: $S_1 \leq s_1 + s'_1$ is an integer and

$$k = (k_1, ..., k_s) \ , \quad k' = (k'_1, ..., k'_s)$$
$$k^{(1)} = (k_1, ..., k_{s_1}) \ , \quad k^{(2)} = (k_{s_1+1}, ..., k_s)$$
$$k'^{(1)} = (k'_1, ..., k'_{s'_1}) \ , \quad k'^{(2)} = (k'_{s'_1+1}, ..., k'_s)$$
$$K = (k_1, ..., k_s, k'_1, ..., k'_s) \ ,$$
$$K^{(1)} = (k_1, ..., k_{s_1}, k'_1, ..., k'_{s'_1}) \ , \quad K^{(2)} = (k_{s_1+1}, ..., k_s, k'_{s'_1+1}, ..., k'_s) \ ,$$

finally $s_2 := s - s_1$ and $s'_2 := s - s'_1$.

First remark that, due to Lemma 6 and (50) one has that σ_f^2 is estimated

by the sum of finitely many terms of the form

$$
\frac{C}{(N+1)^{s-2}} \|f\|_+^2 \sum_{(k,k')\in\mathbb{Z}^{2s}} \delta_{\lfloor \tilde{\tau}\cdot k \rceil} \delta_{\lfloor \tilde{\tau}'\cdot k' \rceil}
$$

$$
\left[\langle \hat{p}_{k_{s_1+1}} \cdots \hat{p}_{k_s} \hat{p}_{k'_{s'_1+1}} \cdots \hat{p}_{k'_s} \rangle \langle r_{l_1}....r_{l_{s_1}} r_{l'_1}...r_{l'_{s_1}} \rangle \right.
$$

$$
\left. - \langle \hat{p}_{k_{s_1+1}} \cdots \hat{p}_{k_s} \rangle \langle \hat{p}_{k'_{s'_1+1}} \cdots \hat{p}_{k'_s} \rangle \langle r_{l_1}....r_{l_{s_1}} \rangle \langle r_{l'_1}...r_{l'_{s_1}} \rangle \right]
\tag{53}
$$

$$
\times \left(\prod_{i=1}^{S_1} \delta_{\lfloor \tau^{(i)}\cdot K^{(1)} \rceil}(N+1) \right) \frac{1}{(N+1)^{(s_1+s'_1)/2}}
$$

where $l_1,...,l_{s_1}, l'_1,...,l'_{s'_1}$ fulfills the selection rule (44) with respect to the partition $\tau^{(i)}$.

According to Lemma 7 one has that (53) can be written as

$$
\Sigma_1 + \Sigma_2 + \Sigma_3 \ ,
$$

where

$$
\Sigma_1 := \frac{C}{(N+1)^{s-2}} \|f\|_+^2 \sum_{(k,k')\in\mathbb{Z}^{2s}} \delta_{\lfloor \tilde{\tau}\cdot k \rceil} \delta_{\lfloor \tilde{\tau}'\cdot k' \rceil} \frac{1}{(2\beta)^{\frac{s_2+s'_2}{2}}}
\tag{54}
$$

$$
\sum_{\substack{\tau'\in\mathcal{T}^{s_2} \\ \tau''\in\mathcal{T}^{s'_2}}} \left(\langle r_{l_1}....r_{l_{s_1}} r_{l'_1}...r_{l'_{s_1}} \rangle - \langle r_{l_1}....r_{l_{s_1}} \rangle \langle r_{l'_1}...r_{l'_{s_1}} \rangle \right) \left(\prod_{i=1}^{s_2/2} \delta_{k^{(2)}\cdot\tau'^{(i)}} \right) \left(\prod_{i=1}^{s'_2/2} \delta_{k^{(2)}\cdot\tau''^{(i)}} \right)
$$

$$
\tag{55}
$$

$$
\times \left(\prod_{i=1}^{S_1} \delta_{\lfloor \tau^{(i)}\cdot K^{(1)} \rceil}(N+1) \right) \frac{1}{(N+1)^{(s_1+s'_1)/2}}
\tag{56}
$$

$$
\Sigma_2 := \frac{C}{(N+1)^{s-2}} \|f\|_+^2 \sum_{(k,k')\in\mathbb{Z}^{2s}} \delta_{\lfloor \tilde{\tau}\cdot k \rceil} \delta_{\lfloor \tilde{\tau}'\cdot k' \rceil} \frac{1}{(2\beta)^{\frac{s_2+s'_2}{2}}}
\tag{57}
$$

$$
\langle r_{l_1}....r_{l_{s_1}} r_{l'_1}...r_{l'_{s_1}} \rangle \sum_{\tau'\in\mathcal{T}^{s,s'}} \prod_{i=1}^{(s_2+s'_2)/2} \delta_{K^{(2)}\cdot\tau'^{(i)}}
\tag{58}
$$

$$
\times \left(\prod_{i=1}^{S_1} \delta_{\lfloor \tau'^{(i)}\cdot K^{(1)} \rceil}(N+1) \right) \frac{1}{(N+1)^{(s_1+s'_1)/2}}
\tag{59}
$$

18

$$\Sigma_3 := \frac{C}{(N+1)^{s-2}} \|f\|_+^2 \sum_{(k,k')\in\mathbb{Z}^{2s}} \delta_{\lfloor \tilde{\tau}\cdot k\rceil}\delta_{\lfloor \tilde{\tau}'\cdot k'\rceil} \frac{1}{(2\beta)^{\frac{s_2+s_2'}{2}}} \tag{60}$$

$$\sum_{\Sigma^{s_2+s_2'}\in\mathcal{S}_4^{s_2+s_2'}} E_{\Sigma^{s_2+s_2'}}\delta_{K^{(2)}}^{\Sigma^{s_2+s_2'}} \tag{61}$$

$$\times \left(\prod_{i=1}^{S_1}\delta_{\lfloor \tau^{(i)}\cdot K^{(1)}\rceil}(N+1)\right)\frac{1}{(N+1)^{(s_1+s_1')/2}} \tag{62}$$

where the indexes $l_1,...,l_{s_1},l_1',...,l_{s_1'}'$ fulfills the selection rule (44) with respect to the collection τ.

Lemma 8 *The following estimate holds*

$$|\Sigma_3| \leq \frac{C(N+1)\|f\|_+^2}{\beta^s} \tag{63}$$

Proof. For this computation we can neglect the delta's in (60). Every δ in (62) reduces by 1 the effective dimension of the lattice over which $K^{(1)}$ runs. Thus, the effective dimension of such a lattice is $s_1+s_1'-S_1\geq 0$. By remark 3.2, $K^{(2)}$ runs over a lattice of dimension at most $(s-s_1+s-s_1')/2-1$. Thus, the number of nonvanishing terms is at most of order

$$(N+1)^{\wedge}\left(s+\frac{s_1+s_1'}{2}-S_1-1\right),$$

while, counting the powers of $(N+1)$, one has that each term has size controlled by a constant times

$$\frac{\|f\|_+^2}{\beta^s}(N+1)^{\wedge}\left(S_1-\frac{s_1+s_1'}{2}-s+2\right),$$

so that the result follows.   Q.E.D

We have now to understand when it can happen that the deltas coming from the zero momentum conditions are not independent of the other deltas (more precisely the corresponding τ vectors). This is analyzed by the forthcoming Lemma 9.

Write $\mathbb{Z}^{2s} = \mathbb{Z}^s \oplus \mathbb{Z}^s$ and denote by P_1 the projection on the first factor and by P_2 the projection on the second one. Then the following Lemma holds.

Lemma 9 *Let $\tau^{(i)}$ be a $(2s)$-admissible collection of vectors and let $\tilde{\tau}\in\mathbb{Z}_3^{2s}$ be a vector with support equal to $(1,...,s)$, namely s.t. $\tilde{\tau}_i \neq 0 \ \forall i=1,...,s$ and $\tilde{\tau}_i=0$ $\forall i=s+1,...,2s$. Assume that there exists $\bar{\imath}$ s.t. $P_1\tau^{(\bar{\imath})}\neq 0$ and $P_2\tau^{(\bar{\imath})}\neq 0$, then $\tilde{\tau}$ is linear independent of the vectors $\tau^{(i)}$.*

19

Proof. Consider the equation

$$c\tilde{\tau} + \sum_i c_i \tau^{(i)} = 0 \; ;$$

applying P_1 and P_2 one gets

$$cP_1\tilde{\tau} + \sum_i c_i P_1 \tau^{(i)} = 0 \; , \tag{64}$$

$$\sum_i c_i P_2 \tau^{(i)} = 0 \; . \tag{65}$$

Since the supports of the $\tau^{(i)}$'s are disjoint, (65) implies $c_i = 0$ for all i's s.t. $P_2 \tau^{(i)} \neq 0$. In particular one has $c_{\bar{\imath}} = 0$. There exists a component of $P_1 \tau^{(\bar{\imath})}$ which is different from zero. Assume for definiteness that it is the first one. It follows that all the other vectors $\tau^{(i)}$ have first component equal to zero. Thus, taking the first component of (64) one gets

$$c\tilde{\tau}_1 + c_{\bar{\imath}} \tau_1^{(\bar{\imath})} = c\tilde{\tau}_1 = 0 \implies c = 0 \; ,$$

which is the claimed independence. Q.E.D

In particular it follows that, in the expression of Σ_2, at least one of the $\tilde{\tau}$'s is independent of all the other τ's. Thus the Following Lemma holds

Lemma 10 *The following estimate holds*

$$|\Sigma_2| \leq \frac{C(N+1)\,\|f\|_+^2}{\beta^s} \tag{66}$$

Proof. Every δ reduces by 1 the effective dimension of the lattice over which K runs, provided the corresponding vectors τ are independent. In the considered case the effective dimension is at most

$$2s - (S_1 + s - (s_1 + s_1')/2),$$

thus, counting the powers of $(N+1)$ as in the proof of Lemma 8 one gets the result. Q.E.D

To estimate Σ_1 one has also to consider the dependent case. This is contained in the proof of the following Lemma.

Lemma 11 *The following estimate holds*

$$|\Sigma_1| \leq \frac{C(N+1)\,\|f\|_+^2}{\beta^s} \tag{67}$$

Proof. The case in which the $\tau's$ are independent is dealt with as in the proof of Lemma 10. Consider now the case in which they are dependent. In such a case, all the elements of τ do not mix k and k', by the selection rule (44), it follows that the indexes $l_1, ..., l_{s_1}$ are all different of the indexes $l_1', ..., l_{s_1'}'$, thus the covariance in (55) can be estimated using eq. (30), which adds a power of N at the denominator. Thus the result follows also in this case. Q.E.D

4   Proof of Lemma 3

As both ν and ω are bounded from above, $h_1(\nu)$ can diverge only when the denominator at the r.h.s. of (10) vanishes. We prove that, under the assumptions of the theorem, the numerator vanishes at the same points and the ratio stays bounded.

First, remark that the hypotheses on g and the explicit form of $\omega(x) = 2\sin(\pi x/2)$ imply that ν has derivative bounded by $K \stackrel{\text{def}}{=} 2c_2 + \pi(c_0 + c_2/2)$. In turn, this implies also that the numerator is bounded by $3K$.

Consider now the case in which $z = x+y$. When $\tau_1 = \tau_2 = \tau_3$, the inequality $\sin(\pi x/2) \geq x$ implies that the denominator is bigger than $3x$. Using $\nu(x) \leq Kx$ one has that the ratio which defines h_1 is bounded by K.

When $\tau_1 = -\tau_2 = \tau_3$, using the fact that ω is a non–decreasing function one can bound the denominator from below by x; in turn, using $|\nu(x+y) - \nu(y)| \leq Kx$ one has that the numerator is smaller than $2Kx$, so that the ratio is smaller than $2K$. The same upper bound holds when $-\tau_1 = \tau_2 = \tau_3$.

The case $\tau_1 = \tau_2 = -\tau_3$ is more complicated. We rewrite the numerator of (10) as a function of $g(x)$ and get

$$\nu(x) + \nu(y) - \nu(x+y) = c_0 \left[\omega(x) + \omega(y) - \omega(x+y) \right] + f(x) + f(y) - f(x+y)$$

where we have put $f(x) \stackrel{\text{def}}{=} \omega(x)(g(x) - c_0)$. Due to its definition and to the hypothesis $g'(0) = 0$, $f(x)$ has a zero of third order at 0. Suppose, without loss of generality, that $y \leq x$, then there exists a constant C, such that one has

$$|f(x) - f(x+y)| \leq Cc_2 x^2 y , \quad |f(y)| \leq Cc_2 y^3 \leq Cc_2 x^2 y .$$

On the other hand, for the denominator one has

$$\omega(x) + \omega(y) - \omega(x+y) \geq 2\sin(\pi y/2)(1 - \cos(\pi x/2)) \geq x^2 y , \qquad (68)$$

where in the first inequality use is made of the addition formulas for the sine, in the second of the inequalities $\sin(\pi x/2) \geq x$ and $\cos(\pi x/2) \leq 1 - x^2/2$. Thus we have

$$\left| \frac{\nu(x) + \nu(y) - \nu(x+y)}{\omega(x) + \omega(y) - \omega(x+y)} \right| \leq C(c_0 + c_2) ,$$

with a suitable redefinition of the constant C. A similar arguments, exchanging upper with lower bounds, shows that if $g'(0) \neq 0$ the ratio defining $h_1(\nu)$ is unbounded thus proving the last statement of the theorem.

Consider now the case of $z = 2 - x - y$. Here, when $\tau_1 = \tau_2 = \tau_3$ everything is trivial, because at least one among x, y and z is greater than $2/3$, so that the denominator is larger than $2\sin(\pi/3) = \sqrt{3}$. There remains the case in which one sign is different from the others: as the role of x, y and z is symmetric, we consider only the possibility $\tau_1 = \tau_2 = -\tau_3$. Since $\omega(2 - \alpha) = \omega(\alpha)$, one has

$$\omega(x) + \omega(y) - \omega(2 - x - y) = \omega(x) + \omega(y) - \omega(y + z) ,$$

so that we can bound from below the denominator making use of inequality (68); assuming again without loss of generality that $y \leq x$. In the present case, however, one has $x + y = 2 - z \geq 1$, so that $x \geq 1/2$ and we get

$$\omega(x) + \omega(y) - \omega(2 - x - y) \geq \frac{y}{4} \ .$$

The numerator is bounded according to

$$|\nu(x) - \nu(2 - x - y)| \leq K|2(1 - x) - y| \leq 3Ky \ , \quad |\nu(y)| \leq Ky \ ,$$

where we used the inequality $x \geq 1 - y$. This suffices to bound uniformly the considered ratio also in this case, and thus to complete the proof.　　　Q.E.D

A　Proof of Lemma 4

As already stated, the main tool needed in the proof of this lemma is an estimate of the error introduced by the fact that the mean values of interest are computed with respect to the measure in which all r's are independent, rather than with respect to the measure in which they are conditioned to have vanishing sum. Indeed, if the r's are independent, estimates (29) and (30) are trivial consequences of the properties of Gaussian integration (the r.h.s. of (30) even vanishes). The estimates of the error, as first pointed out by Khinchin (see [18]), can be obtained by using a local central limit theorem.

We begin by considering the extended configuration space Γ, which coincides with $\mathbb{R}^{N+1}$ endowed with the probability measure μ_γ with density

$$\rho_\gamma^{N+1}(r_0, \ldots, r_N) \stackrel{\text{def}}{=} \frac{1}{(q_\gamma(\beta))^{N+1}} \prod_{j=0}^{N} \exp\left(-\gamma r_j - \beta V(r_j)\right) \ , \tag{69}$$

in which the normalization constant is defined by

$$q_\gamma(\beta) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \exp\left(-\gamma r - \beta V(r)\right) \mathrm{d}r \ .$$

Denoting by

$$\Sigma_x \stackrel{\text{def}}{=} \left\{ r \in \Gamma \ : \ R \stackrel{\text{def}}{=} \sum_j r_j = x \right\} \ , \tag{70}$$

the configuration space for our dynamical system corresponds to Σ_0. Moreover, the probability measure induced on it by the Gibbs measure is exactly the measure μ_γ with density (69) and $\gamma = 0$, conditioned on Σ_0. This means that, if M is a subset of Σ_0, one has

$$\mathbf{P}(M) = \mu_0(M|\Sigma_0) \ .$$

22

This suggests the introduction of the structure function $\Omega_{N+1}(x)$, defined by

$$\Omega_{N+1}(x) \overset{\text{def}}{=} \frac{\mathrm{d}}{\mathrm{d}x} \int_{\sum r_i \leq x} e^{-\beta \sum V(r_i)} \mathrm{d}r_0 \ldots \mathrm{d}r_N =$$

$$= \int e^{-\beta \sum V(x_i - x_{i-1})} \mathrm{d}x_1 \ldots \mathrm{d}x_N , \quad x_0 = 0 , \ x_{N+1} = x ,$$

which is the probability density in Γ that R has value x, if the r's are distributed with the measure μ_0. Notice that $Z(\beta) = \Omega_{N+1}(0)$, which will be used below.

Now the idea is that, since the variable r_i are independently distributed, one can use some kind of central limit theorem to expand Ω_N as a simple function around zero for large N, so that the conditioned measure becomes easily tractable. Indeed, as first pointed out by Cramér, this can be accomplished, if one considers the conjugate distribution

$$U_N^{(\gamma)}(x) = \frac{1}{\Phi_N(\gamma)} e^{-\gamma x} \Omega_N(x) , \tag{71}$$

where $\Phi_N(\gamma) = (q_\gamma(\beta))^N$, which corresponds to the probability distribution of the sum of independently distributed random variables distributed with the density ρ_γ^N introduced above by (69). In fact, while the central limit theorem gives no direct information on $\Omega_N(x)$, for x near 0 (because the mean value of x is very far from zero), such information can be obtained by applying the central limit theorem to $U_N^{(\gamma)}(x)$, if one chooses a value of γ, call it θ, such that[4]

$$\int x U_N^{(\theta)}(x) \mathrm{d}x = 0 .$$

Then the central limit theorem can be locally applied to $U_N^{(\theta)}(x)$ near zero and then translated into a property of $\Omega_N(x)$ by inverting (71).

We will use the following local version of the central limit theorem, in which the conjugate distribution is approximated as a function of the functions $q_j(x)$ defined as

$$q_j(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \sum H_{j+2s}(x) \prod_{m=1}^{j} \frac{1}{k_m!} \left(\frac{\gamma_{m+2}}{(m+2)! b^{m+2}} \right)^{k_m} ,$$

where $H_m(x)$ are Hermite polynomials, γ_m is the m-th cumulant of $u^{(\theta)}(x)$ and b its standard deviation, while the sum should be taken over all the non-negative

---

[4]Notice that the integral equation

$$0 = \int_{-\infty}^{\infty} x U_N^{(\theta)}(x) \, \mathrm{d}x = \sum_{j=1}^{N} \frac{1}{q_\theta(\beta)} \left(\int_{-\infty}^{+\infty} r \exp\left(-\theta r - \beta V(r)\right) \mathrm{d}r \right)$$

$$\Rightarrow \int_{-\infty}^{+\infty} r \exp\left(-\theta r - \beta V(r)\right) \mathrm{d}r = 0 . \tag{72}$$

admits a unique solution for all $\beta > 0$ so that $q_\theta(\beta)$ is well defined. It is then obvious that θ depends on β but not on N, so that $q_\theta(\beta)$ is a function of β only. Furthermore, since we are interested in the large-β regime, we point out that $\lim_{\beta \to \infty} \theta = -\alpha$.

integer solutions $(k_1, \ldots, k_j)$ of the equalities $k_1 + 2k_2 + \ldots + jk_j = j$, and $s = k_1 + k_2 + \ldots k_j$.

Theorem 2 (Local central limit, Theorem VII.15 of [23]) *There exist $C, N_0, \beta_0 > 0$ such that, for $N > N_0$, $\beta > \beta_0$, one has*

$$\left| U_N^{(\theta)}(x) - \frac{1}{\sqrt{2\pi N b^2}} \exp\left(-\frac{x^2}{2Nb^2} \right) - \sum_{j=1}^{2} \frac{q_j(x/(\sqrt{N}b))}{N^{(j+1)/2}b} \right| \leq \frac{1}{N^{3/2}b} , \qquad (73)$$

uniformly in x.

From this theorem we can infer (cf. [19], Corollary 1.4 of Appendix 2) an estimate on the deviation of the expectations taken with respect to the Gibbs measure from that taken with respect to the measure $\mu_\theta \equiv \mu_\gamma|_{\gamma=\theta}$. We denote the expectation of f with respect to the Gibbs measure by $\langle f \rangle$, while that with respect to μ_θ by $\langle f \rangle_\theta$. Moreover, given a vector $j = (j_0, \ldots, j_N) \in \{0,1\}^{N+1}$ and a vector $r \in \Gamma$, we denote by $\tilde{r} \in \mathbb{R}^{|j|}$ the collection r_i, $i \in \text{supp} j$.

Corollary 4 *Fix $\bar{\beta} > 0$ and let $f(\tilde{r}) : \mathbb{R}^{|j|} \to \mathbb{R}$ have a finite second order moment with respect to μ_θ, uniformly for all $\beta > \bar{\beta}$. Then there exist C, N_0 and β_0 such that, for all $N > N_0$, $\beta > \beta_0$, one has*

$$|\langle f \rangle - \langle f \rangle_\theta| \leq C \frac{|j|}{N} \sqrt{\langle f^2 \rangle_\theta - \langle f \rangle_\theta^2}$$

Proof. We denote $J \stackrel{\text{def}}{=} |j|$ and, for the sake of simplicity, we assume $\text{supp} j = \{1, \ldots, J\}$; the general case is dealt with exactly in the same way. The average $\langle f \rangle$ can be written as follows

$$
\begin{aligned}
\langle f \rangle &= \int_{\mathbb{R}^{N+1}} f(\tilde{r}) \frac{\exp\left(-\beta \sum V(x_i - x_{i-1}) \right)}{Z(\beta)} \mathrm{d}x_1 \ldots \mathrm{d}x_N \\
&= \int_{\tilde{\Gamma}} f(\tilde{r}) \frac{\Omega_{N+1-J}(-w)}{\Omega_{N+1}(0)} \mathrm{d}\tilde{v} ,
\end{aligned}
$$

where $w \stackrel{\text{def}}{=} \sum_{i=0}^{J} r_i$, and $\tilde{\Gamma} \equiv \mathbb{R}^{N+1-J}$ is endowed with the measure with volume element $\mathrm{d}\tilde{v} \stackrel{\text{def}}{=} \prod_{i=0}^{J} e^{-\beta V(r_i)} \mathrm{d}r_i$, while Ω_{N+1-J} is the structure function for the system in which the first J relative distances r_i are subtracted.

Now the ratio $\Omega_{N+1-J}(-w)/\Omega_{N+1}(0)$ can be expressed in terms of $U_{N+1}^{(\theta)}(x)$, by a simple inversion of (71), as

$$\frac{\Omega_{N+1-J}(-w)}{\Omega_{N+1}(0)} = \frac{U_{N+1-J}^{(\theta)}(-w)}{U_{N+1}^{(\theta)}(0)} \frac{e^{-\theta w}}{q_\theta(\beta)^J} ,$$

where use has been made of the explicit form of $\Phi_N(\theta)$. So, the difference $|\langle f \rangle - \langle f \rangle_\theta|$ may be written as

$$\left| \int_{\tilde{\gamma}} \mathrm{d}\tilde{v} \frac{e^{-\theta w}}{q_\theta(\beta)^J} f(\tilde{r}) \left(\frac{U_{N+1-J}^{(\theta)}(-w)}{U_{N+1}^{(\theta)}(0)} - 1 \right) \right| . \qquad (74)$$

24

Using the relations

$$\int_{\tilde{\gamma}} \mathrm{d}\tilde{v} \frac{e^{-\theta w}}{q_\theta(\beta)^J} \frac{U_{N+1-J}^{(\theta)}(-w)}{U_{N+1}^{(\theta)}(0)} = \langle 1 \rangle = 1 = \langle 1 \rangle_\theta = \int_{\tilde{\gamma}} \mathrm{d}\tilde{v} \frac{e^{-\theta w}}{q_\theta(\beta)^J}$$

one can rewrite the difference $|\langle f \rangle - \langle f \rangle_\theta|$ as follows

$$|\langle f \rangle - \langle f \rangle_\theta| = \left| \int_{\tilde{\gamma}} \mathrm{d}\tilde{v} \frac{e^{-\theta w}}{q_\theta(\beta)^J} \left(f(\tilde{r}) - \langle f \rangle_\theta \right) \left(\frac{U_{N+1-J}^{(\theta)}(-w)}{U_{N+1}^{(\theta)}(0)} - 1 \right) \right| . \qquad (75)$$

Noting that $|e^{-x^2} - 1| \leq x^2$ and $q_j(x) \leq c_j(\beta_0)$, we obtain from Theorem 2 that

$$\left| \frac{U_{N+1-J}^{(\theta)}(-w)}{U_{N+1}^{(\theta)}(0)} - 1 \right| \leq K(N_0, \beta_0) \frac{J}{N} \left(1 + \frac{w^2}{Jb^2} \right) ,$$

for N large enough. By Schwartz inequality the thesis follows. Q.E.D

In order to conclude the proof of Lemma 4 it is now sufficient to apply corollary 4 to the functions of interest, making use of the properties of Gaussian integration to estimate $\langle \cdot \rangle_\theta$.

References

[1] Bambusi, D., Kappeler, T., Paul, T.: De Toda à KdV. C. R. Math. Acad. Sci. Paris **347**(17-18), 1025–1030 (2009). DOI 10.1016/j.crma.2009.07.002. URL http://dx.doi.org/10.1016/j.crma.2009.07.002

[2] Bambusi, D., Ponno, A.: On metastability in FPU. Comm. Math. Phys. **264**(2), 539–561 (2006). DOI 10.1007/s00220-005-1488-1. URL http://dx.doi.org/10.1007/s00220-005-1488-1

[3] Benettin, G., Christodouli, H., Ponno, A.: The fermi-pasta-ulam problem and its underlying integrable dynamics. J. Stat. Phys. **152**, 195–212 (2013)

[4] Carati, A.: An averaging theorem for Hamiltonian dynamical systems in the thermodynamic limit. J. Stat. Phys. **128**(4), 1057–1077 (2007). DOI 10.1007/s10955-007-9332-y. URL http://dx.doi.org/10.1007/s10955-007-9332-y

[5] Carati, A., Galgani, L., Giorgilli, A., Paleari, S.: Fermi-Pasta-Ulam phenomenon for generic initial data. Phys. Rev. E (3) **76**(2), 022,104, 4 (2007). DOI 10.1103/PhysRevE.76.022104. URL http://dx.doi.org/10.1103/PhysRevE.76.022104

[6] Carati, A., Maiocchi, A.M.: Exponentially long stability times for a nonlinear lattice in the thermodynamic limit. Comm. Math. Phys. **314**(1), 129–161 (2012). DOI 10.1007/s00220-012-1522-z. URL http://dx.doi.org/10.1007/s00220-012-1522-z

[7] De Roeck, W., Huveneers, F.: Asymptotic localization of energy in non-disordered oscillator chains. arXiv:1305.512 (2013)

[8] Feller, W.: An introduction to probability theory and its applications. Vol. II. John Wiley & Sons Inc., New York (1966)

[9] Friesecke, G., Pego, R.L.: Solitary waves on Fermi-Pasta-Ulam lattices. I. Qualitative properties, renormalization and continuum limit. Nonlinearity **12**, 1601–1627 (1999)

[10] Friesecke, G., Pego, R.L.: Solitary waves on Fermi-Pasta-Ulam lattices. II. Linear implies nonlinear stability. Nonlinearity **15**, 1343–1359 (2002)

[11] Friesecke, G., Pego, R.L.: Solitary waves on Fermi-Pasta-Ulam lattices. III. Howland-type Floquet theory. Nonlinearity **17**, 207–227 (2004)

[12] Friesecke, G., Pego, R.L.: Solitary waves on Fermi-Pasta-Ulam lattices. IV. Proof of stability at low energy. Nonlinearity **17**, 229–251 (2004)

[13] Gallavotti, G. (ed.): The Fermi-Pasta-Ulam problem, *Lecture Notes in Physics*, vol. 728. Springer, Berlin (2008). DOI 10.1007/978-3-540-72995-2. URL http://dx.doi.org/10.1007/978-3-540-72995-2. A status report

[14] Giorgilli, A., Galgani, L.: Formal integrals for an autonomous Hamiltonian system near an equilibrium point. Celestial Mech. **17**(3), 267–280 (1978). DOI 10.1007/BF01232832. URL http://dx.doi.org/10.1007/BF01232832

[15] Giorgilli, A., Paleari, S., Penati, T.: Extensive adiabatic invariants for nonlinear chains. J. Stat. Phys. **148**(6), 1106–1134 (2012). DOI 10.1007/s10955-012-0568-9. URL http://dx.doi.org/10.1007/s10955-012-0568-9

[16] Giorgilli, A., Paleari, S., Penati, T.: An extensive adiabatic invariant for the klein-gordon model in the thermodynamic limit. preprint (2013)

[17] Hoffman, A., Wayne, C.E.: Counter-propagating two-soliton solutions in the Fermi-Pasta-Ulam lattice. Nonlinearity **21**(12), 2911–2947 (2008). DOI 10.1088/0951-7715/21/12/011. URL http://dx.doi.org/10.1088/0951-7715/21/12/011

[18] Khinchin, A.I.: Mathematical Foundations of Statistical Mechanics. Dover Publications Inc., New York, N. Y. (1949). Translated by G. Gamow

[19] Kipnis, C., Landim, C.: Scaling limits of interacting particle systems, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 320. Springer-Verlag, Berlin (1999)

[20] Maiocchi, A.M., Carati, A., Giorgilli, A.: A series expansion for the time autocorrelation of dynamical variables. J. Stat. Phys. **148**(6), 1054–1071 (2012). DOI 10.1007/s10955-012-0575-x. URL http://dx.doi.org/10.1007/s10955-012-0575-x

[21] Mizumachi, T.: Asymptotic stability of lattice solitons in the energy space. Comm. Math. Phys. **288**(1), 125–144 (2009). DOI 10.1007/s00220-009-0768-6. URL http://dx.doi.org/10.1007/s00220-009-0768-6

[22] Mizumachi, T.: Asymptotic stability of N-solitary waves of the FPU lattices. Arch. Ration. Mech. Anal. **207**(2), 393–457 (2013). DOI 10.1007/s00205-012-0564-x. URL http://dx.doi.org/10.1007/s00205-012-0564-x

[23] Petrov, V.V.: Sums of independent random variables. Springer-Verlag, New York (1975). Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82

[24] Schneider, G., Wayne, C.E.: Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model. In: International Conference on Differential Equations, Vol. 1, 2 (Berlin, 1999), pp. 390–404. World Sci. Publishing, River Edge, NJ (2000)