

An Extensive Adiabatic Invariant for the Klein–Gordon Model in the Thermodynamic Limit

Antonio Giorgilli, Simone Paleari and Tiziano Penati

Abstract. We construct an extensive adiabatic invariant for a Klein–Gordon chain in the thermodynamic limit. In particular, given a fixed and sufficiently small value of the coupling constant a , the evolution of the adiabatic invariant is controlled up to time scaling as $\beta^{1/a}$ for any large enough value of the inverse temperature β . The time scale becomes a stretched exponential if the coupling constant is allowed to vanish jointly with the specific energy. The adiabatic invariance is exhibited by showing that the variance along the dynamics, i.e. calculated with respect to time averages, is much smaller than the corresponding variance over the whole phase space, i.e. calculated with the Gibbs measure, for a set of initial data of large measure. All the perturbative constructions and the subsequent estimates are consistent with the extensive nature of the system.

1. Introduction and Statement of Results

In the quest for a mathematically rigorous foundation of Statistical Physics in general, and Statistical Mechanics in particular, despite many efforts and recent successes, a lot of work is still to be done. More specifically, if one considers an Hamiltonian system, instead of some ad hoc model, for the microscopic description of large systems, the behaviour over different long time scales is often still a challenge. One of the possible, and natural strategies, is to apply the techniques and results of Hamiltonian perturbation theory to large systems, with particular attention to the thermodynamic limit, i.e. when the number of degrees of freedom grows very large, at fixed, non-vanishing, specific energy. The present paper is concerned with the existence of an adiabatic invariant for an arbitrarily large one-dimensional Klein–Gordon chain, with estimates uniform in the size of the system.

It is well known that results like the KAM and the Nekhoroshev theorems stated for finite-dimensional systems (see e.g. [2, 25, 30–33]) appear to be somewhat useless as the number N of degrees of freedom of the system grows, for the estimated dependence on N of the constants involved is usually very bad, and in particular the (specific) energy thresholds do vanish in the limit $N \rightarrow \infty$. It is worth to mention however that a first theoretical result at finite specific energy, hence with estimates uniform in N , and on an average time scale can be found in [15]. Extensions in the infinite-dimensional case have been made (see e.g. [4, 12, 13, 17, 26, 27, 34, 36] for the case of partial differential equations, or [3, 9, 20] for the case of lattices), but always for finite energy, i.e. for zero specific energy. Our aim is precisely to remove such a drawback, producing a long time estimate for finite *specific* energy.

In the present paper, we consider the Hamiltonian

$$H(x, y) = \frac{1}{2} \sum_{j=1}^N \left[(y_j^2 + x_j^2) + a(x_j - x_{j-1})^2 + \frac{1}{2}x_j^4 \right], \quad (1.1)$$

of a Klein–Gordon chain with N degrees of freedom, periodic boundary conditions $x_0 = x_N$, and coupling constant a .

A previous investigation of a similar model has been made in [24]. In that paper, a first-order (in the sense of perturbation theory) adiabatic invariant has been analytically constructed. Moreover, by numerical investigation it has been shown that the adiabatic invariance persists for times much longer than those predicted by the first-order theoretical analysis. Thus, the model appeared to be worth of further theoretical investigation.

A very recent breakthrough in this direction is represented by the paper [16], which exploits the idea of complementing the perturbation estimates with probabilistic techniques, thus producing a control of the long time evolution in the thermodynamic limit. We will come back later to the relation between that paper and the present work.

Let us give a brief sketch of our procedure. The basic idea of both the quoted works is to avoid the usual procedure of introducing normal modes for the quadratic part of the Hamiltonian (1.1), thus considering the model as a set of identical harmonic oscillators with a coupling which includes a small quadratic term describing a nearest neighbours interaction controlled by the small parameter a . We construct an extensive adiabatic invariant as follows. First, as in [24], we exploit a transformation of the quadratic part of the Hamiltonian into the sum of two terms in involution, one of them including all resonant coupling terms. The relevant fact is that the transformation preserves the extensive nature of the system and produces new coordinates which are each exponentially localized around the corresponding original ones. As a subsequent step, the perturbation process is performed here at higher order. Thus, we produce an adiabatic invariant which still preserves both the extensive nature of the system and the exponential decay of the interaction with the distance. Furthermore, we produce estimates which are uniform in the number N of degrees of freedom.

We stress that our model contains two independent perturbation parameters, namely (i) the coupling parameter a , and (ii) the specific energy ϵ . This is a point that deserves particular consideration. We pay special attention in keeping these two parameters separated, so that we can deal with the physically sound hypothesis that the coupling parameter a is fixed, and the inverse temperature β grows arbitrarily large. Actually, the main Theorem 5.1 is formulated so that one is allowed to play independently with both parameters in suitable ranges.

A second relevant point is concerned with the question how to assess the adiabatic invariance of our quantity. The delicate point is again related to the thermodynamic limit, which was indeed a major obstacle in tackling the problem with perturbation methods, but can be dealt with using a statistical approach. In a simplified description (see e.g. [28]) we can say that as the number of degrees of freedom grows, all the extensive functions appear as essentially constant over the energy surface, in the sense that for increasing N their densities approach a delta function centred around their average value. Clearly an almost constant function is also approximately constant along an orbit, which seems not to give a meaningful information. The idea is thus to compare the dynamical fluctuation with the statistical deviation of the function over the phase space, using the Gibbs measure. A function defined on the phase space will be considered reasonably conserved if its fluctuation along the orbit is significantly smaller than its Gibbs variance, for a large set of initial data.

In the present paper we are able to show that, in the physically sound assumption of fixed coupling constant a , as the specific energy ϵ goes to zero, for a large (asymptotically full) Gibbs measure of initial data, and for time scaling as inverse powers of ϵ , the time variance of our quantity is smaller than the corresponding Gibbs variance; their ratio vanishing as a power of ϵ . The estimates are uniform in the number N of degrees of freedom.

We come now to a formal presentation of the results in a somehow simplified form. A general formulation is given in the main Theorem 5.1, where some parameters appear that may be subjected to a fine tuning (a and β among them). In the statement below we reduce the number of free parameters by making appropriate choices, so as to give more readable, but still physically meaningful results. Whenever, it will be useful we shall denote by z all the coordinates and momenta (x, y) and by $H(z, a)$ the Hamiltonian so as to bring into evidence the dependence on the coupling constant a .

We denote here by dz the $2N$ -dimensional Lebesgue measure in the phase space $\mathcal{M} := \mathbb{R}^{2N}$, by dm the Gibbs measure and by Z the corresponding partition function, namely

$$dm(\beta, a) := \frac{e^{-\beta H(z, a)}}{Z(\beta, a)} dz, \quad Z(\beta, a) := \int_{\mathcal{M}} e^{-\beta H(z, a)} dz;$$

for every function $X : \mathcal{M} \rightarrow \mathbb{R}$ we denote its phase average and its variance¹ respectively by

$$\langle X \rangle := \int_{\mathcal{M}} X dm(\beta, a), \quad \sigma^2[X] := \langle X^2 \rangle - \langle X \rangle^2.$$

For every measurable set $A \in \mathcal{M}$, we will denote $m(A) := \int_A dm(\beta, a)$.

We recall that for β large and a small, β is roughly the inverse of the average specific energy

$$\frac{1}{\beta} \sim \frac{\langle H \rangle}{N}.$$

We also need to define the time average and the time variance, evaluated along the time evolution. Denoting by ϕ^t the Hamiltonian flow, these quantities are naturally defined as

$$\bar{X}(z, t) := \frac{1}{t} \int_0^t (X \circ \phi^s)(z) ds, \quad \sigma_t^2[X] := \bar{X}^2 - \bar{X}^2.$$

We remark that $\sigma_t^2[X]$ is a function of (z, t) , and that all the previously defined averages and variances are clearly functions of β and a , even though we do not write these dependencies explicitly.

We state here a particular version of the main result of the paper giving, for fixed coupling constant and small specific energies, a control for time scales growing as a power of β . In the statements of the present section, the symbols C_1, C_2, \dots denote constants that may have different values in different contexts.

Theorem 1.1. *There exist positive constants $a^*, \beta_0, \beta_1, C_1$ and C_2 such that, for all $0 < a < a^*$, given the integer $r := \lfloor C_1 \sqrt{\frac{1+2a}{a}} \rfloor$, there exists an extensive polynomial $\Phi : \mathcal{M} \rightarrow \mathbb{R}$ of degree $2r + 2$, such that for all $\beta > \max\{\beta_0, \beta_1 r^6\}$ one has*

$$m \left(z \in \mathbb{R}^{2N} : \sigma_t^2[\Phi] \geq \frac{\sigma^2[\Phi]}{\sqrt{\beta}} \right) \leq \frac{C_2}{\beta} \left(\frac{t}{\bar{t}} \right)^2, \quad \bar{t} = \beta^{r/2}$$

Remark 1.2. According to the result stated above, given a system with Hamiltonian (1.1) with a sufficiently small, and fixed, coupling constant a , there exists a quantity whose time variance is smaller than its phase variance for a set of initial data of large Gibbs measure; this holds over long times scaling with $\epsilon^{-C/\sqrt{a}}$, for small enough average specific energy ϵ . Actually, given the relation among β, r and a , the minimal time scale (corresponding to the maximal specific energy allowed) is of order r^r , i.e. $(1/\sqrt{a})^{1/\sqrt{a}}$.

We may state another result, where the time scale is a stretched exponential in β . The price to be paid, again in the hypothesis of a fixed coupling constant, is that the specific energy must be bounded both from below and

¹ We do not use the standard notation σ_X^2 because we reserve the subscript, in particular with a t , for the variance along the dynamics.

from above; otherwise, it is necessary to let the coupling vanish as the specific energy goes to zero.

Theorem 1.3. *There exist positive constants a^* , β_* , C^* , C_1 , C_2 and C_3 such that, for all $\beta \geq \beta_*$ and $0 < a < a^*$ satisfying*

$$\sqrt{a} \sqrt[3]{\beta} \leq C^*,$$

given the integer $r := \lfloor C_1 \sqrt[3]{\beta} \rfloor$, there exist an extensive polynomial $\Phi : \mathcal{M} \rightarrow \mathbb{R}$ of degree $2r + 2$, such that one has

$$m \left(z \in \mathbb{R}^{2N} : \sigma_{\bar{t}}^2[\Phi] \geq \frac{\sigma^2[\Phi]}{\sqrt{\beta}} \right) \leq \frac{C_2}{\beta} \left(\frac{t}{\bar{t}} \right)^2, \quad \bar{t} = e^{C_3 \sqrt[3]{\beta}}$$

The proofs of the theorems stated here are given in Sect. 5 as corollaries of the main Theorem 5.1.

The paper actually consists of two separate parts, namely (i) the construction of an approximate conserved quantity with perturbation methods, and (ii) the control of the dynamical fluctuation using statistical tools.

The first part makes use of the formal perturbation expansion method introduced in [23] and used in subsequent works, but implements a quantitative scheme of estimates that exploits the characteristics of the present system, namely the complete resonance, the extensivity of the model and the exponential decay of interactions with the distance, to produce estimates uniform in N . The role of complete resonance in removing the critical dependencies on N goes back to [9, 10] and has been used later on, e.g. in [3, 24]. The extensivity property has been dealt with in our previous paper [24] exploiting the *cyclic symmetry*; some results are restated here in a more terse way, using the formalism of circulant matrices. The method of control of the exponential decay introduced here is new, up to our knowledge. The quantitative perturbation scheme developed in the present work significantly improves the one in [16]. In this respect, we put emphasis on the different method of solving the homological equation. The problem is to invert a linear operator that depends on the coupling parameter a . We are able to formulate a direct inversion lemma, thus replacing the truncation method used in [16] with a more effective one; the price we pay is that, at variance with their approach, we actually have to control small divisors, thus introducing an upper bound on the number of perturbation steps allowed. The crucial positive outcome of our choice, in this technical point, is the possibility of preserving the independence of the parameters a and ϵ , while they are collectively controlled in [16] as $a + \epsilon$.

In the second part, the statistical control of the fluctuation is reminiscent of the scheme used in [16]. However, we are able to produce improved estimates of the adiabatic invariance and to prove almost everything independently and in a different way. First of all, we do not rely on probabilistic techniques (like marginal probabilities) but we use a more direct approach. In particular, we exploit a mechanism of cancellations of unwanted interaction terms, which allows us to bring into evidence the decay properties of spatial correlations. A second point concerns the fundamental use, besides the decay properties of spatial correlations, of the short-range interaction properties of the system

which are preserved by our perturbative construction, the latter being another outcome of the improvements of the first part.

Let us add a comment on the possible extensions of the present work. Most of the ideas and techniques used here are not restricted to the one-dimensional case. E.g. properties like complete resonance, extensivity and exponential decay of interaction range may be handled essentially in the same way even for a multidimensional lattice, perhaps at the price of more complicated estimates. A trickier formalization may be required in the part concerning the statistical estimates, in particular for the cancellations.

A further comment is devoted to the fact that in both the present result and in [16] the coupling parameter must be small enough: thus, the applicability to the Fermi–Pasta–Ulam model still remains open. This is particularly relevant since the question of the relaxation properties of the latter model at the thermodynamic limit still remains not completely understood (see e.g. [21]), despite some recent advancement in the investigation of the integrability origin of the long time stability exhibited both with long wave initial data and with generic initial data (see e.g. [5–8, 11, 14, 22, 29]).

We close this review of the literature with a very recent² and interesting result: paper [19]. Although they consider different models and deal with a different question, i.e. the problem of heat conduction, it appears as a relevant work since they are able to perform a normal form at the thermodynamic limit.

The paper is organized as follows. In Sect. 2 the general setting is introduced, with the formalization of the extensivity of the system, the interaction range and its relation with the perturbation tools. In Sect. 3 we recall the normal form transformation of the quadratic part of the Hamiltonian, which produces the zeroth-order approximation of the adiabatic invariant; the formal construction is then carried on at higher perturbative orders in Sect. 4. The control on the time evolution of the adiabatic invariant, and the estimates on the measure of the set of initial data for which they hold, is given in Sect. 5, where we actually give the complete and detailed version of the main result of paper. An Appendix with several technical lemmas closes the paper.

2. General Setting

One of the guiding ideas of this work is to exploit some general characteristics of a many particles mechanical system:

- (i) Particles interacting with a two-body potential;
- (ii) The potential is invariant with respect to rotations and translations;
- (iii) The potential is assumed to be a smooth function; actually we consider the stronger condition of being analytic in the coordinates.

With these conditions the Hamiltonian may be given the generic form

$$H(q, p) = \sum_j \frac{1}{2} p_j^2 + \frac{1}{2} \sum_{i \neq j} V(q_i, q_j)$$

² We became aware of it actually during submission.

where the potential V possesses the symmetry and short-range properties above.

These properties are quite general ones. E.g. besides the realm of Statistical Mechanics, they also apply to the Solar System and have actually been used by Lagrange in his theory of secular motions.

Here, we restrict our attention to a system of identical particles on a d -dimensional lattice, with a short- or even finite-range interaction. In this case one needs just to know the local interaction of a particle with its neighbours or with the whole chain, and the complete Hamiltonian is the sum of the contribution of every particle to both the kinetic and the potential energy. This is usually expressed by saying that the Hamiltonian is *extensive*. Functions possessing the same extensivity property of the Hamiltonian are particularly relevant.

2.1. Formalization

We restrict our attention to the simplified model of a finite one-dimensional lattice with periodic boundary conditions and short-range interactions. We denote by x_j, y_j the position and the momentum of a particle, with $x_{j+N} = x_j$ and $y_{j+N} = y_j$ for any j .

Cyclic Symmetry. We give a formal implementation of extensivity by introducing the concept of cyclic symmetry. The *cyclic permutation* operator τ is defined as

$$\tau(x_1, \dots, x_N) = (x_2, \dots, x_N, x_1), \quad \tau(y_1, \dots, y_N) = (y_2, \dots, y_N, y_1). \quad (2.1)$$

We shall denote $(\tau f)(x, y) = f(\tau x, \tau y)$.

Definition 2.1. We say that a function F is *cyclically symmetric* if $\tau F = F$.

Cyclically symmetric functions may be constructed as follows. Let f be given. A new function $F = f^\oplus$ is constructed as

$$F(x, y) = f^\oplus(x, y) = \sum_{l=1}^N \tau^l f(x, y). \quad (2.2)$$

The upper index \oplus should be considered as an operator defining the new function. We shall say that $f^\oplus(x, y)$ is generated by the *seed* $f(x, y)$. Generally speaking, the decomposition of a cyclically symmetric function in the form (2.2) need not be unique. We shall often use the convention³ of denoting extensive functions with capital letters and their seeds with the corresponding lower case letter.

The following properties will be useful:

- (i) If $f = f' + f''$ is a seed of a function F then $\tau^{s'} f' + \tau^{s''} f''$ is also a seed of the same function, for any integers s', s'' ;
- (ii) The Poisson bracket $h^\oplus = \{f^\oplus, g^\oplus\}$ between two cyclically symmetric functions is also cyclically symmetric. A seed is easily constructed as $h = \{f, g^\oplus\}$, but other choices are allowed using the property (i) above.

³ We will also reserve the prime to denote auxiliary variables rather than derivatives.

Norm of an Extensive Function. Assume now that we are equipped with a norm for our functions $\|\cdot\|$, e.g. the supremum norm over a suitable domain. We introduce a norm $\|\cdot\|^\oplus$ for an extensive function $F = f^\oplus$ by defining

$$\|F\|^\oplus = \|f\|,$$

i.e. we actually measure the norm of the seed. An obvious remark is that the norm so defined depends on the choice of the seed, but this will be harmless for the following reason. All the perturbation procedure and the quantitative estimates on the norm, in the rest of the paper, are based on the fact that all algebraic operations, in particular Poisson brackets, induce a natural choice of the seed for the resulting function. Thus, the relevant estimates will be made directly on the seed so that the initial choice is propagated through the whole procedure. For these reasons, all the quantitative estimates in the rest of the paper could be restated as: the function we are considering possesses a seed whose norm satisfies the stated inequality. We do not explicitly mention this fact in every statement. Moreover, we also have the following relevant facts:

- (i) For any s one has $\|\tau^s f\| = \|f\|$;
- (ii) The inequality $\|F\| \leq N\|f\|$ holds true for any choice of the seed.

This is particularly useful if we are able to produce norms of the seed which are independent of N , since this fully exploits the property of the system of being extensive. This is what we plan to do, indeed.

Polynomial Norms. Let $f(x, y) = \sum_{j,k} f_{j,k} x^j y^k$ be a homogeneous polynomial of degree s in x, y . We define its polynomial norm as

$$\|f\| := \sum_{j,k} |f_{j,k}|.$$

Short-Range Interaction. The short-range interaction is characterized by writing the seed f of a function as a sum $f = \sum_m f^{(m)}$, where the decomposition $f^{(m)}$ is explained in Sect. 2.2, formula (2.7). We consider in particular the case of exponential decay of interactions using two positive parameters: we say that a function f expanded as above is of class $\mathcal{D}(C_f, \sigma)$ in case one has $\|f^{(m)}\| \leq C_f e^{-m\sigma}$. Such a characterization of function is particularly useful in statistical calculation. The known quantitative perturbation schemes will be adapted to deal with these classes of functions.

Circulant Matrices. Let us restrict our attention to the harmonic approximation around a stable equilibrium. The Hamiltonian is a quadratic form represented by a matrix A

$$H_0(x, y) = \frac{1}{2}y \cdot y + \frac{1}{2}Ax \cdot x.$$

If the Hamiltonian H is extensive, then the same holds also for its quadratic part $H_0 = h_0^\oplus$. This implies that A commutes with the matrix τ representing the cyclic permutation (2.1)

$$\tau_{ij} = \begin{cases} 1 & \text{if } i = j + 1 \pmod{N}, \\ 0 & \text{otherwise.} \end{cases}$$

We remark that the matrix τ is orthogonal and generates a cyclic group of order N with respect to the matrix product.

We recall the following

Definition 2.2. A matrix $A \in \text{Mat}_{\mathbb{R}}(N, N)$ is said to be *circulant* if

$$A_{j,k} = a_{(k-j) \pmod N}.$$

Actually, the set of circulant matrices is a subset of Toeplitz matrices, i.e. those which are constant on each diagonal. For a comprehensive treatment of circulant matrices, see e.g. [18]. We just recall some properties that will be useful later.

1. The set of $N \times N$ circulant matrices is a real vector space of dimension N , and a basis is given by the cyclic group generated by τ (see 3.1 of [18]).
2. The set of matrices which commute with τ , i.e. those A such that $A\tau = \tau A$, coincides with the set of circulant matrices (see 3.1 of [18]).
3. The set of eigenvalues of a circulant matrix is the Discrete Fourier Transform of the first row of the matrix and vice-versa. This allows to construct the circulant matrix from its spectrum.
4. Let $M^2 = A$, where A is circulant; then M is circulant, too. Moreover, from the definition of $M := \sqrt{A}$, it follows that if A is symmetric, then M is also symmetric.

In our problem the cyclic symmetry of the Hamiltonian implies that the matrix A of the quadratic form is circulant. Obviously it is also symmetric, so that the space of matrices of interest to us has dimension $\lfloor \frac{N}{2} \rfloor + 1$. Indeed, a circulant and symmetric matrix is completely determined by $\lfloor \frac{N}{2} \rfloor + 1$ elements of its first line.

2.2. Interaction Range

We give here a formal characterization of finite-range interaction, pointing out some properties that will be useful in the rest of the paper. We first consider the case of an infinite chain, which is easier to deal with. Then, we shall point out the differences with the periodic case.

The Infinite Chain. We start with some definitions. Let us label the variables as x_l, y_l with $l \in \mathbb{Z}$. Let us consider a monomial $x^j y^k$ (in multi-index notation). We define the *support* $S(x^j y^k)$ of the monomial and the *interaction distance* $\ell(x^j y^k)$ as follows: considering the exponents (j, k) we set

$$S(x^j y^k) = \{l : j_l \neq 0 \text{ or } k_l \neq 0\}, \quad \ell(x^j y^k) = \text{diam}(S(x^j y^k)).$$

We say that the monomial is *left aligned* in case $S(x^j y^k) \subset \{0, \dots, \ell(x^j y^k) - 1\}$.

The definitions above is extended to a homogeneous polynomial f by saying that $S(f)$ is the union of the supports of all the monomials in f , and that f is left aligned if all its monomials are left aligned. The relevant property is that if \tilde{f} is a seed of a cyclically symmetric function F , then there exists also a left aligned seed f of the same function F : just left align all the monomials in \tilde{f} .

For the seed f of a function (using z to collectively denote the x and y variables, and k the corresponding multi-index) consider the decomposition

$$f(z) = \sum_{m \geq 0} f^{(m)}(z), \quad f^{(m)}(z) = \sum_{\ell(k) \leq m} f_k z^k, \tag{2.3}$$

assuming that every $f^{(m)}$ is left aligned. It would be interesting to replace the inequality $\ell(k) \leq m$ with equality, but this is not compatible with the fact that the Poisson bracket can possibly reduce the interaction range. However, for our purposes it is enough to assure two properties, namely (i) in $f^{(m)}$ there are no terms with interaction range longer than m (upper bound); (ii) the size of $f^{(m)}$, estimated with a norm, is of order μ^m , with some positive μ . This is what we are going to do.

For the Poisson bracket between two cyclically symmetric functions we have

$$\begin{aligned} \{f^\oplus, g^\oplus\} &= \sum_{s,s'} \{\tau^s f, \tau^{s'} g\} = \sum_{m,m'} \sum_{s,s'} \{\tau^s f^{(m)}, \tau^{s'} g^{(m')}\} \\ &= \sum_s \tau^s \left(\sum_{m,m'} \sum_{s'} \{f^{(m)}, \tau^{s'} g^{(m')}\} \right). \end{aligned}$$

The last expression immediately suggests to construct a seed by just removing the translation τ^s and the sum over s . However, we remark that the obvious equality

$$\{\tau^{s+j} f^{(m)}, \tau^{s'+j} g^{(m')}\} = \tau^j \{\tau^s f^{(m)}, \tau^{s'} g^{(m')}\}$$

holds true. Thus, we may replace any term $\{f^{(m)}, \tau^{s'} g^{(m')}\}$ for $s \in \mathbb{Z}$ with a translated one.

Let us exploit these facts. Given $s \in \mathbb{Z}$, we concentrate our attention on the expression $\{f^{(m)}, \tau^{s'} g^{(m')}\}$. The following properties hold true.

1. If $s' < -m'$ or $s' > m$ then one has $\{f^{(m)}, \tau^{s'} g^{(m')}\} = 0$, for the two functions that depend on independent sets of variables.
2. If $s' < 0$ we may replace the seed $\{f^{(m)}, \tau^{s'} g^{(m')}\}$ with $\{\tau^{-s'} f^{(m)}, g^{(m')}\}$.
3. A seed for $\{f^\oplus, g^\oplus\}$ is given by the $m + m' + 1$ expressions

$$\begin{aligned} \{f^{(m)}, g^{(m')}\}, \{\tau f^{(m)}, g^{(m')}\}, \dots, \{\tau^{m'} f^{(m)}, g^{(m')}\} \\ \{f^{(m)}, \tau g^{(m')}\}, \dots, \{f^{(m)}, \tau^m g^{(m')}\} \end{aligned} \tag{2.4}$$

letting $m, m' \geq 0$.

4. Between the expressions in (2.4) there are

$$\begin{aligned} |m - m'| + 1 & \text{ with } \ell(\cdot) \leq \max(m, m'), & \text{ plus} \\ 2 & \text{ with } \ell(\cdot) \leq \max(m, m') + 1, & \text{ plus} \\ 2 & \text{ with } \ell(\cdot) \leq \max(m, m') + 2, & \text{ plus} \\ & \vdots \\ 2 & \text{ with } \ell(\cdot) \leq m + m'. \end{aligned} \tag{2.5}$$

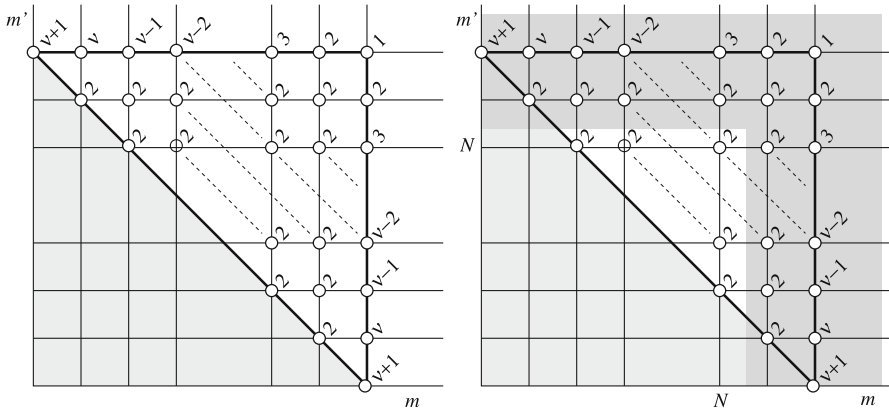


FIGURE 1. Graphical representation of the values of m, m' which concur in forming the seed $h^{(\nu)}$. The *white triangle* contains all nodes that must be selected according to Table 2.5. The *left figure* refers to the case of an infinite chain. The *right figure* shows which nodes are removed in the case of a finite chain. (See text for more details)

The third property follows from the first two, which are obvious. The seed (2.4) so found is left aligned. It may contain duplicated monomials in some expression, but this is harmless because we are only interested in bounding the interaction range. The last property is just matter of counting.

Denote now $h^\oplus = \{f^\oplus, g^\oplus\}$. Letting m and m' to vary, we reorder the seed (2.4) so that we can write $h = \sum_{\nu \geq 0} h^{(\nu)}$. To this end, we collect together in $h^{(\nu)}$ all expressions which according to (2.5) have an estimated upper bound of the interaction range equal to ν . This assures on the one hand that the interaction range of $h^{(\nu)}$ does not exceed ν and on the other hand, that terms with interaction range certainly less than ν are placed in some $h^{(\nu')}$ with $\nu' < \nu$. It is convenient to represent graphically the Table (2.5) as a tridimensional diagram on \mathbb{N}^3 by putting (m, m') on the horizontal plane and the admitted upper bounds for $\ell(\cdot)$ on the vertical axis, i.e. $\max(m, m'), \max(m, m') + 1, \dots, m + m'$. To each non-empty node so identified we attach a weight given by the number of terms in the first column of (2.5). Then we make a section with the horizontal plane of height ν , thus obtaining the left diagram of Fig. 1 which represents schematically all terms in (2.5) that go into $h^{(\nu)}$. The non-empty nodes on the selected plane satisfy $\max(m, m') \leq \nu \leq m + m'$, namely they belong to the white triangle in the diagram. The nodes of the diagram together with their weight contain all the information we need to estimate the norms. The nodes inside the grey triangle have $\ell(\cdot)$ certainly less than ν , so we need not to include them in $h^{(\nu)}$. This rearrangement of seeds assures that the norm of every term in $\eta^{(\nu)}$ has a factor μ^ν , as we shall see later.

We conclude that

$$h^{(\nu)} = \sum_{m,m',s,s'} \{\tau^s f^{(m)}, \tau^{s'} g^{(m')}\} \tag{2.6}$$

the sum being extended to the nodes m, m' in the diagram with the translations s, s' allowed for them according to the property 3.

The Periodic Chain. In view of the periodicity, the labels of the variables may be taken to be $0, \dots, N - 1$, and the definitions of support, interaction range and left alignment are easily adapted. In particular, the infinite sum on (2.3) is truncated at $m = N$. Taking into account the finite limits in the sums, we have

$$h^\oplus = \{f^\oplus, g^\oplus\} = \sum_{s=0}^{N-1} \tau^s \left(\sum_{m,m'=0}^{N-1} \sum_{s'=0}^{N-1} \{f^{(m)}, \tau^{s'} g^{(m')}\} \right).$$

The seed's components $h^{(\nu)}$ are constructed in much the same way with a minor change. Precisely in (2.4) we must distinguish two different case. For $m + m' < N - 1$ we get exactly the same formula. For $m + m' \geq N - 1$ we only have a subset of N elements, namely

$$\{f^{(m)}, g^{(m')}\}, \dots, \{f^{(m)}, \tau^{N-1} g^{(m')}\}.$$

This is represented in the right part of the diagram of Fig. 1, where the part to be omitted is covered in dark grey.

Exponential Decay of Interactions. We recall the definition given in Sect. 2.1. The seed f of a function f is said to be of class $\mathcal{D}(C_f, \sigma)$ in case

$$\|f^{(m)}\| \leq C_f e^{-\sigma m}, \quad C_f > 0, \quad \sigma > 0, \tag{2.7}$$

where $f = \sum_m f^{(m)}$ is the expansion of f in terms of increasing interaction range, as in (2.3).

The following Lemma produces a general estimate of the Poisson bracket specially adapted to the case of cyclically symmetric polynomials. It is crucial for the control of the dependence on N of the norms of extensive functions generated by our perturbation scheme.

Lemma 2.3. *Let $f(x, y)$ and $g(x, y)$ be homogeneous polynomials of degree r and s , respectively. Then, $\{f, g\}$ is a homogeneous polynomial of degree $r + s - 2$, and one has*

$$\|\{f, g\}\| \leq rs \|f\| \|g\|.$$

Moreover, the seed $\{f, g^\oplus\}$ of $\{f^\oplus, g^\oplus\}$ satisfies

$$\|\{f^\oplus, g^\oplus\}\|^\oplus \leq rs \|f\| \|g\|. \tag{2.8}$$

Proof. To prove the first inequality write the Poisson bracket as

$$\{f, g\} = \sum_{j,k,j',k'} f_{j,k} g_{j',k'} \sum_{l=1}^n \frac{j_l k'_l - j'_l k_l}{x_l y_l} x^{j+j'} y^{k+k'}$$

In view of the definition of the norm we may estimate

$$\|\{f, g\}\| \leq \sum_{j,k,j',k'} |f_{j,k}| |g_{j',k'}| \sum_{l=1}^n (j_l k'_l + j'_l k_l).$$

Since $j'_l \leq s$ and $k'_l \leq s$ one has $\sum_{l=1}^n (j_l k'_l + j'_l k_l) \leq s \sum_{l=1}^n (j_l + k_l) = rs$, which readily gives the first inequality. Coming to (2.8), remark that we may write

$$f^\oplus(x, y) = \sum_{j,k} f_{j,k} \sum_{m=1}^N (\tau^m x)^j (\tau^m y)^k,$$

meaning that all monomials $(\tau^m x)^j (\tau^m y)^k$ have the same coefficient. Differentiating with respect to x_l yields

$$\frac{\partial f^\oplus}{\partial x_l} = \sum_{j,k} \sum_{m=1}^N f_{j,k} \frac{(\tau^{-m} j)_l}{x_l} (\tau^m x)^j (\tau^m y)^k.$$

Using the cyclic decomposition of the Poisson bracket $\{f^\oplus, g^\oplus\} = (\{f, g^\oplus\})^\oplus$, one gets

$$\begin{aligned} \{f, g^\oplus\} = \sum_{j,k,j',k'} f_{j,k} g_{j',k'} \sum_{l=1}^N \frac{j_l}{x_l y_l} \left(x^j y^k \sum_{m=1}^N (\tau^{-m} k')_l (\tau^m x)^{j'} (\tau^m y)^{k'} + \right. \\ \left. - x^{j'} y^{k'} \sum_{m=1}^N (\tau^{-m} j')_l (\tau^m x)^j (\tau^m y)^{k'} \right). \end{aligned}$$

The norm is thus estimated as

$$\|\{f, g^\oplus\}\| \leq \sum_{j,k,j',k'} |f_{j,k}| |g_{j',k'}| \sum_{l=1}^N \left(j_l \sum_{m=1}^N (\tau^{-m} k')_l + k'_l \sum_{m=1}^N (\tau^{-m} j')_l \right).$$

Remarking that $\sum_{m=1}^N (\tau^{-m} j')_l = |j'|$ and $\sum_{m=1}^N (\tau^{-m} k')_l = |k'|$, one has

$$\sum_{l=1}^N (j_l |k'| + k'_l |j'|) \leq (|j'| + |k'|) \sum_{l=1}^N (j_l + k_l) = (|j'| + |k'|)(|j| + |k|).$$

In view of the definition of the norm one gets

$$\|\{f, g^\oplus\}\| \leq rs \|f\| \|g\|,$$

from which (2.8) follows. □

The next statements provide the basic estimates for controlling the exponential decay in the framework of perturbation theory.

Lemma 2.4. *Let F, G be cyclically symmetric homogeneous polynomials of degree r', r'' , respectively. Let the seeds f, g be of class $\mathcal{D}(C_f, \sigma')$ and $\mathcal{D}(C_g, \sigma'')$, respectively, and let $\sigma < \min(\sigma', \sigma'')$. Then, there exists $C_h \geq 0$*

such that the seed h of $H = \{F, G\}$ is of class $\mathcal{D}(C_h, \sigma)$. An explicit estimate is

$$C_h = \frac{r' r'' C_f C_g}{(1 - e^{-\max(\sigma', \sigma'')})(1 - e^{-\max(\sigma', \sigma'') + \sigma})}.$$

Proof. According to (2.6), the seed of H may be written as

$$h^{(\nu)} = \sum_{m', m''} \sum_{s', s''} \{\tau^{s'} f^{(m')}, \tau^{s''} g^{(m'')}\},$$

where the sum must be extended to all nodes of the triangle of the diagram 1. In view of the general estimate of the Poisson bracket in Lemma 2.3, we have

$$\left\| \sum_{s', s''} \{\tau^{s'} f^{(m')}, \tau^{s''} g^{(m'')}\} \right\| \leq r' r'' C_f C_g e^{-m' \sigma'} e^{-m'' \sigma''} ;$$

This uses the cyclic symmetry and the fact that the sum over s', s'' is restricted to the values allowed by (2.5). Thus, for all nodes of the diagram we get a common factor $r' r'' C_f C_g$ and must deal only with the exponentials. Possibly exchanging the functions we may suppose that $\sigma' > \sigma'' (> \sigma)$. To get the estimate of $h^{(\nu)}$ we have to sum up all the couples (m', m'') in the white triangle of Fig. 1: we perform the summation by fixing the each diagonal segment $m' + m'' = l$ and increasing $l = \nu, \dots, 2\nu$. Hence, we can write

$$\begin{aligned} & \sum_{l=\nu}^{2\nu} \sum_{m'+m''=l} e^{-m' \sigma'} e^{-m'' \sigma''} \sum_{l=\nu}^{2\nu} \sum_{m'=l-\nu}^{\nu} e^{-l \sigma''} e^{-m'(\sigma' - \sigma'')} \\ &= \sum_{l=0}^{\nu} \sum_{m'=l}^{\nu} e^{-(l+\nu)\sigma''} e^{-m'(\sigma' - \sigma'')} \\ &= \sum_{l=0}^{\nu} \sum_{m=0}^{\nu-l} e^{-(l+\nu)\sigma''} e^{-(m+l)(\sigma' - \sigma'')} \\ &= \sum_{l=0}^{\nu} \sum_{m=0}^{\nu-l} e^{-\nu\sigma''} e^{-l\sigma'} e^{-m(\sigma' - \sigma'')}, \end{aligned}$$

where we have first replaced $m'' = l - m'$, then have shifted back the interval of the running index l (thus exhibiting $e^{-\nu\sigma''}$), and finally have shifted back the interval of the running index $m (= m')$. Then, we estimate

$$\begin{aligned} \sum_{l=0}^{\nu} \sum_{m=0}^{\nu-l} e^{-\nu\sigma''} e^{-l\sigma'} e^{-m(\sigma' - \sigma'')} &= e^{-\nu\sigma} e^{-\nu(\sigma'' - \sigma)} \sum_{l=0}^{\nu} e^{-l\sigma'} \sum_{m=0}^{\nu-l} e^{-m(\sigma' - \sigma'')} \\ &\leq e^{-\nu\sigma} \sum_{l=0}^{\nu} e^{-l\sigma'} \sum_{m=0}^{\nu-l} e^{-m(\sigma' - \sigma)} \\ &< \frac{e^{-\nu\sigma}}{(1 - e^{-\sigma'}) (1 - e^{-(\sigma' - \sigma)})}. \end{aligned}$$

The claim follows by replacing σ' with $\max(\sigma', \sigma'')$. This completes the proof. □

Corollary 2.5. *If in Lemma 2.4 we have $\sigma' \neq \sigma''$ then we may set $\sigma = \min(\sigma', \sigma'')$ and*

$$C_h = \frac{r' r'' C_f C_g}{(1 - e^{-\max(\sigma', \sigma'')})(1 - e^{-|\sigma' - \sigma''|})}.$$

Proof. Just set $\sigma = \sigma''$ and at the end replace $\sigma' - \sigma$ with $|\sigma' - \sigma''|$. □

Corollary 2.6. *If in Lemma 2.4 we have $\sigma' > \sigma''$ and $f^{(0)} = 0$, i.e. $f = \sum_{m \geq 1} f^{(m)} = O(e^{-\sigma'})$ then we may set $\sigma = \sigma''$ and*

$$C_h = \frac{2e^{-(\sigma' - \sigma'')} r' r'' C_f C_g}{(1 - e^{-\sigma'})(1 - e^{-(\sigma' - \sigma'')})}.$$

Proof. Set $\sigma = \sigma''$. Then hypothesis $f = \sum_{m' \geq 1} f^{(m')}$ implies that we must remove the element $(m', m'') = (0, \nu)$ from the elements of the white triangle of Fig. 1: this element gives a factor $e^{-\nu\sigma''}$. Hence, the sum in Lemma 2.4 becomes

$$\begin{aligned} \sum_{l=0}^{\nu} \sum_{m=0}^{\nu-l} e^{-(m+l)\sigma'} e^{-(\nu-m)\sigma''} - 1 &= e^{-\sigma''\nu} \left[\sum_{l=0}^{\nu} e^{-l\sigma'} \sum_{m=0}^{\nu-l} e^{-m(\sigma' - \sigma'')} - 1 \right] \\ &< e^{-\sigma''\nu} \left[\frac{e^{-\sigma'} + e^{-(\sigma' - \sigma'')}}{(1 - e^{-\sigma'})(1 - e^{-(\sigma' - \sigma'')})} \right] \\ &< e^{-\sigma''\nu} \frac{2e^{-(\sigma' - \sigma'')}}{(1 - e^{-\sigma'})(1 - e^{-(\sigma' - \sigma'')})}, \end{aligned}$$

which readily gives the claim. □

3. Normal Form for the Quadratic Hamiltonian

Let us rewrite the Hamiltonian (1.1) as a sum of its quadratic and quartic parts $H = H_0 + H_1$, where

$$H_0(x, y) := \frac{1}{2} \sum_{j=1}^N [y_j^2 + x_j^2 + a(x_j - x_{j-1})^2], \quad H_1(x, y) := \frac{1}{4} \sum_{j=1}^N x_j^4. \quad (3.1)$$

The aim of this section is to give the quadratic part a resonant normal form so that it turns out to be written as $H_0 = H_\Omega + Z_0$ with Z_0 an extensive function exhibiting an exponential decay of the interaction among sites with their distance, and $\{H_0, Z_0\} = 0$. That is, Z_0 is a first integral for H_0 . This result has been already stated in [24]; we give here a different proof. We will then apply the transformation also to the quartic part of our Hamiltonian, showing that it still has an exponential decay of the interactions.

3.1. The Normalizing Transformation

We introduce the positive parameters $\omega(a) > 1$ and $\mu(a) < 1/2$

$$\omega^2(a) := 1 + 2a, \quad \mu := \frac{a}{\omega^2}$$

and rewrite the quadratic part of our Hamiltonian as

$$H_0(x, y) = \frac{1}{2}y \cdot y + \frac{1}{2}x \cdot Ax,$$

where [recalling τ as the permutation matrix generating (2.1)]

$$A = \omega^2 \begin{pmatrix} 1 & -\mu & 0 & \dots & 0 & -\mu \\ -\mu & 1 & -\mu & \dots & 0 & 0 \\ 0 & -\mu & 1 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 & -\mu \\ -\mu & 0 & 0 & \dots & -\mu & 1 \end{pmatrix} = \omega^2 [\mathbb{I} - \mu(\tau + \tau^\top)], \quad (3.2)$$

which is clearly circulant and symmetric, and gives a finite-range interaction. The latter form is particularly useful because it exhibits the perturbation parameter μ that will be assumed to be small. This particular form allows us to look at our model as a system of identical harmonic oscillators with a small linear coupling. The resulting complete resonance is one of the keys of our result. Introduce the constant Ω as the average of the square roots of the eigenvalues of A .

Proposition 3.1. *For $\mu < 1/2$ there exists a canonical linear transformation which gives the Hamiltonian H_0 the particular resonant normal form*

$$H_0 = H_\Omega + Z_0, \quad \{H_\Omega, Z_0\} = 0 \quad (3.3)$$

with H_Ω and Z_0 cyclically symmetric with seeds

$$h_\Omega = \frac{\Omega}{2}(q_1^2 + p_1^2),$$

$$\zeta_0 = \frac{1}{2} \sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} b_j (q_0 q_j + p_0 p_j + q_0 q_{N-j} + p_0 p_{N-j+1}) + \delta b_{\frac{N}{2}+1} (q_0 q_{\frac{N}{2}+1} + p_0 p_{\frac{N}{2}+1})$$

$$|b_j(\mu)| = \mathcal{O}((2\mu)^j), \quad \delta = \begin{cases} 0 & N \text{ odd} \\ 1 & N \text{ even} \end{cases} \quad (3.4)$$

The linear transformation is given by

$$q = A^{1/4}x, \quad p = A^{-1/4}y, \quad (3.5)$$

where the circulant and symmetric matrix $A^{1/4}$ satisfies

$$(A^{1/4})_{1,j} = c_j(\mu)(2\mu)^{j-1}, \quad 1 \leq j \leq \left\lfloor \frac{N}{2} \right\rfloor + 1, \quad |c_j(\mu)| \leq 2\sqrt{\omega}. \quad (3.6)$$

H_1 remains an extensive and cyclically symmetric function once composed with the transformation (3.5).

We remark that all the perturbative construction is performed after the linear transformation (3.5), but all the estimates with the Gibbs measure of Sect. 5 are made in the original variables. We thus need some further properties of the transformation itself, which are given in the following two results. Recalling that, according to the notations of Sect. 2.2, we label the coordinates with indices $0, \dots, N - 1$, and introducing the decay rate σ_0

$$\sigma_0 := -\ln(2\mu) \quad \Rightarrow \quad 2\mu = e^{-\sigma_0}, \tag{3.7}$$

we have

Proposition 3.2. *The linear canonical transformation (3.5) is the flow at time $t = 1$ of the cyclically symmetric quadratic form \mathcal{X}_0*

$$\mathcal{X}_0(x, y) := x \cdot B y, \quad B := \frac{1}{4} \ln(A),$$

where B is a symmetric and circulant matrix characterized by

$$B_{1,j} = c_j(\mu)(2\mu)^j, \quad |c_j(\mu)| \leq \frac{1}{2} C_0(a) := \frac{1}{4} \left| \ln \left(\frac{\omega^2}{1 - 2\mu} \right) \right|,$$

for $1 \leq j \leq \lfloor \frac{N}{2} \rfloor + 1$; the seed of \mathcal{X}_0 satisfies $\chi_0 \in \mathcal{D}(C_0(a), \sigma_0)$ and reads

$$\chi_0 = \sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} B_{1,j}(x_0 y_j + y_0 x_j) + \delta B_{1,N/2+1} x_0 y_{N/2+1} \quad \delta = \begin{cases} 0 & N \text{ odd} \\ 1 & N \text{ even} \end{cases} \tag{3.8}$$

Lemma 3.3. *Let ρ^\oplus an homogeneous polynomial of degree $2r + 2$ in $\mathcal{D}(C_\rho, \sigma_*)$ and assume $\chi_0 \in \mathcal{D}(C_0(a), \sigma')$ with $\sigma_* < \sigma' \leq \sigma_0$, then*

$$T_{\chi_0} \rho \in \mathcal{D}(e^{(r+1)\tilde{C}} C_\rho, \sigma_*), \quad \tilde{C} \leq \frac{2C_0(a)}{(1 - e^{-\sigma'})(1 - e^{-(\sigma' - \sigma_*)})}.$$

A fundamental point is represented by the decay properties of the seeds of Z_0 and H_1 .

Lemma 3.4. *The seeds of the functions Z_0 and H_1 satisfy*

$$\begin{aligned} \zeta_0 \in \mathcal{D}(C_0(a), \sigma_0), \quad C_0(a) = \mathcal{O}(1) \\ h_1 \in \mathcal{D}(C_1(a), \sigma_1), \quad C_1(a) = \mathcal{O}(1) \end{aligned} \quad \text{for } a \rightarrow 0, \quad \sigma_1 := \frac{1}{2} \sigma_0.$$

Remark 3.5. The seed h_1 cannot preserve the same exponential decay rate of the linear transformation (see the corresponding proof in the Appendix); however, it is possible to show that $h_1 \in \mathcal{D}(C_1(a), \sigma_1)$ for any $\sigma_1 < \sigma_0$. We make the choice $\sigma_1 = \sigma_0/2$ to explicitly relate σ_1 to the small natural parameter a of the model, since it will be useful in the estimates of the main Theorems of the paper.

The proofs of all the statements of this section are deferred to Appendix 6.1.1.

4. Construction of an Extensive First Integral

We construct a formal first integral for the Hamiltonian (1.1) using the Lie transform algorithm in the form introduced in [23]. We include a brief description, referring to the quoted paper for proofs.

Given a generating sequence $\{\mathcal{X}_s\}_{s \geq 1}$, we define the linear operator $T_{\mathcal{X}}$ as

$$T_{\mathcal{X}} = \sum_{s \geq 0} E_s, \quad E_0 = \mathbb{I}, \quad E_s = \sum_{j=1}^s L_{\mathcal{X}_j} E_{s-j}, \tag{4.1}$$

where $L_{\mathcal{X}_j} \cdot = \{\mathcal{X}_j, \cdot\}$ is the Lie derivative with respect to the flow generated by \mathcal{X}_j . The operator $T_{\mathcal{X}}$ turns out to be invertible and to possess the relevant properties

$$T_{\mathcal{X}}(f \cdot g) = (T_{\mathcal{X}}f) \cdot (T_{\mathcal{X}}g), \quad T_{\mathcal{X}}\{f, g\} = \{T_{\mathcal{X}}f, T_{\mathcal{X}}g\}. \tag{4.2}$$

Let now Z satisfy the equation $T_{\mathcal{X}}Z = H$, and let Φ_0 commute with Z , i.e. $\{\Phi_0, Z\} = 0$. Then in view of the second of (4.2), one immediately gets $\{T_{\mathcal{X}}\Phi_0, H\} = 0$, i.e. $\Phi = T_{\mathcal{X}}\Phi_0$ is a first integral for the Hamiltonian H .

The operator $T_{\mathcal{X}}$ is defined here at a formal level, and it is known that using it in normal form theory usually produces non-convergent expansions. However, we may well use it in formal sense, as explained by Poincaré (Ch. VIII in [35]). What we actually do is truncate the all expansions at a given order, so that all the equalities above are true up to terms of order larger than r (i.e. of degree larger than $2r + 2$ in our polynomial expansions). E.g. the sentence above “ $\Phi = T_{\mathcal{X}}\Phi_0$ is a first integral for H ” should be interpreted as “having determined $\mathcal{X}_1, \dots, \mathcal{X}_r$, then truncate $\Phi^{(r)} = \Phi_0 + \Phi_1 + \dots + \Phi_r$, so that we have $\{\Phi^{(r)}, H\} = \mathcal{O}(r + 1)$ ”. The statement of Proposition 4.1 below must be interpreted in this sense.

In this section we prove the following

Proposition 4.1. *Consider the Hamiltonian $H = h_{\Omega}^{\oplus} + \zeta_0^{\oplus} + h_1^{\oplus}$ with seeds $h_{\Omega} = \frac{\Omega}{2}(x_0^2 + y_0^2)$, the quadratic term ζ_0 of class $\mathcal{D}(C_0, \sigma_0)$ with $\zeta_0^{(0)} = 0$, and the quartic term h_1 of class $\mathcal{D}(C_1, \sigma_1)$, with $\sigma_0 > \sigma_1 > \ln(4)$. Pick a positive $\sigma_* < \sigma_1$. Then, there exist positive γ, μ_* and C_**

$$\begin{aligned} \mu_* &= \frac{\Omega(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})}{8C_0 e^{\sigma_1}}, \\ \gamma &= 2\Omega \left(1 - \frac{r\mu}{\mu_*}\right), \\ C_* &= \frac{C_1}{\gamma(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})}. \end{aligned} \tag{4.3}$$

such that for any positive integer r satisfying

$$2r\mu < \mu_*, \tag{4.4}$$

there exists a finite generating sequence $\mathcal{X} = \{\chi_1^{\oplus}, \dots, \chi_r^{\oplus}\}$ of a Lie transform such that $T_{\mathcal{X}}Z - H = \mathcal{O}(r + 1)$, i.e. the remainder $\mathcal{O}(r + 1)$ starts with terms of degree bigger than $2r + 4$ and Z is an extensive function of the form

$$Z = h_{\Omega}^{\oplus} + \zeta_0^{\oplus} + \dots + \zeta_r^{\oplus}$$

with $L_{\Omega}Z_s = 0$ for $s = 0, \dots, r$, $Z_s = \zeta_s^{\oplus}$ of degree $2s + 2$.

Moreover, defining

$$C_r := 64r^2C_*; \tag{4.5}$$

$$\sigma_s := \frac{s\sigma_* + (r - s)\sigma_1}{r} \quad \text{for } s = 2, \dots, r, \tag{4.6}$$

the following statements hold true:

- (i) The seed χ_s of \mathcal{X}_s is of class $\mathcal{D}(C_r^{s-1} \frac{C_1}{\gamma_s}, \sigma_s)$;
- (ii) The seed ζ_s of Z_s is of class $\mathcal{D}(C_r^{s-1}C_1, \sigma_s)$;
- (iii) If $\Phi = \varphi^{\oplus}$ is a homogeneous polynomial of degree $2m$ and of class $\mathcal{D}(C_{\varphi}, \sigma_0)$ then for $s = 0, \dots, r$ one has that $E_s\Phi$ is of class $\mathcal{D}(F_r^s C_{\varphi}, \sigma_s)$ with $F_r = 16(m + 2)r^2C_*$;
- (iv) Setting $\Phi = H_{\Omega}$ in the previous point we have that E_sH_{Ω} is of class $\mathcal{D}(F_r^{s-1}C_1, \sigma_s)$;
- (v) Setting $\Phi_0 = H_{\Omega}$ and considering the first $r + 1$ terms in the expansion of $T_{\mathcal{X}}\Phi_0$, namely $\Phi^{(r)} = \Phi_0 + \dots + \Phi_r$ with $\Phi_s = E_s\Phi_0$, we have

$$\dot{\Phi}^{(r)} = \{H_1, \Phi_r\}$$

which is a cyclically symmetric homogeneous polynomial of degree $2r + 4$ and of class $\mathcal{D}(C_{\rho}, \sigma_*)$ with

$$C_{\rho} = \frac{8(r + 2)(16r^2C_*)^{r-1}C_1^2}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})}.$$

The rest of this section is devoted to the proof of the proposition. We first include a formal part, where we illustrate in detail the process of construction of the normal form and introduce an appropriate framework which allows us to control how the interaction range propagates. Then, we give quantitative estimates paying particular attention to the exponential decay of interactions with the distance.

4.1. Formal Algorithm and Solution of the Homological Equation

We now translate the equation $T_{\mathcal{X}}Z = H$ into a formal recursive algorithm that allows us to construct both Z and \mathcal{X} . We take into account that our Hamiltonian has the particular form $H = H_0 + H_1$, where H_1 is a homogeneous polynomial of degree 4.

For $s \geq 1$ the generating function \mathcal{X}_s and the normalized term Z_s must satisfy the recursive set of *homological equations*

$$L_{H_0}\mathcal{X}_s = Z_s + \Psi_s; \tag{4.7}$$

where

$$\begin{aligned} \Psi_1 &= H_1, \\ \Psi_s &= \frac{s-1}{s}L_{\mathcal{X}_{s-1}}H_1 + \sum_{j=1}^{s-1} \frac{j}{s}E_{s-j}Z_j, \quad s \geq 2. \end{aligned} \tag{4.8}$$

A justification of this algorithm is the following. Using the Definition (4.1) of T_χ , we expand the equation $T_\chi Z = H$ into the recursive set of equations

$$\begin{aligned} Z_0 &= H_0, \\ Z_1 + E_1 Z_0 &= H_1, \\ E_s Z_0 + \sum_{l=1}^s E_{s-l} Z_l + Z_s &= 0 \quad \text{for } s > 1 \end{aligned} \tag{4.9}$$

In view of $E_1 Z_0 = L_{\chi_1} H_0$, the second equation is readily written as $L_{H_0} \chi_1 = Z_1 - H_1$, which is the homological equation at order $s = 1$. Then, using the definition of E_s , we replace $E_s Z_0 = \sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} E_{s-l} Z_0$ in the third of (4.9) and get the homological equation $L_{H_0} \chi_s = Z_s + \Psi_s$, where

$$\Psi_s = \sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} E_{s-l} Z_0 + \sum_{l=1}^{s-1} E_{s-l} Z_l.$$

The expression for Ψ_s may be simplified thanks to the equations of the previous orders as follows. Replacing $E_{s-l} Z_0$ as given by (4.9) in the first sum calculate

$$\begin{aligned} \sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} E_{s-l} Z_0 &= \frac{s-1}{s} L_{\chi_{s-1}} H_1 - \sum_{l=1}^{s-1} \frac{l}{s} L_{\chi_l} \sum_{j=1}^{s-l} E_{s-l-j} Z_j \\ &= \frac{s-1}{s} L_{\chi_{s-1}} H_1 - \sum_{j=1}^{s-1} \frac{s-j}{s} \sum_{l=1}^{s-j} \frac{l}{s-j} L_{\chi_l} E_{s-j-l} Z_j \\ &= \frac{s-1}{s} L_{\chi_{s-1}} H_1 - \sum_{j=1}^{s-1} \frac{s-j}{s} E_{s-j} Z_j, \end{aligned}$$

where the definition of the operator E_s has been used in the last equality. Then, replace the latter expression in the r.h.s. of Ψ_s above and get the wanted expression (4.8).

Our aim is to solve the homological Eq. (4.7) with the prescription that $L_\Omega Z_s = 0$ where $L_\Omega \cdot := \{H_\Omega, \cdot\}$ is the Lie derivative along the vector field generated by H_Ω as defined in (3.3). Thus, the next step is to point out the properties of the operator L_Ω and discuss the solution of the homological equation.

4.1.1. The Linear Operator L_Ω . It is an easy matter to check that L_Ω maps the space of homogeneous polynomials into itself. It is also well known that L_Ω may be diagonalized via the canonical transformation

$$x_j = \frac{1}{\sqrt{2}}(\xi_j + i\eta_j), \quad y_j = \frac{i}{\sqrt{2}}(\xi_j - i\eta_j), \quad j = 1, \dots, N, \tag{4.10}$$

where $(\xi, \eta) \in \mathbb{C}^{2n}$ are complex variables. A straightforward calculation gives

$$L_\Omega \xi^j \eta^k = i\Omega (|k| - |j|) \xi^j \eta^k,$$

where $|j| = |j_1| + \dots + |j_N|$ and similarly for $|k|$.

A relevant general property is that if $f(x, y) = \sum_{j,k} c_{j,k} x^j y^k$ (in multi-index notation) is a real polynomial, then the transformation (4.10) produces a polynomial $g(\xi, \eta) = \sum_{j,k} b_{j,k} \xi^j \eta^k$ with complex coefficients $b_{j,k}$ satisfying

$$b_{j,k} = -b_{k,j}^*$$

Conversely, this is the condition that the coefficients of $g(\xi, \eta)$ must satisfy to assure that transforming it back to real variables x, y we get a real polynomial.

Let us denote by $\mathcal{P}^{(s)}$ the (finite) linear space of the homogeneous polynomials of degree s in the $2n$ canonical variables $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n$. The kernel and the range of L_Ω are defined in the usual way, namely

$$\mathcal{N}^{(s)} = L_\Omega^{-1}(0), \quad \mathcal{R}^{(s)} = L_\Omega(\mathcal{P}^{(s)})$$

The property of L_Ω of being diagonal implies

$$\mathcal{N}^{(s)} \cap \mathcal{R}^{(s)} = \{0\}, \quad \mathcal{N}^{(s)} \oplus \mathcal{R}^{(s)} = \mathcal{P}^{(s)}.$$

Thus, the inverse $L_\Omega^{-1} : \mathcal{R}^{(s)} \rightarrow \mathcal{R}^{(s)}$ is uniquely defined on the restriction $\mathcal{R}^{(s)}$ of $\mathcal{P}^{(s)}$. It will also be useful to introduce the projectors on the range and on the kernel defined as

$$\Pi_{\mathcal{R}^{(s)}} = L_\Omega^{-1} L_\Omega, \quad \Pi_{\mathcal{N}^{(s)}} = \mathbb{I} - \Pi_{\mathcal{R}^{(s)}},$$

so that we have $\Pi_{\mathcal{R}^{(s)}} + \Pi_{\mathcal{N}^{(s)}} = \mathbb{I}$.

Lemma 4.2. *Let $f \in \mathcal{P}^{(s)}$ and $g \in \mathcal{P}^{(r)}$. Then, the following composition table applies*

$\{\cdot, \cdot\}$	$\mathcal{N}^{(r)}$	$\mathcal{R}^{(r)}$	(4.11)
$\mathcal{N}^{(s)}$	$\mathcal{N}^{(r+s-2)}$	$\mathcal{R}^{(r+s-2)}$	
$\mathcal{R}^{(s)}$	$\mathcal{R}^{(r+s-2)}$	$\mathcal{P}^{(r+s-2)}$	

Proof. For any pair of functions f, g , by Jacobi’s identity for Poisson brackets we have $L_\Omega\{f, g\} = \{L_\Omega f, g\} + \{f, L_\Omega g\}$. If f, g are in the respective kernels, then $L_\Omega\{f, g\} = 0$, which proves that $\{f, g\} \in \mathcal{N}^{(r+s-2)}$. If $g \in \mathcal{R}^{(r)}$ then we may write $g = L_\Omega L_\Omega^{-1} g$, and so if $f \in \mathcal{N}^{(s)}$ we have, still using Jacobi’s identity, $\{f, L_\Omega L_\Omega^{-1} g\} = L_\Omega\{f, L_\Omega^{-1} g\}$ in view of $L_\Omega f = 0$, which proves that $\{f, g\} \in \mathcal{R}^{(r+s-2)}$. If f, g are in the respective ranges, then nothing can be said in general. This gives the table. □

4.1.2. The Linear Operator L_{H_0} . We come now to the solution of the homological Eq. (4.7). In view of (3.3), we have $L_{H_0} = L_\Omega + L_{Z_0}$, so that we immediately get

$$L_{H_0} = L_\Omega(\mathbb{I} + L_\Omega^{-1} L_{Z_0}).$$

Thus, we have

$$L_{H_0}^{-1} = (\mathbb{I} + K)^{-1} L_\Omega^{-1}, \quad K := L_\Omega^{-1} L_{Z_0}, \tag{4.12}$$

and using the Neumann’s series we can write

$$(\mathbb{I} + K)^{-1} = \sum_{l \geq 0} (-1)^l K^l.$$

Let us consider L_{H_0} on the (finite dimensional) topological space $\mathcal{P}^{(s)}$; with the notation $\|\cdot\|_{\text{op}}$ we mean the dual norm of a linear operator acting on $\mathcal{P}^{(s)}$ (they are all continuous). The following proposition claims that, although we lack information about its Kernel and Range, we can invert L_{H_0} on $\mathcal{R}^{(s)}$. This is one of the crucial technical points of the paper, leading eventually to the independence of the two perturbative parameters a and $1/\beta$. See also the forthcoming Remark 4.6 on the control of small divisors.

Proposition 4.3. *If the restriction of K to $\mathcal{R}^{(s)}$ satisfies*

$$\|K\|_{\text{op}} < 1, \tag{4.13}$$

then for any $g \in \mathcal{R}^{(s)}$, there exists an element $f \in \mathcal{R}^{(s)}$ such that

$$(\mathbb{I} + K)f = g \quad \text{with} \quad f = \sum_{l \geq 0} (-1)^l K^l g.$$

Proof. Let us take $g \in \mathcal{R}^{(s)}$, then from (4.11) one has $L_{Z_0}g \in \mathcal{R}^{(s)}$ and also $L_{\Omega}^{-1}L_{Z_0}g = Kg \in \mathcal{R}^{(s)}$; in other words,

$$K : \mathcal{R}^{(s)} \rightarrow \mathcal{R}^{(s)}.$$

The sequence $\{K^l g\}$ is composed of elements of $\mathcal{R}^{(s)}$ and the same holds for the finite sum

$$f_n = \sum_{l=0}^n (-1)^l K^l g \in \mathcal{R}^{(s)}, \quad n \geq 1.$$

Condition (4.13) provides the convergence of the sequence $f_n \rightarrow f$, with f which belongs to $\mathcal{R}^{(s)}$, since it is a closed subset of $\mathcal{P}^{(s)}$. To prove that f solves the required equation, we consider the sequence

$$g_n = (\mathbb{I} + K)f_n \in \mathcal{R}^{(s)};$$

from the definition of f_n we have $g_n = g + (-1)^n K^{n+1}g$, and since $K^{n+1}g$ vanishes, the sequence g_n converges to g . But the continuity of K implies also $g_n = (\mathbb{I} + K)f_n \rightarrow (\mathbb{I} + K)f$, so the uniqueness of the limit gives the thesis. \square

4.2. Quantitative Estimates and Exponential Decay of Interactions

Here, we complete the formal setting of the previous sections by producing all estimates of the norms of the relevant functions. We also prove the crucial property that the exponential decay of interactions is preserved by our construction.

Recalling the Definition (4.6) of σ_s , so that $\sigma_1 > \dots > \sigma_r = \sigma_*$, our aim is to show that the functions \mathcal{X}_s, Ψ_s and Z_s that are generated by the formal construction are of class $\mathcal{D}(\cdot, \sigma_s)$, with some constant to be evaluated in place of the dot.

We shall repeatedly use the following elementary estimate. By the general inequality

$$1 - e^{-x} \geq x \frac{1 - e^{-a}}{a} \quad \text{for } 0 \leq x \leq a.$$

we have

$$\begin{aligned} 1 - e^{-\sigma_j} &\geq \frac{\sigma_j(1 - e^{-\sigma_0})}{\sigma_0} && \text{for } 1 \leq j \leq r, \\ 1 - e^{-(\sigma_j - \sigma_k)} &\geq \frac{(\sigma_j - \sigma_k)(1 - e^{-(\sigma_0 - \sigma_*)})}{\sigma_0 - \sigma_*} && \text{for } 1 \leq j < k \leq r. \end{aligned}$$

Moreover, in view of the Definition (4.6) of $\sigma_0, \dots, \sigma_r$ for $0 \leq j < s \leq r$, we get

$$\begin{aligned} 1 - e^{-\max(\sigma_j, \sigma_{s-j})} &\geq \frac{1 - e^{-\sigma_0}}{\sigma_0} \max(\sigma_j, \sigma_{s-j}) > \frac{(1 - e^{-\sigma_0})}{2}, \\ 1 - e^{-(\sigma_j - \sigma_k)} &\geq \frac{k-j}{r} (1 - e^{-(\sigma_0 - \sigma_*)}). \end{aligned} \tag{4.14}$$

Estimate of the Homological Equation. We first consider the operator L_Ω^{-1} .

Lemma 4.4. *Let $F = f^\oplus \in \mathcal{R}^{(r)}$ be a cyclically symmetric homogeneous polynomial of degree r of class $\mathcal{D}(C_f, \sigma)$. Then, there exists a cyclically symmetric homogeneous polynomial $\Phi = \varphi^\oplus \in \mathcal{R}^{(s)}$ which solves $L_\Omega \Phi = F$ and is of class $\mathcal{D}(C_\varphi, \sigma)$ with*

$$C_\varphi \leq \frac{C_f}{2\Omega}.$$

The proof is a straightforward consequence of the diagonal form of L_Ω .

Coming to the inversion of L_{H_0} , we state the following

Lemma 4.5. *Let $G = g^\oplus \in \mathcal{R}^{(2s+2)}$ be a cyclically symmetric homogeneous polynomial of degree $2s + 2$ of class $\mathcal{D}(C_g, \sigma_s)$. Let K as defined in (4.12) and assume*

$$C_K := \frac{4C_0 e^{-(\sigma_0 - \sigma_1)}}{\Omega(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} \leq \frac{1}{2r}. \tag{4.15}$$

Then, there exists a cyclically symmetric homogeneous polynomial $\mathcal{X} = \chi^\oplus \in \mathcal{R}^{(2s+2)}$ which solves $L_{H_0} \mathcal{X} = G$; moreover, χ is of class $\mathcal{D}(C_g/\gamma, \sigma_s)$ with

$$\gamma = 2\Omega(1 - rC_K). \tag{4.16}$$

Remark 4.6. In Proposition 4.3, we ask $\|K\|_{\text{op}} < 1$ to simply perform the inversion. In the above Lemma 4.5, condition (4.15) reads as $\|K\|_{\text{op}} < 1/2$, and the stronger requirement is to control the small divisors (4.16).

We emphasize that in view of the first of (4.3) we have $C_K = \mu/\mu_*$. Therefore, condition (4.15) reads $2r\mu < \mu_*$, which is the smallness condition for μ of Proposition 4.1. Furthermore, this gives the value of γ in (4.3).

We also emphasize that the constant γ is evaluated as independent of s , but seems to depend on the degree r of truncation of the first integral. However, in view of the condition on μ we have $\Omega \leq \gamma \leq 2\Omega$.

Proof. Recall that ζ_0 is of class $\mathcal{D}(C_0, \sigma_0)$, as stated in Lemma 3.4. By Corollary 2.6, with Z_0, σ_0 and σ_s in place of f, σ' and σ'' , respectively, we see that $L_{Z_0}g$ is of class $\mathcal{D}(C' C_g, \sigma_s)$ with

$$C' \leq \frac{4(s+1)C_0 e^{-(\sigma_0 - \sigma_s)}}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_s)})} \leq \frac{8rC_0 e^{-(\sigma_0 - \sigma_1)}}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})}$$

where the second of (4.14) has been used. Using Lemma 4.4, we get that Kg is of class $\mathcal{D}(rC_K C_g, \sigma_s)$ with C_K given by (4.15). This also implies that $\|K\|_{\text{op}} \leq rC_K$. In view of condition (4.15) we may apply Proposition 4.3, thus concluding that the inverse of L_{H_0} is well defined. With an explicit calculation we also calculate $\|K^m\|_{\text{op}} \leq (rC_K)^m$, thus concluding that $L_{H_0}^{-1}g$ is of class $\mathcal{D}(C_g/\gamma, \sigma_s)$ with γ as in (4.16), as claimed. \square

Having thus proved that the homological equation can be solved, the statement (i) of Proposition 4.1 follows.

Iterative Estimates on the Generating Sequence. We recall that the generating sequence is found by recursively solving the homological equations $L_{H_0}\chi_s = Z_s + \Psi_s$ for $s = 1, \dots, r$ with

$$\begin{aligned} \Psi_1 &= H_1, \\ \Psi_s &= \frac{s-1}{s} L\chi_{s-1} H_1 + \sum_{l=1}^{s-1} \frac{l}{s} E_{s-l} Z_l, \\ E_s Z_l &= \sum_{j=1}^s \frac{j}{s} L\chi_j E_{s-j} Z_l \quad \text{for } s \geq 1. \end{aligned} \tag{4.17}$$

Our aim is to find positive constants $C_{\psi,1}, \dots, C_{\psi,r}$ so that Ψ_s is of class $\mathcal{D}(C_{\psi,s}, \sigma_s)$. In view of Lemma 4.4 this implies that Z_s is of class $\mathcal{D}(C_{\zeta,s}, \sigma_s)$ with $C_{\zeta,s} = C_{\psi,s}$ and χ_s of class $\mathcal{D}(C_{\chi,s}, \sigma_s)$ with $C_{\chi,s} = C_{\psi,s}/\gamma$. Meanwhile, we also find constants $C_{\zeta,s,l}$ such that $E_s Z_{2l}$ is of class $\mathcal{D}(C_{\zeta,s,l}, \sigma_{s+l})$ whenever $s+l \leq r$.

We look for a constant B_r and two sequences $\{\eta_s\}_{1 \leq s \leq r}$ and $\{\theta_s\}_{1 \leq s \leq r}$ such that

$$\begin{aligned} C_{\psi,1} &\leq \eta_1 C_1, \quad C_{\zeta,0,1} \leq \eta_1 \theta_0 C_1, \\ C_{\psi,s} &\leq \frac{\eta_s}{s} B_r^{s-1} C_1 \quad \text{for } s > 1, \\ C_{\zeta,s,l} &\leq \theta_s \eta_l B_r^{s+l-1} C_1 \quad \text{for } s \geq 1, l \geq 1. \end{aligned} \tag{4.18}$$

In view of $\Psi_1 = H_1$ and of $E_0 Z_1 = Z_1$ we can choose $\eta_1 = \theta_0 = 1$. By (4.17) and using Lemmas 4.4 and 2.4 together with Corollary 2.5, we get the recursive relations

$$\begin{aligned} C_{\zeta,s,l} &\leq \frac{4}{s} \sum_{j=1}^{s-1} \frac{j(s+l-j)\eta_j \eta_l \theta_{s-j}}{(1 - e^{-\max(\sigma_j, \sigma_{s+l-j}) + \sigma_{s+l}})(1 - e^{-\max(\sigma_j, \sigma_{s+l-j})})} \frac{B_r^{s+l-2} C_1^2}{\gamma}, \\ C_{\psi,s} &\leq \left(\frac{8(s-1)\eta_{s-1} C_1}{s(1 - e^{-(\sigma_0 - \sigma_s)})(1 - e^{-\sigma_0})} + \sum_{l=1}^{s-1} \frac{l B_r}{s} \eta_l \theta_{s-l} \right) \frac{B_r^{s-2} C_1}{\gamma}, \end{aligned} \tag{4.19}$$

We observe that, from the first of (4.14), we have

$$\begin{aligned} 1 - e^{-\max(\sigma_j, \sigma_{s+l-j})} &\geq \frac{1 - e^{-\sigma_0}}{\sigma_0} \max\{\sigma_j, \sigma_{s+l-j}\} \\ &= \frac{1 - e^{-\sigma_0}}{\sigma_0} \left(\sigma_0 - \frac{\sigma_0 - \sigma_*}{r} \min\{j, s+l-j\} \right) \\ &> \frac{1 - e^{-\sigma_0}}{\sigma_0} \left(\frac{\sigma_0 + \sigma_*}{2} \right) \end{aligned}$$

since $\min\{j, s-j\} \leq r/2$, thus it gives

$$1 - e^{-\max(\sigma_j, \sigma_{s+l-j})} > \frac{1 - e^{-\sigma_0}}{2}.$$

Using the second of (4.14) in a similar way to deal with

$$1 - e^{-[\max(\sigma_j, \sigma_{s+l-j}) - \sigma_{s+l}]} \geq \frac{s+l - \min\{j, s+l-j\}}{r} \left(1 - e^{-(\sigma_0 - \sigma_*)} \right),$$

and setting

$$B_r = \frac{16C_1 r}{\gamma(1 - e^{-(\sigma_0 - \sigma_*)})(1 - e^{-\sigma_0})}.$$

we get

$$\begin{aligned} C_{\zeta, l, s} &\leq \frac{1}{s} \sum_{j=1}^s j \eta_j \eta_l \theta_{s-j} B_r^{s+l-1} C_1, \\ C_{\psi, s} &\leq \left(\frac{1}{s} \eta_{s-1} + \sum_{l=1}^{s-1} \frac{l}{s} \eta_l \theta_{s-l} \right) B_r^{s-1} C_1, \end{aligned}$$

Therefore, the required inequalities (4.18) are satisfied by the sequences recursively defined as

$$\begin{aligned} \theta_s &= \sum_{j=1}^s \frac{j}{s} \eta_j \theta_{s-j} \quad \text{for } s \geq 1, \\ \eta_s &= \eta_{s-1} + \sum_{j=1}^{s-1} j \eta_j \theta_{s-j} \quad \text{for } s \geq 2. \end{aligned}$$

starting with $\eta_1 = \theta_0 = 1$. Actually, to find an estimate for the generating function it is enough to investigate the sequence η_1, \dots, η_r . To this end, after multiplication by a factor $1/s$, we subtract the second relation from the first one, thus getting

$$\theta_1 = 1, \quad \theta_s = \left(\frac{s+1}{s} \right) \eta_s - \frac{1}{s} \eta_{s-1} < 2\eta_s - \frac{1}{s} \eta_{s-1}.$$

Then, we substitute the latter expression for θ_{s-j} in the second of the relations above and get

$$\eta_s < \eta_{s-1} + \sum_{j=1}^{s-1} 2j \eta_j \eta_{s-j} < 3 \sum_{j=1}^{s-1} j \eta_j \eta_{s-j}.$$

Hence, the wanted inequality (4.18) for $C_{\psi,s}$ is satisfied by the sequence

$$\eta_1 = 1, \quad \eta_s = 3 \sum_{j=1}^{s-1} j \eta_j \eta_{s-j} = 3s \sum_{j=1}^{\lfloor s/2 \rfloor} \eta_j \eta_{s-j}, \quad s \geq 2.$$

By induction it is possible to prove that $\eta_s \leq 9^{s-1} s!$ for all $s = 1, \dots, r$: indeed, it holds

$$x_s \leq 9^{s-1} \frac{s}{3} \sum_{j=1}^{\lfloor s/2 \rfloor} j!(s-j)! \leq 9^{s-1} s!,$$

provided

$$\sum_{j=1}^{\lfloor s/2 \rfloor} j!(s-j)! = \sum_{j=2}^{\lfloor s/2 \rfloor} j!(s-j)! \leq 2(s-1)!;$$

the latter being true since for $4 \leq s \leq r$ and $2 \leq j \leq \lfloor s/2 \rfloor$

$$j! \frac{(s-j)!}{(s-1)!} = \prod_{i=0}^{j-2} \left(\frac{j-i}{s-j-i} \right) \leq \left(\frac{2}{3} \right)^{j-1}.$$

Then, by $s! \leq (\sqrt{e})^{-(s-1)} s^s$ we obtain for $1 \leq s \leq r$

$$\eta_s \leq \left(\frac{9}{\sqrt{e}} \right)^{s-1} s^s < 4^{s-1} r^{s-1}$$

Replacing this and (4.19) in the inequality (4.18) for $C_{\psi,s}$ and recalling that $C_{\chi,s} \leq C_{\psi,s}/\gamma$, we have

$$C_{\chi,s} \leq (64r^2 C_*)^{s-1} \frac{C_1}{\gamma^s}, \quad C_* = \frac{C_1}{\gamma(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})}.$$

This proves the statement (ii) of Proposition 4.1 with the estimated value of C_* in (4.3). The statement (iii) also follows in view of $C_{\zeta,s} \leq C_{\psi,s}$.

Estimate of the Truncated First Integral. We give an estimate for the first r terms of $T_{\mathcal{X}}\Phi$ where Φ is a homogeneous polynomial, as specified in the statement (iv) of Proposition (4.1). We look for a sequence $C_{\varphi,s}$ of constants such that $E_s\Phi$ is of class $\mathcal{D}(C_{\varphi,s}, \sigma_s)$ for $s = 0, \dots, r$. Of course we have $C_{\varphi,0} = C_{\varphi}$, so we look for a recursive estimate for $s > 0$ using Lemma 2.4 and the Definition (4.1) of $T_{\mathcal{X}}$. Recalling (4.14) we have that $E_s\Phi$ is of class $\mathcal{D}(A, \sigma_s)$ with a constant A satisfying

$$\begin{aligned} A &\leq \sum_{j=1}^s \frac{j}{s} \cdot \frac{4(j+1)(s-j+m+1)}{(1 - e^{-\max(\sigma_j, \sigma_{s-j}) + \sigma_s})(1 - e^{-\max(\sigma_j, \sigma_{s-j})})} \cdot \frac{C_1}{j\gamma} (C_r)^{j-1} C_{\varphi, s-j} \\ &\leq \sum_{j=1}^s \frac{(s-j+m+1)}{\max(j, s-j)} \cdot \frac{16C_1 r}{\gamma(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_*)})} (C_r)^{j-1} C_{\varphi, s-j} \end{aligned}$$

Thus, recalling the Definition (4.3) of C_* , we may set

$$C_{\varphi,s} = \frac{1}{4} \sum_{j=1}^s \frac{s-j+m+1}{\max(j, s-j)} (C_r)^j C_{\varphi,s-j}.$$

For $s = 1$ this gives

$$C_{\varphi,1} = (m+1)r^2 C_* C_{\varphi}, \tag{4.20}$$

so that the claim is true with F_r as given in (4.3). For $s > 1$, we extract from the sum the term $j = 1$ and replace the index j with $j + 1$ in the rest of the sum, thus getting

$$\begin{aligned} C_{\varphi,s} &\leq \frac{(s+m)}{4(s-1)} C_r C_{\varphi,s-1} + \frac{C_r}{4} \sum_{j=1}^{s-1} \frac{s-j+m}{\max(j+1, s-1-j)} (C_r)^j C_{\varphi,s-1-j}, \\ &\leq \frac{s+m}{4(s-1)} C_r C_{\varphi,s-1} + \frac{1}{4} C_r C_{\varphi,s-1} \\ &\leq \frac{m+2}{4} C_r C_{\varphi,s-1}. \end{aligned}$$

This proves the statement (iv) of Proposition 4.1.

Concerning the statement (v), we remark that $L_{\mathcal{X}_1} H_{\omega} = -L_{\Omega} \mathcal{X}_1 = -Z_1 - H_1$ in view of the homological equation at order 1. Therefore, we may replace (4.20) with $C_{\varphi,1} = C_1$. For $s > 1$, the argument above for a generic function Φ requires only a minor modification and one obtains the same recursive relation for $C_{\varphi,s}$, where we just replace a different value for $C_{\varphi,1}$. This proves the claim.

Estimate of the Time Derivative of the Approximate First Integral. We come to the statement (vi) of Proposition 4.1. Recall that by construction we have $T_{\mathcal{X}} Z - H = \mathcal{O}(r+2)$, meaning that its expansion starts with terms of degree at least $2(r+2)$. Since $L_{\Omega} Z_s = 0$ for $s = 0, \dots, r$ and recalling the general property $T_{\mathcal{X}}\{f, g\} = \{T_{\mathcal{X}}f, T_{\mathcal{X}}g\}$, we immediately have

$$\{H, T_{\mathcal{X}}\Phi_0\} = T_{\mathcal{X}}\{Z, \Phi_0\} = \mathcal{O}(r+2)$$

On the other hand, since $T_{\mathcal{X}}\Phi_0 - \Phi^{(r)} = \mathcal{O}(r+2)$, we also have $\{H, \Phi^{(r)}\} = \mathcal{O}(r+2)$. Substituting the expansions $H = H_0 + H_1$ and $\Phi^{(r)} = \Phi_0 + \dots + \Phi_r$ we get $\dot{\Phi}^{(r)} = \{H, \Phi^{(r)}\} = \{H_1, \Phi_r\}$, which is an extensive homogeneous polynomial of degree $2r+4$, as claimed. Recalling that H_1 is of class $\mathcal{D}(C_1, \sigma_0)$ and Φ_r is of class $\mathcal{D}(F_r^{r-1}C_1, \sigma_r)$, in view of the statement (v), a straightforward application of Lemma 2.4 gives

$$C_{\rho} \leq \frac{8(r+2)F_r^{r-1}C_1^2}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_r)})},$$

and the result follows by just replacing the estimated value of F_r from statement (iv), with $m = 0$.

This concludes the proof of Proposition 4.1.

5. Long Time Estimates and Statistical Control of Fluctuations

In this section we actually present, and prove, the main result of the paper in its complete and detailed form; the results given in the introduction, i.e. Theorems 1.1 and 1.3, are simplified statements with some particular choices of the parameters involved.

We first stress that although the whole perturbative construction of our conserved quantity $\Phi \equiv \Phi^{(r)}$ is based upon an initial normal form transformation of the quadratic part of the Hamiltonian, i.e. there is a change of coordinates at the very beginning of our procedure, we will state our result and the corresponding proof in the original⁴ variables $z = (x, y)$.

As explained in the introduction, our aim is to show that Φ is a good adiabatic invariant over a long time scale: to this purpose, we introduce its variation over a time interval

$$\Delta_t \Phi(z) := \Phi(\phi^t(z)) - \Phi(z),$$

where $\phi^t(z)$ is the Hamiltonian flow. We will show that $\Delta_t \Phi$ remains small, compared with the phase variance of Φ , over a long time scale, for a set of initial data z of large Gibbs measure. This kind of control is quite weak for all the times between 0 and t , since the set of large measure is in principle allowed to change if we change the final t to control the intermediate times. We thus give two stronger estimates: the first deals with $\Delta_t \Phi$. Its smallness implies that for every large deviation of a given sign at intermediate times must correspond a similar deviation with the opposite sign. An even stronger control is obtained with the smallness of $\sigma_t^2[\Delta_t \Phi]$: in this case, we have that $\Delta_s \Phi$ is small also for all $s \in (0, t)$.

5.1. Main Result

Concerning the time scale over which we are able to control the evolution of our adiabatic invariant, we have actually two types of estimates, as a power law and a stretched exponential, each with its own set of hypothesis and constants, but with a similar formulation; we thus present the two results together. To simplify the statement, we find it convenient to formulate in advance the hypothesis and definitions under which the result holds in those two cases. In particular, we define the time scale \bar{t} and the corresponding bounds on β .

Given the constants⁵ $a_0, \mu_*, \mu_2, \beta^*, \beta_0, \beta_1, \beta_3, \beta_4, K_1$, defining

$$\mu_0 := \frac{a_0}{1 + 2a_0}, \quad \mu_1 := \frac{\left(1 - \frac{3}{4(\max\{\beta_0, 1\})^2}\right)^8}{64K_1^8(1 + 4a_0)^4},$$

we introduce

⁴ We will take care of this via the application of Lemma 3.3 throughout the proof.

⁵ See Propositions 4.1, 5.4, 5.8 and Lemmas 6.2 and 5.9.

HD1 (power law estimate). there exist $\beta_2 > 0$ and $r^*(\mu) = \frac{\mu_*}{2\mu}$ such that for any integer $r \in [1, r^*)$ and for any $\nu \in (0, 1]$, defining

$$\beta_* := \max \left\{ \beta_0, \beta_1, \beta_2, \beta_3, \beta_4, (\beta^* r^3)^{1/\nu}, \sqrt{3/2} \right\},$$

$$\mu^* := \min \left\{ \mu_0, \mu_1, \mu_2, \frac{1}{8} \right\},$$

then

assume $\beta_* \leq \beta$, and define $\lambda := r(1 - \nu) + 1 - \nu/2$,
 $\bar{t} := \beta^\lambda$.

HD2 (exponential estimate). there exists $\mu_3 > 0$ such that defining

$$\beta_* := \max \left\{ \beta_0, \beta_1, \beta_3, \beta_4, 64e\beta^*, \sqrt{3/2} \right\}$$

$$\mu^* := \min \left\{ \mu_0, \mu_1, \mu_2, \mu_3, \mu_* \sqrt[3]{\frac{e\beta^*}{\beta_*}}, \frac{1}{8} \right\},$$

then

assume $\beta_* \leq \beta < e\beta^* \left(\frac{\mu_*}{\mu}\right)^3$ and define $\kappa := \frac{3}{2} \sqrt[3]{\frac{\beta}{e\beta^*}}$,
 $\bar{t} := \frac{\kappa^{9/2} e^{\kappa/2}}{\beta}$.

We are now ready to state the result:

Theorem 5.1. For either the hypothesis and definitions of case **HD1** or those of case **HD2**, there exist constants $K > 1$ such that, for all $\mu < \mu^*$, and for any positive δ one has

$$m\left(z \in \mathbb{R}^{2N} : |\Delta_t \Phi(z)| \geq \delta \sigma[\Phi]\right) \leq \frac{12K}{\delta^2} \left(\frac{t}{\bar{t}}\right)^2,$$

$$m\left(z \in \mathbb{R}^{2N} : |\overline{\Delta_t \Phi}(z)| \geq \delta \sigma[\Phi]\right) \leq \frac{3K}{\delta^2} \left(\frac{t}{\bar{t}}\right)^2,$$

$$m\left(z \in \mathbb{R}^{2N} : \sigma_t^2[\Delta_t \Phi(z)] \geq \delta \sigma^2[\Phi]\right) \leq \frac{4K}{\delta} \left(\frac{t}{\bar{t}}\right)^2.$$

Remark 5.2. The estimates contained in the above theorem can be seen as the generalizations of Propositions 2, 3 and 4 of paper [24]. The result of paper [16] can be compared with the first estimate, with the hypothesis and definition set **HD2** in the case of vanishing coupling constant; the only difference being a slightly improved exponent for the argument of the exponential (1/3 in our case, 1/4 in their result).

Proof. The first two estimates of the theorem are actually Tchebychev estimates, while the third is a Markov estimate. We recall that, in our notations,

given any measurable function f and any real $\eta > 0$, for $p = 1$ and 2 respectively, Markov and Tchebychev estimates are

$$m\left(z \in \mathbb{R}^{2N} : |f(z)| \geq \eta\right) \leq \frac{\langle |f|^p \rangle}{\eta^p}.$$

Choosing $\eta = \delta\sigma[\Phi]$, and using Tchebychev for the first two estimates, and with $\eta = \delta\sigma^2[\Phi]$ and using Markov for the third one, one has to control respectively the following three quantities

$$\frac{\langle (\Delta_t \Phi)^2 \rangle}{\delta^2 \sigma^2[\Phi]}, \quad \frac{\langle (\overline{\Delta_t \Phi})^2 \rangle}{\delta^2 \sigma^2[\Phi]}, \quad \frac{\langle \sigma_t^2[\Delta_t \Phi] \rangle}{\delta \sigma^2[\Phi]}. \tag{5.1}$$

By the rough inequality $\sigma_t^2[\Delta_t \Phi(z)] \leq \overline{(\Delta_t \Phi)^2}(z)$ it is clear that to estimate the three quantities above, we need to control the phase average of $(\Delta_t \Phi)^2$, $(\overline{\Delta_t \Phi})^2$ and $\overline{(\Delta_t \Phi)^2}$. By defining

$$R := \{\Phi^{(r)}, H\} = \{\Phi_r, H_1\}, \tag{5.2}$$

we have $\Delta_t \Phi(z) = -\int_0^t R \circ \phi^s(z) ds$, so that we may write

$$\begin{aligned} \langle (\Delta_t \Phi)^2 \rangle &= \left\langle \int_{[0,t]^2} (R \circ \phi^{s_1})(R \circ \phi^{s_2}) ds_1 ds_2 \right\rangle \\ &= \int_{[0,t]^2} \langle (R \circ \phi^{s_1})(R \circ \phi^{s_2}) \rangle ds_1 ds_2 \\ &\leq \int_{[0,t]^2} \|R \circ \phi^{s_1}\|_{L^2} \|R \circ \phi^{s_2}\|_{L^2} ds_1 ds_2 \\ &= \int_{[0,t]^2} \|R\|_{L^2}^2 ds_1 ds_2 = t^2 \langle R^2 \rangle. \end{aligned}$$

where we used Fubini's theorem, Schwartz inequality and the invariance of the measure under the Hamiltonian flow. For the second quantity we need to add a further double integration over time to perform the time average, but the scheme is the same:

$$\begin{aligned} \langle (\overline{\Delta_t \Phi})^2 \rangle &= \left\langle \frac{1}{t^2} \int_{[0,t]^2} \Delta_{s_1} \Phi \Delta_{s_2} \Phi ds_1 ds_2 \right\rangle \\ &= \frac{1}{t^2} \int_{[0,t]^2} \left\langle \int_{[0,s_1] \times [0,s_2]} (R \circ \phi^{\tau_1})(R \circ \phi^{\tau_2}) d\tau_1 d\tau_2 \right\rangle ds_1 ds_2 \\ &\leq \frac{t^2}{4} \langle R^2 \rangle. \end{aligned}$$

In the third case, we instead have a single time integration from the time average:

$$\begin{aligned} \langle \overline{(\Delta_t \Phi)^2} \rangle &= \left\langle \frac{1}{t} \int_{[0,t]} (\Delta_s \Phi)^2 ds \right\rangle = \frac{1}{t} \int_{[0,t]} \langle (\Delta_s \Phi)^2 \rangle ds \\ &\leq \frac{1}{t} \int_{[0,t]} s^2 \langle R^2 \rangle ds = \frac{t^2}{3} \langle R^2 \rangle. \end{aligned}$$

For all the three quantities (5.1), everything we thus need to control the quotient of $\langle R^2 \rangle$ over $\sigma^2[\Phi]$, with an upper bound for the numerator and a lower bound for the denominator: for the former we apply Proposition 5.4, and for the latter we use Proposition 5.8. In particular, the hypothesis and definition sets **HD1** and **HD2** imply the hypothesis of Proposition 5.4 part 1 and, respectively, part 2.

Indeed if **HD1** holds, from $\mu < \mu^*$ we have that $a < \min\{a_0, 1/6\}$ ($\mu^* < \min\{\mu_0, 1/8\}$) and⁶ $D^2 \mu_b < 1/2$ ($\mu < \mu_1$). The condition $\mu < \mu_2$ is required in Lemma 5.9, for the lower bound of $\sigma^2[\Phi]$. With respect to the hypothesis of Proposition 5.8 we observe that if **HD1** holds, the bounds on β are satisfied (for β large enough, setting β_2 as the threshold), since we need it to be scaling like r^3 and according to **HD1** we have it scaling as $r^{(3/\nu)}$ with $\nu \leq 1$.

If **HD2** holds, since in that case from Proposition 5.4 we set the optimal integer r as $\lfloor \kappa/3 \rfloor$, the condition on β translates in the following inequality

$$1 > \frac{3}{4} e^{\tilde{C}-1} \left(1 + 2 \sqrt[3]{\frac{e\beta^*}{\beta}} \right)$$

which is true since \tilde{C} vanishes with μ (set here μ_3 as the threshold).

Using also the constants Ω (Proposition 3.1), C_1 (Proposition 4.1) and K_2 (Proposition 5.3), setting

$$K := \frac{3^2 5 e^6 K_1^2 K_2 \Omega^4}{2 C_1^2} \cdot \begin{cases} 2\beta^{*3} & \text{case HD1} \\ \frac{3^8}{e^4} & \text{case HD2} \end{cases}$$

we have the thesis. □

Proof of Theorem 1.1. After observing that $\sigma_t^2[\Delta_t \Phi] = \sigma_t^2[\Phi]$, use the third estimate of Theorem 5.1, hypothesis and definitions set **HD1**, with $r = \lfloor r^* \rfloor$, $\nu = \frac{1}{2}$, $\delta = \beta^{-1/2}$ and letting only $\beta^{r/2}$ in the time scale \bar{t} . □

Proof of Theorem 1.3. Applying Theorem 5.1, hypothesis and definitions set **HD2**, third estimate, with $\delta = \beta^{-1/2}$ and letting only $e^{c\kappa}$ in the time scale \bar{t} with a constant c slightly smaller than $1/2$ to get the correct power of β outside the exponential factor. The upper bound on $\sqrt{a} \sqrt[3]{\beta}$ represents last condition in **HD2** using the Definition (3.7) of μ . □

⁶ The constant D is defined in Appendix 6.2 and recalled in Section 5.2, while μ_b in (5.11).

The rest of the section is devoted to the proofs of the upper bound of $\langle R^2 \rangle$, in Subsect. 5.3, and of the lower bound of $\sigma^2[\Phi]$ in Subsect. 5.4. Due to its relevance and to the slightly different techniques involved, we anticipate in Subsect. 5.2 the result on the control of the decay of correlations.

5.2. Decay of Correlations

The main result of this section is an estimate of the correlation between two polynomials with disjoint supports: we show such a correlation to be (at least) small as a^d where d is the distance between the two supports.

Let β_0 , a_0 and K_1 be the constants of Lemma 6.2. Consider also these other constants⁷ defined in Appendix 6.2:

$$A_1 = \sqrt{1 + 4a}, \quad A_2 = \sqrt{1 - 2a}, \quad B = 1 - \frac{3}{4\beta^2}, \quad D = \frac{K_1 A_1(a)}{B(\beta)};$$

Introduce the following:

$$\mu_{\sharp} := a(2 + a), \quad K_2 := \frac{4(2K_1)^{2+8a_0}}{(1 - 2a_0)^4(1 - a_0)^6}. \tag{5.3}$$

Proposition 5.3. *Let N be the length of the periodic chain. Let ϕ and ψ be two homogeneous polynomial of degree $2r$ and $2s$, respectively, and interaction length m and m' , respectively; suppose their supports are disjoint and denote by d their distance,⁸ then for any $\beta > \beta_0$ and $a < a_0$ it holds*

$$|\langle \phi\psi \rangle - \langle \phi \rangle \langle \psi \rangle| \leq K_2 \left[D^{m+m'+2d+4} \right] \mu_{\sharp}^d \left[\frac{2^{r+s} r! s!}{(A_2^2 \beta)^{r+s}} \right] \|\phi\| \|\psi\|.$$

Before entering into the details of the proof it is necessary to introduce another notation for the measure that will be useful also in the sequel of this section. To this purpose, we split the original Hamiltonian in a different way. We recall that H is naturally split in two different terms H_0 and H_1 [see (3.1)] according to the degree, but here we want to put into evidence the coupling terms of H . There are two possible choices, the first being to separate all the terms depending on the coupling constant a , i.e. $\frac{a}{2} \sum_j (x_{j+1} - x_j)^2$. We instead separate the diagonal and the off-diagonal part of A [see (3.2) and (6.1)] like in Proposition 3.1, but maintaining the original variables; in this way, we put into evidence the real coupling terms. Accordingly we define, on a subset of variables, the uncoupled component of the Gibbs measure by

$$dV_s^{(m)} := \prod_{j=s}^{s+m-1} e^{-\beta \left[(1+2a) \frac{x_j^2}{2} + \frac{x_j^4}{4} \right]} dx_j,$$

which depend only on m variables, and the coupling part by

$$[p, q] := e^{\beta a x_p x_{p+1}} \dots e^{\beta a x_{q-1} x_q},$$

⁷ A_j are actually functions of a , B a function of β and D a function of both the parameters, but all these quantities are asymptotically constants as $a \rightarrow 0$ and $\beta \rightarrow \infty$.

⁸ If $p = \min S(\phi)$, $q = \max S(\phi)$, $t = \min S(\psi)$ and $u = \max S(\psi)$, with $q < t$, then $d = \min(t - q - 1, N - u + p - 1)$.

for⁹ $p < q \leq p + N$. We observe that, for any $m < l$, it is possible to factorize both the component of the measure: $dV_s^{(l)} = dV_s^{(m)} dV_{s+m}^{(l-m)}$ and $[s, s + l] = [s, s + m][s + m, s + l]$. Whenever $S(\phi) \subset \{p, \dots, q\}$, we will write $[p, \phi, q] := \phi[q, p]$, to stress the bound on the support. The full Gibbs measure, again ignoring the y variables, is then given by $[0, l]dV_0^{(l)}$ for a system with periodic boundary conditions,¹⁰ so the partition function will be

$$Z_l := \int_{\mathbb{R}^l} [0, l]dV_0^{(l)} = \int_{\mathbb{R}^l} e^{-\beta H(x)} dx.$$

Proof. The proof consists of two main steps: in the first one we show the presence of several cancellations, while in the second one we actually estimate the remaining terms.

Without any loss of generality we may assume ϕ to be left aligned, so that its support is contained in $\{0, \dots, m - 1\}$; denote by t the minimal index in the support of ψ , which will be therefore contained in $\{t, \dots, t + m' - 1\}$.

As a first step, we rescale all the variables by a factor $\sqrt{\beta}$. We need to introduce a corresponding notation for the relevant objects with the rescaled variables, and to this purpose we will systematically add a \diamond as a superscript when needed. We remark that the powers of β appearing as a multiplying factor will not be included in the \diamond -objects. For example, we have

$$[p, q]^\diamond := e^{ax_p x_{p+1}} \dots e^{ax_{q-1} x_q}, \quad dV_s^{(m)\diamond} := \prod_{j=s}^{s+m-1} e^{-\left[(1+2a)\frac{x_j^2}{2} + \frac{x_j^4}{4\beta^2}\right]} dx_j. \tag{5.5}$$

If we perform such a scaling on the correlation, we get

$$\langle \phi \psi \rangle - \langle \phi \rangle \langle \psi \rangle = \frac{1}{\beta^{r+s}} \left(\langle \phi \psi \rangle^\diamond - \langle \phi \rangle^\diamond \langle \psi \rangle^\diamond \right).$$

We need to introduce a further notation related to the coupling terms $[p, q]^\diamond$ of the measure. These terms are products of factors of the form e^α , each being the coupling term between two consecutive sites of the chain; to possibly decouple the chain in several positions we use the trivial identity $e^\alpha = 1 + (e^\alpha - 1)$, so that for example $[0, m]^\diamond$ turns out to be the sum of 2^m terms each of which is the product of m factors: for every $j = 0, \dots, m - 1$ the factor can be either 1 or $e^{ax_j x_{j+1}} - 1$. We will identify each term in the sum with a string of m symbols in $\{0, 1\}$: 0 for the factor $e^{ax_j x_{j+1}} - 1$, and 1 for the factor 1. For example, for $m = 4$, a possible factor is

$$k = 0010 \rightarrow [0, m]_k^\diamond = (e^{ax_0 x_1} - 1) \cdot (e^{ax_1 x_2} - 1) \cdot 1 \cdot (e^{ax_3 x_4} - 1).$$

⁹ We remark that in this notation, given the generic dimension l of the space, the indexes must be considered modulo l .

¹⁰ for a system with free boundary conditions the measure is $[0, l - 1]dV_0^{(l)}$, so the partition function will be

$$Z_l := \int_{\mathbb{R}^l} [0, l - 1]dV_0^{(l)} = \int_{\mathbb{R}^l} e^{-\beta \mathcal{H}(x)} dx. \tag{5.4}$$

Thus, we may write $[0, m]^\diamond = \sum_{\mathbf{k}} [0, m]_{\mathbf{k}}^\diamond$: we will use this kind of expansion for the coupling terms involving sites outside the support of the polynomials. We remark that a factor of the type $e^\alpha - 1$ is roughly of order α , i.e. in our case of order a , so the number of “zeros” in the sequence \mathbf{k} can be used to quantify the smallness of the corresponding term.

Let us now rewrite the correlation collecting the partition function in the denominators

$$\langle \phi \psi \rangle^\diamond - \langle \phi \rangle^\diamond \langle \psi \rangle^\diamond = \frac{\langle \langle \phi \psi \rangle \rangle^\diamond Z^\diamond - \langle \langle \phi \rangle \rangle^\diamond \langle \langle \psi \rangle \rangle^\diamond}{Z^{\diamond 2}}, \tag{5.6}$$

where we used the notation $\langle \langle \phi \rangle \rangle := Z \cdot \langle \phi \rangle = \int \phi(x) e^{-\beta H(x)} dx$, and let us concentrate our attention on the numerator.

According to the supports of ϕ and ψ , we split the coupling part of the measure in the following way

$$[0, N]^\diamond = [0, m - 1]^\diamond \cdot [m - 1, t]^\diamond \cdot [t, t + m' - 1]^\diamond \cdot [t + m' - 1, N]^\diamond;$$

moreover, in every integral, we will expand the “holes” between the supports:

$$[m - 1, t]^\diamond = \sum_{\mathbf{j}} [m - 1, t]_{\mathbf{j}}^\diamond, \quad [t + m' - 1, N]^\diamond = \sum_{\mathbf{k}} [t, t + m' - 1]_{\mathbf{k}}^\diamond.$$

We rewrite the two addenda of the numerator of (5.6) as

$$\begin{aligned} \langle \langle \phi \psi \rangle \rangle^\diamond Z_N^\diamond &= \sum_{\mathbf{j}} \sum_{\mathbf{k}} \sum_{\mathbf{j}'} \sum_{\mathbf{k}'} \\ &\int_{\mathbb{R}^N} [0, \phi, m - 1]^\diamond [m - 1, t]_{\mathbf{j}}^\diamond [t, \psi, t + m' - 1]^\diamond [t + m' - 1, N]_{\mathbf{k}}^\diamond dV_0^{(N)\diamond} \\ &\times \int_{\mathbb{R}^N} [0, m - 1]^\diamond [m - 1, t]_{\mathbf{j}'}^\diamond [t, t + m' - 1]^\diamond [t + m' - 1, N]_{\mathbf{k}'}^\diamond dV_0^{(N)\diamond}, \\ \langle \langle \phi \rangle \rangle^\diamond \langle \langle \psi \rangle \rangle^\diamond &= \sum_{\mathbf{j}} \sum_{\mathbf{k}} \sum_{\mathbf{j}'} \sum_{\mathbf{k}'} \\ &\int_{\mathbb{R}^N} [0, \phi, m - 1]^\diamond [m - 1, t]_{\mathbf{j}}^\diamond [t, t + m' - 1]^\diamond [t + m' - 1, N]_{\mathbf{k}}^\diamond dV_0^{(N)\diamond} \\ &\times \int_{\mathbb{R}^N} [0, m - 1]^\diamond [m - 1, t]_{\mathbf{j}'}^\diamond [t, \psi, t + m' - 1]^\diamond [t + m' - 1, N]_{\mathbf{k}'}^\diamond dV_0^{(N)\diamond} \end{aligned} \tag{5.7}$$

Given such a decomposition of the correlation $\langle \phi \psi \rangle^\diamond - \langle \phi \rangle^\diamond \langle \psi \rangle^\diamond$, and using the bitwise “and” operator \wedge [see (6.8) in Appendix 6.3 for a formal definition], we give the main claim of the proof:

1. All the terms such that $\mathbf{j} \wedge \mathbf{j}' \neq 0$ AND $\mathbf{k} \wedge \mathbf{k}' \neq 0$ cancel;
 2. All the terms such that $\mathbf{j} \wedge \mathbf{j}' = 0$ OR $\mathbf{k} \wedge \mathbf{k}' = 0$ are (at least) of order a^d .
- The idea behind the cancellations is that $\mathbf{j} \wedge \mathbf{j}' \neq 0$ ensure the presence (see Remark 6.16) of at least a “1” in the same position in both \mathbf{j} and \mathbf{j}' : this corresponds to the absence of the coupling term so that both the integrals in

each of the expressions in (5.7) can be split in the same position; the same happens with \mathbf{k} and \mathbf{k}' . This opportunity to cut the integrals in the holes between the supports of ϕ and ψ allows us to rearrange the terms to show that actually all these terms cancel. The formal proof is deferred to Appendix 6.3.

Concerning the second part instead, the idea is that $j \wedge j' = 0$ ensure the presence of enough “zeros”, each contributing with an order in a . To make the argument more precise, let us first assume that the strings j and j' are not longer than \mathbf{k} and \mathbf{k}' ; thus, according to the statement of the proposition, the length l of both j and j' is equal to $d + 2$.

Let us consider first the case in which $j \wedge j' = 0$. Since we must consider all the cases for \mathbf{k} and \mathbf{k}' , we actually do not expand the second hole between the supports of ϕ and ψ . Moreover, instead of expanding the whole term $[m - 1, t]$, we will write

$$[m - 1, t]^\diamond = [m - 1, m]^\diamond \sum_j [m, t - 1]_j^\diamond [t - 1, t]^\diamond,$$

and similarly for the same hole in the other integral. Please note that, despite the use of the same letter, now the string j has length exactly d : expanding only on the “interior” of the hole simply means that we will include in our estimates some terms that actually could be avoided because they cancel.

The strategy is thus to apply Lemma 6.8 to $\langle \phi \psi \rangle_N^\diamond$ and $\langle 1 \rangle_N^\diamond$, cutting the chain in four parts for the first average and into two parts for the second one:

$$\begin{aligned} & \langle \phi \psi \rangle_N^\diamond \langle 1 \rangle_N^\diamond \\ & \leq K_1^{m+m'+d} \left\langle \phi e^{a(x_0^2+x_{m-1}^2)} \right\rangle_m \frac{1}{Z_d^\diamond} \sum_j \int_{\mathbb{R}^d} e^{\frac{a}{2}(x_m^2+x_{t-1}^2)} [m, t - 1]_j^\diamond dV_m^{(d)\diamond} \\ & \quad \left\langle \psi e^{a(x_t^2+x_{t+m'-1}^2)} \right\rangle_{m'} \left\langle e^{a(x_{t+m'}^2+x_{N-1}^2)} \right\rangle_{N-m-m'-d}^\diamond \\ & \quad K_1^d \left\langle e^{a(x_t^2+x_{m-1}^2)} \right\rangle_{N-d} \frac{1}{Z_d^\diamond} \sum_{j'} \int_{\mathbb{R}^d} e^{\frac{a}{2}(x_m^2+x_{t-1}^2)} [m, t - 1]_{j'}^\diamond dV_m^{(d)\diamond}. \end{aligned} \tag{5.8}$$

We first deal with the sum given by the expansion of the smaller holes; for the other terms we will apply some Lemmas proven in the Appendix. Introducing the notation $|j|$ to count the number of “zeros” in the string j , according to Remark 6.16 we have $d \leq |j| + |j'|$; clearly it also holds $|j| + |j'| \leq 2d$.

We need to estimate the generic term $[m, t - 1]_j^\diamond$; we will use the inequality $e^\alpha - 1 \leq ae^\alpha$ on each of the “zero” factors, and then apply the estimate which decouple the measure:

$$[m, t - 1]_j^\diamond \leq a^{|j|} \prod_{j_n=0} x_n x_{n+1} e^{ax_n x_{n+1}} \leq a^{|j|} \prod_{j_n=0} x_n x_{n+1} e^{\alpha \frac{1}{2}(x_n^2+x_{n+1}^2)}.$$

With the use of the previous estimate, the measure within $(m, t - 1)$ can be factorized so that we have, for $b \in \{0, 1, 2\}$, the following integrals:

$$\int_{\mathbb{R}} |x|^b e^{-\left(\alpha_b \frac{x^2}{2} + \frac{x^4}{4\beta^2}\right)} \leq \left(\frac{2}{\alpha_b}\right)^{\frac{b+1}{2}} \Gamma\left(\frac{b+1}{2}\right) \quad \alpha_b = 1 + (2-b)a,$$

with the exception of the boundary sites m and $t-1$ where $\alpha_b = 1 + (1-b)a$. Since it could be not completely trivial to control, for all the possible strings j , how many factors have $b = 0$ or $b = 1$ or $b = 2$, the idea is to evaluate at the same time the terms coming from $[m, t-1]_j^\diamond$ with those coming from $[m, t-1]_{j'}^\diamond$. Let us consider a site n with $m < n < t-1$, and suppose that from $[m, t-1]_j^\diamond$ we have a term with $b = 0$: this is possible if and only if $j_n = 1$ and $j_{n-1} = 1$. But then, from $j \wedge j' = 0$, one has $j'_n = 0$ and $j'_{n-1} = 0$ which implies that the corresponding term from $[m, t-1]_{j'}^\diamond$ will be an integral with $b = 2$. Thus, the product of the two terms will be bounded by 2π . If we start $b = 1$, this is compatible with both 01 and 10 as substrings of j , which implies respectively 10 or 00 , and 01 or 00 for j' : thus, from $[m, t-1]_j^\diamond$, the contribution will be with $b = 1$ or $b = 2$, and the product can be bounded in the same way. As a last case, starting from $b = 2$, this requires 00 for j which is compatible with all the cases in j' : again, the product of the two terms will be bounded by 2π . For the two boundary sites, one has the same kind of control¹¹ simply with a factor $(1-a)^{\frac{3}{2}}$ on each site for the worst case.

We thus have the following estimate

$$\begin{aligned} & \int_{\mathbb{R}^d} e^{\frac{a}{2}(x_m^2+x_{t-1}^2)} [m, t-1]_j^\diamond dV_m^{(d)\diamond} \int_{\mathbb{R}^d} e^{\frac{a}{2}(x_m^2+x_{t-1}^2)} [m, t-1]_{j'}^\diamond dV_m^{(d)\diamond} \\ & \leq \frac{a^{|j|+|j'|}}{(1-a)^6} (2\pi)^d. \end{aligned}$$

We need now to control the sum over all the possible strings such that $j \wedge j' = 0$. We exploit the fact that $d \leq |j| + |j'| \leq 2d$ and count the number of configurations for the couple of strings j and j' with a given value of $|j| + |j'|$. It is easy to verify that

$$\#\{(j, j') : |j| + |j'| = d + i\} = \binom{d}{i} 2^{d-i},$$

so we end up with

$$\begin{aligned} & \sum_{j, j'} \int_{\mathbb{R}^d} e^{\frac{a}{2}(x_m^2+x_{t-1}^2)} [m, t-1]_j^\diamond dV_m^{(d)\diamond} \int_{\mathbb{R}^d} e^{\frac{a}{2}(x_m^2+x_{t-1}^2)} [m, t-1]_{j'}^\diamond dV_m^{(d)\diamond} \\ & \leq \frac{(2\pi a)^d}{(1-a)^6} \sum_{i=0}^d \binom{d}{i} 2^{d-i} a^i = \frac{[2\pi a(2+a)]^d}{(1-a)^6}. \end{aligned} \tag{5.9}$$

We are now ready to go on with the estimate (5.8); using Lemma 6.9 for the averages involving ϕ and ψ , Lemma 6.10 for the averages without

¹¹ The fact that we need to control also the position $n-1$ is not a problem because, even if we expand only over $(m, t-1)$, the condition $j \wedge j' = 0$ actually holds for the strings over $(m-1, t)$.

polynomials, Lemmas 6.4 and 6.5 for the remaining partition functions and (5.9) for the remaining terms, one has

$$\begin{aligned} \langle \phi\psi \rangle_N^\diamond &\leq K_1^{m+m'+2d} \left(\frac{A_1}{B}\right)^m \frac{2^r r!}{A_2^{2+2r}} \|\phi\| \left(\frac{A_1}{B}\right)^{m'} \frac{2^s s!}{A_2^{2+2s}} \|\psi\| \\ &\quad \times (2K_1)^{2+8a_0} \left(\frac{D}{A_2}\right)^4 \left(\frac{A_1}{\sqrt{2\pi}B}\right)^{2d} \frac{[2\pi a(2+a)]^d}{(1-a)^6}. \end{aligned}$$

The same strategy can be applied to $\langle \phi \rangle^\diamond \langle \psi \rangle^\diamond$, cutting, for each average, the term containing the polynomial, the term containing the hole to be expanded, and the rest of the chain: it is easy to realize that the same factors present in (5.8) arise, with the same estimates, so that we will simply add a factor 2.

We are now left only with the case $j \wedge j' \neq 0$, which implies $k \wedge k' = 0$. By our assumption, the strings k are longer, so these remaining terms are even smaller as powers of a , and we may simply repeat the same procedure working on k , but expanding a substring of length d : we get the same results so we will close our proof adding another factor 2 to the estimate. \square

5.3. Upper Bound of $\langle R^2 \rangle$

In this part, we prove that $\langle R^2 \rangle$ is of order $\mathcal{O}(N/\beta^{\lambda_1})$, with $\lambda_1 \in [4, 2r + 4)$, if we impose a suitable lower bound for β , or of order $\mathcal{O}(Ne^{-\frac{3}{\sqrt{\beta}}}/\beta^3)$ if we have both a lower and an upper bound for β . One remarkable point is the proportionality to N instead of N^2 , and the other relevant aspect is the dependence on the specific energy via the parameter β . Both these points are a joint consequence of the control of the decay of correlations, as given by Proposition 5.3, and the decay of the interaction range preserved throughout the whole perturbative construction. Concerning the dependence on β , the fact that the exponential estimate does not hold for vanishing specific energies is due to condition (4.4) which gives an upper bound to the (optimal) perturbative order we can reach at fixed coupling μ .

In (5.2) we defined the remainder as

$$R = \{\Phi^{(r)}, H\} = \{\Phi_r, H_1\},$$

which represents the rate of time variation of the almost conserved integral $\Phi^{(r)}$. By recalling that in our perturbative construction we maintain the cyclical symmetry, have that $\Phi_r = \varphi_r^\oplus$ and $H_1 = h_1^\oplus$, and thus can write $R = \rho^\oplus$ where $\rho := \{\varphi_r, h_1^\oplus\}$. In the proof we need to exploit the decay properties of ρ [see again Proposition 4.1, point (vi)]; with the choice

$$\sigma_* := \sigma_0/4 < \sigma_1, \tag{5.10}$$

the decomposition (2.3) is written in this case as:

$$\rho = \sum_{l=1}^N \rho^{(l)}, \quad \|\rho^{(l)}\| \leq C_\rho \mu_b^l, \quad \mu_b := e^{-\sigma_*}. \tag{5.11}$$

Introduce the following constant¹² quantities

$$K_4 := K_3 \frac{3^3 e^6 \Omega^6}{2^9 C_1^2}, \quad \beta^* := \frac{2^6 e^{\bar{C}-1}}{(1 - \mu_b^2)(1 - \mu_b)A_2^2} \frac{C_1}{\Omega}, \quad \lambda_1 := 2r(1 - \nu) + 4 - \nu. \tag{5.12}$$

Proposition 5.4. *The following different estimates hold:*

1. for any $\nu \in (0, 1]$, $a < \min\{a_0, 1/4\}$ and $\beta > \max\{\beta_0, (\beta^* r^3)^{1/\nu}, \sqrt{3/2}\}$ such that $D^2 \mu_b < 1$, for any integer $r < \mu_*/(2\mu)$ one has

$$\langle R^2 \rangle \leq \frac{N}{\beta^{\lambda_1}} \left[\frac{K_4 \beta^{*3}}{(1 - D^2 \mu_b)^2} \right];$$

2. for any $a < \min\{a_0, 1/4\}$ and $\beta \geq \max\{\beta_0, 64e\beta^*, \sqrt{3/2}\}$ such that $D^2 \mu_b < 1$, there exists κ such that taking $r = \lfloor \kappa/3 \rfloor$, then

$$\langle R^2 \rangle \leq N \frac{e^{-\kappa}}{\kappa^9} \left[\frac{3^8 K_4}{2e^4(1 - D^2 \mu_b)^2} \right], \quad \begin{cases} \kappa := \frac{3}{2} \sqrt[3]{\frac{\beta}{e\beta^*}}, & \beta < e\beta^* \left(\frac{\mu_*}{\mu}\right)^3, \\ \kappa := \frac{3}{2} \frac{\mu_*}{\mu}, & \beta \geq e\beta^* \left(\frac{\mu_*}{\mu}\right)^3. \end{cases}$$

The proof will be carried out in several steps, actually working first on the dependence on N and then on the scaling in β ; in particular, for the first point, it is useful to rewrite $\langle R^2 \rangle$ as the sum of two term to be dealt with separately. To this purpose, we exploit the cyclic symmetry of R , i.e. using the decomposition (2.2):

$$R = \rho^\oplus \implies R = \sum_{j=0}^{N-1} \rho_j, \quad \rho_j := \rho \circ \tau^j.$$

It is thus possible to write $R^2 = \sum_{j=0}^{N-1} \rho_j^2 + 2 \sum_{0 \leq i < j \leq N-1} \rho_i \rho_j$, which implies, by translational invariance,¹³ the following expression

$$\langle R^2 \rangle = N \langle \rho^2 \rangle + 2 \sum_{0 \leq i < j \leq N-1} \langle \rho_i \rho_j \rangle. \tag{5.13}$$

While the first addendum is clearly proportional to N , the second appears to be proportional to N^2 : we will now show that this is actually not the case. As a first step we recall that $\langle \rho \rho_j \rangle = \langle \rho \rho_{N-j} \rangle$, thus we may write

$$\sum_{0 \leq i < j \leq N-1} \langle \rho_i \rho_j \rangle = \sum_{j=1}^{N-1} (N-j) \langle \rho \rho_j \rangle = N \sum_{j=1}^{\lfloor N/2 \rfloor} \langle \rho \rho_j \rangle, \tag{5.14}$$

where the last equality, in the case N odd, comes from Lemma 6.19. We remark that, although the expression for the even case (see Lemma 6.19) is slightly different, in the subsequent estimates it will be bounded from above by the odd one.

¹² More precisely β^* is asymptotically constant with $a \rightarrow 0$: we have $2^6 e^{-1} C_1 / \Omega \leq \beta^* < 2^9 e^2 C_1 / \Omega$ for $0 \leq a < 1/4$.

¹³ All the terms $\langle \rho_j^2 \rangle$ are equivalent because of the same symmetry of the measure and of ρ itself.

The strategy is first to give an estimate of the generic term $\langle \rho \rho_j \rangle$ for all $j \leq [N/2]$, and then perform the sum. Defining the following constant

$$K_3 := 2^8 K_1^4 K_2 > 4D^4 K_2, \tag{5.15}$$

the inequality holds since $a < 1/4$ and $\beta > \sqrt{3/2}$. We have

Lemma 5.5. *For any $a < \min\{a_0, 1/4\}$ and $\beta > \max\{\beta_0, \sqrt{3/2}\}$ such that $D^2\mu_b < 1$, one has*

$$\langle \rho \rho_j \rangle < \frac{3}{4} K_3 C_\rho^2 g_2(r) \left[\frac{(D^2\mu_b)^j + (D^2\mu_b)^{N-j}}{1 - D^2\mu_b} \right] \quad j \leq N/2, \tag{5.16}$$

with

$$g(r) := \frac{4^{2r+4}(r+2)!^2}{(A_2^2\beta)^{2r+4}}. \tag{5.17}$$

Proof. Using (5.11) we may write $\langle \rho \rho_j \rangle = \sum_{i=0}^{2N} \sum_{l+l'=i} \langle \rho^{(l)} \rho_j^{(l')} \rangle$. The idea is to split such a sum into two parts: in the first one, choosing l and l' small enough, we will include (part of the) terms for which $\rho^{(l)}$ and $\rho_j^{(l')}$ have disjoint supports and will exploit the decay of correlations as given by Proposition 5.3; in the second term, with l and l' bounded from below, we will instead exploit the decay (5.11) of ρ itself. We thus define

$$A := \sum_{l+l' < j} \langle \rho^{(l)} \rho_j^{(l')} \rangle, \quad B := \sum_{l+l' \geq j} \langle \rho^{(l)} \rho_j^{(l')} \rangle.$$

A term: We clearly have $l < j$; moreover, since we restrict to $j \leq N/2$ due to Lemma 6.19, we also have $j + l' \leq N - 1$. Thus, $\rho^{(l)}$ and $\rho_j^{(l')}$ have disjoint supports. Using Proposition 5.3, with $d = \min\{j - l, N - (j + l')\}$, we obtain

$$\left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| \leq K_2 \mu_\#^d g(r) D^{2d+l+l'+4} \left\| \rho^{(l)} \right\| \left\| \rho^{(l')} \right\| ;$$

we remark that in the dependence on r (recall, from (5.2), that R has degree $2r + 4$) we could expect a factor 2^{2r+4} in $g(r)$, but we overestimate with 4^{2r+4} , because the B term needs exactly such a dependence.

Since¹⁴ $\mu_\# < \mu_b$ for (at least) $a < 1/4$, using also (5.11) one has

$$\left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| \leq K_2 C_\rho^2 g(r) \mu_b^{d+l+l'} D^{2d+l+l'+4}.$$

It is useful to distinguish the cases $d = j - l$ and $d = N - (j + l')$, that we call A1 and A2, respectively. In the sub-case A1, one has $l + l' + d = j + l'$ and $l + l' + 2d = j + l' + d$; thus we get

$$\sum_{l+l' < j, A1} \left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| \leq K_2 D^4 C_\rho^2 g(r) (D\mu_b)^j \sum_{l+l' < j} D^{l'+d} \mu_b^{l'}.$$

¹⁴ Although this does not “sounds” good ...

We deal with the remaining sum in this way

$$\begin{aligned} \sum_{l+l' < j} D^{l'+d} \mu_b^{l'} &= \sum_{i=0}^{j-1} \sum_{l'=0}^i (D\mu_b)^{l'} D^{j+l'-i} = D^j \sum_{i=0}^{j-1} D^{-i} \sum_{l'=0}^i (D^2\mu_b)^{l'} \\ &< \frac{D^j}{1 - D^2\mu_b} \sum_{i=0}^{j-1} D^{-i} < \frac{2D^j}{1 - D^2\mu_b}, \end{aligned}$$

where we exploit also that $C > 2$ [see (5.3)]. Then,

$$\sum_{l+l' < j, A1} \left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| < 2K_2 D^4 C_\rho^2 g(r) \frac{(D^2\mu_b)^j}{1 - D^2\mu_b}. \tag{5.18}$$

In the other sub-case A2, just replacing l with l' and j with $N - j$, we get

$$\left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| \leq K_2 D^4 C_\rho^2 g(r) \mu_b^{N-j+l} D^{N-j+l+d}.$$

With the same approach used above one obtains

$$\sum_{l+l' \leq j-1} D^{l'+d} \mu_b^{l'} = D^{N-j} \sum_{i=0}^{j-1} \sum_{l=0}^i D^{-i} (D^2\mu_b)^l < \frac{2D^{N-j}}{1 - D^2\mu_b},$$

hence the contribution coming from A2 is

$$\sum_{l+l' < j, A2} \left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| < 2K_2 D^4 C_\rho^2 g(r) \frac{(D^2\mu_b)^{N-j}}{1 - D^2\mu_b}, \tag{5.19}$$

and combining A1 with A2 we get

$$\sum_{l+l' < j} \left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| < \frac{K_3}{2} C_\rho^2 g(r) \left[\frac{(D^2\mu_b)^j + (D^2\mu_b)^{N-j}}{1 - D^2\mu_b} \right]. \tag{5.20}$$

B term: Since we aim at exploiting the decay of ρ , by Schwartz inequality we rewrite the correlation as

$$\sum_{l+l' \geq j} \left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| \leq \sum_{l+l' \geq j} \sqrt{\langle (\rho^{(l)})^2 \rangle \langle (\rho^{(l')})^2 \rangle};$$

using Proposition 6.11 it holds

$$\langle (\rho^{(l)})^2 \rangle < \frac{(2K_1)^{1+4a_0}}{A_2^{4r+12}} D^{2+l} \left(\frac{2}{\beta} \right)^{2r+4} (2r+4)! \left\| \rho^{(l)} \right\|^2;$$

the analogous estimate holds for $\langle (\rho^{(l')})^2 \rangle$. Then, by observing that

$$\frac{(2K_1)^{1+4a_0}}{A_2^4} < K_2, \quad (2r+4)! \leq 2^{2r+4} (r+2)!^2,$$

one obtains

$$\sqrt{\langle (\rho^{(l)})^2 \rangle \langle (\rho^{(l')})^2 \rangle} < K_2 D^2 D^{(l+l')/2} \left(\frac{4}{A_2^2 \beta} \right)^{2r+4} (r+2)!^2 \left\| \rho^{(l)} \right\| \left\| \rho^{(l')} \right\|.$$

If we notice that $K_2 D^2 < K_3/4$, and using (5.17), then

$$\sqrt{\langle (\rho^{(l)})^2 \rangle \langle (\rho^{(l')})^2 \rangle} < \frac{K_3}{4} C_\rho^2 g(r) D^{(l+l')/2} \mu_b^{l+l'} ; \tag{5.21}$$

the sum over l and l' gives

$$\sum_{l+l' \geq j} \left| \langle \rho^{(l)} \rho_j^{(l')} \rangle \right| < \frac{K_3}{4} C_\rho^2 g(r) \frac{(D\mu_b)^j}{1 - D\mu_b}. \tag{5.22}$$

To get (5.16), simply add (5.22) to (5.20). □

We are now ready to state and prove an estimate showing how $\langle R^2 \rangle$ is proportional to N .

Lemma 5.6. *For $a < \min\{a_0, 1/4\}$ and $\beta > \max\{\beta_0, \sqrt{3/2}\}$ such that $D^2\mu_b < 1$, one has*

$$\langle R^2 \rangle \leq N \frac{K_3 C_\rho^2 g(r)}{(1 - D^2\mu_b)^2}. \tag{5.23}$$

Proof. As showed before [see (5.13) and the subsequent paragraphs], we have

$$\langle R^2 \rangle \leq N \langle \rho^2 \rangle + N \sum_{j=1}^{[N/2]} |\langle \rho \rho_j \rangle|. \tag{5.24}$$

Concerning the first term we may proceed like in the proof Lemma 5.5; in the case of B term, the only difference being the index of the sum; we thus follow such a proof up to formula (5.21), and then

$$\begin{aligned} \langle \rho^2 \rangle &< \sum_{i=0}^{N-1} \sum_{l+l'=i} \left| \langle \rho^{(l)} \rho^{(l')} \rangle \right| < \sum_{i=0}^{N-1} \sum_{l+l'=i} \frac{1}{4} K_3 C_\rho^2 g(r) \left(D^{\frac{1}{2}} \mu_b \right)^{l+l'} \\ &< \frac{1}{4} K_3 \frac{C_\rho^2 g(r)}{(1 - D^2\mu_b)}. \end{aligned} \tag{5.25}$$

To control the second term of (5.24), from Lemma 5.5 we have to estimate the following sum

$$\sum_{j=1}^{[N/2]} \left[(D^2\mu_b)^j + (D^2\mu_b)^{N-j} \right] = \frac{(D^2\mu_b) - (D^2\mu_b)^N}{1 - D^2\mu_b} < \frac{(D^2\mu_b)}{1 - D^2\mu_b}, \tag{5.26}$$

which leads us to

$$\sum_{j=1}^{[N/2]} |\langle \rho \rho_j \rangle| \leq \frac{3}{4} K_3 C_\rho^2 g(r) \left[\frac{(D^2\mu_b)}{(1 - D^2\mu_b)^2} \right].$$

Adding (5.25) and then multiplying by N , one has (5.23). □

The final step in the proof of Proposition 5.4 consists in showing the correct dependence on β : to this purpose, we estimate the numerator of (5.23).

Lemma 5.7. *The following two estimates hold:*

1. For any $\nu \in (0, 1]$, $a < \min\{a_0, 1/4\}$ and $\beta > \max\{\beta_0, (\beta^* r^3)^{1/\nu}, \sqrt{3/2}\}$ one has

$$K_3 C_\rho^2 g(r) \leq K_4 \frac{\beta^{*3}}{\beta^{\lambda_1}}. \tag{5.27}$$

2. For any $a < \min\{a_0, 1/4\}$ and $\beta \geq \max\{\beta_0, 64e\beta^*, \sqrt{3/2}\}$, there exists κ such that taking $r = \lfloor \kappa/3 \rfloor$, then

$$K_3 C_\rho^2 g(r) \leq \frac{3^8 K_4 e^{-\kappa}}{2e^4 \kappa^9}, \quad \begin{cases} \kappa := \frac{3}{2} \sqrt[3]{\frac{\beta}{e\beta^*}}, & \beta < e\beta^* \left(\frac{\mu_*}{\mu}\right)^3, \\ \kappa := \frac{3}{2} \frac{\mu_*}{\mu}, & \beta \geq e\beta^* \left(\frac{\mu_*}{\mu}\right)^3. \end{cases} \tag{5.28}$$

Proof. As a first step, we have to work on C_ρ : from Proposition 4.1 (vi)

$$C_\rho \leq \frac{8(r+2)F_r^{r-1}C_1^2 e^{(r+2)\tilde{C}}}{(1-\mu_b^2)(1-\mu_b)},$$

where the correction $e^{(r+2)\tilde{C}}$ comes from the transformation back to the original variables (see Lemma 3.3), and for $F_r = 16r^2 C_*$, using (4.3) and (4.4), it holds

$$\frac{8r^2 C_1}{\Omega(1-\mu_b^2)(1-\mu_b)} \leq F_r \leq \frac{16r^2 C_1}{\Omega(1-\mu_b^2)(1-\mu_b)}.$$

Using the above estimates and rearranging some terms, we have

$$C_\rho^2 g(r) \leq \frac{\Omega^6(1-\mu_b^2)^4(1-\mu_b)^4}{2^{12}C_1^2} \frac{(r+2)^2}{r^{12}} \left(\frac{4F_r e^{\tilde{C}}}{A_2^2 \beta}\right)^{2r+4} [(r+2)!]^2;$$

this can be further simplified as

$$C_\rho^2 g(r) \leq \frac{3^3 \Omega^6}{2^9 C_1^2} a_r^2 \quad a_r := \left(e\beta^* \frac{r^2}{\beta}\right)^{r+2} \frac{r!}{r^3}. \tag{5.29}$$

We now proceed in two different ways to get the two different estimates (5.27) and (5.28). For the first one, with a power law dependence on β , using the standard upper bound $r! \leq e\sqrt{r}\left(\frac{r}{e}\right)^r$, we have

$$a_r \leq e^3 \beta^{*2} \frac{r^{3/2}}{\beta^2} \left(\beta^* \frac{r^3}{\beta}\right)^r = e^3 \beta^{*3/2} \left(\beta^* \frac{r^3}{\beta^\nu}\right)^{r+1/2} \beta^{-[r(1-\nu)+2-\frac{\kappa}{2}]},$$

so that

$$K_3 C_\rho^2 g(r) \leq K_3 \frac{3^3 e^6 \Omega^6}{2^9 C_1^2} \left[\beta^{*3/2} \left(\beta^* \frac{r^3}{\beta^\nu}\right)^{r+1/2} \beta^{-\lambda_1/2}\right]^2 \leq K_4 \frac{\beta^{*3}}{\beta^{\lambda_1}},$$

where we used $\beta > (\beta^* r^3)^{1/\nu}$ to estimate from above $\left(\beta^* \frac{r^3}{\beta^\nu}\right)^{r+1/2}$.

For the exponential estimate (5.28), we work on (5.29) in a different way, trying to optimize the order r . As usual, the sequence a_r will be initially

decreasing, and then definitively increasing: we are interested in the larger $r \geq 2$ such that $a_r < a_{r-1}$, which translates in

$$\frac{e\beta^*}{\beta} r^3 \left(\frac{r}{r-1}\right)^{2r-1} < 1 \iff r < r_{\text{opt}} := \left\lfloor \sqrt[3]{\frac{\beta}{8e\beta^*}} \right\rfloor.$$

The lower bound on β in the hypothesis implies that $r_{\text{opt}} \geq 2$; moreover, the corresponding upper bound is enough to conclude that $2r_{\text{opt}}\mu \leq \mu_*$ as required in condition (4.4) of Proposition 4.1. Using again $r! \leq e\sqrt{r}\left(\frac{r}{e}\right)^r$, and using $r_{\text{opt}}^3 \leq \frac{\beta}{8e\beta^*}$, we have

$$a_r \leq \left(\frac{r^2}{8r_{\text{opt}}^3}\right)^{r+2} \frac{e\sqrt{r}r^r e^{-r}}{r^3} = \frac{er\sqrt{r}}{64r_{\text{opt}}^6} \left(\frac{r^3}{8er_{\text{opt}}^3}\right)^r,$$

so that

$$K_3 C_\rho^2 g(r_{\text{opt}}) \leq \frac{K_4}{2^{12}e^4} \frac{1}{r_{\text{opt}}^9} (8e)^{-2r_{\text{opt}}},$$

and the thesis follows using the lower bound $(2 + 2 \ln 8)r_{\text{opt}} \geq F(\beta)$.

If instead $\beta > e\beta^* (\mu_*/\mu)^3$, then $r_{\text{opt}} > r_{\text{max}} := \left\lfloor \frac{\mu_*}{2\mu} \right\rfloor$ allowed by Proposition 4.1; in this case it holds $r_{\text{max}}^3 \leq \frac{\beta}{8e\beta^*}$, so that the previous estimates on a_r holds true with r_{max} instead of r_{opt} and the thesis follows using the lower bound $(2 + 2 \ln 8)r_{\text{max}} \geq 3/2(\mu_*/\mu)$. \square

5.4. Lower Bound of $\sigma^2[\Phi]$

In this section, we give the required lower bound for $\sigma^2[\Phi]$: besides the fact that an estimate from below can be less trivial than one from above, the main point here is that Φ is non-homogeneous.

As for the estimate of the previous subsection, the one relevant aspect is the proportionality with N , and the other one is the scaling with β .

Given the constants μ_2 and β_4 of Lemma 5.9, we have the following

Proposition 5.8. *There exists a positive constant β_3 , such that if $\mu < \mu_2$, $D^2\mu_b < 1$ and $\beta > \max\{\beta_0, \beta_3, \beta_4, 6e^{\tilde{C}}\beta^*r^2(r+1), \sqrt{3/2}\}$, then*

$$\sigma^2[\Phi] \geq \frac{N}{\beta^2} \frac{\Omega^2}{10}.$$

Proof. We exploit once again the cyclic symmetry of Φ , i.e. $\Phi = \varphi^\oplus = \sum_j \varphi \circ \tau^j$. Recalling the standard notation $\sigma(f, g) = \langle fg \rangle - \langle f \rangle \langle g \rangle$ for the covariance, we thus write¹⁵

$$\begin{aligned} \sigma^2[\Phi] &= \sum_j \sigma^2[\varphi \circ \tau^j] + 2 \sum_{h < k} \sigma(\varphi \circ \tau^h, \varphi \circ \tau^k) \\ &\geq N\sigma^2[\varphi] - 2 \sum_{h < k} |\sigma(\varphi \circ \tau^h, \varphi \circ \tau^k)|, \end{aligned}$$

¹⁵ We avoid the notation $\varphi_j = \varphi \circ \tau^j$ because we will need the index to indicate the homogeneity degree.

where we used the translation invariance to extract the factor N from the first sum. The proof will be carried out in the two forthcoming Lemmas, whose application immediately gives the estimate (5.8): the first dealing with the variance $\sigma^2[\varphi]$ and the second for the estimate of the covariance terms. We exploit the different scalings in β to ensure the second term being smaller than the first one, provided β is large enough (setting β_2 as the threshold). \square

Lemma 5.9. *There exist positive constants μ_2 and β_4 , such that if $\mu < \mu_2$, $D^2\mu_b \leq 1/2$ and $\beta > \max\{\beta_0, \beta_4, 6e^{\tilde{C}}\beta^*r^2(r+1), \sqrt{3/2}\}$, then*

$$\sigma^2[\varphi] \geq \frac{\Omega^2}{9\beta^2}.$$

Proof. We exploit the expansion in homogeneous parts of $\varphi = \sum_{s=0}^r \varphi_s$: the strategy is to show that, under the assumed hypothesis, the leading term of the variance of φ is given by the variance of φ_0 . We thus first expand

$$\sigma^2[\varphi] = \sum_{s,s'=0}^r \sigma(\varphi_s, \varphi_{s'}) = \sigma^2[\varphi_0] + R_\sigma, \quad R_\sigma := \sum_{i=1}^{2r} \sum_{s+s'=i} \sigma(\varphi_s, \varphi_{s'}), \quad (5.30)$$

so that we will get the thesis via $\sigma^2[\varphi] \geq \sigma^2[\varphi_0] - |R_\sigma|$, by giving a lower bound for $\sigma^2[\varphi_0]$ and an upper bound for $|R_\sigma|$.

Step 1 (lower bound for $\sigma^2[\varphi_0]$). From Proposition 3.1 and Proposition 4.1 we have $\varphi_0 = h_\Omega = \Omega(q_0^2 + p_0^2)/2$, which has interaction length equal to one in the variables (q, p) ; once we transform it back to (x, y) variables, from a direct calculation using the decay properties of $A^{1/4}$ (see formula (3.6) of Proposition 3.1) we have that $\varphi_0 \in \mathcal{D}(C_{\varphi_0}, -\ln(2\tilde{\mu}))$ with $C_{\varphi_0} \leq \frac{\Omega}{1-2\tilde{\mu}}$; hence the decomposition $\varphi_0 = \sum_{l \geq 0} \varphi_0^{(l)}$ holds, where

$$\varphi_0^{(0)}(x, y) = \frac{\Omega}{2}(x_0^2 + y_0^2);$$

thus

$$\sigma^2[\varphi_0] = \sigma^2[\varphi_0^{(0)}] + 2 \sum_{l+l' \geq 1} \sigma(\varphi_0^{(l)}, \varphi_0^{(l')}) \geq \sigma^2[\varphi_0^{(0)}] - 2 \sum_{l+l' \geq 1} |\sigma(\varphi_0^{(l)}, \varphi_0^{(l')})|.$$

We will now show that the first term is different from zero, while the second is of order $\mathcal{O}(a)$ and asymptotically vanishing. Indeed, from the factorization of the Gibbs measure with respect to the x and y variables, one has $\sigma^2[x_0^2 + y_0^2] = \sigma^2[x_0^2] + \sigma^2[y_0^2]$; it thus follows

$$\sigma^2[\varphi_0^{(0)}] = \frac{\Omega^2}{4}(\sigma^2[x_0^2] + \sigma^2[y_0^2]) > \frac{\Omega^2}{4}\sigma^2[y_0^2] = \frac{\Omega^2}{8\beta^2}.$$

The covariance part is dealt with exploiting the exponential decay of the homogeneous polynomials $\varphi_0^{(l)}$; in particular, each covariance is estimated via Schwartz inequality as $|\sigma(\varphi, \psi)| \leq |\langle \varphi \psi \rangle| + |\langle \varphi \rangle \langle \psi \rangle| \leq \sqrt{\langle \varphi^2 \rangle \langle \psi^2 \rangle} + |\langle \varphi \rangle \langle \psi \rangle|$, and then applying Lemmas 6.13 and 6.14 we have

$$\begin{aligned}
 \left| \sigma(\varphi_0^{(l)}, \varphi_0^{(l')}) \right| &\leq \sqrt{\langle (\varphi_0^{(l)})^2 \rangle \langle (\varphi_0^{(l')})^2 \rangle} + \left| \langle \varphi_0^{(l)} \rangle \langle \varphi_0^{(l')} \rangle \right| \\
 &\leq \frac{1}{\beta^2} \left\{ \frac{2(2K_1)^{1+4a_0}}{A_2^4} D^{2+l+l'} \sqrt{\|(\varphi_0^{(l)})^2\| \|(\varphi_0^{(l')})^2\|} \right. \\
 &\quad \left. + \frac{(2K_1)^{2+8a_0}}{A_2^8} D^{4+l+l'} \|\varphi_0^{(l)}\| \|\varphi_0^{(l')}\| \right\} \left(\frac{2}{A_2^2} \right)^2 \\
 &\leq \frac{K_3}{\beta^2} C_{\varphi_0}^2 (De^{-\sigma_0})^{l+l'};
 \end{aligned}$$

the sum over l and l' gives

$$\sum_{l+l' \geq 1} (De^{-2\sigma_0})^{l+l'} \leq \sum_{i \geq 1} i(De^{-2\sigma_0})^i < \frac{De^{-2\sigma_0}}{1 - De^{-2\sigma_0}} = \frac{2\tilde{\mu}D}{1 - 2\tilde{\mu}D}.$$

If we collect the two estimates, we have

$$\sigma^2[\varphi_0] \geq \frac{\Omega^2}{\beta^2} \left(\frac{1}{8} - \frac{2K_3D\tilde{\mu}}{(1 - 2\tilde{\mu})^2(1 - 2\tilde{\mu}D)} \right). \tag{5.31}$$

Step 2 (upper bound for $|R_\sigma|$). For the remainder, we will estimate each term in the sum: we first control the covariance terms again via Schwartz inequality

$$\left| \sigma(\varphi_s, \varphi_{s'}) \right| \leq \left| \langle \varphi_s \varphi_{s'} \rangle \right| + \left| \langle \varphi_s \rangle \langle \varphi_{s'} \rangle \right| \leq \sqrt{\langle \varphi_s^2 \rangle \langle \varphi_{s'}^2 \rangle} + \left| \langle \varphi_s \rangle \langle \varphi_{s'} \rangle \right|.$$

We apply Lemma 6.13 and Corollary 6.15 to both φ_s and $\varphi_{s'}$; in particular, using estimate (6.7) and using $\sqrt{(2s+2)!(2s'+2)!} < 2^{s+s'+2}(s+s'+2)!$, we have

$$\sqrt{\langle \varphi_s^2 \rangle \langle \varphi_{s'}^2 \rangle} \leq \left[\frac{4D^2(2K_1)^{1+4a_0}}{A_2^4} \right] \frac{C_{\varphi_s} C_{\varphi_{s'}}}{1 - \mu_b} \left[\left(\frac{4}{A_2^2\beta} \right)^{s+s'+2} (s+s'+2)! \right];$$

where we used $\sigma_s > \sigma_*$, and the Definition (5.11) of μ_b ; in a similar way, using estimate (6.6)

$$\left| \langle \varphi_s \rangle \langle \varphi_{s'} \rangle \right| \leq \left[\frac{4D^4(2K_1)^{2+8a_0}}{A_2^8} \right] C_{\varphi_s} C_{\varphi_{s'}} \left[\left(\frac{2}{A_2^2\beta} \right)^{s+s'+2} (s+s'+2)! \right].$$

Summing up the previous formulas, and using the Definitions (5.3) and (5.15) of K_2 and K_3 ,

$$\sqrt{\langle \varphi_s^2 \rangle \langle \varphi_{s'}^2 \rangle} + \left| \langle \varphi_s \rangle \langle \varphi_{s'} \rangle \right| \leq K_3 \frac{C_{\varphi_s} C_{\varphi_{s'}}}{1 - \mu_b} \left[\left(\frac{4}{A_2^2\beta} \right)^{s+s'+2} (s+s'+2)! \right]. \tag{5.32}$$

As we did in the proof of Lemma 5.7, we now make explicit the dependence of C_{φ_s} from the index s : from points (iv), (v) and (vi) of Proposition 4.1

$$C_{\varphi_s} = e^{(s+1)\tilde{C}} F_r^{s-1} C_1, \quad \frac{24r^2 C_1}{\Omega(1 - \mu_b^2)(1 - \mu_b)} \leq F_r \leq \frac{48r^2 C_1}{\Omega(1 - \mu_b^2)(1 - \mu_b)},$$

where the factor $e^{(s+1)\tilde{C}}$ comes from the transformation back to the original variables (see Lemma 3.3). Using (5.32) in (5.30), we have

$$\begin{aligned} |R_\sigma| &\leq \frac{K_3 C_1^2}{1 - \mu_b} \sum_{i=1}^{2r} \sum_{s+s'=i} e^{(s+s'+2)\tilde{C}} F_r^{s+s'-2} \left[\left(\frac{4}{A_2^2 \beta} \right)^{s+s'+2} (s+s'+2)! \right] \\ &\leq \frac{K_3 C_1^2}{F_r^4 (1 - \mu_b)} \sum_{i=1}^{2r} \left[\left(\frac{4F_r e^{\tilde{C}}}{A_2^2 \beta} \right)^{i+2} (i+2)! \right] \\ &\leq \frac{K_3}{2^{12} 3^4} \left(\frac{\Omega}{C_1} \right)^4 \frac{1}{r^8} \sum_{i=1}^{2r} \left[\left(\frac{2^6 3 C_1 e^{\tilde{C}} r^2}{\Omega (1 - \mu_b^2) (1 - \mu_b) A_2^2 \beta} \right)^{i+2} (i+2)! \right]. \end{aligned} \tag{5.33}$$

The sum in the previous formula is of the form $\sum_{i=1}^{2r} a_i$ with $a_i = X^{i+2}(i+2)!$; we have that a_i is monotone decreasing, for $i = 1, \dots, 2r$, if $X(2r+2) < 1$. But for hypothesis we have $\beta > 6e^{\tilde{C}}\beta^*r^2(r+1)$, so [recalling the Definition (5.12) of β^*] it holds $a_i \leq a_1$ and thus $\sum_{i=1}^{2r} a_i \leq 2ra_1 = 12X^3r$. Inserting such an estimate in (5.33), we have

$$|R_\sigma| \leq \frac{2^8 K_3 e^{3\tilde{C}}}{(1 - \mu_b^2)^3 (1 - \mu_b)^3 A_2^6} \left(\frac{\Omega}{C_1} \right) \frac{1}{r\beta^3} \tag{5.34}$$

Step 3. With such an upper bound for $|R_\sigma|$ proportional to β^{-3} and with the lower bound (5.31) proportional to β^{-2} , there exist β_4 and μ_2 such that if $\mu < \mu_2$ and $\beta > \beta_4$ the thesis follows. \square

Lemma 5.10. *Given $\beta > \max\{\beta_0, 6e^{\tilde{C}}\beta^*r^2(r+1), \sqrt{3/2}\}$ and μ such that $D^2\mu_b < 1$, one has*

$$\left| \sum_{h < k} \sigma(\varphi \circ \tau^h, \varphi \circ \tau^k) \right| \leq N \left[\frac{(D^2\mu_b)}{(1 - D^2\mu_b)^2} \right] \frac{2^8 K_3 e^{3\tilde{C}}}{(1 - \mu_b^2)^3 (1 - \mu_b)^3 A_2^6} \left(\frac{\Omega}{C_1} \right) \frac{1}{r\beta^3}.$$

Proof. We first proceed like in formula (5.14) exploiting the translation invariance

$$\sum_{h < k} \sigma(\varphi \circ \tau^h, \varphi \circ \tau^k) = N \sum_{j=1}^{[N/2]} \sigma(\varphi, \varphi \circ \tau^j),$$

where we used Lemma 6.19 for the part $\langle (\varphi \circ \tau^h)(\varphi \circ \tau^k) \rangle$ and an easy direct calculation for the terms $\langle \varphi \circ \tau^h \rangle \langle \varphi \circ \tau^k \rangle$. We will follow the strategy of the proof of Lemma 5.5: at variance with that situation, here φ is not homogeneous so, besides the expansion (2.3) on the interaction lengths, we also need to expand over the different degrees $\varphi = \sum_{s=0}^r \varphi_s$. We will use the notation $\varphi_{s,j} := \varphi_s \circ \tau^j$ to indicate at the same time the degree and the translation.

$$\sigma(\varphi, \varphi \circ \tau^j) = \sum_{s,s'=0}^r \sigma(\varphi_s, \varphi_{s',j}) = \sum_{s,s'=0}^r \sum_{i=0}^{2N} \sum_{l+l'=i} \sigma(\varphi_s^{(l)}, \varphi_{s'}^{(l')}).$$

We split again the sum over i into two parts:

$$A := \sum_{l+l' < j} \sigma\left(\varphi_s^{(l)}, \varphi_{s',j}^{(l')}\right), \quad B := \sum_{l+l' \geq j} \sigma\left(\varphi_s^{(l)}, \varphi_{s',j}^{(l')}\right).$$

In the A term, $\varphi_s^{(l)}$ and $\varphi_{s',j}^{(l')}$ have disjoint supports, so we can apply Proposition 5.3

$$\begin{aligned} |A| &\leq \sum_{l+l' < j} K_2 D^{l+l'+2d+4} \mu_{\sharp}^d \left(\frac{2}{A_2^2 \beta}\right)^{s+s'+2} (s+1)!(s'+1)! \left\| \varphi_s^{(l)} \right\| \left\| \varphi_{s',j}^{(l')} \right\| \\ &\leq 2K_2 D^4 C_{\varphi_s} C_{\varphi_{s'}} \left(\frac{2}{A_2^2 \beta}\right)^{s+s'+2} (s+s'+2)! \sum_{l+l' < j} D^d (D\mu_b)^{l+l'+d} \\ &\leq K_3 C_{\varphi_s} C_{\varphi_{s'}} \left(\frac{2}{A_2^2 \beta}\right)^{s+s'+2} (s+s'+2)! \left[\frac{(D^2 \mu_b)^j + (D^2 \mu_b)^{N-j}}{1 - D^2 \mu_b} \right], \end{aligned}$$

where we used also the Definitions (5.3) and (5.15) of K_2 and K_3 , the fact that $\mu_{\sharp} < \mu_b$ for $a < 1/4$, and $e^{-\sigma_s} < \mu_b$, and we worked out the sum like in (5.18), (5.19) and (5.20).

In the B term, where $l+l'$ is relatively large, we exploit the exponential decay; after the usual application of Schwartz inequality, by Proposition 6.11 we have

$$\begin{aligned} \left| \sigma\left(\varphi_s^{(l)}, \varphi_{s',j}^{(l')}\right) \right| &\leq \sqrt{\left\langle \left(\varphi_s^{(l)}\right)^2 \right\rangle \left\langle \left(\varphi_{s',j}^{(l')}\right)^2 \right\rangle} + \left| \left\langle \varphi_s^{(l)} \right\rangle \left\langle \varphi_{s',j}^{(l')} \right\rangle \right| \\ &\leq \left\{ \frac{(2K_1)^{1+4a_0}}{A_2^4} D^{2+\frac{l+l'}{2}} \sqrt{(2s+2)!(2s'+2)!} \right. \\ &\quad \times \sqrt{\left\| \left(\varphi_s^{(l)}\right)^2 \right\| \left\| \left(\varphi_{s',j}^{(l')}\right)^2 \right\|} + \left\| \varphi_s^{(l)} \right\| \left\| \varphi_{s',j}^{(l')} \right\| \\ &\quad \left. + \frac{(2K_1)^{2+8a_0}}{A_2^8} D^{4+l+l'} (s+1)!(s'+1)! \right\} \left(\frac{2}{A_2^2 \beta}\right)^{s+s'+2} \\ &\leq K_3 (s+s'+2)! \left(\frac{4}{A_2^2 \beta}\right)^{s+s'+2} C_{\varphi_s} C_{\varphi_{s'}} (D\mu_b)^{l+l'} \end{aligned}$$

where we used again $\sqrt{(2s+2)!(2s'+2)!} < 2^{s+s'+2} (s+s'+2)!$, $e^{-\sigma_s} < \mu_b$ and the Definitions (5.3) and (5.15) of K_2 and K_3 . We thus have

$$|B| \leq K_3 (s+s'+2)! \left(\frac{4}{A_2^2 \beta}\right)^{s+s'+2} C_{\varphi_s} C_{\varphi_{s'}} \frac{(D\mu_b)^j}{1 - D\mu_b}.$$

To sum over the various degrees, we observe that the terms depending on s and s' are the same as in formula (5.32), the only difference being here that the sum starts from 0; at the same time, since the dependence on j is factorized, we deal with the sum over the translation like in (5.26):

$$\begin{aligned}
 \left| \sum_{h < k} \sigma(\varphi \circ \tau^h, \varphi \circ \tau^k) \right| &\leq N \sum_{j=1}^{[N/2]} |\sigma(\varphi, \varphi \circ \tau^j)| \leq N \sum_{j=1}^{[N/2]} \sum_{i=0}^{2r} \sum_{s+s'=i} |A| + |B| \\
 &\leq N \sum_{j=1}^{[N/2]} \frac{(D^2\mu_b)^j + (D^2\mu_b)^{N-j}}{1 - D^2\mu_b} 2K_3 C_1^2 \sum_{i=1}^{2r} e^{(i+2)\tilde{C}} F_r^{i-2} \left[\left(\frac{4}{A_2^2\beta} \right)^{i+2} (i+2)! \right] \\
 &\leq N \left[\frac{(D^2\mu_b)}{(1 - D^2\mu_b)^2} \right] \frac{2K_3 C_1^2}{F_r^4} \sum_{i=1}^{2r} \left[\left(\frac{4F_r e^{\tilde{C}}}{A_2^2\beta} \right)^{i+2} (i+2)! \right] \\
 &\leq N \left[\frac{(D^2\mu_b)}{(1 - D^2\mu_b)^2} \right] \frac{K_3}{2^{12}3^4} \left(\frac{\Omega}{C_1} \right)^4 \frac{1}{r^8} \sum_{i=1}^{2r} \left[\left(\frac{2^6 3 C_1 e^{\tilde{C}} r^2}{\Omega(1 - \mu_b^2)(1 - \mu_b) A_2^2\beta} \right)^{i+2} (i+2)! \right].
 \end{aligned}$$

Also in this case, we have a sum of the form $\sum_{i=0}^{2r} a_i$ just like in (5.33), with i starting from 0 instead of 1. The same conclusions hold, and thus using (5.34) we have the thesis. □

Acknowledgements

We thank Dario Bambusi, Giancarlo Benettin, Andrea Carati, Luigi Galgani and Antonio Ponno for useful discussions and comments. This research is partially supported by MIUR-PRIN program under project 2007 B3RBEY (“Teoria delle perturbazioni ed applicazioni alla Meccanica Statistica ed all’Elettrodinamica”) and project 2010 JJ4KPA (“Teorie geometriche e analitiche dei sistemi Hamiltoniani in dimensioni finite e infinite”).

6. Appendix

6.1. Proofs of Section 3

We provide here the proofs of all the statements claimed in Sect. 3 concerning the linear normalizing transformation $A^{1/4}$.

6.1.1. Proof of Proposition 3.1. If we apply the canonical linear change of coordinates $q = A^{1/4}x$ and $y = A^{1/4}p$, then H_0 reads

$$H_0(q, p) = \frac{1}{2}p \cdot A^{1/2}p + \frac{1}{2}q \cdot A^{1/2}q.$$

Since A is circulant and symmetric, the same holds also for both $A^{1/2}$ and $A^{1/4}$ (see property 4. stated after Definition 2.2); thus, we can isolate the diagonal¹⁶ part of $A^{1/2}$

$$A^{1/2} = \Omega \mathbb{I} + B,$$

¹⁶ The coefficient of the diagonal is also the first element of the first row, and thus, from property 3 of circulant matrices stated after Definition 2.2, it is the average of the eigenvalues of $A^{1/2}$.

and write

$$H_0 = H_\Omega + Z_0, \quad \text{where} \quad H_\Omega = \left[\frac{\Omega}{2} (q_1^2 + p_1^2) \right]^\oplus, \quad Z_0 = \frac{1}{2} p \cdot Bp + \frac{1}{2} q \cdot Bq. \tag{6.1}$$

Hence it follows (3.4), with $b_j(\mu)$ being the first half-row of the circulant and symmetric matrix B . A direct computation gives.¹⁷ $\{H_\Omega, Z_0\} = 0$. The exponential decay of the elements $b_j \sim \mu^j$ follows from the same argument and so we are going to use this to show the exponential decay of the elements of the matrix $A^{1/4}$.

Introducing $T = \tau + \tau^\top$, recalling (3.2), and by expanding with respect to μ , we get

$$A^{1/4} = \sqrt{\omega} \left[\mathbb{I} + \sum_{k \geq 1} \binom{1/4}{k} (-1)^k (\mu T)^k \right].$$

The exponential decay follows from the analysis of the first $[N/2] - 1$ powers of the symmetric matrix T ; indeed, for $k \leq [N/2] - 1$ one has

$$T_{j,j+i}^k = 0, \quad |i - [N/2] - 1| \leq [N/2] - k - 1,$$

which allows to claim immediately

$$(A^{1/4})_{1,j} = c_j(\mu) \mu^{j-1}, \quad k = 1, \dots, [N/2] + 1.$$

We want to prove that $|c_j(\mu)| = \mathcal{O}(2^j)$; we first observe that for any $1 \leq j \leq [N/2] + 1$ it holds

$$(A^{1/4})_{1,j} = \sqrt{\omega} \sum_{h \geq j-1} (-1)^h \mu^h T_{1,j}^h \binom{1/4}{h},$$

so that by inserting $\binom{1/4}{j} = \frac{(-1)^j (4j-5)!!}{4^j j!}$ it reads

$$\begin{aligned} (A^{1/4})_{1,1} &= \sqrt{\omega} \left[1 + \sum_{h \geq 1} \frac{\mu^h T_{1,1}^h (4h-5)!!}{h! 4^h} \right], \\ (A^{1/4})_{1,j} &= \sqrt{\omega} \sum_{h \geq j-1} \frac{\mu^h T_{1,j}^h (4h-5)!!}{h! 4^h} < 0, \quad j \geq 2. \end{aligned}$$

We stress that for $j \geq 2$ the element $(A^{1/4})_{1,j}$ is negative. From the decomposition

$$T^k = \sum_{l=0}^k \binom{k}{l} \tau^{k-l} \tau^{-l} = \sum_{l=0}^k \binom{k}{l} \tau^{k-2l},$$

¹⁷ The same conclusion follows also from the theory of the linear centralizer unfolding (see [1]) or from the normal form construction performed in Section 2 of [24].

it is possible to give an estimate uniform in i and j

$$0 < T_{ij}^k = \sum_{l=0}^k \binom{k}{l} \tau_{ij}^{k-2l} < \sum_{l=0}^k \binom{k}{l} = 2^k,$$

since the elements of τ^{k-2l} are only 1 and 0. We can thus estimate

$$\begin{aligned} |(A^{1/4})_{1,j}| &\leq -\sqrt{\omega}(2\mu)^{j-1} \sum_{k \geq 0} \frac{(2\mu)^k}{(k+j-1)!} \frac{(4k+4j-9)!!}{4^{k+j-1}} \\ &< \sqrt{\omega}(2\mu)^{j-1} \left[1 - \sum_{k \geq 1} \frac{(2\mu)^k}{k!} \frac{(4k-5)!!}{4^k} \right] < 2\sqrt{\omega}(2\mu)^{j-1}. \end{aligned}$$

The estimate holds also for $j = 1$. We can furthermore apply the same argument to $A^{1/2}$

$$A^{1/2} = \Omega \mathbb{I} + B, \quad |b_j| \leq 2\omega(2\mu)^{j-1}, \quad j \geq 2$$

thus showing that B is a μ perturbation of $\Omega \mathbb{I}$. The behaviour of the coefficients b_j also justifies (3.4). The shape of the seed ζ_0 is a direct consequence of (6.1): the vector Bq is determined by its first element, being $(Bq)_j = \tau^{j-1}(Bq)_1$, and since we can write $q_j = \tau^{j-1}q_1$ for any j , then it follows that $q_0(Bq)_1$ can be chosen as a seed for Z_0 . The same holds for the conjugated variables p . \square

6.1.2. Proof of Proposition 3.2. The flow of the Hamiltonian χ of (3.8) is given by the exponentials $x(t) = e^{Bt}x_0$ and $y(t) = e^{-Bt}y_0$, so the time $t = 1$ flow gives $A^{1/4} = e^B$ and hence the definition $B := \frac{1}{4} \ln(A)$ which, due to (3.2), reads

$$B := \frac{1}{4} \ln(\omega^2)\mathbb{I} + \frac{1}{4} \ln(\mathbb{I} - \mu T), \quad \ln(\mathbb{I} - \mu T) := - \sum_{k \geq 1} \frac{1}{k} \mu^k T^k.$$

From the series definition, we immediately get that B is circulant and symmetric. The exponential decay of its elements and the subsequent upper bound can be obtained with the same estimates used in the proof of Proposition 3.1. The final decomposition and the claim $\chi_0 \in \mathcal{D}(C_0(a), \sigma_0)$ follow from the same argument used for (6.2): the first choice for the seed is $\chi_0 = x_0(By)_1$, which gives $\chi_0^{(m)} = B_{1,m}(x_0y_m + y_0x_m)$. Then, one uses translations to get (3.8). \square

6.1.3. Proof of Lemma 3.3. We apply Corollary 2.5 with $f = \mathcal{X}_0$, thus obtaining $L_{\mathcal{X}_0}\rho \in \mathcal{D}(CC_\rho, \sigma_*)$ with

$$C \leq \frac{(r+1)2C_0}{(1 - e^{-\sigma'})(1 - e^{-(\sigma' - \sigma_*)})}.$$

Then, we recall that $T_{\mathcal{X}_0}\rho = \sum_{l \geq 0} \frac{1}{l!} L^l_{\mathcal{X}_0}\rho$; hence for any homogeneous term in the series it holds

$$\|(L^l_{\mathcal{X}_0}\rho)^{(m)}\| \leq C^l C_\rho e^{-m\sigma_*}.$$

\square

6.1.4. Proof of lemma 3.4. The claim involving the seed of Z_0 is based on the choice of ζ_0 in Proposition 3.1, namely formula (3.4). One has to observe that any term of the form $b_j(q_0q_{N-1-j} + p_0p_{N-1-j})$ in the seed ζ_0 can be replaced with the corresponding $b_j(q_0q_j + p_0p_j)$, obtained with the translation τ^j . Hence, another seed for Z_0 (in the odd case, for example) reads

$$\zeta_0 = \sum_{m=1}^{[N/2]} \zeta_0^{(m)}, \quad \zeta_0^{(m)} = 2 \sum_{m=1}^{[N/2]} b_m(q_0q_m + p_0p_m). \tag{6.2}$$

The estimate for ζ_0 then follows from

$$\left\| \zeta_0^{(m)} \right\| = 4|b_m| \leq 8\omega (2\mu)^m.$$

To prove the claim involving $H_1 = h_1^\oplus$, we adapt the proof of Lemma 3.3, working on the series expansion of the seed $T_{\mathcal{X}_0}h_1$. The proof exploits some peculiarities of the Lie derivatives $L_{\mathcal{X}_0}^k h_1$, which are derived by $\chi_0^{(0)} = 0$ and $h_1^{(m)} = 0$ for all $m \geq 1$.

We have the following:

$L_{\mathcal{X}_0}h_1$. we can write explicitly

$$L_{\mathcal{X}_0}h_1^{(m)} = \sum_{l=1}^m B_{1,l}x_0^3x_l, \quad m = 1, \dots, [N/2],$$

where $B_{1,l}$ are the coefficients of the matrix B defining \mathcal{X}_0 in Proposition 3.2. Then, for any $m \geq 1$ it holds

$$\left\| L_{\mathcal{X}_0}h_1^{(m)} \right\| = \left\| \chi_0^{(m)} \right\| \leq C_0(a)e^{-\sigma_0 m} = C_0(a)e^{-(\sigma_0 - \sigma'')m} e^{-\sigma''m},$$

thus $L_{\mathcal{X}_0}h_1 \in \mathcal{D}(C_{1,1} := C_0e^{-(\sigma_0 - \sigma'')}, \sigma'')$ for any $\sigma'' < \sigma_0$. In particular, $L_{\mathcal{X}_0}h_1 \in \mathcal{D}(C_{1,1}, \sigma_1)$.

$L_{\mathcal{X}_0}^2h_1$. the second Lie derivative represents the Poisson bracket between \mathcal{X}_0 with $\chi_0 \in \mathcal{D}(C_0, \sigma_0)$ and the previous Lie derivative seed $L_{\mathcal{X}_0}h_1$, hence $L_{\mathcal{X}_0}^2h_1 \in \mathcal{D}(C_{1,2}, \sigma_1)$. Remarkably, the expansion of both the functions starts with $m = 1$, being $\chi_0^{(0)} = L_{\mathcal{X}_0}h_1^{(0)} = 0$, thus it is possible to apply Corollary 2.6 to obtain

$$C_{1,2} = \frac{16e^{-(\sigma_0 - \sigma_1)}C_{1,1}C_0}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_1)})}.$$

$L_{\mathcal{X}_0}^3h_1$. the third derivative represents the Poisson bracket between $L_{\mathcal{X}_0}^2h_1 \in \mathcal{D}(C_{1,2}, \sigma_1)$ and \mathcal{X}_0 . Thus, surely we have $L_{\mathcal{X}_0}^3h_1 \in \mathcal{D}(C_{1,3}, \sigma_1)$. Corollary 2.6 provides

$$C_{1,3} = \frac{16e^{-(\sigma_0 - \sigma_1)}C_0}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0 - \sigma_1)})}C_{1,2}C_0 ;$$

$L_{\mathcal{X}_0}^k h_1$. the iterated Lie derivative is still the Poisson bracket between an homogeneous polynomial $L_{\mathcal{X}_0}^{k-1}h_1 \in \mathcal{D}(C_{1,k-1}, \sigma_1)$ of degree 4 and \mathcal{X}_0 . Corollary 2.6 provides

$$C_{1,k} = \frac{16e^{-(\sigma_0-\sigma_1)}C_0}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0-\sigma_1)})}C_{1,k-1}C_0,$$

which iteratively yields the estimate

$$C_{1,k} \leq \left[\frac{16e^{-(\sigma_0-\sigma_1)}C_0}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0-\sigma_1)})} \right]^k C_0. \tag{6.3}$$

By collecting all the Lie derivatives in the definition of the Lie transform, we have

$$\begin{aligned} \left\| (T_{\mathcal{X}_0} h_1)^{(m)} \right\| &\leq \sum_{k \geq 0} \frac{1}{k!} \left\| (L_{\mathcal{X}_0}^k h_1)^{(m)} \right\| \leq e^{-\sigma_1 m} \sum_{k \geq 0} \frac{1}{k!} C_{1,k} \\ &\leq C_0 e^{-\sigma_1 m} \sum_{k \geq 0} \frac{1}{k!} \left[\frac{16e^{-(\sigma_0-\sigma_1)}C_0}{(1 - e^{-\sigma_0})(1 - e^{-(\sigma_0-\sigma_1)})} \right]^k \\ &= C_1 e^{-\sigma_1 m}, \end{aligned}$$

where using $\sigma_0 = 2\sigma_1$, we have defined

$$C_1 := C_0 \exp \left(\frac{16e^{-\sigma_0/2}C_0}{(1 - e^{-\sigma_0})(1 - e^{-\sigma_0/2})} \right).$$

The claim $C_1(a) = \mathcal{O}(1)$ comes from $(a \rightarrow 0) \Rightarrow (\sigma_0 \rightarrow +\infty)$. □

6.2. Technical Lemmas on Gibbs Averages

6.2.1. Useful Integrals. We recall a couple of useful formulas¹⁸ for integrals. The first is

$$I_k := \int_{\mathbb{R}} x^k e^{-x^2} dx \implies I_{2k} = \frac{2k-1}{2} I_{2(k-1)} = \frac{(2k-1)!!}{2^k} \sqrt{\pi}. \tag{6.4}$$

The second formula gives an estimate for

$$G_k(\gamma) := \int_{\mathbb{R}} x^k e^{-(x^2+\gamma^2 x^4)} dx.$$

We have the following

Lemma 6.1. *Let $\gamma > 0$ satisfy $\gamma^2(2k+1)(2k+3) < 4$. Then one has*

$$1 - \gamma^2 \frac{(2k+1)(2k+3)}{4} \leq \frac{G_{2k}(\gamma)}{G_{2k}(0)} \leq 1.$$

Proof. There exists $\xi \in [0, \gamma]$, such that

$$G_k(\gamma) - G_k(0) = \gamma \frac{\partial G_k}{\partial \gamma}(\xi) = -2\gamma \xi G_{k+4}(\xi).$$

Since G_k is a monotonic function of γ , we can estimate $\xi G_{k+4}(\xi)$ both from above and from below as $0 \leq \xi G_{k+4}(\xi) \leq \gamma G_{k+4}(0)$ for all $\xi \in [0, \gamma]$. In view of $G_k(0) = I_k$ and using twice the recursive relation (6.4), the claim follows. □

¹⁸ We also use the notation $k!! := k[(k-2)!!]$ if $k > 1$, and $k!! := 1$ if $k = 1$ or $k = 0$.

6.2.2. Partition Function. Throughout the present section, we consider a periodic chain of length l and drop the y depending part of H because the corresponding integrals are trivially factorized. We also split H_0 as

$$H_0 = H_d + H_a, \quad H_d = \frac{1}{2} \sum_j x_j^2, \quad H_a = \frac{a}{2} \sum_j (x_{j+1} - x_j)^2 \tag{6.5}$$

where a plays the role of the small coupling parameter. We also recall that $H_1 = \frac{1}{4} \sum_j x_j^4$. Finally, we emphasize the role of the parameter a by denoting the partition function as $Z(\beta, a)$.

We present three Lemmas investigating the role of boundary conditions and that of the parameters N , a and β on the partition function; we start recalling the following

Lemma 6.2 (Lemma 5 and 6 of [16]). *There exist constants $\beta_0 > 0$, $a_0 > 0$, $K_1 > 2$ such that for any $\beta > \beta_0$ and $0 < a < a_0$, one has*

$$Z_l(\beta) \geq Z_{l-1}(\beta) \frac{1}{K_1} \sqrt{\frac{2\pi}{\beta}}.$$

Denoting by \mathcal{Z} the partition function for the system with free boundary conditions,¹⁹ one has

$$Z_l(\beta) \leq Z_l(\beta) (2K_1)^{1+4a_0}.$$

Proof. For the proof we refer to [16]; please note that our variables x_j correspond to their variables q_j after a rescaling: $q_j = \sqrt{\omega}x_j$. □

Corollary 6.3. *Under the same hypothesis of Lemma 6.2, for the scaled²⁰ quantities one has*

$$Z_l^\diamond(\beta) \geq Z_{l-1}^\diamond(\beta) \frac{\sqrt{2\pi}}{K_1} \quad Z_l^\diamond(\beta) \leq Z_l^\diamond(\beta) (2K_1)^{1+4a_0}.$$

In the following Lemma, we control the complete partition function $Z(\beta, a)$ with the totally uncoupled one $Z(\beta, 0)$, exploiting the decomposition (6.5). We remark that it is independent of the boundary conditions.

Lemma 6.4. *Let $A_1 := \sqrt{1 + 4a}$. If $a > 0$ then*

$$\frac{1}{A_1^l} Z_l(\beta, 0) \leq Z_l(\beta, a) \leq Z_l(\beta, 0).$$

Proof. Using the trivial inequality $(v - w)^2 \leq 2(v^2 + w^2)$, when $a > 0$ we use the estimate $H_a \leq 4aH_d$ to get the following inequalities

$$H_d + H_1 \leq H \leq (1 + 4a)H_d + (1 + 4a)^2 H_1.$$

¹⁹ See (5.4).

²⁰ See the paragraph before definitions (5.5) in the proof of Proposition 5.3.

Inserting the previous formula in the explicit expressions for $Z(\beta, a)$ and $Z(\beta, 0)$, we have

$$\begin{aligned} \int_{\mathbb{R}^l} e^{-\beta(H_d+H_1)} dx &\geq \int_{\mathbb{R}^l} e^{-\beta H} dx \\ &\geq \left[\int_{\mathbb{R}} e^{-\beta\left((1+4a)\frac{x^2}{2}+(1+4a)^2\frac{x^4}{4}\right)} dx \right]^l = \frac{1}{A_1^l} \int_{\mathbb{R}^l} e^{-\beta(H_d+H_1)} dx, \end{aligned}$$

using, in the last equality, the change of variable $y = \sqrt{1 + 4a}x$. □

As a corollary of Lemma 6.1, in the uncoupled case it is possible to control the partition function $Z(\beta, 0)$ with the corresponding harmonic²¹ one, which can be denoted by $Z_l(\infty, 0) := \int_{\mathbb{R}^l} e^{-\beta H_d} dx = \left(\frac{2\pi}{\beta}\right)^{l/2}$:

Lemma 6.5. *For $\beta > \sqrt{3/4}$, denoting $B := 1 - \frac{3}{4\beta^2} < 1$, one has*

$$B^l \leq \frac{Z_l(\beta, 0)}{Z_l(\infty, 0)} \leq 1.$$

Proof. Rescale Z by $\sqrt{\beta}$ and then apply Lemma 6.1. □

As a corollary of the previous Lemmas, one has

Corollary 6.6. *Let β_0, a_0 and K_1 be the constants of Lemma 6.2, then for any $\beta > \beta_0, 0 < a < a_0$ and $l' < l$ one has*

$$\frac{Z_{l-l'}(\beta, a)Z_{l'}(\beta, a)}{Z_l(\beta, a)} \leq K_1^{l'}.$$

Proof. Apply Lemma 6.2 l' times, then apply Lemma 6.4 and Lemma 6.5. □

Remark 6.7. The previous statements, Lemma 6.4, Lemma 6.5 and Corollary 6.6, hold also for the scaled quantities Z^\diamond , with identical proofs.

6.2.3. Gibbs Averages of Monomial. We first describe the possibility of cutting a (periodic) chain into two (periodic) subchains, showing how the respective averages are related. We consider in general two functions f and g with disjoint supports; we have the following

Lemma 6.8. *In a periodic chain of length l , let f and g be two functions with disjoint supports, in particular $\text{supp}(f) \subseteq \{0, \dots, l' - 1\}$ and $\text{supp}(g) \subseteq \{l', \dots, l - 1\}$. Let β_0, a_0 and K_1 be the constants of Lemma 6.2, then for any $\beta > \beta_0$ and $0 < a < a_0$ one has*

$$\langle fg \rangle_l \leq K_1^{l'} \left\langle f e^{\beta a(x_{l'-1}^2 + x_0^2)} \right\rangle_{l'} \left\langle g e^{\beta a(x_{l-1}^2 + x_{l'}^2)} \right\rangle_{l-l'},$$

and an identical inequality, but without β in the exponential factors, for the scaled case.

²¹ The corresponding scaled version is clearly $Z_l^\diamond(\infty, 0) := (2\pi)^{l/2}$.

Proof. We start writing the averaged quantity splitting the measure in the following way:

$$\langle fg \rangle_l = \frac{1}{Z_l} \int_{\mathbb{R}^l} [0, f, l' - 1][l' - 1, l'][l', g, l - 1][l - 1, l] dV_0^{(l')} dV_{l'}^{(l-l')}.$$

If we could remove the terms $[l' - 1, l']$ and $[l - 1, l]$, we would end up in two separated chain with free boundaries; to recover periodicity in both chains we should multiply by $e^{\beta a x_{l'-1} x_0}$ and $e^{\beta a x_{l-1} x_{l'}}$; now observe that, using $|bc| < \frac{1}{2}(b^2 + c^2)$, one has

$$\begin{aligned} [l' - 1, l'] \frac{e^{\beta a x_{l'-1} x_0}}{e^{\beta a x_{l'-1} x_0}} [l - 1, l] \frac{e^{\beta a x_{l-1} x_{l'}}}{e^{\beta a x_{l-1} x_{l'}}} \\ \leq e^{\beta a x_{l'-1} x_0} e^{\beta a (x_{l'-1}^2 + x_0^2)} e^{\beta a x_{l-1} x_{l'}} e^{\beta a (x_{l-1}^2 + x_{l'}^2)}, \end{aligned}$$

so we have

$$\langle fg \rangle_l \leq \frac{Z_{l'} Z_{l-l'}}{Z_l} \left\langle f e^{\beta a (x_{l'-1}^2 + x_0^2)} \right\rangle_{l'} \left\langle g e^{\beta a (x_{l-1}^2 + x_{l'}^2)} \right\rangle_{l-l'};$$

the thesis now follows from Corollary 6.6. For scaled case, simply add \diamond everywhere and remove β as needed in the proof above. \square

In the next Lemma we are interested in the phase average of a monomial: in this case the dimension of the space is not a critical aspect of the estimate, so we can simply perform our estimate passing to the uncoupled quantities paying the price of the power of a constant to the dimension of the space.

Lemma 6.9. *Let x^k be a monomial of degree $2r$, $A_2 := \sqrt{1 - 2a}$, $a < 1/2$ and $\beta > \sqrt{3/4}$, then*

$$\begin{aligned} \langle x^k \rangle_l &\leq A_1^l \langle x^k \rangle_{l,0}, \\ \left\langle x^k e^{\beta a (x_0^2 + x_{l-1}^2)} \right\rangle_l &\leq \frac{A_1^l}{A_2^{2+k_0+k_{l-1}}} \langle x^k \rangle_{l,0} \\ \langle x^k \rangle_{l,0} &\leq \frac{1}{B^l} \frac{2^r r!}{\beta^r}. \end{aligned}$$

Proof. For the first inequality, one simply use $H_a > 0$ and Lemma 6.4 to get

$$\langle x^k \rangle_l = \frac{1}{Z_l(\beta, a)} \int_{\mathbb{R}^l} x^k e^{-\beta(H_d + H_a + H_1)} \leq \frac{A_1^l}{Z_l(\beta, 0)} \int_{\mathbb{R}^l} x^k e^{-\beta(H_d + H_1)},$$

which is actually $A_1^l \langle x^k \rangle_{l,0}$. For the second inequality, we further use $e^{z^4/4} \leq e^{(z^4/A_2)^4/4}$ in the previous estimate, and then scale by A_2 to the “boundary” variables appearing in the exponential factor.

For the last inequality, using $H_1 > 0$ and Lemma 6.5, one can first exploit the control of the (decoupled) nonlinear Gibbs measure by the (decoupled) harmonic one:

$$\langle x^k \rangle_{l,0} \leq \frac{1}{B^l Z_l(\infty, 0)} \int_{\mathbb{R}^l} x^k e^{-\beta H_d}.$$

Then, concerning the numerator we have

$$\begin{aligned} \int_{\mathbb{R}^l} x^k e^{-\beta H_a} &= \prod_{j=1}^l \int_{\mathbb{R}} x^{k_j} e^{-\beta \frac{x_j^2}{2}} dx = \prod_{j=1}^l \left(\frac{2}{\beta}\right)^{\frac{k_j+1}{2}} \Gamma\left(\frac{k_j+1}{2}\right) \\ &\leq \left(\frac{2}{\beta}\right)^r \left(\sqrt{\frac{2\pi}{\beta}}\right)^l \prod_{j=1}^l \left(\frac{k_j}{2}\right)!, \end{aligned}$$

where we used, when k_j is even, the relation $\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!}{4^n n!} \sqrt{\pi} \leq n! \sqrt{\pi}$; if k_j is odd, and we actually estimate $\int |x^{k_j}| e^{-\beta H_a}$, then $\Gamma\left(\frac{k_j+1}{2}\right) = \left(\frac{k_j-1}{2}\right)!$ which bounded by the estimate of the other case. As a last step we recall that $\prod_j (m_j!) \leq \left(\sum_j m_j\right)!$ to get the thesis. \square

In the next Lemma, despite its similarity with the previous one, we need a different approach, since it will be applied in a case with the dimension of the space growing with N ; thus, it is not possible to pay the price of the constant A_1^l .

Lemma 6.10. *Let β_0 , a_0 and K_1 be the constants of Lemma 6.2; for every $0 \leq i < l$ it holds*

$$\langle e^{\beta a(x_i^2 + x_{i+1}^2)} \rangle_l \leq (2K_1)^{(1+4a_0)} \left(\frac{D}{A_2}\right)^2, \quad D := \frac{K_1 A_1}{B}.$$

Proof. Let us write first the Hamiltonian isolating the energy of two chains with free boundaries,²² the first defined on the two sites i and $i + 1$ and the second on the remaining sites, plus the connecting springs:

$$\begin{aligned} H(x) &= \mathcal{H}^{(2)}(x_i, x_{i+1}) + \mathcal{H}^{(l-2)}(x_{i+2}, x_{i+3}, \dots, x_{i-1}) \\ &\quad + \frac{a}{2}(x_{i-1} - x_i)^2 + \frac{a}{2}(x_{i+1} - x_{i+2})^2; \end{aligned}$$

then, estimating by above with 1 the contribution of the connecting springs

$$\begin{aligned} \langle e^{\beta a(x_i^2 + x_{i+1}^2)} \rangle_l &\leq \frac{1}{Z_l} \int_{\mathbb{R}^{l-2}} e^{-\beta \mathcal{H}^{(l-2)}} \int_{\mathbb{R}^2} e^{\beta a(x_i^2 + x_{i+1}^2)} e^{-\beta \mathcal{H}^{(2)}} dx_1 \dots dx_l \\ &= \frac{Z_{l-2} Z_2}{Z_l} \frac{Z_{l-2}}{Z_{l-2}} \frac{\int_{\mathbb{R}^2} e^{\beta a(x_i^2 + x_{i+1}^2)} e^{-\beta \mathcal{H}^{(2)}} dx_i dx_{i+1}}{Z_2} \\ &\leq K_1^2 (2K_1)^{(1+4a_0)} \frac{\int_{\mathbb{R}^2} e^{-\beta(1-2a)(x_i^2/2 + x_{i+1}^2/2)} dx_i dx_{i+1}}{A_1^{-2} Z_2(\beta, 0)}, \end{aligned}$$

where we used Corollary 6.6, Lemma 6.2 and Lemma 6.4; then using also Lemma 6.5 and a rescaling like in Lemma 6.9 one has the thesis. \square

In the next proposition, we consider the average of monomial whose support is not the entire space. It is actually a trivial consequence of the previous results.

²² We recall the use of the calligraphic letters \mathcal{H} and \mathcal{Z} for the free boundary cases [see (5.4)].

Proposition 6.11. *Let β_0, a_0 and K_1 be the constants of Lemma 6.2, and let x^k be a left aligned even monomial of degree $2r$ and interaction range $l' < l$. Then, for any $\beta > \beta_0$ and $a < a_0$ one has*

$$\langle x^k \rangle_l \leq \frac{(2K_1)^{(1+4a_0)}}{A_2^{4+k_0+k_{l'-1}}} D^{2+l'} \frac{2^r r!}{\beta^r}.$$

Proof. Use Lemma 6.8 to cut the chain at the boundaries of the support of x^k ; then apply Lemma 6.9 and 6.10 to the resulting terms. \square

Remark 6.12. In all the previous results, i.e. Lemmas 6.8, 6.9 and 6.10 and Proposition 6.11, the same estimates hold true if one substitutes every average $\langle \cdot \rangle$ with the corresponding scaled version $\langle \cdot \rangle^\diamond$ and remove β . The proofs are almost, if not completely, identical.

6.2.4. Gibbs Averages of Polynomials of Class $\mathcal{D}(C, \sigma)$. In this subsection we consider polynomials instead of monomials, but we restrict to those with an exponential decay of the interaction range, as described in Sect. 2.2 (we recall in particular the decomposition (2.3) and the Definition (2.7) of class $\mathcal{D}(C, \sigma)$). A first result is the following

Lemma 6.13. *If $\varphi \in \mathcal{D}(C_\varphi, \sigma_s)$ is a polynomial of degree $2s + 2$ with $s \leq r$, if $De^{-\sigma_s} \leq 1/2$ then*

$$\langle \varphi \rangle \leq \left[\frac{2D^2(2K_1)^{1+4a_0}}{A_2^4} \right] C_\varphi \left[\left(\frac{2}{A_2^2 \beta} \right)^{s+1} (s+1)! \right] \tag{6.6}$$

Proof. Using the decomposition (2.3), let us write $\langle \varphi \rangle = \sum_l \langle (\varphi)^{(l)} \rangle$. By applying Proposition 6.11 we can estimate each addendum as

$$\langle \varphi^{(l)} \rangle \leq \frac{(2K_1)^{1+4a_0}}{A_2^4} D^{2+l} \left(\frac{2}{A_2^2 \beta} \right)^{s+1} (s+1)! \|\varphi^{(l)}\|;$$

performing the sum over l , we have

$$\sum_l \langle \varphi^{(l)} \rangle \leq \frac{(2K_1)^{1+4a_0}}{A_2^4} \left(\frac{2}{A_2^2 \beta} \right)^{s+1} (s+1)! D^2 C_\varphi \sum_{l=0}^{N-1} (De^{-\sigma_s})^l,$$

and from the estimate $\sum_{l=0}^{N-1} (De^{-\sigma_s})^l \leq 2$ (the latter being true due to condition $D\mu_b \leq 1/2$) the thesis follows. \square

We are interested in a similar estimate for the square of a given polynomial; to this purpose we first give the following

Lemma 6.14. *If $\varphi \in \mathcal{D}(C_\varphi, \sigma_s)$ then $\varphi^2 \in \mathcal{D}(2C_\varphi^2/(1 - e^{-\sigma_s}), \sigma_s)$.*

Proof. We write $\varphi^2 = \sum_{l=0}^{N-1} (\varphi^2)^{(l)}$, with

$$(\varphi^2)^{(l)} = \left(\varphi^{(l)} \right)^2 + 2\varphi^{(l)} \sum_{l'=0}^{l-1} \varphi^{(l')},$$

hence

$$\|(\varphi^2)^{(l)}\| \leq 2 \|\varphi^{(l)}\| \sum_{l'=0}^l \|\varphi^{(l')}\| \leq 2C_\varphi^2 e^{-\sigma_s l} \sum_{l'=0}^l e^{-\sigma_s l'} < \frac{2C_\varphi^2}{1 - e^{-\sigma_s}} e^{-\sigma_s l}.$$

□

A trivial consequence of the two previous lemmas is then the following

Corollary 6.15. *If $\varphi \in \mathcal{D}(C_\varphi, \sigma_s)$ is a polynomial of degree $2s + 2$, if $De^{-\sigma_s} \leq 1/2$ then*

$$\langle \varphi^2 \rangle \leq \left[\frac{4D^2(2K_1)^{1+4a_0}}{A_2^4} \right] \frac{C_\varphi^2}{1 - e^{-\sigma_s}} \left[\left(\frac{2}{A_2^2 \beta} \right)^{2s+2} (2s + 2)! \right] \tag{6.7}$$

6.3. Cancellations

In this appendix, we provide the formal proof of the cancellations that take place in the correlation terms of Proposition 5.3.

We first recall a few basic facts about binary strings used in the proof. Define $\Sigma_m := \{\mathbf{k} = \mathbf{k}_1 \cdots \mathbf{k}_m \text{ s.t. } \mathbf{k}_l \in \{0, 1\} \forall l\}$ the space of binary strings of finite length. We shall denote by $\ell(\mathbf{k})$ the length of the string \mathbf{k} . Recall that the usual bitwise “and” operator $\wedge : \Sigma_m \times \Sigma_m \rightarrow \Sigma_m$ is defined as

$$(\mathbf{i}, \mathbf{j}) \mapsto \mathbf{k} = \mathbf{i} \wedge \mathbf{j}, \quad \mathbf{k}_l = \begin{cases} 1 & \mathbf{i}_l = 1 \text{ and } \mathbf{j}_l = 1 \\ 0 & \text{otherwise} \end{cases} \tag{6.8}$$

Remark 6.16. An elementary property is the following: given two strings \mathbf{j} and \mathbf{k} in Σ_m , if $\mathbf{j} \wedge \mathbf{k} = 0$, then the total number of zeros contained collectively in \mathbf{j} and \mathbf{k} is at least m .

A relevant point is that the presence of a 1 allows us to decouple the measure as follows. Expand $[p, r] = \sum_{\mathbf{j}} [p, r]_{\mathbf{j}}$ and select a string $\mathbf{j} = \mathbf{v}1\mathbf{w}$ with a 1 in the position corresponding to the site q . Then, we have $[p, r]_{\mathbf{j}} = [p, q]_{\mathbf{v}}[q + 1, r]_{\mathbf{w}}$ in place of the general decomposition $[p, r] = [p, q][q, r]$. We shall use this property in dealing with expressions (5.7) in cases of having $\mathbf{j} \wedge \mathbf{j}' \neq 0$ and $\mathbf{k} \wedge \mathbf{k}' \neq 0$. For, if $\mathbf{j} \wedge \mathbf{j}' \neq 0$, then both \mathbf{j} and \mathbf{j}' have a 1 in the same position, so that a decomposition as above may take place in the corresponding site of the chain.

Consider the set

$$\Sigma := \{ \{\mathbf{j}, \mathbf{k}, \mathbf{j}', \mathbf{k}'\} : \mathbf{j} \wedge \mathbf{j}' \neq 0, \mathbf{k} \wedge \mathbf{k}' \neq 0, \ell(\mathbf{j}) = \ell(\mathbf{j}') = l, \ell(\mathbf{k}) = \ell(\mathbf{k}') = l' \}.$$

We introduce now an equivalence relation as follows. Let $\mathbf{j} \wedge \mathbf{j}' \neq 0$ and look for the first digit 1 in $\mathbf{j} \wedge \mathbf{j}'$. Then, split $\mathbf{j} = \mathbf{v}1\mathbf{w}$ and $\mathbf{j}' = \mathbf{v}'1\mathbf{w}'$ with $\ell(\mathbf{v}) = \ell(\mathbf{v}')$ and with $\mathbf{v} \wedge \mathbf{v}' = 0$. Similarly, split $\mathbf{k} = \mathbf{x}1\mathbf{y}$ and $\mathbf{k}' = \mathbf{x}'1\mathbf{y}'$ with $\ell(\mathbf{x}) = \ell(\mathbf{x}')$ and $\mathbf{x} \wedge \mathbf{x}' = 0$. We say that $\{\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}'\} \in \Sigma$ is equivalent to $\{\mathbf{j}, \mathbf{k}, \mathbf{j}', \mathbf{k}'\}$ in case it can be obtained by an arbitrary exchange of the pairs $(\mathbf{v}, \mathbf{v}')$, $(\mathbf{w}, \mathbf{w}')$,

$(\mathbf{x}, \mathbf{x}')$ and $(\mathbf{y}, \mathbf{y}')$. More precisely, $(\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}')$ must be equal to one of

$$\begin{aligned} & \{v1w, x1y, v'1w', x'1y'\}, \quad \{v'1w, x1y, v1w', x'1y'\}, \\ & \{v1w, x1y', v'1w', x'1y'\}, \quad \{v'1w, x1y', v1w', x'1y'\}, \\ & \{v1w, x'1y, v'1w', x1y'\}, \quad \{v'1w, x'1y, v1w', x1y'\}, \\ & \{v1w, x'1y', v'1w', x1y'\}, \quad \{v'1w, x'1y', v1w', x1y'\}, \\ & \{v1w', x1y, v'1w, x'1y'\}, \quad \{v'1w', x1y, v1w, x'1y'\}, \\ & \{v1w', x1y', v'1w, x'1y'\}, \quad \{v'1w', x1y', v1w, x'1y'\}, \\ & \{v1w', x'1y, v'1w, x1y'\}, \quad \{v'1w', x'1y, v1w, x1y'\}, \\ & \{v1w', x'1y', v'1w, x1y'\}, \quad \{v'1w', x'1y', v1w, x1y'\}, \end{aligned}$$

where some of the above combinations may well coincide, e.g. if $v = v'$. The above list of 16 combinations actually describes the equivalence classes generated by our relation.

Remark 6.17. It is clearly sufficient to show that cancellations occurs within each equivalence class, and this is what we are going to do.

Let us now pick an equivalence class, i.e. choose one of its elements denoted again by $\{j, k, j', k'\} = \{v1w, x1y, v'1w', x'1y'\}$, and exploit the decomposition corresponding to the 1 digits. For the term $\langle\langle\phi\psi\rangle\rangle Z$ in (5.7), writing only the integrand function, we get

$$\begin{aligned} & ([0, \phi, m-1] \cdot [m-1, t]_j \cdot [t, \psi, t+m'-1] \cdot [t+m'-1, N]_k) \\ & \quad \times ([0, m-1] \cdot [m-1, t]_{j'} \cdot [t, t+m'-1] \cdot [t+m'-1, N]_{k'}) \\ & = ([u'-1, N]_y \cdot [0, \phi, m-1] \cdot [m-1, u-1]_v) \\ & \quad \times ([u, t]_w \cdot [t, \psi, t+m'-1] \cdot [t+m'-1, u'-1]_x) \\ & \quad \times ([u'-1, N]_{y'} \cdot [0, m-1] \cdot [m-1, u-1]_{v'}) \\ & \quad \times ([u, t]_{w'} \cdot [t, t+m'-1] \cdot [t+m'-1, u'-1]_{x'}), \end{aligned}$$

where u and u' denote the sites corresponding to the 1, i.e. $[m-1, t]_j = [m-1, u-1]_v [u, t]_w$ and $[t+m'-1, N]_k = [t+m'-1, u'-1]_x [u'-1, N]_y$, and the same for the primed strings. We do not write explicitly the analogous equality for the term in $\langle\langle\phi\rangle\rangle \langle\langle\psi\rangle\rangle$.

The argument that follows becomes more transparent by introducing a compact notation for each of the lines in the above formula, namely

$$\begin{aligned} [y : \phi : v] & := [u'-1, N]_y \cdot [0, \phi, m-1] \cdot [m-1, u-1]_v, \\ [w : \psi : x] & := [u, t]_w \cdot [t, \psi, t+m'-1] \cdot [t+m'-1, u'-1]_x, \\ [y' :: v'] & := [u'-1, N]_{y'} \cdot [0, m-1] \cdot [m-1, t]_{v'} \cdot [m-1, u-1]_{w'}, \\ [w' :: x'] & := [u, t]_{w'} \cdot [t, t+m'-1] \cdot [t+m'-1, u'-1]_{x'}. \end{aligned}$$

Using such a notation and writing the corresponding terms for the whole expression $\langle\langle\phi\psi\rangle\rangle Z - \langle\langle\phi\rangle\rangle \langle\langle\psi\rangle\rangle$, one has

$$[y : \phi : v] \cdot [w : \psi : x] \cdot [y' :: v'] \cdot [w' :: x'] - [y : \phi : v] \cdot [w :: x] \cdot [y' :: v'] \cdot [w' : \psi : x'];$$

we remark that each factor in the above expression has a support disjoint from the supports of the other factors, so that the integral can be factorized.

Select now $\{v1w, x1y, v'1w', x'1y'\}$ and $\{v1w', x'1y, v'1w, x1y'\}$, i.e. two elements which belong to the same equivalence class being obtained via a permutation of both pairs (w, w') and (x, x') and add together the corresponding contributions, namely

$$\begin{aligned}
 & [y : \phi : v] \cdot [w : \psi : x] \cdot [y' : v'] \cdot [w' : x'] - [y : \phi : v] \cdot [w : x] \cdot [y' : v'] \cdot [w' : \psi : x'] \\
 & + [y : \phi : v] \cdot [w' : \psi : x'] \cdot [y' : v'] \cdot [w : x] - [y : \phi : v] \cdot [w' : x'] \cdot [y' : v'] \cdot [w : \psi : x];
 \end{aligned}$$

they clearly compensate each other and add up to zero. Performing the permutation only in the pair (w, w') and then only in the pair (x, x') , we have instead

$$\begin{aligned}
 & [y : \phi : v] \cdot [w' : \psi : x] \cdot [y' : v'] \cdot [w : x'] - [y : \phi : v] \cdot [w' : x] \cdot [y' : v'] \cdot [w : \psi : x'] \\
 & + [y : \phi : v] \cdot [w : \psi : x'] \cdot [y' : v'] \cdot [w' : x] - [y : \phi : v] \cdot [w : x'] \cdot [y' : v'] \cdot [w' : \psi : x],
 \end{aligned}$$

which again add up to zero. The previous expressions show the cancellations among 4 of 16 terms in the equivalence class; it is not difficult to verify that the other cancellations take place working also, jointly or separately, with the permutations of the other pairs (v, v') and (y, y') . This completes the proof that all contributions to $\langle\langle\phi\psi\rangle\rangle Z - \langle\langle\phi\rangle\rangle\langle\langle\psi\rangle\rangle$ coming from elements of every equivalence class in Σ cancel out in pairs.

6.4. Structure of the Remainder

We present here a couple of lemmas concerning the structure of the remainder of our perturbative construction. We recall that $\Phi = \varphi^\oplus$ and $R = \rho^\oplus$, so that in particular $R = \sum_{j=0}^{N-1} \rho_j$ with $\rho_j := \rho \circ \tau^j$.

Lemma 6.18. *In the original variables x, y , the seed $\varphi(x, y)$ is even in the momenta y and consequently the seed of the remainder $\rho(x, y)$ is odd both in the momenta y and in the coordinates x .*

Proof. We recall that the original Hamiltonian $H(x, y) = T(y) + V(x)$ is obviously of even degree in the momenta and that also the iterative scheme gives polynomials $\Phi_r(x, y)$ which are of even degree in the momenta. Hence, the Poisson bracket defining ρ produces a polynomial of odd degree both in the momenta and being the whole degree even, in the configuration variables. \square

Lemma 6.19. *It holds*

$$\sum_{j=1}^{N-1} (N - j) \langle \rho \rho_j \rangle = \begin{cases} \frac{N}{2} \langle \rho \rho_{N/2} \rangle + N \sum_{j=1}^{N/2-1} \langle \rho \rho_j \rangle, & N \text{ even} \\ N \sum_{j=1}^{\lfloor N/2 \rfloor} \langle \rho \rho_j \rangle, & N \text{ odd} \end{cases}$$

Proof. We provide the proof in the odd case, the other one being very similar. We have the following equalities:

$$\sum_{j=1}^{N-1} \langle \rho \rho_j \rangle = \sum_{j=1}^{[N/2]} \langle \rho \rho_j \rangle + \sum_{j=[N/2]+1}^{N-1} \langle \rho \rho_{N-j} \rangle = 2 \sum_{j=1}^{[N/2]} \langle \rho \rho_j \rangle ;$$

$$\sum_{j=1}^{N-1} j \langle \rho \rho_j \rangle = \sum_{j=1}^{[N/2]} [j + (N - j)] \langle \rho \rho_j \rangle = N \sum_{j=1}^{[N/2]} \langle \rho \rho_j \rangle.$$

Thus, subtracting the second line to N times the first one gives the thesis. \square

References

- [1] Arnold, V.: *Chapitres supplémentaires de la théorie des équations différentielles ordinaires*. “Mir”, Moscow (1984) (Translated from the Russian by Djilali Embarek, Reprint of the 1980 edition)
- [2] Arnold, V.I.: Proof of a theorem of A N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspehi Mat. Nauk* **18**(5 (113)), 13–40 (1963)
- [3] Bambusi, D., Giorgilli, A.: Exponential stability of states close to resonance in infinite-dimensional Hamiltonian systems. *J. Stat. Phys.* **71**(3–4), 569–606 (1993)
- [4] Bambusi, D., Nekhoroshev, N.N.: A property of exponential stability in nonlinear wave equations near the fundamental linear mode. *Phys. D* **122**(1–4), 73–104 (1998)
- [5] Bambusi, D., Ponno, A.: On metastability in FPU. *Commun. Math. Phys.* **264**(2), 539–561 (2006)
- [6] Benettin, G., Christodoulidi, H., Ponno, A.: The Fermi–Pasta–Ulam problem and its underlying integrable dynamics. *J. Stat. Phys.* **152**, 195–212 (2013)
- [7] Benettin, G., Livi, R., Ponno, A.: The Fermi–Pasta–Ulam problem: scaling laws vs. initial conditions. *J. Stat. Phys.* **135**(5–6), 873–893 (2009)
- [8] Benettin, G., Ponno, A.: Time-scales to equipartition in the Fermi–Pasta–Ulam problem: finite-size effects and thermodynamic limit. *J. Stat. Phys.* **144**, 793–812 (2011). doi:[10.1007/s10955-011-0277-9](https://doi.org/10.1007/s10955-011-0277-9)
- [9] Benettin, G., Fröhlich, J., Giorgilli, A.: A Nekhoroshev-type theorem for Hamiltonian systems with infinitely many degrees of freedom. *Commun. Math. Phys.* **119**(1), 95–108 (1988)
- [10] Benettin, G., Galgani, L., Giorgilli, A.: Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. II. *Commun. Math. Phys.* **121**(4), 557–601 (1989)
- [11] Berchiulla, L., Giorgilli, A., Paleari, S.: Exponentially long times to equipartition in the thermodynamic limit. *Phys. Lett. A* **321**(3–4), 167–172 (2004)
- [12] Bourgain, J.: Hamiltonian methods in nonlinear evolution equations. In: *Fields Medallists’ lectures*, pp. 542–554. World Scientific Publishing, River Edge, NJ, (1997)
- [13] Bourgain, J.: Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. *Ann. Math. (2)* **148**(2), 363–439 (1998)
- [14] Carati, A., Galgani, L., Giorgilli, A., Paleari, S.: Fermi–Pasta–Ulam phenomenon for generic initial data. *Phys. Rev. E* **76**(2), 022104 (2007)

- [15] Carati, A.: An averaging theorem for Hamiltonian dynamical systems in the thermodynamic limit. *J. Stat. Phys.* **128**, 1057–1077 (2007)
- [16] Carati, A., Maiocchi, A.M: Exponentially long stability times for a nonlinear lattice in the thermodynamic limit. *Commun. Math. Phys.* **314**, 129–161 (2012). doi:[10.1007/s00220-012-1522-z](https://doi.org/10.1007/s00220-012-1522-z)
- [17] Craig, W.: KAM theory in infinite dimensions. In: *Dynamical systems and probabilistic methods in partial differential equations* (Berkeley, CA, 1994), pp. 31–46. American Mathematical Society, Providence, RI (1996)
- [18] Davis, Philip J.: *Circulant matrices*. A Wiley-Interscience Publication, Pure and Applied Mathematics. Wiley, New York (1979)
- [19] De Roeck, W., Huveneers, F.: Asymptotic localization of energy in non-disordered oscillator chains. ArXiv e-prints, May 2013. 1305.5127
- [20] Fröhlich, J., Spencer, T., Wayne, C.E.: An invariant torus for nearly integrable Hamiltonian systems with infinitely many degrees of freedom. In: *Stochastic processes in classical and quantum systems* (Ascona, 1985), pp. 256–268. Springer, Berlin (1986)
- [21] Gallavotti, G. (ed): *The Fermi–Pasta–Ulam problem*. Volume 728 of *Lecture Notes in Physics*. Springer, Berlin (2008) (A status report)
- [22] Genta, T., Giorgilli, A., Paleari, S., Penati, T.: Packets of resonant modes in the Fermi–Pasta–Ulam system. *Phys. Lett. Sect. A Gen. At. Solid State Phys.* **376**(28–29), 2038–2044 (2012)
- [23] Giorgilli, A., Galgani, L.: Formal integrals for an autonomous Hamiltonian system near an equilibrium point. *Celest. Mech.* **17**(3), 267–280 (1978)
- [24] Giorgilli, A., Paleari, S., Penati, T.: Extensive adiabatic invariants for nonlinear chains. *J. Stat. Phys.* **148**(6), 1106–1134 (2012)
- [25] Kolmogorov, A.N.: On conservation of conditionally periodic motions for a small change in Hamilton’s function. *Dokl. Akad. Nauk SSSR (N.S.)* **98**, 527–530 (1954)
- [26] Kuksin, S.B.: KAM-theory for partial differential equations. In: *First European Congress of Mathematics*, vol. II (Paris, 1992), pp. 123–157. Birkhäuser, Basel, (1994)
- [27] Kuksin, S.B.: *Analysis of Hamiltonian PDEs*. The Clarendon Press, Oxford University Press, New York (2000)
- [28] Lanford, O.E.: Entropy and equilibrium states in classical statistical mechanics. In: *Statistical mechanics and mathematical problems* (Battelle Seattle 1971 Rencontre), volume 20 of *Lecture Notes in Physics*, pp. 1–113. Springer, Berlin (1973)
- [29] Lorenzoni, P., Paleari, S.: Metastability and dispersive shock waves in the Fermi–Pasta–Ulam system. *Phys. D* **221**(2), 110–117 (2006)
- [30] Moser, J.K.: A new technique for the construction of solutions of nonlinear differential equations. *Proc. Nat. Acad. Sci. USA* **47**, 1824–1831 (1961)
- [31] Moser, J.K.: On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1962**, 1–20 (1962)
- [32] Nekhoroshev, N.N.: An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. *Uspehi Mat. Nauk* **32** (6(198)), 5–66, 287 (1977). [English translation: *Russian Math. Surveys* **32** no. 6, 1–65 (1977)]

- [33] Nekhoroshev, N.N.: An exponential estimate of the time of stability of nearly integrable Hamiltonian systems. II. Trudy Sem. Im. G. Petrovskogo **5**, 5–50 (1979). [English translation: Topics in modern Mathematics, Petrovskij Semin. **5**, 1–58 (1985)]
- [34] Paleari, S., Bambusi, D., Cacciatori, S.: Normal form and exponential stability for some nonlinear string equations. Z. Angew. Math. Phys. **52**(6), 1033–1052 (2001)
- [35] Poincaré, H.: Les méthodes nouvelles de la mécanique céleste. Tome II. Librairie Scientifique et Technique Albert Blanchard, Paris. Méthodes de MM. Newcomb, Gylden, Lindstedt et Bohlin (1987) [The methods of Newcomb, Gylden, Lindstedt and Bohlin, Reprint of the 1893 original, Bibliothèque Scientifique Albert Blanchard. Albert Blanchard Scientific Library]
- [36] Pöschel, J.: On Nekhoroshev estimates for a nonlinear Schrödinger equation and a theorem by Bambusi. Nonlinearity **12**(6), 1587–1600 (1999)

Antonio Giorgilli, Simone Paleari and Tiziano Penati

Dipartimento di Matematica

Università degli Studi di Milano

Via Saldini, 50

20133 Milan, Italy

e-mail: antonio.giorgilli@unimi.it;

simone.paleari@unimi.it;

tiziano.penati@unimi.it

Communicated by Dmitry Dolgopyat.

Received: July 17, 2013.

Accepted: March 20, 2014.