Metastability phenomena in two-dimensional rectangular lattices with nearest-neighbour interaction

View the article online for updates and enhancements.

Metastability phenomena in two-dimensional rectangular lattices with nearest-neighbour interaction

M Gallone[1] and S Pasquali[2,*]

[1] Dipartimento di Matematica, Università degli Studi di Milano, via Cesare Saldini 50, 20133, Milan, Italy
[2] Matematikcentrum, Lunds Universitet, Sölvegatan 18, 223 62 Lund, Sweden

E-mail: matteo.gallone@unimi.it and stefano.pasquali@math.lu.se

Abstract

We study analytically the dynamics of two-dimensional rectangular lattices with periodic boundary conditions. We consider anisotropic initial data supported on one low-frequency Fourier mode. We show that, in the continuous approximation, the resonant normal form of the system is given by integrable PDEs. We exploit the normal form in order to prove the existence of metastability phenomena for the lattices. More precisely, we show that the energy spectrum of the normal modes attains a distribution in which the energy is shared among a packet of low-frequencies modes; such distribution remains unchanged up to the time-scale of validity of the continuous approximation.

Keywords: continuous approximation, metastability, energy localization
Mathematics Subject Classification numbers: 37K10, 37K60, 70H08, 70K45.

(Some figures may appear in colour only in the online journal)

1. Introduction

In this paper we present an analytical study of the dynamics of two-dimensional rectangular lattices with nearest-neighbour interaction and periodic boundary conditions, for initial data with only one low-frequency Fourier mode initially excited. We give some rigorous results

concerning the relaxation to a metastable state, in which energy sharing takes place among low-frequency modes only.

The study of metastability phenomena for lattices started with the numerical result by Fermi, Pasta and Ulam (FPU) [19951965], who investigated the dynamics of a one-dimensional chain of particles with nearest neighbour interaction. In the original simulations all the energy was initially given to a single low-frequency Fourier mode with the aim of measuring the time of relaxation of the system to the 'thermal equilibrium' by looking at the evolution of the Fourier spectrum. Classical statistical mechanics prescribes that the energy spectrum corresponding to the thermal equilibrium is a plateau (the so-called theorem of equipartition of energy). Despite the authors believed that the approach to such an equilibrium would have occurred in a short time-scale, the outcoming Fourier spectrum was far from being flat and they observed two features of the dynamics that were in contrast with their expectations: the lack of thermalization displayed by the energy spectrum and the recurrent behaviour of the dynamics.

Both from a physical and a mathematical point of view, the studies on FPU-like systems have a long and active history: a concise survey of this vast literature is discussed in the monograph [2007]. For a more recent account on analytic results on the 'FPU paradox' we refer to [2015].

In particular, we mention the papers [Bam08, BP06], in which the authors used the techniques of canonical perturbation theory for PDEs in order to show that the FPU α model (respectively, β model) can be rigorously described by a system of two uncoupled KdV (resp. mKdV) equations, which are obtained as a resonant normal form of the continuous approximation of the FPU model; moreover, this result allowed to deduce a rigorous result about the energy sharing among the Fourier modes, up to the time-scales of validity of the approximation. If we denote by N the number of degrees of freedom for the lattice and by $\mu \sim \frac{1}{N} \ll 1$ the wave-number of the initially excited mode, if we assume that the specific energy $\epsilon \sim \mu^4$ (resp. $\epsilon \sim \mu^2$ for the FPU β model), then the dynamics of the KdV (resp. mKdV) equations approximates the solutions of the FPU model up to a time of order $\mathcal{O}(\mu^{-3})$. However, the relation between the specific energy and the number of degrees of freedom implies that the result does not hold in the thermodynamic limit regime, namely for large N and for fixed specific energy ϵ (such a regime is the one which is relevant for statistical mechanics).

Unlike the extensive research concerning one-dimensional systems, it seems to the authors that the behaviour of the dynamics of two-dimensional lattices is far less clear; it is expected that the interplay between the geometry of the lattice and the specific energy regime could lead to different results.

Benettin and collaborators [Ben05, BG08, BVT80] studied numerically a two-dimensional FPU lattice with triangular cells and different boundary conditions in order to estimate the equipartition time-scale. They found out that in the thermodynamic limit regime the equipartition is reached faster than in the one-dimensional case. The authors decided not to consider model with square cells in order to have a spectrum of linear frequencies which is different with respect to the one of the one-dimensional model; they also added (see [BG08], section B (iii)):

There is a good chance, however, that models with square lattice, and perhaps a different potential so as to avoid instability, behave differently from models with triangular lattice, and are instead more similar to one-dimensional models. This would correspond to an even stronger lack of universality in the two-dimensional FPU problem.

Up to the authors' knowledge, the only analytical results on the dynamics of two-dimensional lattices in this framework concern the existence of breathers [BPP10, BW06, BW07, 2015, 2005].

In this paper we study two-dimensional rectangular lattices with $(2N_1 + 1) \times (2N_2 + 1)$ sites, square cell, nearest-neighbour interaction and periodic boundary conditions, and we show the existence of metastability phenomena as in [BP06]. More precisely, if we denote by $\mu \ll 1$ the wave-number of the Fourier mode initially excited and by σ the ratio between the sides of the lattice, we obtain for a 2D electrical transmission lattice (ETL) either a system of two uncoupled KP-II equations for $\mu \ll 1$ and $\sigma = 2$, or a system of two uncoupled KdV equations for $\mu \ll 1$ and $2 < \sigma < 7$ as a resonant normal form for the continuous approximation of the lattice, while for the 2D Klein–Gordon lattice with quartic defocusing nonlinearity we obtain a one-dimensional cubic defocusing NLS equation for $\mu \ll 1$ and $1 < \sigma < 7$. Since all the above PDEs are integrable, we can exploit integrability to deduce a mathematically rigorous result on the formation of the metastable packet.

Up to the authors' knowledge, this is the first analytical result about metastable phenomena in two-dimensional Hamiltonian lattices with periodic boundary conditions; in particular, this is the first rigorous result for two-dimensional lattices in which the dynamics of the lattice in a two-dimensional regime is described by a system of two-dimensional integrable PDEs.

Some comments are in order:

(a) The time-scale of validity of our result is of order $\mathcal{O}(\mu^{-3})$ for the 2D ETL lattice, and of order $\mathcal{O}(\mu^{-2})$ for the 2D Klein–Gordon lattice;

(b) The ansatz about the small amplitude solutions gives a relation between the specific energy of the system ϵ and the wave-number $\mu \sim \frac{1}{N_1}$ of the Fourier mode initially excited. More precisely, we obtain $\epsilon \sim \mu^4$ for the 2D ETL lattice as in [BP06], and $\epsilon \sim \mu^2$ for the 2D Klein–Gordon lattice. This implies that the result does not hold in the thermodynamic limit regime;

(c) Our result can be easily generalized to higher-dimensional lattices, such as the physical case of three-dimensional rectangular lattices with cubic cells;

(d) Depending on the geometry of the lattice which is encoded in the parameter σ, the effective dynamics is described by one-dimensional PDEs for highly anisotropic lattices and by two-dimensional PDEs for low values of σ. The normal form equation of the ETL lattices is not integrable for $1 \leqslant \sigma < 2$ and they are integrable if $\sigma > 2$. The edge case $\sigma = 2$ is very sensitive to the potential: if the cubic term of the potential is present, then the normal form equation it is the integrable two-dimensional KP-II equation, otherwise it is a non-integrable modification of the mKdV equation. On the other hand, the normal form equation for KG lattices is integrable for $\sigma > 1$, thus also for less anisotropic lattices;

(e) The upper bounds for σ in the KdV regime and in the NLS regime come from a technical assumption in the approximation results (see propositions 6.5, 6.2 and 6.9). The approximation of solutions for the lattice with solutions of integrable PDEs in one-dimensional lattices was obtained through a detailed analysis in order to bound the error, and this is also the case for two-dimensional lattices, (see propositions 6.5, 6.9, appendices C and E), where one has to do very careful estimates in order to bound the different contributions to the error.

To prove our results we follow the strategy of [BP06]. The first step consists in the approximation of the dynamics of the lattice with the dynamics of a continuous system. This step gives also a natural perturbative order and, since $\sigma > 1$, the leading term is given by a PDE in only one space variable. The effect of the second dimension is of order μ^σ, and thus it appears at the second perturbative order for low values of σ and at higher perturbative orders for higher values. In this sense one expects that for higher values of σ the normal form equations are one-dimensional, but it is not trivial that the effect of the second dimension do not destroy

integrability. For this reason, as a second step we perform a normal form canonical transformation and we obtain that the effective dynamics is given by a system of integrable PDEs (KdV, KP-II, NLS depending on the lattice and the relation between N_1 and N_2). Next, we exploit the dynamics of these integrable PDEs in order to construct approximate solutions of the original discrete lattices and we estimate the error with respect to a true solution with the corresponding initial datum. Finally, we use the known results about the dynamics of the above mentioned integrable PDEs in order to estimate the specific energies for the approximate solutions of the original lattices.

The novelties of this work are: on the one side, a mathematically rigorous proof of the approximation of the dynamics of the 2D ETL lattice by the dynamics of certain integrable PDEs (among these integrable PDEs, there is one which is *genuinely* two-dimensional, the KP-II equation) and of the dynamics of the 2D KG lattice by the dynamics of the one-dimensional nonlinear Schrödinger (NLS) equation; on the other side, there are two technical differences with respect to previous works, namely the normal form theorem (which is a variant of the technique used in [Bam05, 2005, Pas19]) and the estimates for bounding the error between the approximate solution and the true solution of the lattice (which need a more careful study than the ones appearing in [BP06, SW00] for the one-dimensional case).

The paper is organized as follows: in section 2 we introduce the mathematical setting of the models and we state our main results, theorems 2.1, 2.5 and 2.7. In section 3 we state an abstract averaging theorem, which we prove in section 3.2. In section 4 we apply the averaging theorem to the two-dimensional lattices, deriving the integrable approximating PDEs in the different regimes. In section 5 we review some results about the dynamics of the normal form equation. In section 6 we use the normal form equations in order to construct approximate solutions (see propositions 6.5, 6.2 and 6.9), and we estimate the difference with respect to the true solutions with corresponding initial data in propositions 6.6, 6.3 and 6.10. In appendix A we prove the technical lemma 3.6; in appendix B we prove propositions 6.2 and 6.5; in appendix C we prove propositions 6.6 and 6.3; in appendix D we prove proposition 6.9; in appendix E we prove proposition 6.10.

Relevant notations. For the sake of clarity we provide a short explanation of some of the symbols we use in the paper. With $\mathbb{Z}^2_{N_1,N_2}$ we denote $\frac{\mathbb{Z}}{2N_1+1} \times \frac{\mathbb{Z}}{2N_2+1} = \{(j_1,j_2) \in \mathbb{Z}^2 : |j_1| \leqslant N_1, |j_2| \leqslant N_2\}$. We will denote by $k = (k_1,k_2) \in \mathbb{Z}^2$ the index of the wave-vector; we denote by $\kappa = (\kappa_1,\kappa_2)$ the specific wave vector defined in (9). In section 6 and in the appendices we will denote by $K = (K_1,K_2)$ the index of the wave-vector with $|K_1| \leqslant N_1$ and $|K_2| \leqslant N_2$. With N_r ($r = 1,2$) we denote the half-length of the rectangular lattice and with $N := (2N_1 + 1)(2N_2 + 1)$ we denote the total number of sites.

Space variables are denoted by x_r (with $r = 1,2$) with $x_r \in [-N_r, N_r]$ and the rescaled space variables are denoted by y_r with $y_r \in [-1,1]$.

The Hilbert space $\ell^2_{\rho,s}$ is defined in definition 3.1 and $B_{\rho,s}(R)$ denotes the open ball of radius R centred in 0, $B_{\rho,s}(R) \subset \ell^2_{\rho,s}$. We also use $\mathcal{H}^{\rho,s} := \ell^2_{\rho,s} \times \ell^2_{\rho,s}$ and $\mathcal{B}_{\rho,s}(R) := B_{\rho,s}(R) \times B_{\rho,s}(R) \subset \mathcal{H}^{\rho,s}$.

We also denote with $|(a,b)| := \sqrt{a^2 + b^2}$ the Euclidean norm of the vector (a,b). We denote with $[f] := \int_0^T f \circ \Phi^t_\mathsf{h}(y) \frac{\mathrm{d}t}{T}$ the time-average of f with respect to Φ^t_h.

In section 4 we introduce interpolating functions and we denote with $Q(t,x)$ the interpolating function for the lattice variable $Q_j(t)$ and with $P(t,x)$ the interpolating function for the lattice variable $P_j(t)$. We use this little abuse of notation because the former is actually an extension of the latter in the sense that $P_j(t) = P(t,j)$ for every t and j.

The component k of the Fourier transform of the continuous approximation at time t will be denoted by $\hat{Q}(t,k)$ to distinguish it from the kth component of the Fourier transform of the

reticular variable $\hat{Q}_k(t)$. Since for the rescaled functions such as q and p this ambiguity do not hold, we prefer the short notation $\hat{q}_k(\tau)$ and $\hat{p}_k(\tau)$ for their Fourier transform.

In section 6 we make use of the following abuse of notation: $Q(t) = \xi(\tau)$. This is to avoid the long-writing $Q(t) = \xi(\tau(t))$.

2. Main results

We consider a periodic two-dimensional rectangular lattice, called ETL lattice, which in the non-periodic setting has been studied in [BW06], and which can be regarded as a simpler version of a 2D rectangular FPU model. We denote

$$\mathbb{Z}^2_{N_1,N_2} := \{(j_1, j_2) : j_1, j_2 \in \mathbb{Z}, \quad |j_1| \leqslant N_1, \quad |j_2| \leqslant N_2\}; \tag{1}$$

we also write $e_1 := (1, 0)$, $e_2 := (0, 1)$ and we denote by $N := (2N_1 + 1)(2N_2 + 2)$ the total number of sites of the lattice.

The Hamiltonian describing the ETL lattice is given by

$$H_{\mathrm{ETL}}(Q, P) = \sum_{j \in \mathbb{Z}^2_{N_1,N_2}} -\frac{1}{2} P_j (\Delta_1 P)_j + (V(Q))_j, \tag{2}$$

$$(\Delta_1 P)_j := (P_{j+e_1} - 2P_j + P_{j-e_1}) + (P_{j+e_2} - 2P_j + P_{j-e_2}), \tag{3}$$

$$(V(Q))_j = \frac{Q_j^2}{2} + \alpha \frac{Q_j^3}{3} + \beta \frac{Q_j^4}{4}. \tag{4}$$

We refer to (2) as $\alpha + \beta$ model (respectively, β model) if $\alpha \neq 0$ (respectively $\alpha = 0$). With the above Hamiltonian formulation the equations of motion associated to (2) are given by

$$\begin{cases} \dot{Q}_j = -(\Delta_1 P)_j \\ \dot{P}_j = -(V'(Q))_j \end{cases} ;$$

$$\ddot{Q}_j = (\Delta_1 V'(Q))_j. \tag{5}$$

We also introduce the Fourier coefficients of Q via the following standard relation,

$$Q_j(t) := \frac{1}{\sqrt{N}} \sum_{k \in \mathbb{Z}^2_{N_1,N_2}} \hat{Q}_k(t) \, e^{2\pi i \frac{jk}{N}} \quad j \in \mathbb{Z}^2_{N_1,N_2}, \tag{6}$$

and similarly for P_j. We denote by

$$E_k := \frac{\omega_k^2 |\hat{P}_k|^2 + |\hat{Q}_k|^2}{2}, \tag{7}$$

$$\omega_k^2 := 4 \sin^2 \left(\frac{k_1 \pi}{2N_1 + 1} \right) + 4 \sin^2 \left(\frac{k_2 \pi}{2N_2 + 1} \right), \tag{8}$$

the energy and the square of the frequency of the mode at site $k = (k_1, k_2) \in \mathbb{Z}^2_{N_1,N_2}$ (see figure 1). For states described by real functions, one has $E_{(k_1,k_2)} = E_{(-k_1,k_2)}$ and $E_{(k_1,k_2)} = E_{(k_1,-k_2)}$ for all $k = (k_1, k_2)$, so we will consider only indexes in

$$\mathbb{Z}^2_{N_1,N_2,+} := \{(k_1, k_2) \in \mathbb{Z}^2_{N_1,N_2} : k_1, \quad k_2 \geqslant 0\}.$$
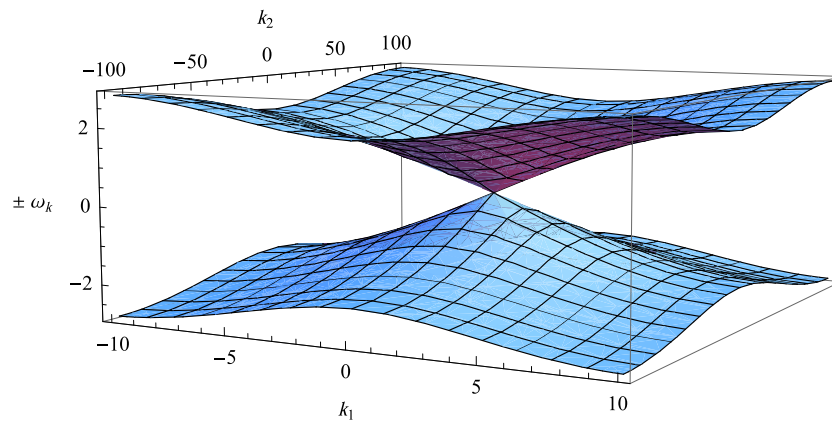
Figure 1. The dispersion relation for the ETL lattice (2), namely $\pm\omega_k$, $k = (k_1, k_2)$, vs the integer coordinates, for $N_1 = 10$ and $N_2 = 100$.

As is customary in lattices with a large number of degrees of freedom, especially in relation with statistical mechanics, we introduce the *specific wave vector* κ as

$$\kappa := \kappa(k) = \left(\frac{k_1}{N_1 + \frac{1}{2}}, \frac{k_2}{N_2 + \frac{1}{2}} \right), \tag{9}$$

and the *specific energy of the specific normal mode* $\kappa = \kappa(k)$ as

$$\mathcal{E}_\kappa := \frac{E_k}{\left(N_1 + \frac{1}{2} \right) \left(N_2 + \frac{1}{2} \right)}. \tag{10}$$

We want to study the behaviour of small amplitude solutions of (5), with initial data in which only one low-frequency Fourier mode is excited.

We assume $N_1 \leqslant N_2$, and we introduce the quantities

$$\mu := \frac{2}{2N_1 + 1}, \tag{11}$$

$$\sigma := \log_{N_1 + \frac{1}{2}} \left(N_2 + \frac{1}{2} \right), \tag{12}$$

which play the role of parameters in our construction: we will use them in the asymptotic expansion of the dispersion relation of the continuous approximation of the lattice (see (88), (89), (105) and (106)) in order to derive the integrable approximating PDEs in the regimes we are considering.

We study the $\alpha + \beta$ model of (5) in the following regime:

(KP) The *weakly transverse regime*, where the effective dynamics is described by a system of two *uncoupled* Kadomtsev–Petviashvili (KP) equation. This corresponds to taking $\mu \ll 1$ and $\sigma = 2$.

From now on, we denote by $\kappa_0 := \left(\frac{1}{N_1 + \frac{1}{2}}, \frac{1}{(N_1 + \frac{1}{2})^\sigma} \right) = (\mu, \mu^\sigma)$. Our main result is the following:

Theorem 2.1. *Consider* (5) *with* $\alpha \neq 0$, $\sigma = 2$.

Fix $0 < \gamma < \frac{1}{2}$ *and two positive constants* C_0 *and* T_0, *then there exist positive constants* μ_0, C_1 *and* C_2 *(depending only on* γ, C_0 *and on* T_0*) such that the following holds. Consider an initial datum with*

$$\mathcal{E}_{\kappa_0}(0) = C_0\mu^4, \quad \mathcal{E}_\kappa(0) = 0 \quad \forall \, \kappa = (\kappa_1, \kappa_2) \neq \kappa_0, \tag{13}$$

and assume that $\mu < \mu_0$. *Then there exists* $\rho > 0$ *such that along the corresponding solution one has*

$$\mathcal{E}_\kappa(t) \leqslant C_1 \, \mu^4 e^{-\rho|(\kappa_1/\mu, \kappa_2/\mu^\sigma)|} + C_2 \, \mu^{4+\gamma}, \quad \forall \, |t| \leqslant \frac{T_0}{\mu^3} \tag{14}$$

for all κ.

Remark 2.2.   Theorem 2.1 is the first rigorous result for two-dimensional lattices in which the dynamics of the lattice in a genuinely two-dimensional regime is described by a system of two-dimensional integrable PDEs. Moreover, in theorem 2.1 we do not mention the existence of a sequence of almost-periodic functions approximating the specific energies of the modes, and this is a difference with respect to theorem 5.3 in [BP06]. This is related to the construction of action-angle/Birkhoff coordinates for the KP equation, which is an open problem in the theory of integrable PDEs.

Remark 2.3.   For the sake of simplicity, we have proved theorem 2.1 for initial data in which only one low-frequency Fourier mode is excited. One can also prove that a variant of theorem 2.1 holds also in the case the higher harmonics of a low-frequency Fourier mode are excited, provided that the energy decreases exponentially with respect to $|(\kappa_1/\mu, \kappa_2/\mu^\sigma)|$, and also for initial data in which the symmetrical modes of a given low-frequency Fourier mode are excited. To summarize, we are only able to prove stability of the solutions we constructed for initial data with vanishing specific energy for a time-scale $\mathcal{O}(\mu^{-3})$.

Remark 2.4.   Here and in the following theorems we decided to write the bound on the specific energy with respect to the intensive variables $\kappa = (\kappa_1, \kappa_2)$. This writing has the advantage to be consistent with the previous literature (e.g. compare (13) above with (3.8) in [BP06]). Anyway, let us emphasize that, using the definition of κ, (13) can be written equivalently as

$$\mathcal{E}_\kappa(t) \leqslant C_1\mu^4 \, e^{-\rho|(k_1, k_2)|} + C_2\mu^{4+\gamma}, \quad \forall \, t \leqslant \frac{T_0}{\mu^3}. \tag{15}$$

This implies that, as the number of sites increase, the Fourier spectrum is exponentially localized (and remains so, as $\mu \to 0$) around $\kappa = 0$, apart from an error of order $\mu^{4+\gamma}$.

We also point out that there are also other regimes in which the dynamics of a two-dimensional lattice can be approximated by integrable PDEs. For example, we can consider $\alpha + \beta$ model of (5) in the following regime:

(KdV) The *very weakly transverse regime*, where the effective dynamics is described by a system of two *uncoupled* Korteweg–de Vries (KdV) equations. This corresponds to taking $\mu \ll 1$ and $2 < \sigma < 7$.

The corresponding result one can prove in such a regime is the following.

Theorem 2.5.   *Consider* (5) *with* $\alpha \neq 0$, $2 < \sigma < 7$. *Define* $\gamma_0(\sigma) := \frac{3}{2}\left(\sigma - \frac{5}{3}\right)$ *for* $2 < \sigma < 3$, *and* $\gamma_0(\sigma) := \frac{1}{2}(7 - \sigma)$ *for* $3 \leqslant \sigma < 7$.

Fix $0 < \gamma < \gamma_0(\sigma)$ and two positive constants C_0 and T_0, then there exist positive constants μ_0, C_1 and C_2 (depending only on γ, C_0, σ and on T_0) such that the following holds. Consider an initial datum with

$$\mathcal{E}_{\kappa_0}(0) = C_0\mu^4, \qquad \mathcal{E}_\kappa(0) = 0, \quad \forall\, \kappa = (\kappa_1, \kappa_2) \neq \kappa_0, \tag{16}$$

and assume that $\mu < \mu_0$. Then there exists $\rho > 0$ such that along the corresponding solution one has

$$\mathcal{E}_\kappa(t) \leqslant C_1\mu^4 e^{-\rho|(\kappa_1/\mu, \kappa_2/\mu^\sigma)|} + C_2\mu^{4+\gamma}, \quad \forall\, |t| \leqslant \frac{T_0}{\mu^3} \tag{17}$$

for all κ. Moreover, for any n_2 with $0 \leqslant n_2 \leqslant N_2$ there exists a sequence of almost-periodic functions $(F_n)_{n=(n_1,n_2)\in\mathbb{Z}^2_{N_1,N_2,+}}$ such that, if we define

$$\mathcal{F}_{\kappa_0} = \mu^4 F_n, \qquad \mathcal{F}_\kappa = 0 \quad \forall\, \kappa \neq n\kappa_0 \tag{18}$$

then one has, for the specific energy distribution $\mathcal{E}_\kappa(t)$,

$$|\mathcal{E}_\kappa(t) - \mathcal{F}_\kappa(t)| \leqslant C_2\mu^{4+\gamma}, \quad \forall\, |t| \leqslant \frac{T_0}{\mu^3}. \tag{19}$$

Remark 2.6.   We point out that in the statement of theorem 2.5 the assumption $\sigma > 2$ comes from an asymptotic expansion of the dispersion relation of the continuous approximation of the lattice (see (88) and (89)), while the assumption $\sigma < 7$ comes from a technical assumption under which we can approximate the dynamics of the lattice with the dynamics of the system of uncoupled KdV equations (see the statement of proposition 6.6).

We can also consider two-dimensional KG lattices, which combine the nearest-neighbour potential with an on-site one: the scalar model is described by

$$H_{\mathrm{KG}}(Q,P) = \sum_{j\in\mathbb{Z}^2_{N_1,N_2}} \frac{P_j^2}{2} + \frac{1}{2}\sum_{j\in\mathbb{Z}^2_{N_1,N_2}} Q_j(-\Delta_1 Q)_j + \sum_{j\in\mathbb{Z}^2_{N_1,N_2}} U(Q_j), \tag{20}$$

$$U(x) = m^2\frac{x^2}{2} + \beta\frac{x^{2p+2}}{2p+2}, \quad m > 0, \quad \beta > 0, \quad p \geqslant 1, \tag{21}$$

(see [Ros03] for a physical interpretation of the model). The associated equations of motion are

$$\ddot{Q}_j = (\Delta_1 Q)_j - m^2 Q_j - \beta Q_j^{2p+1}, \quad j \in \mathbb{Z}^2_{N_1,N_2}. \tag{22}$$

If we take $p = 1$, we obtain a generalization of the one-dimensional ϕ^4 model.

We now introduce the Fourier coefficients of Q as in (6), and similarly for P, and denote by

$$E_k := \frac{|\hat{P}_k|^2 + \omega_k^2|\hat{Q}_k|^2}{2}, \tag{23}$$

$$\omega_k^2 := m^2 + 4\sin^2\left(\frac{k_1\pi}{2N_1+1}\right) + 4\sin^2\left(\frac{k_2\pi}{2N_2+1}\right), \tag{24}$$

the energy and the square of the frequency of the mode at site $k = (k_1, k_2) \in \mathbb{Z}^2_{N_1,N_2}$ (see figure 2).

In the rest of the paper we will assume that $m = 1$.

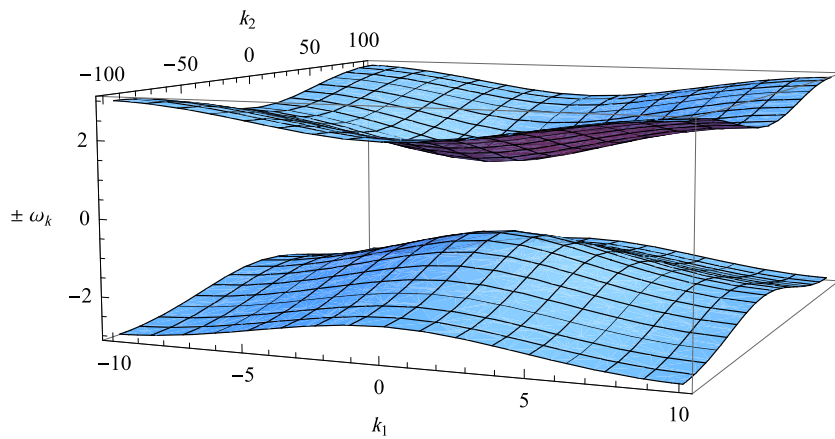We consider the two-dimensional KG lattice (20) in the following regime:

Figure 2. The dispersion relation for the KG lattice (20), namely $\pm\omega_k$, $k = (k_1, k_2)$, vs the integer coordinates, for $N_1 = 10$ and $N_2 = 100$.

(1D NLS) The *very weakly transverse regime*, where the effective dynamics is described by a cubic one-dimensional NLS equation. This corresponds to taking $\mu \ll 1$ and $1 < \sigma < 7$.

Theorem 2.7. *Consider* (20) *with* $\beta > 0$, $1 < \sigma < 7$. *Define* $\gamma_1(\sigma) := \frac{3}{2}\left(\sigma - \frac{1}{3}\right)$ *for* $1 < \sigma < 2$, *and* $\gamma_1(\sigma) := \frac{1}{2}(7 - \sigma)$ *for* $2 \leqslant \sigma < 7$.

Fix $0 < \gamma \leqslant \gamma_1(\sigma)$ *and two positive constants* C_0 *and* T_0, *then there exist positive constants* μ_0, C_1 *and* C_2 *(depending only on* γ, C_0, σ *and on* T_0*) such that the following holds. Consider an initial datum with*

$$\mathcal{E}_{\kappa_0}(0) = C_0\mu^2, \qquad \mathcal{E}_\kappa(0) = 0, \quad \forall\, \kappa = (\kappa_1, \kappa_2) \neq \kappa_0, \tag{25}$$

and assume that $\mu < \mu_0$. *Then there exists* $\rho > 0$ *such that along the corresponding solution one has*

$$\mathcal{E}_\kappa(t) \leqslant C_1\,\mu^2 e^{-\rho|(\kappa_1/\mu, \kappa_2/\mu^\sigma)|} + C_2\mu^{2+\gamma}, \quad |t| \leqslant \frac{T_0}{\mu^2} \tag{26}$$

for all κ. *Moreover, for any* n_2 *with* $0 \leqslant n_2 \leqslant N_2$ *there exists a sequence of almost-periodic functions* $(F_n)_{n=(n_1, n_2)\in\mathbb{Z}^2_{N_1, N_2, +}}$ *such that, if we denote*

$$\mathcal{F}_{\kappa_0} = \mu^2 F_n, \qquad \mathcal{F}_\kappa = 0 \quad \forall\, \kappa \neq n\kappa_0 \tag{27}$$

then we have for the specific energy distribution

$$|\mathcal{E}_\kappa(t) - \mathcal{F}_\kappa(t)| \leqslant C_2\mu^{2+\gamma}, \quad |t| \leqslant \frac{T_0}{\mu^2}. \tag{28}$$

Remark 2.8. In theorem 2.7 we are able to prove stability of the solutions we constructed for initial data with vanishing specific energy for a time-scale $\mathcal{O}(\mu^{-2})$.

Remark 2.9. As for theorem 2.5, in the statement of theorem 2.7 the assumption $\sigma > 1$ comes from an asymptotic expansion of the dispersion relation of the continuous approximation of the lattice (see (105) and (106)), while the assumption $\sigma < 7$ comes from a technical

assumption under which we can approximate the dynamics of the lattice with the dynamics of the system of uncoupled NLS equations (see the statement of theorem 6.10).

2.1. Further remarks

Remark 2.10. The specific choice of the direction of longitudinal propagation in the regimes that we have considered is not relevant.

Remark 2.11. We point out that the time of validity of theorem 2.7 for the KG lattice is of order $\mathcal{O}(\mu^{-2})$, which is different from the time of validity of theorems 2.5 and 2.1 for the FPU lattice. In the one-dimensional case it has been observed that, for a fixed value of specific energy ϵ and for long-wavelength modes initially excited, the ϕ^4 model reached equipartition faster than the FPU β model (see [2007], section 2.1.8).

Remark 2.12. Theorems 2.1, 2.5 and 2.7 can be generalized to higher dimensional lattices. Indeed, let $d \geqslant 2$, define

$$\mathbb{Z}^d_{N_1,\ldots,N_d} := \{(j_1,\ldots,j_d) : j_1,\ldots,j_d \in \mathbb{Z}, |j_1| \leqslant N_1,\ldots,|j_d| \leqslant N_d\}, \tag{29}$$

and consider the d-dimensional ETL

$$H^{(d)}_{\text{ETL}}(Q,P) = \sum_{j \in \mathbb{Z}^d_{N_1,\ldots,N_d}} \left(-\frac{1}{2} P_j (\Delta_1 P)_j + (V(Q))_j \right),$$

$$(V(Q))_j = \frac{Q_j^2}{2} + \alpha \frac{Q_j^3}{3} + \beta \frac{Q_j^4}{4}, \quad j \in \mathbb{Z}^d_{N_1,\ldots,N_d}, \tag{30}$$

and the d-dimensional nonlinear Klein–Gordon (NLKG) lattice

$$H^{(d)}_{\text{KG}}(Q,P) = \sum_{j \in \mathbb{Z}^d_{N_1,\ldots,N_d}} \frac{P_j^2}{2} + \frac{1}{2} \sum_{\substack{j,k \in \mathbb{Z}^d_{N_1,\ldots,N_d} \\ |j-k|=1}} \frac{(Q_j - Q_k)^2}{2} + \sum_{j \in \mathbb{Z}^d_{N_1,\ldots,N_d}} U(Q_j),$$

$$U(x) = \frac{x^2}{2} + \beta \frac{x^4}{4}, \quad \beta > 0. \tag{31}$$

We assume $N_1 \leqslant N_2,\ldots,N_d$, and we introduce the quantities

$$\mu := \frac{2}{2N_1 + 1}, \quad \sigma_i := \log_{N_1 + \frac{1}{2}}\left(N_{i+1} + \frac{1}{2}\right), \quad i = 1,\ldots,d-1. \tag{32}$$

Then we can describe the following regimes:

(KdV-d) The model (30) for $d \leqslant 4$, in the very weakly transverse regime with $\mu \ll 1$ and $2 < \sigma_1,\ldots,\sigma_{d-1} < 7$;

(KP-d) The model (30) for $d \leqslant 4$, in the weakly transverse regime with $\mu \ll 1$ and $\sigma_1 = 2$, $2 < \sigma_2,\ldots,\sigma_{d-1} < 7$;

(1DNLS-d) The model (31) for $d \leqslant 6$, in the very weakly transverse regime, with $\mu \ll 1$, $1 < \sigma_1,\ldots,\sigma_{d-1} < 7$.

Remark 2.13. There are other interesting regimes for (5) and (22) especially for their relation with the modified KdV equation and two-dimensional NLS equation respectively. These will be discussed in remarks 4.6 and 4.9 respectively.

3. Galerkin averaging

In the next section we will show that for large N the dynamics of both ETL and KG lattices can be approximated by an infinite dimensional Hamiltonian, which can be written as the sum of an integrable part and a non-integrable perturbation. For this kind of systems it is often possible to analyze the dynamics taking into account only the leading terms of the integrable part and the 'average effect' of the perturbation, which describe the relevant qualitative behaviour of the system for a sufficiently long time-scale.

In this section we prove an abstract averaging theorem whose assumptions are satisfied by the systems we have introduced in section 2. This is the crucial technical result that allows to rigorously approximate the infinite dimensional Hamiltonian system with the leading terms of the integrable part and the average of the perturbation, up to the time-scales we are interested in. Since the average has to be computed along the solutions of the unperturbed system (see (47) below), the vector field of the averaged perturbation commutes with the vector field of the unperturbed system, thus resulting in a system in normal form.

The idea of its proof (following [Bam05, BP06, Pas19]) is to make a Galerkin cutoff, namely to approximate the original infinite dimensional system by a finite dimensional one, to put in normal form the cutoffed system, and then to choose the dimension of the cutoffed system in such a way that the error due to the Galerkin cutoff and the error due to the truncation in the normalization procedure are of the same order of magnitude. The system one gets is composed by a part which is in normal form, and by a remainder which is a smaller singular perturbation.

If we neglect the remainder, we obtain a system whose solutions are approximate solutions of the original system. In section 6 we will show how to control the error with respect to a true solution of the original system.

This section is divided in two parts. In the first part we introduce the analytic setting we are working with. This includes the definition functional setting of the problem and the average theorem 3.3. In the second part we give a concise proof of the average theorem, deferring the proof of the technical lemma 3.6 to appendix A.

3.1. An averaging theorem

To define the function spaces we are working with, we introduce a topology in the phase space. This is conveniently done in terms of Fourier coefficients.

Definition 3.1. Fix two constants $\rho \geqslant 0$ and $s \geqslant 0$. We will denote by $\ell_{\rho,s}^2$ the Hilbert space of complex sequences $v = (v_n)_{n \in \mathbb{Z}^2 \setminus \{0\}}$ with obvious vector space structure and with scalar product

$$\langle v, w \rangle_{\rho,s} := \sum_{n \in \mathbb{Z}^2 \setminus \{0\}} \overline{v_n} w_n \, \mathrm{e}^{2\rho|n|} |n|^{2s}. \tag{33}$$

and such that

$$\|v\|_{\rho,s}^2 := \langle v, v \rangle_{\rho,s} = \sum_{n \in \mathbb{Z}^2 \setminus \{0\}} |v_n|^2 \, \mathrm{e}^{2\rho|n|} |n|^{2s} \tag{34}$$

is finite. We will denote by ℓ^2 the space $\ell_{0,0}^2$.

We will identify a two-periodic function f with the sequence of its Fourier coefficients $(\hat{f}_n)_{n\in\mathbb{Z}^2\setminus\{0\}}$,

$$f(y) = \frac{1}{2}\sum_{n\in\mathbb{Z}^2}\hat{f}_n\, e^{i\pi n\cdot y},$$

and, with a small abuse of notation, we will say that $f \in \ell^2_{\rho,s}$ if the sequence of its Fourier coefficients belong to $\ell^2_{\rho,s}$.

Now fix $\rho \geqslant 0$ and $s \geqslant 1$, and consider the scale of Hilbert spaces $\mathcal{H}^{\rho,s} := \ell^2_{\rho,s} \times \ell^2_{\rho,s} \ni \zeta = (\xi, \eta)$, endowed with one of the following symplectic forms:

$$\Omega_1 := \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \qquad \Omega_2 := \begin{pmatrix} -\partial_{y_1}^{-1} & 0 \\ 0 & \partial_{y_1}^{-1} \end{pmatrix}. \tag{35}$$

Observe that $\Omega_\Gamma : \mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s+\Gamma-1}$ ($\Gamma = 1, 2$) is a well-defined operator. Moreover, Ω_2 is well-defined on the space of functions with zero-average with respect to the y_1-variable, i.e. on those functions $\zeta(y_1, y_2)$ such that for every y_2 we have $\int_{-1}^{1} \zeta(y_1, y_2)\mathrm{d}y_1 = 0$.

If we fix $\Gamma \in \{1, 2\}$, s and $U_s \subset \ell^2_{\rho,s}$ open, we define the gradient of $\mathsf{K} \in C^\infty(U_s, \mathbb{R})$ with respect to $\xi \in \ell^2_{\rho,s}$ as the unique function *s.t.*

$$\langle \nabla_\xi \mathsf{K}, h \rangle = \mathrm{d}_\xi \mathsf{K}(h), \quad \forall\, h \in \ell^2_{\rho,s}.$$

Similarly, for an open set $\mathcal{U}_s \subset \mathcal{H}^{\rho,s}$ the Hamiltonian vector field of the Hamiltonian function $\mathsf{H} \in C^\infty(\mathcal{U}_s, \mathbb{R})$ is given by

$$X_{\mathsf{H}}(\zeta) = \Omega_\Gamma^{-1}\nabla_\zeta \mathsf{H}(\zeta).$$

The open ball of radius R and center 0 in $\ell^2_{\rho,s}$ will be denoted by $B_{\rho,s}(R)$; we write $\mathcal{B}_{\rho,s}(R) := B_{\rho,s}(R) \times B_{\rho,s}(R) \subset \mathcal{H}^{\rho,s}$.

Now, we introduce the Fourier projection operators $\hat{\pi}_j : \ell^2_{\rho,s} \to \ell^2_{\rho,s}$

$$\hat{\pi}_j((v_n)_{n\in\mathbb{Z}^2\setminus\{0\}}) := \begin{cases} v_n & \text{if} \quad j-1 \leqslant |n| < j \\ 0 & \text{otherwise} \end{cases}, \quad j \geqslant 1, \tag{36}$$

the operators $\pi_j : \mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}$

$$\pi_j((\zeta_n)_{n\in\mathbb{Z}^2\setminus\{0\}}) := \begin{cases} \zeta_n & \text{if} \quad j-1 \leqslant |n| < j \\ 0 & \text{otherwise} \end{cases}, \quad j \geqslant 1, \tag{37}$$

and the operators $\Pi_M : \mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}$

$$\Pi_M((\zeta_n)_{n\in\mathbb{Z}^2\setminus\{0\}}) := \begin{cases} \zeta_n & \text{if} \quad |n| \leqslant M \\ 0 & \text{otherwise} \end{cases}, \quad M \geqslant 0. \tag{38}$$

Last, we define the operator $\overline{\pi}_j := id - \pi_j$ that will be used in Appendix C and Appendix E.

Lemma 3.2. *The projection operators defined in* (37) *and* (38) *satisfy the following properties for any* $\zeta \in \mathcal{H}^{\rho,s}$:

(a) *For any $j \geqslant 0$*

$$\zeta = \sum_{j \geqslant 0} \pi_j \zeta;$$

(b) *For any $j \geqslant 0$*

$$\|\Pi_M \zeta\|_{\mathcal{H}^{\rho,s}} \leqslant \|\zeta\|_{\mathcal{H}^{\rho,s}};$$

(c) *The following equality holds*

$$\|\zeta\|_{\mathcal{H}^{\rho,s}} = \left\| \left[\sum_{j \in \mathbb{N}} j^{2s} |\pi_j \zeta|^2 \right]^{1/2} \right\|_{\mathcal{H}^{\rho,0}} \tag{39}$$

where $|\zeta|$, for $\zeta \in \mathcal{H}^{\rho,s}$ is the element $|\zeta| \in \mathcal{H}^{\rho,s}$ whose nth element is

$$|\zeta|_n := (|\xi_n|, |\eta_n|)$$

and $(\zeta^\alpha)_n := (\xi_n^\alpha, \eta_n^\alpha)$.

Now we consider a Hamiltonian system of the form

$$\mathsf{H} = \mathsf{h}_0 + \delta \mathsf{F}, \tag{40}$$

where we assume that

(PER) h_0 generates a linear periodic flow $\Phi_{\mathsf{h}_0}^\tau$ with period T,

$$\Phi_{\mathsf{h}_0}^{\tau+T} = \Phi_{\mathsf{h}_0}^\tau \quad \forall \tau,$$

which is analytic as a map from $\mathcal{H}^{\rho,s}$ into itself for any $s \geqslant 1$. Furthermore, the flow is an isometry for any $s \geqslant 1$.

(INV) For any $s \geqslant 1$, $\Phi_{\mathsf{h}_0}^\tau$ leaves invariant the space $\Pi_j \mathcal{H}^{\rho,s}$ for any $j \geqslant 0$. Furthermore, for any $j \geqslant 0$

$$\pi_j \circ \Phi_{\mathsf{h}_0}^\tau = \Phi_{\mathsf{h}_0}^\tau \circ \pi_j.$$

Next, we assume that there exists $\nu \in [0, 1]$ such that the vector field of F admits an asymptotic expansion in δ of the form

$$\mathsf{F} \sim \sum_{j \geqslant 1} \delta^{\nu(j-1)} \mathsf{F}_j, \tag{41}$$

$$X_{\mathsf{F}} \sim \sum_{j \geqslant 1} \delta^{\nu(j-1)} X_{\mathsf{F}_j}, \tag{42}$$

and that the following property is satisfied

(HVF) There exists $R^* > 0$ such that for any $j \geqslant 1$
  - X_{F_j} is analytic from $\mathcal{B}_{\rho, s+2j+\Gamma}(R^*)$ to $\mathcal{H}^{\rho,s}$.
    Moreover, for any $r \geqslant 1$ we have that
  - $X_{\mathsf{F} - \sum_{j=1}^r \delta^{\nu(j-1)} \mathsf{F}_j}$ is analytic from $\mathcal{B}_{\rho, s+2(r+1)+\Gamma}(R^*)$ to $\mathcal{H}^{\rho,s}$.

The main result of this section is the following theorem.

Theorem 3.3. *Fix $R > 0$, $s_1 \gg 1$. Consider* (40), *and assume* (PER), (INV) *and* (HVF). *Then $\exists s_0 > 0$ with the following properties: for any $s \geqslant s_1$ there exists $\delta_s \ll 1$ such that for any $\delta < \delta_s$ there exists $\mathcal{T}_\delta : \mathcal{B}_{\rho,s}(R/2) \to \mathcal{B}_{\rho,s}(R)$ analytic canonical transformation such that*

$$\mathsf{H}_1 := \mathsf{H} \circ \mathcal{T}_\delta = \mathsf{h}_0 + \delta \mathcal{Z}_1 + \delta^{1+\nu} \mathcal{R}^{(1)}, \tag{43}$$

where $\mathcal{Z}_1$ is in normal form, namely

$$\{\mathcal{Z}_1, \mathsf{h}_0\} = 0, \tag{44}$$

and there exists a positive constant C_s' (that depends on s) such that

$$\sup_{\mathcal{B}_{\rho,s+s_0}(R/2)} \|X_{\mathcal{Z}_1}\|_{\mathcal{H}^{\rho,s}} \leqslant C_s',$$

$$\sup_{\mathcal{B}_{\rho,s+s_0}(R/2)} \|X_{\mathcal{R}^{(1)}}\|_{\mathcal{H}^{\rho,s}} \leqslant C_s', \tag{45}$$

$$\sup_{\mathcal{B}_{\rho,s}(R/2)} \|\mathcal{T}_\delta - \mathrm{i}d\|_{\mathcal{H}^{\rho,s}} \leqslant C_s' \delta. \tag{46}$$

In particular,

$$\mathcal{Z}_1(\zeta) = \langle \mathsf{F}_1 \rangle (\zeta), \tag{47}$$

where $\langle \mathsf{F}_1 \rangle (\zeta) := \int_0^T \mathsf{F}_1 \circ \Phi_{\mathsf{h}_0}^\tau (\zeta) \frac{\mathrm{d}\tau}{T}$.

3.2. Proof of the averaging theorem

The proof of theorem 3.3 is actually an application of the techniques used in [BP06, Pas19].

First notice that by assumption (INV) the Hamiltonian vector field of h_0 generates a continuous flow $\Phi_{\mathsf{h}_0}^\tau$ which leaves $\Pi_M \mathcal{H}^{\rho,s}$ invariant.

Now we set $\mathsf{H} = \mathsf{H}_{1,M} + \mathcal{R}_{1,M} + \mathcal{R}_1$, where

$$\mathsf{H}_{1,M} := \mathsf{h}_0 + \delta \, \mathsf{F}_{1,M}, \tag{48}$$

$$\mathsf{F}_{1,M} := \mathsf{F}_1 \circ \Pi_M, \tag{49}$$

and

$$\mathcal{R}_{1,M} := \mathsf{h}_0 + \delta \mathsf{F}_1 - \mathsf{H}_{1,M}, \tag{50}$$

$$\mathcal{R}_1 := \delta \, (\mathsf{F} - \mathsf{F}_1). \tag{51}$$

The system described by the Hamiltonian (48) is the one that we will put in normal form.

In the following we will use the notation $a \lesssim b$ to mean: there exists a positive constant K independent of M and R (but eventually on s), such that $a \leqslant Kb$. We exploit the following intermediate results:

Lemma 3.4. *For any $s \geqslant s_1$ there exists $R > 0$ such that $\forall\, \sigma > 0, M > 0$*

$$\sup_{\zeta \in \mathcal{B}_{\rho,s+\Gamma+\sigma+2}(R)} \|X_{\mathcal{R}_{1,M}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \lesssim \frac{\delta}{(M+1)^\sigma}, \tag{52}$$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s+\Gamma+4}(R)} \|X_{\mathcal{R}_1}(\zeta)\|_{\mathcal{H}^{\rho,s}} \lesssim \delta^{1+\nu}. \tag{53}$$

Proof. We recall that $\mathcal{R}_{1,M} = h_0 + \delta F_1 - H_{1,M} = \delta(F_1 - F_{1,M})$.

We first notice that $\|id - \Pi_M\|_{\mathcal{H}^{\rho,s+\sigma} \to \mathcal{H}^{\rho,s}} = (M+1)^{-\sigma}$: indeed, using (39) we obtain

$$
\left\| \sum_{j \geqslant M+1} \pi_j f \right\|_{\mathcal{H}^{\rho,s}} = \left\| \left[\sum_{j \geqslant M+1} |j^s \pi_j f|^2 \right]^{1/2} \right\|_{\mathcal{H}^{\rho,0}}
$$

$$
\leqslant (M+1)^{-\sigma} \left\| \left[\sum_{j \geqslant M+1} |j^{s+\sigma} \pi_j f|^2 \right]^{1/2} \right\|_{\mathcal{H}^{\rho,0}}
$$

$$
\leqslant (M+1)^{-\sigma} \|f\|_{\mathcal{H}^{\rho,s+\sigma}},
$$

whereas the inequality $\|id - \Pi_M\|_{\mathcal{H}^{\rho,s+\sigma} \to \mathcal{H}^{\rho,\sigma}} \geqslant (M+1)^{-\sigma}$ is obtained with a function which has non zero components only for $|j| = M + 1$, i.e. $f = \pi_{M+1} f$.

Inequality (52) follows from

$$
\sup_{\zeta \in \mathcal{B}_{\rho,s+\Gamma+2+\sigma}(R)} \|X_{\mathcal{R}_{1,M}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \lesssim \|dX_{\delta F_1}\|_{L^\infty(\mathcal{B}_{\rho,s+2+\Gamma}(R), \mathcal{H}^{\rho,s})} \|id - \Pi_M\|_{L^\infty(\mathcal{B}_{\rho,s+2+\Gamma+\sigma}(R), \mathcal{B}_{\rho,s+2+\Gamma}(R))}
$$

$$
\lesssim \delta(M+1)^{-\sigma},
$$

while estimate (53) is an immediate consequence of (HVF). $\qquad\square$

Lemma 3.5. *For any $s \geqslant s_1$*

$$
\sup_{\zeta \in \mathcal{B}_{\rho,s}(R^*)} \|X_{F_{1,M}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant K_{1,s}^{(F)} M^{2+\Gamma},
$$

where

$$
K_{1,s}^{(F)} := \sup_{\zeta \in \mathcal{B}_{\rho,s}(R^*)} \|X_{F_1}(\zeta)\|_{\mathcal{H}^{\rho,s-2-\Gamma}} < +\infty.
$$

Proof. Using (39) we have

$$
\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \left\| \sum_{h \leqslant M} \pi_h X_{F_{1,M}}(\zeta) \right\|_{\mathcal{H}^{\rho,s}}
$$

$$
= \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \left\| \left[\sum_{h \leqslant M} |h^s \pi_h X_{F_{1,M}}(\zeta)|^2 \right]^{1/2} \right\|_{\mathcal{H}^{\rho,0}} \tag{54}
$$

$$
\leqslant M^{2+\Gamma} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \left\| \left[\sum_{h \leqslant M} |h^{s-2-\Gamma} \pi_h X_{F_{1,M}}(\zeta)|^2 \right]^{1/2} \right\|_{\mathcal{H}^{\rho,0}} \tag{55}
$$

$$
\leqslant M^{2+\Gamma} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{F_{1,M}}(\zeta)\|_{\mathcal{H}^{\rho,s-2-\Gamma}} = K_{1,s}^{(F)} M^{2+\Gamma}, \tag{56}
$$

where the last quantity is finite for $R \leqslant R^*$ by property (HVF). $\qquad\square$

To normalize (48) we need to prove a reformulation of theorem 4.4 in [Bam99]. Here we report a statement of the result adapted to our context which is proved in appendix A.

Lemma 3.6. *Let* $s \geqslant s_1 + 2 + \Gamma$, $R > 0$, *and consider the system* (48). *Assume that* $\delta < \frac{1}{30^{1/\nu}}$, *and that*

$$12TK_{1,s}^{(\mathsf{F})}M^{2+\Gamma}\delta < R \tag{57}$$

where

$$K_{1,s}^{(\mathsf{F})} := \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{F}_1}(\zeta)\|_{\mathcal{H}^{\rho,s-2-\Gamma}}.$$

Then there exists an analytic canonical transformation $\mathcal{T}_{\delta,M}^{(0)} : \mathcal{B}_{\rho,s}(R/2) \to \mathcal{B}_{\rho,s}(R)$ *such that*

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R/2)} \|\mathcal{T}_{\delta,M}^{(0)}(\zeta) - \zeta\|_{\mathcal{H}^{\rho,s}} \leqslant 2TK_{1,s}^{(F)}M^{2+\Gamma}\delta, \tag{58}$$

and that puts (48) *in normal form up to a small remainder,*

$$\mathsf{H}_{1,M} \circ \mathcal{T}_{\delta,M}^{(0)} = \mathsf{h}_0 + \delta \mathsf{Z}_M^{(1)} + \delta^{1+\nu} \mathcal{R}_M^{(1)}, \tag{59}$$

with $Z_M^{(1)}$ *in normal form, namely* $\{h_{0,M}, Z_M^{(1)}\} = 0$, *and*

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R/2)} \|X_{\mathsf{Z}_M^{(1)}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant K_{1,s}^{(\mathsf{F})}M^{2+\Gamma} \tag{60}$$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R/2)} \|X_{\mathcal{R}_M^{(1)}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant 15K_{1,s}^{(\mathsf{F})}M^{2+\Gamma}. \tag{61}$$

Now we conclude with the proof of theorem 3.3.

Proof. If we define $\delta_s := \min\{\frac{1}{30^{1/\nu}}, \frac{R}{12\,T\,K_{1,s}^{(\mathsf{F})}M^{2+\Gamma}}\}$ and we choose

$$s_0 = \sigma + 2 + \Gamma,$$

$$\sigma \geqslant 2,$$

then the transformation $\mathcal{T}_\delta := \mathcal{T}_{\delta,M}^{(0)}$ defined by lemma 3.6 satisfies (43) because of (59).

Next, equation (44) follows from lemma 3.6, equation (45) follows from (60) and (61), while (46) is precisely (58). Finally, (47) can be deduced by applying lemma A.6 to $\mathsf{G} = \mathsf{F}_1$. $\qquad\square$

4. Applications to two-dimensional lattices

4.1. The KP regime for the ETL lattice

We want to study the behaviour of small amplitude solutions of (5) with initial data in which only one low-frequency Fourier mode is excited.

As a first step, we introduce an interpolating function $Q = Q(t, x)$ such that

(A1) $Q(t, j) = Q_j(t)$, for all $j \in \mathbb{Z}_{N_1,N_2}^2$;
(A2) Q is periodic with period $2N_1 + 1$ in the x_1-variable, and periodic with period $2N_2 + 1$ in the x_2-variable;

(A3) Q has zero average, $\int_{\left[-\left(N_1+\frac{1}{2}\right),N_1+\frac{1}{2}\right]\times\left[-\left(N_2+\frac{1}{2}\right),N_2+\frac{1}{2}\right]} Q(t,x)\,\mathrm{d}x = 0 \ \forall t$;

(A4) Q fulfills

$$\ddot{Q} = \Delta_1(Q + \alpha Q^2 + \beta Q^3), \tag{62}$$

$$\Delta_1 := 4\sinh^2\left(\frac{\partial_{x_1}}{2}\right) + 4\sinh^2\left(\frac{\partial_{x_2}}{2}\right). \tag{63}$$

One can check that the operator (63) is equivalent to (3) by using functional calculus. It is easy to verify that (62) is Hamiltonian with Hamiltonian function

$$\mathsf{H}_{\mathrm{ETL}}(Q,P) = \int_{\left[-\frac{1}{\mu},\frac{1}{\mu}\right]\times\left[-\frac{1}{\mu^2},\frac{1}{\mu^2}\right]} \left(\frac{-P\,\Delta_1 P + Q^2}{2} + \alpha\frac{Q^3}{3} + \beta\frac{Q^4}{4}\right)\mathrm{d}x, \tag{64}$$

where P is a periodic function which has zero average and is canonically conjugated to Q

The existence of such an interpolating function is obvious, indeed such function can be taken to be a Fourier polynomial with N terms. However, by construction, items (A1)–(A4) do not provide a *unique* interpolating function. Among all possible choices of interpolating functions it is convenient to choose the Fourier polynomial supported on $k \in [-N_1,N_1] \times [-N_2,N_2]$ (i.e. the analytic function supported in the lowest possible number of Fourier modes). Also let us emphasize that (A4) ensures that if (A1) is valid for the initial time $t = 0$, then is valid for all $t \in \mathbb{R}$.

We consider (62), with $\alpha \neq 0$, and we look for small amplitude solutions of the form

$$Q(t,x) = \mu^2\, q(\mu t, \mu x_1, \mu^2 x_2), \tag{65}$$

with μ as in (11). We introduce the rescaled time $\tau = \mu t$ and the rescaled space variables $y_1 = \mu x_1, y_2 = \mu^2 x_2$.

Plugging (65) into (62), leads to

$$q_{\tau\tau} = \frac{\Delta_{\mu,y_1}}{\mu^2}\left(q + \mu^2\alpha q^2 + \mu^4\beta q^3\right), \tag{66}$$

$$\Delta_{\mu,y_1} := 4\sinh^2\left(\frac{\mu\partial_{y_1}}{2}\right) + 4\sinh^2\left(\mu^2\frac{\partial_{y_2}}{2}\right), \tag{67}$$

which is a Hamiltonian PDE corresponding to the Hamiltonian functional,

$$\mathsf{K}_1(q,p) = \int_I \left(\frac{-p\Delta_{\mu,y_1}p}{2\mu^2} + \frac{q^2}{2} + \alpha\mu^2\frac{q^3}{3} + \beta\mu^4\frac{q^4}{4}\right)\mathrm{d}y, \tag{68}$$

where

$$I := [-1,1]^2, \tag{69}$$

and p is the variable canonically conjugated to q.

Now, observe that the operator Δ_{μ,y_1} admits the following asymptotic expansion up to terms of order $\mathcal{O}(\mu^4)$,

$$\frac{\Delta_{\mu,y_1}}{\mu^2} \sim \partial_{y_1}^2 + \mu^2\partial_{y_2}^2 + \frac{\mu^2}{12}\partial_{y_1}^4 + \mathcal{O}(\mu^4), \tag{70}$$

therefore the Hamiltonian (68) admits the following asymptotic expansion

$$\mathsf{K}_1(q,p) \sim \hat{\mathsf{h}}_0(q,p) + \mu^2 \hat{\mathsf{F}}_1(q,p) + \mu^4 \hat{\mathcal{R}}(q,p), \tag{71}$$

$$\hat{\mathsf{h}}_0(q,p) = \int_I \left[\frac{p(-\partial_{y_1}^2 p) + q^2}{2} \right] \mathrm{d}y, \tag{72}$$

$$\hat{\mathsf{F}}_1(q,p) = \int_I \left(-\frac{p\partial_{y_1}^4 p}{24} - \frac{p\partial_{y_2}^2 p}{2} + \alpha \frac{q^3}{3} \right) \mathrm{d}y. \tag{73}$$

Following the approach of [BP06], we can introduce the following non-canonical change of coordinates

$$\xi := \frac{1}{\sqrt{2}}(q + \partial_{y_1} p), \qquad \eta := \frac{1}{\sqrt{2}}(q - \partial_{y_1} p). \tag{74}$$

which transforms the Poisson tensor into

$$J = \partial_{y_1} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{75}$$

and Hamilton equations associated to a Hamiltonian K are

$$\partial_\tau \xi = -\partial_{y_1} \frac{\delta \mathsf{K}}{\delta \xi},$$

$$\partial_\tau \eta = \partial_{y_1} \frac{\delta \mathsf{K}}{\delta \eta}.$$

Remark 4.1. By the explicit expression of the Poisson tensor (75) we can compute straightforwardly Casimir invariants associated to J, which are

$$C(\xi, \eta) = A(y_2) + B(y_2) \int_{-1}^{1} \xi(\tau, y_1, y_2) \mathrm{d}y_1 + C(y_2) \int_{-1}^{1} \eta(\tau, y_1, y_2) \mathrm{d}y_1, \tag{76}$$

where A, B and C are arbitrary functions of y_2.

Since Casimir invariants are constants of motion, we can restrict our analysis on the subspace defined by

$$\int_{-1}^{1} [\xi(\tau, y_1, y_2) - \eta(\tau, y_1, y_2)] \mathrm{d}y_1 = 0, \quad \forall \tau \in \mathbb{R}, \quad |y_2| \leqslant 1. \tag{77}$$

However, by recalling (74) one sees that (77) is equivalent to

$$\int_{-1}^{1} \partial_{y_1} p(\tau, y_1, y_2) \mathrm{d}y_1 = 0, \quad \forall \tau \in \mathbb{R}, \quad |y_2| \leqslant 1, \tag{78}$$

which is trivially true due to periodic boundary conditions.

In the new coordinates the Hamiltonian takes the form

$$\mathsf{K}_1(\xi, \eta) \sim \mathsf{h}_0(\xi, \eta) + \mu^2 \mathsf{F}_1(\xi, \eta) + \mu^4 \mathcal{R}(\xi, \eta), \tag{79}$$

$$\mathsf{h}_0(\xi, \eta) = \int_I \frac{\xi^2 + \eta^2}{2} \mathrm{d}y, \tag{80}$$

$$\mathsf{F}_1(\xi, \eta) = \int_I \left(-\frac{[\partial_{y_1}(\xi - \eta)]^2}{48} + \frac{[\partial_{y_2}\partial_{y_1}^{-1}(\xi - \eta)]^2}{4} + \alpha \frac{(\xi + \eta)^3}{3 \cdot 2^{3/2}} \right) dy, \quad (81)$$

where (81) is well defined because of (92).

Now we apply the averaging theorem 3.3 to the Hamiltonian (79), with $\delta = \mu^2$, $\nu = 1$. Observe that the equations of motion of h_0 have the following simple form:

$$\begin{cases} \xi_\tau = -\partial_{y_1}\xi \\ \eta_\tau = \partial_{y_1}\eta \end{cases} ; \quad \begin{cases} \xi(\tau, y) = \xi_0(y_1 - \tau, y_2) \\ \eta(\tau, y) = \eta_0(y_1 + \tau, y_2) \end{cases} . \quad (82)$$

Proposition 4.2. *The average of F_1 in (79) with respect to the flow of h_0 in (79) is given by*

$$\langle \mathsf{F}_1 \rangle (\xi, \eta) = \int_I \left(-\frac{(\partial_{y_1}\xi)^2 + (\partial_{y_1}\eta)^2}{48} + \frac{(\partial_{y_2}\partial_{y_1}^{-1}\xi)^2 + (\partial_{y_2}\partial_{y_1}^{-1}\eta)^2}{4} \right) dy$$

$$+ \frac{\alpha}{3 \times 2^{3/2}}([\xi^3] + [\eta^3]) \quad (83)$$

where we denote by $[f^j]$ the average $\int_I f^j(y)\frac{dy}{4}$.

Proof. For the computation of $\langle \mathsf{F}_1 \rangle(\xi, \eta)$ one can exchange the order of the integrations in y and s. To compute averages we use the following elementary facts:

(a) Let $u, v \in L^2([-1, 1])$, then

$$\int_{-1}^1 dy \int_{-1}^1 ds \, u(y \pm s)v(y \mp s) = \int_{-1}^1 u(y)dy \int_{-1}^1 v(y)dy;$$

(b) Let $u \in L^1([-1, 1])$, then

$$\frac{1}{2}\int_{-1}^1 ds \int_{-1}^1 dy \, u(y \pm s) = \int_{-1}^1 u(x)dx.$$

Since we assume our functions to be periodic, then also $\int_I \partial_{y_r} u(y)dy = 0$, $r = 1, 2$, and hence only terms with only ξ or η are not cancelled by the average procedure. $\square$

Corollary 4.3. *The equations of motion associated to $\mathsf{h}_0(\xi, \eta) + \mu^2 \langle \mathsf{F}_1 \rangle (\xi, \eta)$ are given by*

$$\begin{cases} \xi_\tau = -\partial_{y_1}\xi - \frac{\mu^2}{24}\partial_{y_1}^3\xi - \frac{\mu^2}{2}\partial_{y_1}^{-1}\partial_{y_2}^2\xi - \frac{\alpha\mu^2}{2\sqrt{2}}\partial_{y_1}(\xi^2) \\ \eta_\tau = \partial_{y_1}\eta + \frac{\mu^2}{2}\partial_{y_1}^{-1}\partial_{y_2}^2\eta + \frac{\mu^2}{24}\partial_{y_1}^3\eta + \frac{\alpha\mu^2}{2\sqrt{2}}\partial_{y_1}(\eta^2) \end{cases} . \quad (84)$$

More explicitly, we observe that (84) is a system of two uncoupled KP equations on a two-dimensional torus in translating frames.

4.2. The KdV regime for the ETL lattice

For this regime we consider (62), with $\alpha \neq 0$, and we look for small amplitude solutions of the form

$$Q(t, x) = \mu^2 q(\mu t, \mu x_1, \mu^\sigma x_2), \quad (85)$$

where $q : \mathbb{R} \times \mathbb{T}^2 \to \mathbb{R}$ is a periodic function and μ, $\sigma > 2$ are defined in (11) and (12). We introduce the rescaled variables $\tau = \mu t$, $y_1 = \mu x_1$, $y_2 = \mu^\sigma x_2$, and we denote I is as in (69). Plugging (85) into (62), we get

$$q_{\tau\tau} = \frac{\Delta_{\mu,y_1,\sigma}}{\mu^2} \left(q + \mu^2 \alpha q^2 + \mu^4 \beta q^3 \right),$$

$$\Delta_{\mu,y_1,\sigma} := 4 \sinh^2 \left(\frac{\mu \partial_{y_1}}{2} \right) + 4 \sinh^2 \left(\mu^\sigma \frac{\partial_{y_2}}{2} \right), \tag{86}$$

which is a Hamiltonian PDE corresponding to the Hamiltonian functional

$$\mathsf{K}_2(q, p) = \int_I \left(\frac{-p \Delta_{\mu,y_1,\sigma} p}{2\mu^2} + \frac{q^2}{2} + \alpha \mu^2 \frac{q^3}{3} + \beta \mu^4 \frac{q^4}{4} \right) \mathrm{d}y, \tag{87}$$

and p is the variable canonically conjugated to q.

Now, observe that the operator $\Delta_{\mu,y_1,\sigma}$ admits the following asymptotic expansion,

$$\frac{\Delta_{\mu,y_1,\sigma}}{\mu^2} \sim \partial_{y_1}^2 + \frac{\mu^2}{12} \partial_{y_1}^4 + \mu^{2(\sigma-1)} \partial_{y_2}^2 + \frac{\mu^{2(2\sigma-1)}}{12} \partial_{y_2}^4$$

$$+ \sum_{m \geqslant 2} c_m \left(\mu^{2m} \partial_{y_1}^{2(m+1)} + \mu^{2[(m+1)\sigma-1]} \partial_{y_2}^{2(m+1)} \right), \tag{88}$$

$$c_m := \frac{2}{(2m+2)!},$$

and by recalling that $\sigma > 2$ we have

$$\frac{\Delta_{\mu,y_1,\sigma}}{\mu^2} \sim \partial_{y_1}^2 + \frac{\mu^2}{12} \partial_{y_1}^4 + \mathcal{O}(\mu^{\min(2\sigma-2,4)}). \tag{89}$$

Therefore the Hamiltonian (87) admits the following asymptotic expansion

$$\mathsf{K}_2(q, p) \sim \hat{\mathsf{h}}_0(q, p) + \mu^2 \hat{\mathsf{F}}_2(q, p) + \mu^{\min(2\sigma-2,4)} \hat{\mathcal{R}}(q, p), \tag{90}$$

$$\hat{\mathsf{h}}_0(q, p) = \int_I \frac{-p(\partial_{y_1}^2 p) + q^2}{2} \, \mathrm{d}y, \qquad \hat{\mathsf{F}}_2(q, p) = \int_I \left(-\frac{p \partial_{y_1}^4 p}{24} + \alpha \frac{q^3}{3} \right) \mathrm{d}y. \tag{91}$$

Note that the nonlinearity of degree 4 does not affect the Hamiltonian up to order $\mathcal{O}(\mu^4)$.

By exploiting again the non-canonical change of coordinates $(q, p) \mapsto (\xi, \eta)$ introduced in (74) and the Poisson tensor (75), and

$$\int_{-1}^1 [\xi(\tau, y_1, y_2) - \eta(\tau, y_1, y_2)] \, \mathrm{d}y_1 = 0, \quad \forall \, \tau \in \mathbb{R}, \quad |y_2| \leqslant 1, \tag{92}$$

we obtain

$$\mathsf{K}_2(\xi, \eta) \sim \mathsf{h}_0(\xi, \eta) + \mu^2 \mathsf{F}_2(\xi, \eta) + \mu^{\min(2\sigma-2,4)} \mathcal{R}(\xi, \eta), \tag{93}$$

$$\mathsf{h}_0(\xi, \eta) = \int_I \frac{\xi^2 + \eta^2}{2} \mathrm{d}y, \qquad \mathsf{F}_2(\xi, \eta) = \int_I \left(-\frac{[\partial_{y_1}(\xi - \eta)]^2}{48} + \alpha \frac{(\xi + \eta)^3}{3 \cdot 2^{3/2}} \right) \mathrm{d}y. \tag{94}$$

Now we apply the averaging theorem 3.3 to the Hamiltonian (93), with $\delta = \mu^2$, $\nu = \min(\sigma - 2, 1)$.

Proposition 4.4. *The average of* F_2 *in* (93) *with respect to the flow of* h_0 *in* (94) *is given by*

$$\langle \mathsf{F}_2 \rangle (\xi, \eta) = -\int_I \frac{(\partial_{y_1} \xi)^2 + (\partial_{y_1} \eta)^2}{48} \, \mathrm{d}y + \frac{\alpha}{3 \times 2^{3/2}} ([\xi^3] + [\eta^3]), \tag{95}$$

where we denote by $[f^j]$ *the average* $\int_I f^j(y) \frac{\mathrm{d}y}{4}$.

Corollary 4.5. *The equations of motion associated to* $\mathsf{h}_0(\xi, \eta) + \mu^2 \langle \mathsf{F}_1 \rangle (\xi, \eta)$ *are given by*

$$\begin{cases} \xi_\tau = -\partial_{y_1} \xi - \dfrac{\mu^2}{24} \partial_{y_1}^3 \xi - \dfrac{\mu^2 \alpha}{2\sqrt{2}} \partial_{y_1} (\xi^2) \\[2mm] \eta_\tau = \partial_{y_1} \eta + \dfrac{\mu^2}{24} \partial_{y_1}^3 \eta + \dfrac{\mu^2 \alpha}{2\sqrt{2}} \partial_{y_1} (\eta^2) \end{cases}. \tag{96}$$

The latter is a system of two uncoupled KdV equations in translating frames with respect to the y_1-direction, for each fixed value of the coordinate y_2.

Remark 4.6. One can also study the β model (namely, (62) with $\alpha = 0$, $\beta \neq 0$) in the following regime,

(mKdV) The β model in the very weakly transverse regime, $Q(t, x) = \mu q(\mu t, \mu x_1, \mu^\sigma x_2)$, where $\mu \ll 1$, $\sigma > 2$.

Let us introduce again the rescaled variables $\tau = \mu t$, $y_1 = \mu x_1$, $y_2 = \mu^\sigma x_2$, and the domain I as in (69); plugging the ansatz for Q into (62), we get

$$q_{\tau\tau} = \frac{\Delta_{\mu, y_1, \sigma}}{\mu^2} \left(q + \mu^2 \beta q^3 \right), \tag{97}$$

where $\Delta_{\mu, y_1, \sigma}$ is the operator introduced in (86). Equation (97) is a Hamiltonian PDE with the following corresponding Hamiltonian,

$$\mathsf{K}_3(q, p) = \int_I \left(\frac{-p \Delta_{\mu, y_1, \sigma} p}{2\mu^2} + \frac{q^2}{2} + \beta \mu^2 \frac{q^4}{4} \right) \mathrm{d}y, \tag{98}$$

where p is the variable canonically conjugated to q. Recalling that (77) holds true, we exploit again the non-canonical change of coordinates (74) and the Poisson tensor (75), obtaining that

$$\mathsf{K}_3(\xi, \eta) \sim \mathsf{h}_0(\xi, \eta) + \mu^2 \mathsf{F}_3(\xi, \eta) + \mu^{\min(2\sigma - 2, 4)} \mathcal{R}(\xi, \eta), \tag{99}$$

where h_0 is the same as in (94), while

$$\mathsf{F}_3(\xi, \eta) = \int_I \left(-\frac{[\partial_{y_1} (\xi - \eta)]^2}{48} + \beta \frac{(\xi + \eta)^4}{2^4} \right) \mathrm{d}y. \tag{100}$$

Applying theorem 3.3 to the Hamiltonian (99) with $\delta = \mu^2$ and $\nu = \min(\sigma - 2, 1)$, we get that the equations of motion associated to $\mathsf{h}_0(\xi, \eta) + \mu^2 \langle \mathsf{F}_3 \rangle (\xi, \eta)$ are given by

$$\begin{cases} \xi_\tau = -\left(1 + \dfrac{3}{4}[\eta^2]\right) \partial_{y_1} \xi - \dfrac{\mu^2}{24} \partial_{y_1}^3 \xi - \dfrac{\mu^2 \beta}{4} \partial_{y_1} (\xi^3) \\[2mm] \eta_\tau = \left(1 + \dfrac{3}{4}[\xi^2]\right) \partial_{y_1} \eta + \dfrac{\mu^2}{24} \partial_{y_1}^3 \eta + \dfrac{\mu^2 \beta}{4} \partial_{y_1} (\eta^3) \end{cases} \tag{101}$$

which is a system of two uncoupled mKdV equations in translating frames with respect to the y_1-direction. The integrability properties of the mKdV equation and the existence of Birkhoff coordinates for this model have been proved in [KST08].

4.3. The one-dimensional NLS regime for the KG lattice

We want to study small amplitude solutions of (22), with initial data in which only one low-frequency Fourier mode is excited.

Analogously to the procedure of the previous sections, the first step is to introduce an interpolating function $Q = Q(t, x)$ such that

(B1) $Q(t, j) = Q_j(t)$, for all $j \in \mathbb{Z}^2_{N_1, N_2}$;
(B2) Q is periodic with period $2N_1 + 1$ in the x_1-variable, and periodic with period $2N_2 + 1$ in the x_2-variable;
(B3) Q fulfills

$$\ddot{Q} = \Delta_1 Q - Q - \beta Q^{2p+1}, \tag{102}$$

where Δ_1 is the operator defined in (63) (recall that we also assumed $m = 1$ in (21)).

It is easy to verify that (102) is Hamiltonian with Hamiltonian function

$$\mathsf{H}_{\mathrm{KG}}(Q, P) = \int_{\left[-\frac{1}{\mu}, \frac{1}{\mu}\right] \times \left[-\frac{1}{\mu^\sigma}, \frac{1}{\mu^\sigma}\right]} \left(\frac{P^2}{2} + \frac{Q^2}{2} - \frac{Q\Delta_1 Q}{2} + \beta \frac{Q^{2p+2}}{2p+2} \right) \mathrm{d}x, \tag{103}$$

where P is a periodic function and is canonically conjugated to Q.

Starting from the Hamiltonian (20), where $p = 1$, we look for small amplitude solutions of the form

$$Q(t, x) = \mu q(\mu^2 t, \mu x_1, \mu^\sigma x_2) \tag{104}$$

where $q : \mathbb{R} \times \mathbb{T}^2 \to \mathbb{R}$ is a periodic function and σ, μ are defined respectively in (12) and (11).

We introduce the rescaled variables $\tau = \mu^2 t$, $y_1 = \mu x_1$ and $y_2 = \mu^\sigma x_2$, and we define I as in (69). The Hamiltonian (20) in the rescaled variables is given by

$$\mathsf{K}_4(q, p) = \int_I \left(\frac{p^2}{2} + \frac{q^2}{2} - \frac{q\Delta_{\mu, y_1, \sigma} q}{2} + \beta \mu^2 \frac{q^4}{4} \right) \mathrm{d}y, \tag{105}$$

with the operator $\Delta_{\mu, y_1, \sigma}$ as in (86), and p is the variable canonically conjugated to q. The corresponding equation of motion is given by

$$q_{\tau\tau} = -q + \Delta_{\mu, y_1, \sigma} q - \beta \mu^2 q^3. \tag{106}$$

Recalling (89), we have that the Hamiltonian (105) admits the following asymptotic expansion

$$\mathsf{K}_4(q, p) \sim \hat{\mathsf{h}}_4(q, p) + \mu^2 \hat{\mathsf{F}}_4(q, p) + \mu^{\min(2\sigma, 4)} \hat{\mathcal{R}}(q, p), \tag{107}$$

$$\hat{\mathsf{h}}_4(q, p) = \int_I \frac{p^2 + q^2}{2} \, \mathrm{d}y, \qquad \hat{\mathsf{F}}_4(q, p) = \int_I \left(-\frac{q\partial_{y_1}^2 q}{2} + \beta \frac{q^4}{4} \right) \mathrm{d}y, \tag{108}$$

and the equation of motion associated to $\hat{\mathsf{h}}_4 + \mu^2 \hat{\mathsf{F}}_4$ is given by the following cubic one-dimensional NLKG equation,

$$q_{\tau\tau} = -(q - \mu^2 \partial_{y_1}^2 q) - \mu^2 \beta q^3. \tag{109}$$

We now exploit the change of coordinates $(q, p) \mapsto (\psi, \bar{\psi})$ given by

$$\psi = \frac{1}{\sqrt{2}}(q - \mathrm{i}p), \tag{110}$$

therefore the inverse change of coordinates is given by

$$q = \frac{1}{\sqrt{2}}(\psi + \bar\psi), \qquad p = \frac{i}{\sqrt{2}}(\psi - \bar\psi), \tag{111}$$

while the symplectic form is given by $-i d\psi \wedge d\bar\psi$. With this change of variables the Hamiltonian takes the form

$$K_4(\psi, \bar\psi) \sim h_4(\psi, \bar\psi) + \mu^2 F_4(\psi, \bar\psi) + \mu^{\min(2\sigma, 4)} \mathcal{R}(\psi, \bar\psi), \tag{112}$$

$$h_4(\psi, \bar\psi) = \int_I \psi \bar\psi \, dy,$$

$$F_4(\psi, \bar\psi) = \int_I \left(-\frac{(\psi + \bar\psi)[-\partial_{y_1}^2(\psi + \bar\psi)]}{4} + \beta \frac{(\psi + \bar\psi)^4}{16} \right) dy. \tag{113}$$

Now we apply the averaging theorem 3.3 to the Hamiltonian (112), with $\delta = \mu^2$, $\nu = \min(\sigma - 1, 1)$. Observe that h_4 generates a periodic flow,

$$-i\partial_\tau \psi = \psi; \quad \psi(\tau, y) = e^{i\tau} \psi_0(y). \tag{114}$$

Proposition 4.7. *The average of F_4 in (112) with respect to the flow of h_4 (108) is given by*

$$\langle F_4 \rangle (\psi, \bar\psi) = \int_I \frac{\bar\psi(-\partial_{y_1}^2 \psi)}{2} \, dy + \frac{3\beta}{8} \int_I |\psi|^4 \, dy. \tag{115}$$

Corollary 4.8. *The equations of motion associated to $h_4(\psi, \bar\psi) + \mu^2 \langle F_4 \rangle (\psi, \bar\psi)$ are given by a cubic one-dimensional NLS equation for each fixed value of y_2,*

$$-i\psi_\tau = \psi - \mu^2 \partial_{y_1}^2 \psi + \mu^2 \frac{3\beta}{4} |\psi|^2 \psi. \tag{116}$$

Remark 4.9. If we apply the above argument to the Hamiltonian (20) in the regime

(2-D NLS) The scalar model (20) with $m = 1$, $p = 1$ and $Q(t, x) = \mu q(\mu^2 t, \mu x)$, where $\mu \ll 1$ and $\sigma = 1$,
  We can obtain as a normal form equation the cubic NLS equation

$$-i\psi_\tau = \psi - \mu^2 \Delta \psi + \mu^2 \frac{3\beta}{4} |\psi|^2 \psi. \tag{117}$$

5. Dynamics of the normal form equation

5.1. The KP equation

In this section we recall some known facts on the dynamics of the KP equation on the two-dimensional torus

$$\xi_\tau = -\frac{1}{24} \partial_{y_1}^3 \xi - \frac{1}{2} \partial_{y_1}^{-1} \partial_{y_2}^2 \xi - \frac{\alpha}{2\sqrt{2}} \partial_{y_1}(\xi^2), \quad \alpha = \pm 1, \quad y \in \mathbb{T}^2 := \mathbb{R}^2/(2\pi\mathbb{Z})^2. \tag{118}$$

The KP equation has been introduced in order to describe weakly-transverse solutions of the water waves equations; it has been considered as a two-dimensional analogue of the KdV equation, since also the KP equation admits an infinite number of constants of motions [CL87,

CLL83, LC82]. It is customary to refer to (118) as KP-I equation when $\alpha = -1$, and as KP-II equation when $\alpha = 1$.

The global-well posedness for the KP-II equation on the two-dimensional torus has been discussed by Bourgain in [Bou93]. The main point of the result by Bourgain consists in extending the local well-posedness result to a global one, even though the L^2-norm is the only constant of motion for the KP-II equation that allows an *a priori* bound for the solution (see theorems 8.10 and 8.12 in [Bou93]).

Theorem 5.1. *Consider* (118) *with* $\alpha = 1$.

Let $\rho \geqslant 0$ *and* $s \geqslant 0$, *and assume that the initial datum* $\xi(0, \cdot, \cdot) = \xi_0 \in \ell^2_{\rho,s}$. *Then* (118) *is globally well-posed in* $\ell^2_{\rho,s}$. *Moreover, the* ℓ^2-norm of the solution is conserved,

$$\|\xi(t)\|_{\ell^2} = \|\xi_0\|_{\ell^2}, \tag{119}$$

while

$$\|\xi(t)\|_{\ell^2_{0,s}} \leqslant e^{C|t|} \|\xi_0\|_{\ell^2_{0,s}}, \tag{120}$$

where C depends on s.

Remark 5.2. As pointed out by Bourgain in section 10.2 of [Bou93], a global well-posedness result for sufficiently smooth solution of the KP-I equation (namely, (118) with $\alpha = -1$) on the two-dimensional torus can be obtained by generalizing the argument in [SJ87] for small data and by using the *a priori* bounds given by the constants of motion for the KP-I equation.

Last, we mention that for the KP equation the construction of action-angle/Birkhoff coordinates is still an open problem.

5.2. The KdV equation

In this section we recall some known facts on the dynamics of the KdV equation with periodic boundary conditions. The interested reader can find more detailed explanations and proofs in [2003].

Consider the KdV equation

$$\xi_\tau = -\frac{1}{24}\partial_{y_1}^3 \xi - \frac{\alpha}{2\sqrt{2}}\partial_{y_1}(\xi^2), \quad y_1 \in [0, 2]. \tag{121}$$

Through the Lax pair formulation of the evolution problem (121) one get that the periodic eigenvalues $(\lambda_n)_{n \in \mathbb{N}}$ of the Sturm–Liouville operator

$$L_\xi := -\partial_{y_1}^2 + 6\sqrt{2}\xi(\tau, y_1) \tag{122}$$

are conserved quantities under the evolution of the KdV equation (121). Moreover, if we define the gaps of the spectrum $\gamma_m := \lambda_{2m} - \lambda_{2m-1}$ ($m \geqslant 1$), it is well known that the squared spectral gaps $(\gamma_m^2)_{m \geqslant 1}$ form a complete set of constants of motion for (121).

The following relation between the sequence of the spectral gaps and the regularity of the corresponding solution to the KdV equation holds (see theorems 9–11 in [2008]; see also [Pös11])

Theorem 5.3. *Assume that* $\xi \in L^2$, *then* $\xi \in \ell^2_{0,s}$ *if and only if its spectral gaps satisfy*

$$\sum_{m \geqslant 1} m^{2s}|\gamma_m|^2 < +\infty.$$

Moreover if $\xi \in \ell^2_{\rho,s}$, then

$$\sum_{m \geqslant 1} m^{2s} e^{2\rho m} |\gamma_m|^2 < +\infty; \tag{123}$$

conversely, if (123) *holds, then $\xi \in \ell^2_{\rho',0}$ for some $\rho' > 0$.*

Kappeler and Pöschel constructed the following global Birkhoff coordinates (see theorem 1.1 in [2003]).

Theorem 5.4. *There exists a diffeomorphism $\Omega : L^2 \to \ell^2_{0,1/2} \times \ell^2_{0,1/2}$ such that:*

- Ω *is bijective, bianalytic and canonical;*
- *For each $s \geqslant 0$, the restriction of Ω to $\ell^2_{0,s}$, namely the map*

$$\Omega : \ell^2_{0,s} \to \ell^2_{0,s+1/2} \times \ell^2_{0,s+1/2}$$

  is bijective, bianalytic and canonical;
- *The coordinates $(x, y) \in \ell^2_{0,3/2} \times \ell^2_{0,3/2}$ are Birkhoff coordinates for the KdV equation, namely they form a set of canonically conjugated coordinates in which the Hamiltonian of the KdV equation* (121) *depends only on the action $I_m := \frac{x_m^2 + y_m^2}{2}$ ($m \geqslant 1$).*

The dynamics of the KdV equation (121) in terms of the variables (x, y) is trivial: it can be immediately seen that any solution is periodic, quasiperiodic or almost periodic, depending on the number of spectral gaps (equivalently, depending on the number of actions) initially different from zero.

5.3. The one-dimensional cubic NLS equation

In this section we recall some known facts on the dynamics of the one-dimensional cubic defocusing NLS equation with periodic boundary conditions. The interested reader can find more detailed explanations and proofs in [GKK14, Mol14].

Consider the cubic defocusing NLS equation

$$i\psi_\tau = -\partial_{y_1}^2 \psi + 2|\psi|^2 \psi, \quad y_1 \in \mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z}). \tag{124}$$

Equation (124) is a PDE admitting a Hamiltonian structure: indeed, we can set $\mathcal{H}^{\rho,s} = \ell^2_{\rho,s} \times \ell^2_{\rho,s}$ as the phase space with elements denoted by $\phi = (\phi_1, \phi_2)$, while the associated Poisson bracket and the Hamiltonian are given by

$$\{F, G\} := -i \int_{\mathbb{T}} \left(\partial_{\phi_1} F \partial_{\phi_2} G - \partial_{\phi_1} G \partial_{\phi_2} F \right) dy_1, \tag{125}$$

$$H_{\text{NLS}}(\phi_1, \phi_2) := \int_{\mathbb{T}} \left(\partial_{y_1} \phi_1 \partial_{y_1} \phi_2 + \phi_1^2 \phi_2^2 \right) dy_1. \tag{126}$$

The defocusing NLS equation (124) is obtained by restricting (126) to the invariant subspace of states of real type,

$$\mathcal{H}_r^{\rho,s} := \{\phi \in \mathcal{H}^{\rho,s} : \phi_2 = \bar{\phi}_1\}. \tag{127}$$

The above Hamiltonian (126) is well-defined on $\mathcal{H}^{\rho,s}$ with $s \geqslant 1$ and $\rho \geqslant 0$, while the initial value problem for the NLS equation (124) is well-posed on $\mathcal{H}^{0,0} = \ell^2 \times \ell^2$.

It is well known from the work by Zakharov and Shabat that the NLS equation (124) has a Lax pair, and that it admits infinitely many constants of motion in involution. More precisely, for any $\phi \in \mathcal{H}^{0,0}$ consider the Zakharov–Shabat operator

$$L(\phi) = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix} \partial_{y_1} + \begin{pmatrix} 0 & \phi_1 \\ \phi_2 & 0 \end{pmatrix}, \tag{128}$$

where we call ϕ the potential of the operator $L(\phi)$. The spectrum of $L(\phi)$ on the interval $[0, 2]$ with periodic boundary conditions is pure point, and it consists of the following sequence of periodic eigenvalues

$$\ldots < \lambda_{-1}^- \leqslant \lambda_{-1}^+ < \lambda_0^- \leqslant \lambda_0^+ < \lambda_1^- \leqslant \lambda_1^+ < \ldots, \tag{129}$$

where the quantities $\gamma_m := \lambda_m^+ - \lambda_m^-$ ($m \in \mathbb{Z}$) are called gap lengths. It has been proved that the squared spectral lengths $(\gamma_m^2)_{m \in \mathbb{Z}}$ form a complete set of analytic constants of motion for (124).

Grébert *et al* proved the following relation between the sequence of the squared spectral gaps and the regularity of the corresponding potential (see theorem in [GKM98]).

Theorem 5.5. *Let $\rho \geqslant 0$ and $s > 0$, then for any bounded subset $\mathcal{B} \subset \ell_{\rho,s}^2 \times \ell_{\rho,s}^2$ there exists $n_0 \geqslant 1$ and $M \geqslant 1$ such that for any $|k| \geqslant n_0$ and any $(\phi_1, \phi_2) \in \mathcal{B}$, the following estimate holds*

$$\sum_{|k| \geqslant n_0} (1 + |k|)^{2s} \, \mathrm{e}^{2\rho|k|} |\gamma_m|^2 \leqslant M. \tag{130}$$

Moreover, Grébert and Kappeler constructed the following global Birkhoff coordinates (see theorems 20.1–20.3 in [2014])

Theorem 5.6. *There exists a diffeomorphism $\Omega : L_r^2 \to \mathcal{H}_r^{0,0}$ such that:*

- *Ω is bianalytic and canonical;*
- *For each $s \geqslant 0$, the restriction of Ω to $\mathcal{H}_r^{0,s}$, namely the map*

$$\Omega : \mathcal{H}_r^{0,s} \to \mathcal{H}_r^{0,s}$$

  is again bianalytic and canonical;
- *The coordinates $(x, y) \in \mathcal{H}_r^{0,1}$ are Birkhoff coordinates for the NLS equation, namely they form a set of canonically conjugated coordinates in which the Hamiltonian of the NLS equation (124) depends only on the action $I_m := \frac{x_m^2 + y_m^2}{2}$ ($m \in \mathbb{Z}$).*

The dynamics of the NLS equation (124) in terms of the variables (x, y) is trivial: it can be immediately seen that any solution is periodic, quasiperiodic or almost periodic, depending on the number of spectral gaps (equivalently, depending on the number of actions) initially different from zero.

6. Approximation results

In this section we show how to use the normal form equations in order to construct approximate solutions of (5) and (22), and we estimate the difference with respect to the true solutions with corresponding initial data.

The approach is the same for all regimes (65), (85) and (104). First, we point out a relation between the specific energy of normal mode $\mathcal{E}_\kappa = 2E_k/N$ (where the energy of normal mode E_k is defined in (7) for (5), and in (23) for (22)), $k \in \mathbb{Z}_{N_1,N_2}^2$, and the Fourier coefficients of the solutions of the normal form equations. This procedure has to be done carefully, since

all wavevectors $k + (N_1 L_1, N_2 L_2)$ contributes to the specific energy $\mathcal{E}_\kappa$ of the discrete system. Then we have to prove that the approximate solutions approximate the specific energy of the true normal mode $\mathcal{E}_\kappa$ up to the time-scale for which the continuous approximation is valid, and finally we can deduce the result about the dynamics of the lattice. For simplicity we present in this section only the main part of the argument, and we defer the proof of technical results to appendices B–E.

6.1. The KP regime

Let $I = [-1, 1]^2$ be as in (69) and let $k = (k_1, k_2) \in \mathbb{Z}^2$. We define the Fourier coefficients of the function $q : I \to \mathbb{R}$ by

$$\hat{q}_k := \frac{1}{2} \int_I q(y_1, y_2) e^{-i\pi(k_1 y_1 + k_2 y_2)} \, dy_1 \, dy_2, \tag{131}$$

and similarly for the Fourier coefficients of the function p. Note that in this definition we omitted an eventual time dependence because we are interested in relations among quantities for points in the phase space, i.e. pairs of functions $(q, p) : I \to \mathbb{R}$. These relations extend automatically to trajectories and, in fact, this will be the subject of the last result of each subsection.

Lemma 6.1. *Consider the lattice* (2) *in the regime* (KP) *and with interpolating function* (65). *Then for a state corresponding to* (q, p) *one has*

$$\mathcal{E}_\kappa = \frac{\mu^4}{2} \sum_{L=(L_1, L_2) \in \mathbb{Z}^2 : \mu L_1, \mu^2 L_2 \in 2\mathbb{Z}} \left(|\hat{q}_{K+L}|^2 + \omega_k^2 \left| \frac{\hat{p}_{K+L}}{\mu} \right|^2 \right),$$

$$\forall \, k : \kappa(k) = (\mu K_1, \mu^2 K_2) \tag{132}$$

(where the ω_k are defined as in (8)), *and $\mathcal{E}_\kappa = 0$ otherwise.*

Proof. Take a smooth $(2N_1 + 1, 2N_2 + 1)$-periodic interpolating function Q for Q_j, and similarly for P_j. We denote by

$$\hat{Q}(k) := \frac{1}{\sqrt{N}} \int_{\left[-\left(N_1 + \frac{1}{2}\right), \left(N_1 + \frac{1}{2}\right) \right] \times \left[-\left(N_2 + \frac{1}{2}\right), \left(N_2 + \frac{1}{2}\right) \right]} Q(x) e^{-2\pi i \frac{k \cdot x}{N}} \, dx, \tag{133}$$

so that by the interpolation property we obtain

$$Q_j = Q(j) = \frac{1}{\sqrt{N}} \sum_{k \in \mathbb{Z}^2} \hat{Q}(k) e^{2\pi i \frac{j \cdot k}{N}}$$

$$= \frac{1}{\sqrt{N}} \sum_{k=(k_1, k_2) \in \mathbb{Z}^2_{2N_1+1, 2N_2+1}} \left[\sum_{h=(h_1, h_2) \in \mathbb{Z}^2} \hat{Q}(k_1 + (2N_1 + 1)h_1, k_2 + (2N_2 + 1)h_2) \right] e^{2\pi i \frac{j \cdot k}{N}},$$

hence

$$\hat{Q}_K = \sum_{h \in \mathbb{Z}^2} \hat{Q}(K_1 + (2N_1 + 1)h_1, K_2 + (2N_2 + 1)h_2). \tag{134}$$

The relation between $\hat{Q}(k)$ and $\hat{q}_k$ can be deduced from (65) and from the rescalings $\tau = \mu t$, $y_1 = \mu x_1$, $y_2 = \mu^2 x_2$,

$$Q(j) = \mu^2 q(\mu j_1, \mu^2 j_2);$$

$$\hat{Q}(k) = \frac{1}{2}\mu^{3/2} \int_{\left[-\frac{1}{\mu},\frac{1}{\mu}\right] \times \left[-\frac{1}{\mu^2},\frac{1}{\mu^2}\right]} Q(x_1, x_2) e^{-i\pi(k_1 x_1 \mu + k_2 x_2 \mu^2)} \, dx_1 \, dx_2$$

$$= \frac{1}{2}\mu^{3/2} \int_{\left[-\frac{1}{\mu},\frac{1}{\mu}\right] \times \left[-\frac{1}{\mu^2},\frac{1}{\mu^2}\right]} \mu^2 q\left(\mu x_1, \mu^2 x_2\right)$$

$$\times \ e^{-i\pi(k_1 x_1 \mu + k_2 x_2 \mu^2)} \, dx_1 \, dx_2$$

$$\overset{(85)}{=} \frac{1}{2}\mu^{1/2} \int_I q(y) e^{-i\pi(k_1 y_1 + k_2 y_2)} \, dy = \mu^{1/2} \hat{q}_k, \tag{135}$$

and similarly

$$\hat{P}(k) = \mu^{-1/2} \hat{p}_k. \tag{136}$$

By using (7), (10) and (134)–(136) we have

$$\mathcal{E}_\kappa \overset{(10)}{=} \mu^3 \frac{1}{2} \sum_{L=(L_1,L_2)\in\mathbb{Z}^2 : \mu L_1, \mu^2 L_2 \in 2\mathbb{Z}} \left(|\hat{Q}(K+L)|^2 + \omega_k^2 |\hat{P}(K+L)|^2 \right)$$

$$\overset{(135)}{=} \mu^4 \frac{1}{2} \sum_{L=(L_1,L_2)\in\mathbb{Z}^2 : \mu L_1, \mu^2 L_2 \in 2\mathbb{Z}} \left(|\hat{q}_{K+L}|^2 + \omega_k^2 \left| \frac{\hat{p}_{K+L}}{\mu} \right|^2 \right)$$

for all k such that $\kappa(k) = (\mu K_1, \ \mu^2 K_2)$, and this leads to (132). $\qquad\square$

Proposition 6.2. *Fix $\rho > 0$ and $0 < \delta \ll 1$. Consider the normal form system* (84)*, and define the Fourier coefficients of (ξ, η) through the following formula*

$$\xi(y) = \frac{1}{2} \sum_{h\in\mathbb{Z}^2} \hat{\xi}_h \, e^{i\pi h \cdot y}, \tag{137}$$

$$\eta(y) = \frac{1}{2} \sum_{h\in\mathbb{Z}^2} \hat{\eta}_h \, e^{i\pi h \cdot y}. \tag{138}$$

Suppose that $(\xi, \eta) \in \mathcal{H}^{\rho,0}$, and denote by $\mathcal{E}_\kappa$ the specific energy of the normal mode with index κ as defined in (9) *and* (10)*. Then for any positive μ sufficiently small*

$$\left| \mathcal{E}_\kappa - \mu^4 \frac{|\hat{\xi}_K|^2 + |\hat{\eta}_K|^2}{2} \right| \leqslant C\mu^{4+3/2} \|(\xi, \eta)\|_{\mathcal{H}^{\rho,0}}^2 \tag{139}$$

for all k such that $\kappa(k) = (\mu K_1, \ \mu^2 K_2)$ and $|K_1| + |K_2| \leqslant \frac{(2+\delta)|\log\mu|}{\rho}$. Moreover,

$$|\mathcal{E}_\kappa| \leqslant C\mu^8 \|(\xi, \eta)\|_{\mathcal{H}^{\rho,0}}^2 \tag{140}$$

for all k such that $\kappa(k) = (\mu K_1, \ \mu^2 K_2)$ and $|K_1^2 + K_2^2|^{1/2} > \frac{(2+\delta)|\log\mu|}{\rho}$, and $\mathcal{E}_\kappa = 0$ otherwise.

The proof of the above proposition is deferred to appendix B.

Now, consider the following system of uncoupled KP equations

$$\xi_\tau = -\frac{1}{24}\partial_{y_1}^3 \xi - \frac{1}{2}\partial_{y_1}^{-1}\partial_{y_2}^2 \xi - \frac{\alpha}{2\sqrt{2}}\partial_{y_1}(\xi^2), \tag{141}$$

$$\eta_\tau = \frac{1}{2}\partial_{y_1}^{-1}\partial_{y_2}^2 \eta + \frac{1}{24}\partial_{y_1}^3 \eta + \frac{\alpha}{2\sqrt{2}}\partial_{y_1}(\eta^2). \tag{142}$$

and consider a solution $(\tau, y) \mapsto (\widetilde{\xi}_a(\tau, y), \widetilde{\eta}_a(\tau, y))$ such that it belongs to $\mathcal{H}^{\rho,n}$, for some $n \geqslant 1$.

We consider the approximate solutions (Q_a, P_a) of the ETL model (62) defined by formulas (143) and (144).

$$Q_a(t, y) := \frac{\mu^2}{\sqrt{2}}\left[\widetilde{\xi}_a(\mu^2\tau, y_1 - \tau, y_2) + \widetilde{\eta}_a(\mu^2\tau, y_1 + \tau, y_2)\right] \tag{143}$$

$$\partial_{y_1}P_a(t, y) := \frac{\mu}{\sqrt{2}}\left[\widetilde{\xi}_a(\mu^2\tau, y_1 - \tau, y_2) - \widetilde{\eta}_a(\mu^2\tau, y_1 + \tau, y_2)\right], \tag{144}$$

where we made a little abuse of notation since the left-hand side depends on t and right-hand side on τ that are related as $\tau = \mu t$.

We need to compare the difference between the approximate solution (143) and (144) and the true solution of (5). Let us consider (5), and take an initial datum (Q_0, P_0) with corresponding Fourier coefficients $(\hat{Q}_{0,k}, \hat{P}_{0,k})$ given by (6). Observe that

$$\hat{Q}_{0,k} \neq 0 \quad \text{only if } \kappa(k) = (\mu K_1, \mu^2 K_2) \tag{145}$$

and that, as a consequence of the analyticity of Q_0 and P_0, there exist $C, \rho > 0$ such that

$$\frac{|\hat{Q}_{0,k}|^2 + \omega_k^2|\hat{P}_{0,k}|^2}{N} \leqslant C\,e^{-2\rho|(\kappa_1(k)/\mu, \kappa_2(k)/\mu^2)|}. \tag{146}$$

Moreover, we define an interpolating function for the initial datum (Q_0, P_0) by

$$Q_0(y) = \frac{1}{\sqrt{N}}\sum_{K:(\mu^2|K_1|^2 + \mu^4|K_2|^2)^{1/2}=|\kappa(k)|\leqslant 1}\hat{Q}_{0,k}\,e^{i\pi(\mu K_1 y_1 + \mu^2 K_2 y_2)},$$

and similarly for $y \mapsto P_0(y)$.

Next we show that we can exploit the analyticity of the solution of the approximating integrable PDEs to prove the vicinity between the approximate solution and the true solution.

Proposition 6.3. *Consider (5) with $\sigma = 2$, and fix $0 < \gamma < \frac{1}{2}$. Let us assume that the initial datum for (5) satisfies (145) and (146), and denote by $(Q_j(t), P_j(t))_{j \in \mathbb{Z}^2_{N_1,N_2}}$ the corresponding solution. Consider the approximate solution $(\widetilde{\xi}_a, \widetilde{\eta}_a)$ with the corresponding initial datum. Assume that $(\widetilde{\xi}_a, \widetilde{\eta}_a) \in \mathcal{H}^{\rho,n}$ for some $\rho > 0$ and for some $n \geqslant 1$ for all times, and fix $T_0 > 0$ and $0 < \delta \ll 1$.*

Then there exists $\mu_0 = \mu_0(T_0, \|(\widetilde{\xi}_a(0), \widetilde{\eta}_a(0))\|_{\mathcal{H}^{\rho,n}})$ such that, if $\mu < \mu_0$, we have that there exists $C > 0$ such that

$$\sup_j |Q_j(t) - Q_a(t, j)| + |P_j(t) - P_a(t, j)| \leqslant C\mu^\gamma, \quad \forall\, |t| \leqslant \frac{T_0}{\mu^3}, \tag{147}$$

where (Q_a, P_a) are given by (143) *and* (144)*. Moreover,*

$$\left| \mathcal{E}_\kappa(t) - \mu^4 \frac{|\hat{\xi}_K(\tau)|^2 + |\hat{\eta}_K(\tau)|^2}{2} \right| \leqslant C\mu^{4+\gamma} \quad \forall |t| \leqslant \frac{T_0}{\mu^3} \tag{148}$$

for all k such that $\kappa(k) = (\mu K_1, \ \mu^2 K_2)$ and $|K_1| + |K_2| \leqslant \frac{(2+\delta)|\log \mu|}{\rho}$. Moreover,

$$|\mathcal{E}_\kappa(\tau)| \leqslant \mu^{4+\gamma} \quad \forall |t| \leqslant \frac{T_0}{\mu^3} \tag{149}$$

for all k such that $\kappa(k) = (\mu K_1, \ \mu^2 K_2)$ and $|K_1| + |K_2| > \frac{(2+\delta)|\log \mu|}{\rho}$, and $\mathcal{E}_\kappa = 0$ otherwise.

The proof of the above proposition is deferred to appendix C.

Proof of theorem 2.1. First we prove (14).

We consider an initial datum as in (13); when passing to the continuous approximation (62), this initial datum corresponds to an initial data $(\xi_0, \eta_0) \in \mathcal{H}^{\rho_0, n}$ for some $\rho_0 > 0$ and $n \geqslant 1$. By theorem 5.1 the corresponding solution $(\xi(\tau), \eta(\tau))$ is analytic in a complex strip of width $\rho(t)$. Taking the minimum of such quantities one gets the coefficient ρ appearing in the statement of theorem (2.1). Applying proposition 6.3, we can deduce the corresponding result for the discrete model (5) and the specific quantities (10). □

6.2. The KdV regime

Similarly to lemma 6.1 and proposition 6.2 we can prove the following results

Lemma 6.4. *Consider the lattice* (2) *in the regime* (KdV) *and with interpolating function* (85)*. Then for a state corresponding to (q, p) one has*

$$\mathcal{E}_\kappa = \frac{\mu^4}{2} \sum_{L=(L_1, L_2) \in \mathbb{Z}^2 : \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}} \left(|\hat{q}_{K+L}|^2 + \omega_k^2 \left| \frac{\hat{p}_{K+L}}{\mu} \right|^2 \right), \tag{150}$$

$$\forall \ k : \kappa(k) = (\mu K_1, \mu^\sigma K_2)$$

(where the ω_k are defined as in (8) *and the $\mathcal{E}_\kappa$ in* (10))*, and $\mathcal{E}_\kappa = 0$ otherwise.*

Proof. As in lemma 6.4 we introduce a $(2N_1 + 1, 2N_2 + 1)$-periodic interpolating function for Q_j and P_j. We denote $\hat{Q}(t, k)$ and $\hat{Q}_K(t)$ as in (133) and (134).

The relation between $\hat{Q}(k)$ and $\hat{q}_k$ can be deduced from (85) and from the rescalings $y_1 = \mu x_1, y_2 = \mu^\sigma x_2,$

$$Q(j) = \mu^2 q(\mu j_1, \mu^\sigma j_2);$$

$$\hat{Q}(k) = \frac{1}{2} \mu^{(\sigma+1)/2} \int_{\left[-\frac{1}{\mu}, \frac{1}{\mu}\right] \times \left[-\frac{1}{\mu^\sigma}, \frac{1}{\mu^\sigma}\right]} Q(x_1, x_2) e^{-i\pi(k_1 x_1 \mu + k_2 x_2 \mu^\sigma)} \, dx_1 \, dx_2$$

$$= \frac{1}{2} \mu^{(\sigma+1)/2} \int_{\left[-\frac{1}{\mu}, \frac{1}{\mu}\right] \times \left[-\frac{1}{\mu^\sigma}, \frac{1}{\mu^\sigma}\right]} \mu^2 q\left(\mu x_1, \mu^\sigma x_2\right)$$

$$\times e^{-i\pi(k_1 x_1 \mu + k_2 x_2 \mu^\sigma)} \, dx_1 \, dx_2$$

$$\overset{(85)}{=} \frac{1}{2} \mu^{(3-\sigma)/2} \int_I q(y) e^{-i\pi(k_1 y_1 + k_2 y_2)} \, dy = \mu^{(3-\sigma)/2} \hat{q}_k, \tag{151}$$

and similarly

$$\hat{P}(k) = \mu^{(1-\sigma)/2}\hat{p}_k. \tag{152}$$

By using (7), (10), (151) and (152) we have

$$\mathcal{E}_\kappa \overset{(10)}{=} \mu^{\sigma+1}\frac{1}{2}\sum_{L=(L_1,L_2)\in\mathbb{Z}^2:\mu L_1,\mu^\sigma L_2\in 2\mathbb{Z}}\left(|\hat{Q}(K+L)|^2 + \omega_k^2|\hat{P}(K+L)|^2\right)$$

$$\overset{(151),(151)}{=} \mu^{\sigma+1}\,\mu^{3-\sigma}\frac{1}{2}\sum_{L=(L_1,L_2)\in\mathbb{Z}^2:\mu L_1,\mu^\sigma L_2\in 2\mathbb{Z}}\left(|\hat{q}_{K+L}|^2 + \omega_k^2\left|\frac{\hat{p}_{K+L}}{\mu}\right|^2\right)$$

for all k such that $\kappa(k) = (\mu K_1,\ \mu^\sigma K_2)$, and this leads to (150). $\qquad\square$

Proposition 6.5. *Fix $\rho > 0$ and $0 < \delta \ll 1$. Consider the normal form system* (96)*, and define the Fourier coefficients of (ξ,η) through the following formula*

$$\xi(y) = \frac{1}{2}\sum_{h\in\mathbb{Z}^2}\hat{\xi}_h\,\mathrm{e}^{\mathrm{i}\pi h\cdot y}, \tag{153}$$

$$\eta(y) = \frac{1}{2}\sum_{h\in\mathbb{Z}^2}\hat{\eta}_h\,\mathrm{e}^{\mathrm{i}\pi h\cdot y}. \tag{154}$$

Suppose that $(\xi,\eta)\in\mathcal{H}^{\rho,0}$, and denote by $\mathcal{E}_\kappa$ the specific energy of the normal mode with index κ as defined in (9) *and* (10)*. Then for any positive μ sufficiently small*

$$\left|\mathcal{E}_\kappa - \mu^4\frac{|\hat{\xi}_K|^2 + |\hat{\eta}_K|^2}{2}\right| \leqslant C\mu^{4+3/2}\|(\xi,\eta)\|_{\mathcal{H}^{\rho,0}}^2 \tag{155}$$

for all k such that $\kappa(k) = (\mu K_1,\ \mu^\sigma K_2)$ and $|K_1| + |K_2| \leqslant \frac{(2+\delta)|\log\mu|}{\rho}$. Moreover,

$$|\mathcal{E}_\kappa| \leqslant C\mu^8\|(\xi,\eta)\|_{\mathcal{H}^{\rho,0}}^2 \tag{156}$$

for all k such that $\kappa(k) = (\mu K_1,\ \mu^\sigma K_2)$ and $|K_1| + |K_2| > \frac{(2+\delta)|\log\mu|}{\rho}$, and $\mathcal{E}_\kappa = 0$ otherwise.

We defer the proof of the above proposition to appendix B.

Now, consider the following system of uncoupled KdV equations

$$\xi_\tau = -\frac{1}{24}\partial_{y_1}^3\xi - \frac{\alpha}{2\sqrt{2}}\partial_{y_1}(\xi^2), \tag{157}$$

$$\eta_\tau = \frac{1}{24}\partial_{y_1}^3\eta + \frac{\alpha}{2\sqrt{2}}\partial_{y_1}(\eta^2), \tag{158}$$

and consider a solution $(\tau, y) \mapsto (\widetilde{\xi}_a(\tau, y), \widetilde{\eta}_a(\tau, y))$ such that it belongs to $\mathcal{H}^{\rho,n}$, for some $n \geqslant 1$.

We consider the approximate solutions (Q_a, P_a) of the ETL model (62) defined by formulas (143) and (144).

We need to compare the difference between the approximate solution (143) and (144) and the true solution of (5). Let us consider (5), and take an initial datum (Q_0, P_0) with corresponding Fourier coefficients $(\hat{Q}_{0,k}, \hat{P}_{0,k})$ given by (6); observe that

$$Q_{0,k} \neq 0 \quad \text{only if } \kappa(k) = (\mu K_1, \mu^\sigma K_2). \tag{159}$$

Since Q_0 and P_0 are analytic functions, there exist C, $\rho > 0$ such that

$$\frac{|\hat{Q}_{0,k}|^2 + \omega_k^2 |\hat{P}_{0,k}|^2}{(2N_1 + 1)(2N_2 + 1)} \leqslant C \, \mathrm{e}^{-2\rho|(\kappa_1(k)/\mu, \, \kappa_2(k)/\mu^\sigma)|}. \tag{160}$$

Moreover, we define an interpolating function for the initial datum (Q_0, P_0) by

$$Q_0(y, t) = \frac{1}{\sqrt{N}} \sum_{K:(\mu^2|K_1|^2 + \mu^{2\sigma}|K_2|^2)^{1/2} = |\kappa(k)| \leqslant 1} \hat{Q}_{0,k}(t) \mathrm{e}^{\mathrm{i}\pi(\mu K_1 y_1 + \mu^\sigma K_2 y_2)},$$

and similarly for $y \mapsto P_0(y)$.

Proposition 6.6. *Consider (5) with $\sigma > 2$ and $\gamma \geqslant 1$ such that $\sigma + 2\gamma < \min(4\sigma - 5, 7)$. Let us assume that the initial datum satisfies (159) and (160), and denote by $(Q(t), P(t))$ the corresponding solution. Consider the approximate solution $(\widetilde{\xi}_a(t, x), \widetilde{\eta}_a(t, x))$ with the corresponding initial datum. Assume that $(\widetilde{\xi}_a, \widetilde{\eta}_a) \in \mathcal{H}^{\rho, n}$ for some $\rho > 0$ and for some $n \geqslant 1$ for all times, and fix $T_0 > 0$ and $0 < \delta \ll 1$.*

Then there exists $\mu_0 = \mu_0(T_0, \sigma, \|(\widetilde{\xi}_a(0), \widetilde{\eta}_a(0))\|_{\mathcal{H}^{\rho, n}})$ such that, if $\mu < \mu_0$, we have that there exists $C > 0$ such that

$$\sup_j |Q_j(t) - Q_a(t, j)| + |P_j(t) - P_a(t, j)| \leqslant C\mu^\gamma, \quad \forall \, |t| \leqslant \frac{T_0}{\mu^3}, \tag{161}$$

where (Q_a, P_a) are given by (143) and (144). Moreover,

$$\left| \mathcal{E}_\kappa(t) - \mu^4 \frac{|\hat{\xi}_K(\tau)|^2 + |\hat{\eta}_K(\tau)|^2}{2} \right| \leqslant C\mu^{4+\gamma}, \quad \forall \, |t| \leqslant \frac{T_0}{\mu^3} \tag{162}$$

for all k such that $\kappa(k) = (\mu K_1, \, \mu^\sigma K_2)$ and $|K_1| + |K_2| \leqslant \frac{(2+\delta)|\log \mu|}{\rho}$. Moreover,

$$|\mathcal{E}_\kappa(t)| \leqslant \mu^{4+\gamma}, \quad \forall \, |t| \leqslant \frac{T_0}{\mu^3} \tag{163}$$

for all k such that $\kappa(k) = (\mu K_1, \, \mu^\sigma K_2)$ and $|K_1| + |K_2| > \frac{(2+\delta)|\log \mu|}{\rho}$, and $\mathcal{E}_\kappa = 0$ otherwise.

We defer the proof to appendix C.

Remark 6.7. The conditions $\sigma + 2\gamma < \min(4\sigma - 5, 7)$, which, together with $\gamma > 0$, implies the upper bound $\sigma < 7$ found in the statement of theorem 2.5, is the consequence of a technical condition which allows to estimate the error in the proof of proposition 6.6 (see claim 2, together with (238) and (239)).

Proof of theorem 2.5. First we prove (17).

We consider an initial datum as in (16); when passing to the continuous approximation (62), this initial datum corresponds to an initial data $(\xi_0, \eta_0) \in \mathcal{H}^{\rho_0, n}$ for some $\rho_0 > 0$ and $n \geqslant 1$. By theorem 5.3 the corresponding sequence of gaps belongs to $\mathcal{H}^{\rho_0, n}$, and that the solution $(\xi(\tau), \eta(\tau))$ is analytic in a complex strip of width $\rho(t)$. Taking the minimum of such quantities one gets the coefficient ρ appearing in the statement of theorem 2.5. Applying proposition 6.6, we can deduce the corresponding result for the discrete model (5) and the specific quantities (10).

Next, we prove (19). In order to do so, we exploit the Birkhoff coordinates (x, y) introduced in theorem 5.4; indeed, by rewriting the normal form system (96) in Birkhoff coordinates we get that every solution is almost-periodic in time. Now, let us introduce the quantities

$$E_K^{(1)} := \left|\hat{\xi}_K\right|^2,$$

$$E_K^{(2)} := |\hat{\eta}_K|^2,$$

then $\tau \mapsto E_K^{(1)}(x(\tau), y(\tau))$ and $\tau \mapsto E_K^{(2)}(x(\tau), y(\tau))$ are almost-periodic. If we set $E_K := \frac{1}{2}\left(E_K^{(1)} + E_K^{(2)}\right)$, we can exploit (162) in proposition 6.6 to translate the results in terms of the specific quantities $\mathcal{E}_\kappa$, and we get the thesis. $\square$

6.3. The one-dimensional NLS regime

Let $\beta > 0$ and let I be as in (69), we define the Fourier coefficients of the function $q : I \to \mathbb{R}$ by

$$\hat{q}_k := \frac{1}{2} \int_I q(y_1, y_2) e^{-i\pi(k_1 y_1 + k_2 y_2)} \, \mathrm{d}y_1 \, \mathrm{d}y_2, \tag{164}$$

and similarly for the Fourier coefficients of the function p.

Lemma 6.8. *Consider the lattice* (20) *in the regime* (1D NLS) *and with interpolating function* (104). *Then for a state corresponding to* (q, p) *one has*

$$\mathcal{E}_\kappa = \frac{\mu^2}{2} \sum_{L=(L_1, L_2)\in\mathbb{Z}^2 : \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}} \left(|\hat{p}_{K+L}|^2 + \omega_k^2 |\hat{q}_{K+L}|^2\right), \tag{165}$$

$$\forall \, k : \kappa(k) = (\mu K_1, \mu^\sigma K_2)$$

(*where the* ω_k *are defined as in* (24)), *and* $\mathcal{E}_\kappa = 0$ *otherwise.*

Proof. We introduce a $(2N_1 + 1, 2N_2 + 1)$-periodic interpolating function for Q_j and P_j. We denote $\hat{Q}(t, k)$ and $\hat{Q}_k(t)$ as in (133) and (134). By the interpolation property we obtain (134).

The relation between $\hat{Q}(t, k)$ and $\hat{q}_k(t)$ can be deduced from (104) and from the rescalings $\tau = \mu t$, $y_1 = \mu x_1$, $y_2 = \mu^\sigma x_2$,

$$Q(j) = \mu q(\mu j_1, \mu^\sigma j_2);$$

$$\hat{Q}(k) = \frac{1}{2} \mu^{(\sigma+1)/2} \int_{\left[-\frac{1}{\mu}, \frac{1}{\mu}\right] \times \left[-\frac{1}{\mu^\sigma}, \frac{1}{\mu^\sigma}\right]} Q(x_1, x_2) e^{-i\pi(k_1 x_1 \mu + k_2 x_2 \mu^\sigma)} \, \mathrm{d}x_1 \, \mathrm{d}x_2$$

$$\stackrel{(104)}{=} \frac{1}{2} \mu^{(\sigma+1)/2} \int_{\left[-\frac{1}{\mu}, \frac{1}{\mu}\right] \times \left[-\frac{1}{\mu^\sigma}, \frac{1}{\mu^\sigma}\right]} \mu q\left(\mu x_1, \mu^\sigma x_2\right)$$

$$\times e^{-i\pi(k_1 x_1 \mu + k_2 x_2 \mu^\sigma)} \, \mathrm{d}x_1 \, \mathrm{d}x_2$$

$$= \frac{1}{2} \mu^{(1-\sigma)/2} \int_I q(y) e^{-i\pi(k_1 y_1 + k_2 y_2)} \mathrm{d}y = \mu^{(1-\sigma)/2} \hat{q}_k, \tag{166}$$

and similarly

$$\hat{P}(k) = \mu^{(1-\sigma)/2} \hat{p}_k. \tag{167}$$

By using (23), (10), (134) and (167) we have

$$
\mathcal{E}_\kappa \overset{(10)}{=} \mu^{\sigma+1} \frac{1}{2} \sum_{L=(L_1,L_2)\in\mathbb{Z}^2:\mu L_1,\mu^\sigma L_2\in 2\mathbb{Z}} \left(|\hat{P}(K+L)|^2 + \omega_k^2 |\hat{Q}(K+L)|^2 \right)
$$

$$
\overset{(166)}{=} \mu^{\sigma+1} \mu^{1-\sigma} \frac{1}{2} \sum_{L=(L_1,L_2)\in\mathbb{Z}^2:\mu L_1,\mu^\sigma L_2\in 2\mathbb{Z}} \left(|\hat{p}_{K+L}|^2 + \omega_k^2 |\hat{q}_{K+L}|^2 \right)
$$

for all k such that $\kappa(k) = (\mu K_1, \mu^\sigma K_2)$, and this leads to (165). $\qquad\square$

Proposition 6.9. *Fix $\rho > 0$ and $0 < \delta \ll 1$. Consider the normal form equation (116), and define the Fourier coefficients of $(\psi, \bar\psi)$ through the following formula*

$$
\psi(y) = \frac{1}{2} \sum_{h\in\mathbb{Z}^2} \hat\psi_h \, e^{i\pi h\cdot y}. \tag{168}
$$

Suppose that $(\psi, \bar\psi) \in \mathcal{H}^{\rho,0}$, and denote by $\mathcal{E}_\kappa$ the specific energy of the normal mode with index κ as defined in (9) and (10). Then for any positive μ sufficiently small

$$
\left| \mathcal{E}_\kappa - \mu^2 \frac{|\hat\psi_K|^2}{2} \right| \leqslant C \mu^{2+3/2} \|(\psi, \bar\psi)\|^2_{\mathcal{H}^{\rho,0}} \tag{169}
$$

for all k such that $\kappa(k) = (\mu K_1, \mu^\sigma K_2)$ and $|K_1| + |K_2| \leqslant \frac{(2+\delta)|\log\mu|}{\rho}$. Moreover,

$$
|\mathcal{E}_\kappa| \leqslant C\mu^6 \|(\psi, \bar\psi)\|^2_{\mathcal{H}^{\rho,1}} \tag{170}
$$

for all k such that $\kappa(k) = (\mu K_1, \mu^\sigma K_2)$ and $|K_1| + |K_2| > \frac{(2+\delta)|\log\mu|}{\rho}$, and $\mathcal{E}_\kappa = 0$ otherwise.

We defer the proof of the above proposition to appendix D.

Now, consider the normal form equation, namely the following cubic defocusing one-dimensional NLS

$$
-i\psi_\tau = -\partial_{y_1}^2 \psi + \frac{3\beta}{4} |\psi|^2 \psi. \tag{171}
$$

and consider a solution $(\widetilde{\psi_a}, \overline{\widetilde{\psi}_a})$ such that it belongs to $\mathcal{H}^{\rho,n}$, for some $n > 0$.

We consider the approximate solutions (Q_a, P_a) of the KG lattice (20) (in the following $\tau = \mu^2 t$)

$$
Q_a(t,y) := \frac{\mu}{\sqrt{2}} \left[e^{i\tau} \widetilde{\psi_a}(\tau, y_1, y_2) + e^{-i\tau} \overline{\widetilde{\psi}_a}(\tau, y_1, y_2) \right], \tag{172}
$$

$$
P_a(t,y) := \frac{\mu}{\sqrt{2}i} \left[e^{i\tau} \widetilde{\psi_a}(\tau, y_1, y_2) + e^{-i\tau} \overline{\widetilde{\psi}_a}(\tau, y_1, y_2) \right]. \tag{173}
$$

We need to compare the difference between the approximate solution (172) and (173) and the true solution of (20). Let us consider (20), and take an initial datum (Q_0, P_0) with corresponding Fourier coefficients $(\hat{Q}_{0,k}, \hat{P}_{0,k})$ given by (6); observe that

$$
\hat{Q}_{0,k} \neq 0 \quad \text{only if } \kappa(k) = (\mu K_1, \mu^\sigma K_2). \tag{174}
$$

Since P_0 and Q_0 are analytic functions, there exist $C, \rho > 0$ such that

$$
\frac{|\hat{P}_{0,k}|^2 + \omega_k^2 |\hat{Q}_{0,k}|^2}{N} \leqslant C e^{-2\rho|(\kappa_1(k)/\mu, \kappa_2(k)/\mu^\sigma)|}. \tag{175}
$$

Moreover, we define an interpolating function for the initial datum (Q_0, P_0) by

$$Q_0(t, y) = \frac{1}{\sqrt{N}} \sum_{K:(\mu^2|K_1|^2 + \mu^{2\sigma}|K_2|^2)^{1/2} = |\kappa(k)| \leqslant 1} \hat{Q}_{0,K}(t) e^{i\pi(\mu K_1 y_1 + \mu^\sigma K_2 y_2)},$$

and similarly for $y \mapsto P_0(y)$.

Proposition 6.10. *Consider* (20) *with* $\sigma > 1$ *and* $\gamma > 0$ *such that* $\sigma + 2\gamma < \min(4\sigma - 1, 7)$. *Let us assume that the initial datum satisfies* (174) *and* (175), *and denote by* $(Q_j(t), P_j(t))_{j \in \mathbb{Z}_{N_1, N_2}}$ *the corresponding solution. Consider the approximate solution* $(\widetilde{\psi_a}(t, x), \widetilde{\psi_a}(t, x))$ *with the corresponding initial datum. Assume that* $(\widetilde{\psi_a}, \widetilde{\psi_a}) \in \mathcal{H}^{\rho, n}$ *for some* $\rho > 0$ *and for some* $n \geqslant 0$ *for all times, and fix* $T_0 > 0$ *and* $0 < \delta \ll 1$.

Then there exists $\mu_0 = \mu_0(T_0, \sigma, \|(\widetilde{\psi_a}(0), \widetilde{\psi_a}(0))\|_{\mathcal{H}^{\rho, n}})$ *such that, if* $\mu < \mu_0$, *we have that there exists* $C > 0$ *such that*

$$\sup_j |Q_j(t) - Q_a(t, j)| + |P_j(t) - P_a(t, j)| \leqslant C\mu^\gamma, \quad \forall |t| \leqslant \frac{T_0}{\mu^2}, \tag{176}$$

where (Q_a, P_a) *are given by* (172) *and* (173). *Moreover,*

$$\left| \mathcal{E}_\kappa(t) - \mu^2 \frac{|\hat{\psi}_K(\tau)|^2}{2} \right| \leqslant C\mu^{2+\gamma}, \quad |t| \leqslant \frac{T_0}{\mu^2} \tag{177}$$

for all k such that $\kappa(k) = (\mu K_1, \mu^\sigma K_2)$ *and* $|K_1| + |K_2| \leqslant \frac{(2+\delta)|\log \mu|}{\rho}$. *Moreover,*

$$|\mathcal{E}_\kappa(t)| \leqslant \mu^{2+\gamma}, \quad |t| \leqslant \frac{T_0}{\mu^2} \tag{178}$$

for all k such that $\kappa(k) = (\mu K_1, \mu^\sigma K_2)$ *and* $|K_1| + |K_2| > \frac{(2+\delta)|\log \mu|}{\rho}$, *and* $\mathcal{E}_\kappa = 0$ *otherwise.*

We defer the proof to appendix E.

Remark 6.11. The condition $\sigma + 2\gamma < \min(4\sigma - 1, 7)$, which, together with $\gamma > 0$, implies the upper bound $\sigma < 7$ found in the statement of theorem (2.7), is the consequence of a technical condition which allows to estimate the error in the proof of proposition 6.10 (see claim 2, together with (268)–(270)).

Proof of theorem 2.7. First we prove (26).

We consider an initial datum as in (25); when passing to the continuous approximation (103), this initial datum corresponds to an initial data $(\xi_0, \eta_0) \in \mathcal{H}^{\rho_0, n}$. By theorem 5.5 the corresponding sequence of gaps belongs to $\mathcal{H}^{\rho_0, n}$, and that the solution $(\xi(\tau), \eta(\tau))$ is analytic in a complex strip of width $\rho(t)$. Taking the minimum of such quantities one gets the coefficient ρ appearing in the statement of theorem (2.7). Applying proposition 6.10, we can deduce the corresponding result for the discrete model (22) and the specific quantities (10).

Next, we prove (28). In order to do so, we exploit the Birkhoff coordinates (x, y) introduced in theorem 5.6; indeed, by rewriting the normal form system (116) in Birkhoff coordinates we get that every solution is almost-periodic in time. Now, let us introduce the quantity

$$E_K := \frac{1}{2} \left| \hat{\psi}_K \right|^2,$$

then $\tau \mapsto E_K(x(\tau), y(\tau))$ is almost-periodic. Hence we can exploit (177) in proposition 6.10 to translate the results in terms of the specific quantities $\mathcal{E}_\kappa$, and we get the thesis. □

Acknowledgments

Appendix A. Proof of lemma 3.6

This appendix is devoted to the proof of the lemma 3.6, which is a key step to normalize the system (48). Its proof is an adaptation of theorem 4.4 in [Bam99] and it is based on the method of Lie transform, briefly recalled in the following. Throughout this section, we consider $s \geqslant s_1$ and $\rho \geqslant 0$ to be fixed quantities.

Given an auxiliary function χ analytic on $\mathcal{H}^{\rho,s}$, we consider the auxiliary differential equation

$$\dot{\zeta} = X_\chi(\zeta) \tag{179}$$

and denote by Φ_χ^t its flow at time t.

Lemma A.1. *Let χ and its vector field be analytic in $\mathcal{B}_{\rho,s}(R)$. Fix $d < R$, and assume that*

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant d.$$

Then, if we consider the time-t flow Φ_χ^t of X_χ we have that for $|t| \leqslant 1$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|\Phi_\chi^t(\zeta) - \zeta\|_{\mathcal{H}^{\rho,s}} \leqslant \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}.$$

Definition A.2. The map $\Phi_\chi := \Phi_\chi^1$ is called the *Lie transform* generated by χ.

Given G analytic on $\mathcal{H}^{\rho,s}$, let us consider the differential equation

$$\dot{\zeta} = X_\mathsf{G}(\zeta), \tag{180}$$

where by X_G we denote the vector field of G. Now define

$$\Phi_\chi^* \mathsf{G}(\tilde{\zeta}) := \mathsf{G} \circ \Phi_\chi(\tilde{\zeta}).$$

By exploiting the fact that Φ_χ is a canonical transformation, we have that in the new variable $\tilde{\zeta}$ defined by $\zeta = \Phi_\chi(\tilde{\zeta})$ equation (180) is equivalent to

$$\dot{\tilde{\zeta}} = X_{\Phi_\chi^* \mathsf{G}}(\tilde{\zeta}). \tag{181}$$

Using the relation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_\chi^* \mathsf{G} = \Phi_\chi^*\{\chi, \mathsf{G}\}, \tag{182}$$

and the Poisson bracket formalism $\{\mathsf{G}_1, \mathsf{G}_2\}(\zeta) := \mathrm{d}\mathsf{G}_1(\zeta)[X_{\mathsf{G}_2}(\zeta)]$ we formally get

$$\Phi_\chi^*\mathsf{G} = \sum_{\ell=0}^\infty \mathsf{G}_\ell,$$
$$\mathsf{G}_0 := \mathsf{G}, \tag{183}$$
$$\mathsf{G}_\ell := \frac{1}{\ell}\{\chi, \mathsf{G}_{\ell-1}\}, \quad \ell \geqslant 1.$$

In order to estimate the vector field of the terms appearing in (183), we exploit the following results

Lemma A.3. *Let $R > 0$, and assume that χ, G are analytic on $\mathcal{B}_{\rho,s}(R)$ as well as their vector fields. Then, for any $d \in (0, R)$ we have that $\{\chi, \mathsf{G}\}$ is analytic on $\mathcal{B}_{\rho,s}(R - d)$, and*

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|X_{\{\chi,\mathsf{G}\}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant \frac{2}{d}\left(\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}\right)\left(\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\mathsf{G}(\zeta)\|_{\mathcal{H}^{\rho,s}}\right). \tag{184}$$

Proof. Observe that

$$\|X_{\{\chi,\mathsf{G}\}}(\zeta)\|_{\mathcal{H}^{\rho,s}} = \|\mathrm{d}X_\chi(\zeta)X_\mathsf{G}(\zeta) - \mathrm{d}X_\mathsf{G}(\zeta)X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}$$
$$\leqslant \|\mathrm{d}X_\chi(\zeta)X_\mathsf{G}(\zeta)\|_{\mathcal{H}^{\rho,s}} + \|\mathrm{d}X_\mathsf{G}(\zeta)X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}},$$

and since for any $d \in (0, R)$ Cauchy inequality gives

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|\mathrm{d}X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}} \leqslant \frac{1}{d}\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}},$$

we finally get

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|\mathrm{d}X_\chi(\zeta)X_\mathsf{G}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant \frac{1}{d}\left(\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}\right)$$
$$\times \left(\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\mathsf{G}(\zeta)\|_{\mathcal{H}^{\rho,s}}\right).$$

With a similar estimate for the other term we obtain the thesis. $\square$

Lemma A.4. *Let $R > 0$, and assume that χ, G are analytic on $\mathcal{B}_{\rho,s}(R)$ as well as their vector fields. Let $\ell \geqslant 1$, and consider G_ℓ as defined in (183); for any $d \in (0, R)$, G_ℓ is analytic on $\mathcal{B}_{\rho,s}(R - d)$ as well as it vector field, and*

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|X_{\mathsf{G}_\ell}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant \left(\frac{2e}{d}\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}\right)^\ell \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\mathsf{G}(\zeta)\|_{\mathcal{H}^{\rho,s}}.$$

$$\tag{185}$$

Proof. Fix $\ell \geqslant 1$. We look for a sequence $C_m^{(\ell)}$ such that

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-md/\ell)} \|X_{\mathsf{G}_m}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant C_m^{(\ell)}, \quad \forall\, m \leqslant \ell.$$

Lemma A.3 ensures that the following sequence satisfies this property.

$$C_0^{(\ell)} := \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{G}}(\zeta)\|_{\mathcal{H}^{\rho,s}},$$

$$C_m^{(\ell)} = \frac{2\ell}{\mathrm{d}m} C_{m-1}^{(\ell)} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}.$$

One has

$$C_\ell^{(\ell)} = \frac{1}{\ell!} \left(\frac{2\ell}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \right)^\ell \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{G}}(\zeta)\|_{\mathcal{H}^{\rho,s}},$$

and by using the inequality $\ell^\ell < \ell! e^\ell$ one obtains the estimate (185). $\qquad\square$

Before stating the next lemma, we point out that the Poisson tensor Ω_2^{-1}, obtained by inversion from the associated symplectic form Ω_2 in (35), is not a bounded operator on $\mathcal{H}^{\rho,s}$. We thus have to weaken the hypothesis of theorem 4.4 in [Bam99]; indeed, we just assume that

$$\|\Omega^{-1} f\|_{\mathcal{H}^{\rho,s}} \leqslant \|f\|_{\mathcal{H}^{\rho,s+1}}.$$

This property is satisfied by both Ω_1^{-1} and Ω_2^{-1}.

Lemma A.5. *Let χ and F be analytic on $\mathcal{B}_{\rho,s}(R)$ as well as their vector fields. Fix $d \in (0, R)$, and assume also that*

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant d/3.$$

Then for $|t| \leqslant 1$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|X_{(\Phi_\chi^t)^* \mathsf{F} - \mathsf{F}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant \frac{9}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{F}}(\zeta)\|_{\mathcal{H}^{\rho,s}}. \quad (186)$$

Proof. Since the bound on the norm of X_χ implies that $\Phi_\chi^t(\zeta) \in \mathcal{B}_{\rho,s}(R)$ when $\zeta \in \mathcal{B}_{\rho,s}(R - d/3)$, using Cauchy inequality and lemma A.1

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|\mathrm{d}\Phi_\chi^{-t}(\Phi_\chi^t(\zeta)) - id\|_{\mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}}$$

$$\leqslant \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-2d/3)} \|\mathrm{d}\Phi_\chi^{-t}(\zeta) - id\|_{\mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{3}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d/3)} \|\Phi_\chi^{-t}(\zeta) - \zeta\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{3}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}.$$

Since Φ_χ^t is a canonical transformation, a direct computation shows

$$\Omega^{-1}\mathrm{d}(\mathsf{F}\circ\Phi_\chi^t)(\zeta) = (\mathrm{d}\Phi_\chi^{-t}(\Phi_\chi^t(\zeta)) - id)\Omega^{-1}\mathrm{d}\mathsf{F}(\Phi_\chi^t) + \Omega^{-1}\,\mathrm{d}\mathsf{F}(\Phi_\chi^t(\zeta))$$

hence

$$\sup_{\zeta\in\mathcal{B}_{\rho,s}(R-d)}\|X_{(\Phi_\chi^t)^*\mathsf{F}-\mathsf{F}}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$= \sup_{\zeta\in\mathcal{B}_{\rho,s}(R-d)}\|\Omega^{-1}\mathrm{d}\left(\mathsf{F}\left(\Phi_\chi^t(\zeta)\right) - \mathsf{F}(\zeta)\right)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \sup_{\zeta\in\mathcal{B}_{\rho,s}(R-d)}\|\left(\mathrm{d}\Phi_\chi^{-t}\left(\Phi_\chi^t(\zeta)\right) - id\right)\Omega^{-1}\,\mathrm{d}\mathsf{F}\left(\Phi_\chi^t\right) + \Omega^{-1}\,\mathrm{d}\left(\mathsf{F}\left(\Phi_\chi^t(\zeta)\right) - \mathsf{F}(\zeta)\right)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \sup_{\zeta\in\mathcal{B}_{\rho,s}(R-d)}\|\mathrm{d}\Phi_\chi^{-t}\left(\Phi_\chi^t(\zeta)\right) - id\|_{\mathcal{H}^{\rho,s}\to\mathcal{H}^{\rho,s}}$$

$$\times \sup_{\zeta\in\mathcal{B}_{\rho,s}(R-d)}\|X_\mathsf{F}\left(\Phi_\chi^t(\zeta)\right)\|_{\mathcal{H}^{\rho,s}} + \sup_{\zeta\in\mathcal{B}_{\rho,s}(R-d)}\|X_\mathsf{F}\left(\Phi_\chi^t(\zeta)\right) - X_\mathsf{F}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{3}{d}\sup_{\zeta\in\mathcal{B}_{\rho,s}(R)}\|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}\sup_{\zeta\in\mathcal{B}_{\rho,s}(R)}\|X_\mathsf{F}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$+ \sup_{\zeta\in\mathcal{B}_{\rho,s}(R-d)}\left\|\int_0^t [X_\chi, X_\mathsf{F}]\left(\Phi_\chi^s(\zeta)\right)\mathrm{d}s\right\|_{\mathcal{H}^{\rho,s}}.$$

To estimate the last term we use Cauchy inequality

$$\sup_{\zeta\in\mathcal{B}_{\rho,s}(R-d)}\left\|\int_0^t [X_\chi, X_\mathsf{F}](\Phi_\chi^s(\zeta))\mathrm{d}s\right\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant 2\sup_{\zeta\in\mathcal{B}_{\rho,s}(R-2d/3)}\|[X_\chi, X_\mathsf{F}](\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{6}{d}\sup_{\zeta\in\mathcal{B}_{\rho,s}(R)}\|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}\sup_{\zeta\in\mathcal{B}_{\rho,s}(R)}\|X_\mathsf{F}(\zeta)\|_{\mathcal{H}^{\rho,s}}.$$

Then the thesis follows. $\qquad\square$

Lemma A.6. *Assume that* G *is analytic on* $\mathcal{B}_{\rho,s}(R)$ *as well as its vector field, and that* h_0 *satisfies* (PER). *Then there exists* χ *analytic on* $\mathcal{B}_{\rho,s}(R)$ *and* Z *analytic on* $\mathcal{B}_{\rho,s}(R)$ *with* Z *in normal form, namely* $\{\mathsf{h}_0, \mathsf{Z}\} = 0$, *such that*

$$\{\chi, \mathsf{h}_0\} + \mathsf{G} = \mathsf{Z}. \tag{187}$$

Such Z *and* χ *are given explicitly by*

$$\mathsf{Z}(\zeta) = \frac{1}{T}\int_0^T \mathsf{G}(\Phi_{\mathsf{h}_0}^t(\zeta))\mathrm{d}t, \tag{188}$$

$$\chi(\zeta) = \frac{1}{T}\int_0^T t\left[\mathsf{Z}(\Phi_{\mathsf{h}_0}^t(\zeta)) - \mathsf{G}(\Phi_{\mathsf{h}_0}^t(\zeta))\right]\mathrm{d}t. \tag{189}$$

Furthermore, we have that the vector fields of χ and Z are analytic on $\mathcal{B}_{\rho,s}(R)$, and satisfy

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{Z}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{G}}(\zeta)\|_{\mathcal{H}^{\rho,s}},$$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\chi}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant 2T \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{G}}(\zeta)\|_{\mathcal{H}^{\rho,s}}. \tag{190}$$

Proof. We check directly that the solution of (187) is (189). Indeed,

$$\begin{aligned}
\{\chi, \mathsf{h}_0\}(\zeta) &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \chi(\Phi_{\mathsf{h}_0}^s(\zeta)) \\
&= \frac{1}{T}\int_0^T t \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \left[\mathsf{Z}(\Phi_{\mathsf{h}_0}^{t+s}(\zeta)) - \mathsf{G}(\Phi_{\mathsf{h}_0}^{t+s}(\zeta)) \right] \mathrm{d}t \\
&= \frac{1}{T}\int_0^T t \frac{\mathrm{d}}{\mathrm{d}t} \left[\mathsf{Z}(\Phi_{\mathsf{h}_0}^t(\zeta)) - \mathsf{G}(\Phi_{\mathsf{h}_0}^t(\zeta)) \right] \mathrm{d}t \\
&= \frac{1}{T}\left[t\mathsf{Z}(\Phi_{\mathsf{h}_0}^t(\zeta)) - t\mathsf{G}(\Phi_{\mathsf{h}_0}^t(\zeta)) \right]_{t=0}^T - \frac{1}{T}\int_0^T \left[\mathsf{Z}(\Phi_{\mathsf{h}_0}^t(\zeta)) - \mathsf{G}(\Phi_{\mathsf{h}_0}^t(\zeta)) \right] \mathrm{d}t \\
&= \mathsf{Z}(\zeta) - \mathsf{G}(\zeta).
\end{aligned}$$

In the last step we used the explicit expression of Z provided in (188). Finally, the first estimate in (190) follows from the explicit expression of Z in (188) while for the second estimate we write explicitly the vector field X_χ:

$$X_\chi(\zeta) = \frac{1}{T}\int_0^T tD\Phi_{\mathsf{h}_0}^{-t}(\Phi_{\mathsf{h}_0}^t(\zeta)) \circ X_{\mathsf{Z}-\mathsf{G}}(\Phi_{\mathsf{h}_0}^t(\zeta))\mathrm{d}t.$$

Hypothesis (PER) guarantees that $\Phi_{\mathsf{h}_0}^t$ as well as its derivatives and the inverses are uniformly bounded as operators from $\mathcal{H}^{\rho,s}$ into itself. Moreover, for any $t \in \mathbb{R}$, the map $\zeta \mapsto \Phi_{\mathsf{h}_0}^t(\zeta)$ is a diffeomorphism of $\mathcal{B}_{\rho,s}(R)$ into itself. Using the fact that $\Phi_{\mathsf{h}_0}^t$ is an isometry, we have

$$\begin{aligned}
\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} &\leqslant T \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \left(\|X_{\mathsf{Z}}(\zeta)\|_{\mathcal{H}^{\rho,s}} + \|X_{\mathsf{G}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \right) \\
&\leqslant 2T \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{G}}(\zeta)\|_{\mathcal{H}^{\rho,s}}
\end{aligned}$$

where in the last step we used the first inequality in (190). $\qquad\square$

Lemma A.7. *Assume that G and its vector fields are analytic on $\mathcal{B}_{\rho,s}(R)$, and that h_0 satisfies* (PER). *Let χ and its vector field be analytic on $\mathcal{B}_{\rho,s}(R)$, and assume that χ solves* (187). *For any $\ell \geqslant 1$ denote by $\mathsf{h}_{0,\ell}$ the functions defined recursively as in* (183) *from h_0. Then for any $d \in (0, R)$ one has that $\mathsf{h}_{0,\ell}$ and its vector field are analytic on $\mathcal{B}_{\rho,s}(R - d)$, and*

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|X_{\mathsf{h}_{0,\ell}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant 2 \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_{\mathsf{G}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \left(\frac{9}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \right)^\ell. \tag{191}$$

Proof. By using (187) one gets that $\mathsf{h}_{0,1} = \mathsf{Z} - \mathsf{G}$ is analytic on $\mathcal{B}_{\rho,s}(R)$. Then by exploiting (186) one gets the result. $\qquad\square$

Lemma A.8.　*Assume that G and its vector field are analytic on $\mathcal{B}_{\rho,s}(R)$, and that h_0 satisfies (PER). Let χ be the solution of (187), denote by Φ_χ^t the flow of the Hamiltonian vector field associated to χ and by Φ_χ the corresponding time-one map. Moreover, denote by*

$$\mathcal{F}(\zeta) := \mathsf{h}_0(\Phi_\chi(\zeta)) - \mathsf{h}_0(\zeta) - \{\chi, \mathsf{h}_0\}(\zeta).$$

Let $d < R$, and assume that

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant d/3.$$

Then we have that $\mathcal{F}$ and its vector field are analytic on $\mathcal{B}_{\rho,s}(R-d)$, and

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|X_\mathcal{F}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant \frac{18}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \|X_\mathsf{G}(\zeta)\|_{\mathcal{H}^{\rho,s}}. \qquad (192)$$

Proof.　Since

$$\mathsf{h}_0(\Phi_\chi(\zeta)) - \mathsf{h}_0(\zeta) = \int_0^1 \{\chi, \mathsf{h}_0\} \circ \Phi_\chi^t(\zeta) \mathrm{d}t$$

$$\overset{(187)}{=} \int_0^1 \mathsf{Z}(\Phi_\chi^t(\zeta)) - \mathsf{G}(\Phi_\chi^t(\zeta)) \mathrm{d}t,$$

if we define $\mathsf{F}(\zeta) := \mathsf{Z}(\zeta) - \mathsf{G}(\zeta)$, we get

$$\mathcal{F}(\zeta) = \int_0^1 \left(\mathsf{F}(\Phi_\chi^t(\zeta)) - \mathsf{F}(\zeta)\right) \mathrm{d}t.$$

Now, we have

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|X_\mathcal{F}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$= \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \left\| \Omega^{-1} \mathrm{d} \left[\int_0^1 \left(\mathsf{F}(\Phi_\chi^t(\zeta)) - \mathsf{F}(\zeta)\right) \mathrm{d}t \right] \right\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \left\| \int_0^1 \left[(\mathrm{d}\Phi_\chi^{-t}(\Phi_\chi^t(\zeta)) - \mathrm{i}d)\Omega^{-1} \, \mathrm{d}\mathsf{F}(\Phi_\chi^t) \right. \right.$$

$$\left. \left. + \Omega^{-1} \, \mathrm{d}(\mathsf{F}(\Phi_\chi^t(\zeta)) - \mathsf{F}(\zeta)) \right] \mathrm{d}t \right\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \left\| \int_0^1 (\mathrm{d}\Phi_\chi^{-t}(\Phi_\chi^t(\zeta)) - \mathrm{i}d)\Omega^{-1} \, \mathrm{d}\mathsf{F}(\Phi_\chi^t) \mathrm{d}t \right\|_{\mathcal{H}^{\rho,s}}$$

$$+ \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \left\| \int_0^1 \left(X_\mathsf{F}(\Phi_\chi^t(\zeta)) - X_\mathsf{F}(\zeta)\right) \mathrm{d}t \right\|_{\mathcal{H}^{\rho,s}}$$

and by dominated convergence we can bound the last quantity by

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \sup_{t \in [0,1]} \|\mathrm{d}\Phi_\chi^{-t}(\Phi_\chi^t(\zeta)) - \mathrm{i}d\|_{\mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}}$$

$$\times \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \|X_\mathsf{F}(\Phi_\chi^t(\zeta))\|_{\mathcal{H}^{\rho,s}}$$

$$+ \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \sup_{t \in [0,1]} \| X_{\mathsf{F}}(\Phi_\chi^t(\zeta)) - X_{\mathsf{F}}(\zeta) \|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \sup_{t \in [0,1]} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \| \mathrm{d}\Phi_\chi^{-t}(\Phi_\chi^t(\zeta)) - \mathrm{id} \|_{\mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}}$$

$$\times \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \| X_{\mathsf{F}}(\Phi_\chi^t(\zeta)) \|_{\mathcal{H}^{\rho,s}}$$

$$+ \sup_{t \in [0,1]} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \| X_{\mathsf{F}}(\Phi_\chi^t(\zeta)) - X_{\mathsf{F}}(\zeta) \|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{3}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \| X_\chi(\zeta) \|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \| X_{\mathsf{F}}(\zeta) \|_{\mathcal{H}^{\rho,s}}$$

$$+ \sup_{t \in [0,1]} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \left\| \int_0^t [X_\chi, X_{\mathsf{F}}](\Phi_\chi^s(\zeta)) \mathrm{d}s \right\|_{\mathcal{H}^{\rho,s}},$$

where we can estimate the last term by Cauchy inequality

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \left\| \int_0^t [X_\chi, X_{\mathsf{F}}](\Phi_\chi^s(\zeta)) \mathrm{d}s \right\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant 2 \sup_{\zeta \in \mathcal{B}_{\rho,s}(R-2d/3)} \| [X_\chi, X_{\mathsf{F}}](\zeta) \|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{6}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \| X_\chi(\zeta) \|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \| X_{\mathsf{F}}(\zeta) \|_{\mathcal{H}^{\rho,s}}.$$

By the above computations and (190) we obtain

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R-d)} \| X_{\mathcal{F}}(\zeta) \|_{\mathcal{H}^{\rho,s}} \leqslant \frac{9}{d} \sup_{\mathcal{B}_{\rho,s}(R)} \| X_\chi(\zeta) \|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \| X_{\mathsf{F}}(\zeta) \|_{\mathcal{H}^{\rho,s}}$$

$$\overset{(190)}{\leqslant} \frac{18}{d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \| X_\chi(\zeta) \|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \| X_{\mathsf{G}}(\zeta) \|_{\mathcal{H}^{\rho,s}}.$$

$\square$

Lemma A.9. *Let $s \geqslant s_1 \gg 1, R > 0, m \geqslant 0, 0 < \nu \leqslant 1$, and consider the Hamiltonian*

$$H^{(m)}(\zeta) = \mathsf{h}_0(\zeta) + \delta \mathsf{Z}^{(m)}(\zeta) + \delta^{\nu m+1} \mathsf{F}^{(m)}(\zeta). \tag{193}$$

Assume that h_0 satisfies (PER) and (INV), and that

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R)} \| X_{\mathsf{F}^{(0)}}(\zeta) \|_{\mathcal{H}^{\rho,s}} \leqslant F.$$

Fix $d < \frac{R}{m+1}$, and set $R_m := R - md \ (m \geqslant 1)$.
Assume also that $\mathsf{Z}^{(m)}$ is analytic on $\mathcal{B}_{\rho,s}(R_m)$, and that

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \| X_{\mathsf{Z}^{(0)}}(\zeta) \|_{\mathcal{H}^{\rho,s}} = 0,$$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \| X_{\mathsf{Z}^{(m)}}(\zeta) \|_{\mathcal{H}^{\rho,s}} \leqslant F \sum_{i=0}^{m-1} \delta^{\nu i} K_0^i, \quad m \geqslant 1, \tag{194}$$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\mathsf{F}^{(m)}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant FK_0^m, \quad m \geqslant 1, \tag{195}$$

with $K_0 \geqslant 15$ and $d > 3T\delta F$.

Then, if $\delta^\nu K_0 < 1/2$ there exists a canonical transformation $\mathcal{T}_\delta^{(m)}$ analytic on $\mathcal{B}_{\rho,s}(R_{m+1})$ such that

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_{m+1})} \|\mathcal{T}_\delta^{(m)}(\zeta) - \zeta\|_{\mathcal{H}^{\rho,s}} \leqslant 2T\delta^{\nu m+1} K_0^m F, \tag{196}$$

$\mathsf{H}^{(m+1)} := \mathsf{H}^{(m)} \circ \mathcal{T}^{(m)}$ *has the form* (193) *and satisfies* (195) *with m replaced by* $m + 1$.

Proof. The key point of the proof is to look for $\mathcal{T}_\delta^{(m)}$ as the time-one map of the Hamiltonian vector field of an analytic function $\delta^{\nu m+1}\chi_m$. Hence, consider the differential equation

$$\dot{\zeta} = X_{\delta^{\nu m+1}\chi_m}(\zeta). \tag{197}$$

By standard theory we have that, if $\|X_{\delta^{\nu m+1}\chi_m}\|_{\mathcal{B}_{\rho,s}(R_m)}$ is small enough (e.g. $\|X_{\delta^{\nu m+1}\chi_m}\|_{\mathcal{B}_{\rho,s}(R_m)} \leqslant \frac{md}{m+1}$) and $\zeta_0 \in \mathcal{B}_{\rho,s}(R_{m+1})$, then the solution of (197) exists for $|t| \leqslant 1$.

Therefore we can define $\mathcal{T}_{m,\delta}^t : \mathcal{B}_{\rho,s}(R_{m+1}) \to \mathcal{B}_{\rho,s}(R_m)$, and in particular the corresponding time-one map $\mathcal{T}_\delta^{(m)} := \mathcal{T}_{m,\delta}^1$, which is an analytic canonical transformation, $\delta^{\nu m+1}$-close to the identity. We have

$$\begin{aligned}
(\mathcal{T}_\delta^{(m)})^*(\mathsf{h}_0 + \delta\mathsf{Z}^{(m)} + \delta^{m+1}\mathsf{F}^{(m)}) &= \mathsf{h}_0 + \delta\mathsf{Z}^{(m)} + \delta^{\nu m+1}\left[\{\chi_m, \mathsf{h}_0\} + \mathsf{F}^{(m)}\right] \\
&\quad + \left(\mathsf{h}_0 \circ \mathcal{T}_\delta^{(m)} - \mathsf{h}_0 - \delta^{\nu m+1}\{\chi_m, \mathsf{h}_0\}\right) \\
&\quad + \delta\left(\mathsf{Z}^{(m)} \circ \mathcal{T}_\delta^{(m)} - \mathsf{Z}^{(m)}\right) \tag{198}
\end{aligned}$$

$$+ \delta^{\nu m+1}\left(\mathsf{F}^{(m)} \circ \mathcal{T}_\delta^{(m)} - \mathsf{F}^{(m)}\right). \tag{199}$$

It is easy to see that the first two terms on the right-hand side are already normalized, that the third term is the non-normalized part of order $m + 1$ that can be normalized through the choice of a suitable χ_m, and that (198) and (199) contain all the terms of order higher than $m + 1$.

In order to normalize the third term on the right-hand side we solve the homological equation

$$\{\chi_m, \mathsf{h}_0\} + \mathsf{F}^{(m)} = \mathsf{Z}_{m+1},$$

with Z_{m+1} in normal form. Lemma A.6 ensures the existence of χ_m and Z_{m+1} as well as their explicit expressions:

$$\mathsf{Z}_{m+1}(\zeta) = \frac{1}{T}\int_0^T \mathsf{F}^{(m)}(\Phi_{\mathsf{h}_0}^t(\zeta))\mathrm{d}t,$$

$$\chi_m(\zeta) = \frac{1}{T}\int_0^T t[\mathsf{F}^{(m)}(\Phi_{\mathsf{h}_0}^t(\zeta)) - \mathsf{Z}_{m+1}(\Phi_{\mathsf{h}_0}^t(\zeta))]\mathrm{d}t.$$

The explicit expression of X_{χ_m} can be computed following the argument of lemma A.6. Using this explicit expression, the analyticity of the flow $\Phi_{\mathsf{h}_0}^t$ ensured by (PER) and (190) one has

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\chi_m}(\zeta)\|_{\mathcal{H}^{\rho,\sigma}} \leqslant 2T \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\mathsf{F}^{(m)}}\|_{\mathcal{H}^{\rho,\sigma}} \leqslant 2TK_0^m F. \tag{$*$}$$

Straightforwardly, from the explicit expression of $Z_{m+1}(\zeta)$ and (195) one has

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{Z_{m+1}}\|_{\mathcal{H}^{\rho,s}} \leqslant K_0^m F.$$

Now define $\mathsf{Z}^{(m+1)} := \mathsf{Z}^{(m)} + \delta^{\nu m} \mathsf{Z}_{m+1}$ and notice that as a consequence of the latter estimate and (194) we have

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_{m+1})} \|X_{\mathsf{Z}^{(m+1)}}(\zeta)\| \leqslant \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_{m+1})} \|X_{\mathsf{Z}^{(m)}}(\zeta)\|_{\mathcal{H}^{\rho,s}} + \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_{m+1})} \|X_{\delta^{\nu m} \mathsf{Z}_{m+1}}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant F \left(\sum_{j=0}^{m-1} \delta^{\nu j} K_0^j + \delta^{\nu m} K_0^m \right).$$

Defining now $\mathcal{T}_\delta^{(m)}(\zeta) := \Phi^1_{\delta^{\nu m+1} \chi_m}(\zeta)$ we can apply lemma A.1 and $(*)$ to obtain

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_{m+1})} \|\mathcal{T}_\delta^{(m)}(\zeta) - \zeta\|_{\mathcal{H}^{\rho,s}} = \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_{m+1})} \|\Phi^1_{\delta^{\nu m+1} \chi_m}(\zeta) - \zeta\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\delta^{\nu m+1} \chi_m}\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant 2T \delta^{\nu m+1} K_0^m F.$$

Let us set now $\delta^{\nu(m+1)+1} \mathsf{F}^{(m+1)} := (198) + (199)$. Using lemma A.5 one can estimate separately the three pieces. We notice that $\sup_{\mathcal{B}_{\rho,s}(R_m)} \|X_{\delta^{\nu m+1} \chi_m}\|_{\mathcal{H}^{\rho,s}} \leqslant 2T \delta^{\nu m+1} K_0^m F$ and since $\delta^\nu K_0 < \frac{1}{2}$ we have $\sup_{\mathcal{B}_{\rho,s}(R_m)} \|X_{\delta^{\nu m+1} \chi_m}\|_{\mathcal{H}^{\rho,s}} < T\delta F < \frac{d}{3} \leqslant \frac{(m+1)d}{3}$. We can thus apply lemmas A.5 and A.8 to get

$$\sup_{\zeta \in \mathcal{B}(R_{m+1})} \|X_{\mathsf{Z}^{(m)} \circ \mathcal{T}_\delta^{(m)} - \mathsf{Z}^{(m)}}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{27 \delta^{\nu m+1}}{(m+1)d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\chi_m}(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\mathsf{Z}^{(m)}}\|_{\mathcal{H}^{\rho,s}},$$

$$\sup_{\zeta \in \mathcal{B}(R_{m+1})} \|X_{\mathsf{F}^{(m)} \circ \mathcal{T}_\delta^{(m)} - \mathsf{F}^{(m)}}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{27 \delta^{\nu m+1}}{(m+1)d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\chi_m}(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\mathsf{F}^{(m)}}\|_{\mathcal{H}^{\rho,s}},$$

$$\sup_{\zeta \in \mathcal{B}(R_{m+1})} \|X_{\mathsf{h}_0 \circ \mathcal{T}_\delta^{(m)} - \mathsf{h}_0 - \delta^{\nu m+1} \{\chi_m, \mathsf{h}_0\}}\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \frac{18 \, \delta^{2m\nu+2}}{(m+1)d} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\chi_m}(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\mathsf{F}^{(m)}}(\zeta)\|_{\mathcal{H}^{\rho,s}}.$$

By means of these inequalities and by exploiting $\|X_{\delta^{\nu m+1} \chi_m}\|_{\mathcal{H}^{\rho,s}} \leqslant \frac{(m+1)d}{3}$ together with assumptions (194) and (195), we can estimate

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_{m+1})} \|X_{\delta^{\nu(m+1)+1} \mathsf{F}^{(m+1)}}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant 9 \delta^{\nu(m+1)+1} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\mathsf{Z}^{(m)}}(\zeta)\|_{\mathcal{H}^{\rho,s}} + 9 \delta^{2m\nu+2} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\mathsf{F}^{(m)}}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$+ 6\delta^{2m\nu+2} \sup_{\zeta \in \mathcal{B}_{\rho,s}(R_m)} \|X_{\mathsf{F}^{(m)}}(\zeta)\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant 9\delta^{\nu(m+1)+1} F \sum_{i=0}^{m-1} \delta^i K_0^i + 9\delta^{2m\nu+2} F K_0^m + 6\delta^{2m\nu+2} F K_0^m$$

$$= \delta^{\nu(m+1)+1} \left(9F \sum_{i=0}^{m-1} \delta^i K_0^i + 9\delta^{\nu(m-1)+1} F K_0^m + 6\delta^{\nu(m-1)+1} F K_0^m \right).$$

If $m = 0$ then we have

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_1)} \|X_{\delta^{1+\nu}\mathsf{F}^{(1)}}\|_{\mathcal{H}^{\rho,s}} \leqslant \delta^{1+\nu}(9F + 6F).$$

If $m \geqslant 1$ we exploit the smallness condition $\delta^\nu K_0 < \frac{1}{2}$ to get $\sum_{i=0}^{m-1} \delta^{\nu i} K_0^i < 2$ and

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(R_{m+1})} \|X_{\delta^{\nu(m+1)+1}\mathsf{F}^{(m+1)}}\|_{\mathcal{H}^{\rho,s}}$$

$$\leqslant \delta^{\nu(m+1)+1} \left(6F + 9\delta K_0 \frac{F}{2^{m-1}} + 6\delta K_0 \frac{F}{2^{m-1}} \right), \tag{200}$$

and since $0 < \nu \leqslant 1$ the right-hand side of (200) can be bounded by

$$\delta^{\nu(m+1)+1} \left(6F + 9\delta^{1-\nu} \frac{F}{2^m} + 6\delta^{1-\nu} \frac{F}{2^m} \right) \leqslant 15\delta^{\nu(m+1)+1} F.$$

$$\square$$

Proof of lemma 3.6.　The Hamiltonian (48) satisfies the assumptions of lemma A.9 with $m = 0$, $\mathsf{F}_{1,M}$ in place of $\mathsf{F}^{(0)}$, $F = K_{1,s}^{(F)} M^{2+\Gamma}$. So we apply lemma A.9 with $d = R/4$, provided that

$$\delta < \frac{R}{12TF} = \frac{R}{12TK_{1,s}^{(F)} M^{2+\Gamma}}$$

which is true due to (57). Hence there exists an analytic canonical transformation $\mathcal{T}_{\delta,M}^{(0)} : \mathcal{B}_{\rho,s}(3R/4) \to \mathcal{B}_{\rho,s}(R)$ with

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(3R/4)} \|\mathcal{T}_{\delta,M}^{(0)}(\zeta) - \zeta\|_{\mathcal{H}^{\rho,s}} \leqslant 2TF\delta,$$

such that

$$\mathsf{H}_{1,M} \circ \mathcal{T}_{\delta,M}^{(0)} = \mathsf{h}_0 + \delta\mathsf{Z}_M^{(1)} + \delta^{1+\nu}\mathcal{R}_M^{(1)}, \tag{201}$$

$$\mathsf{Z}_M^{(1)} := \langle \mathsf{F}_{1,M} \rangle, \tag{202}$$

$$\delta^{1+\nu}\mathcal{R}_M^{(1)} := \delta^{1+\nu}\mathsf{F}^{(1)}$$

$$= \left(\mathsf{h}_0 \circ \mathcal{T}_{\delta,M}^{(0)} - \mathsf{h}_0 - \delta\{\chi_1, \mathsf{h}_0\} \right) + \delta \left(\mathsf{Z}_M^{(1)} \circ \mathcal{T}_{\delta,M}^{(0)} - \mathsf{Z}_M^{(1)} \right)$$

$$+ \delta^{1+\nu} \left(\mathsf{F}_{1,M} \circ \mathcal{T}_{\delta,M}^{(0)} - \mathsf{F}_{1,M} \right), \tag{203}$$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(3R/4)} \|X_{Z_M^{(1)}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant F, \tag{204}$$

$$\sup_{\zeta \in \mathcal{B}_{\rho,s}(3R/4)} \|X_{\mathcal{R}_N^{(1)}}(\zeta)\|_{\mathcal{H}^{\rho,s}} \leqslant 15F. \tag{205}$$

and $K_0 = 15$, hence $\delta < \frac{1}{30^{1/\nu}}$. $\qquad\qquad\square$

Appendix B. Proof of propositions 6.5 and 6.2

Proof of proposition 6.5. In order to prove proposition 6.5 we first discuss the specific energies associated to the high modes, and then the ones associated to the low modes.

First we remark that for all k such that $\kappa(k) = (\mu K_1, \mu^\sigma K_2)$ we have

$$
\begin{aligned}
\left|\frac{\omega_k^2}{\mu^2}\right| &\overset{(8)}{=} \frac{4}{\mu^2}\left[\sin^2\left(\frac{k_1\pi}{2N_1+1}\right) + \sin^2\left(\frac{k_2\pi}{2N_2+1}\right)\right] \\
&= \frac{4}{\mu^2}\left[\sin^2\left(\frac{\mu K_1\pi}{2}\right) + \sin^2\left(\frac{\mu^\sigma K_2\pi}{2}\right)\right] \\
&\leqslant \pi^2(K_1^2 + \mu^{2(\sigma-1)}K_2^2);
\end{aligned}
\tag{206}
$$

moreover, for $K_1 \neq 0$

$$
\begin{aligned}
&\frac{|\hat{q}_K|^2 + \pi^2(K_1^2 + \mu^{2(\sigma-1)}K_2^2)|\hat{p}_K|^2}{2} \\
&\leqslant \pi^2\, e^{-2\rho|K|}\frac{|\hat{q}_K|^2 + (K_1^2 + \mu^{2(\sigma-1)}K_2^2)|\hat{p}_K|^2}{2}\, e^{2\rho|K|} \\
&\leqslant \pi^2\, e^{-2\rho|K|}\left(1 + \mu^{2(\sigma-1)}\frac{K_2^2}{K_1^2}\right)\|(\xi,\eta)\|_{\mathcal{H}^{\rho,0}}^2,
\end{aligned}
\tag{207}
$$

while for $|K_2| \leqslant |K_1|$

$$\frac{|\hat{q}_K|^2 + \pi^2(K_1^2 + \mu^{2(\sigma-1)}K_2^2)|\hat{p}_K|^2}{2} \overset{|K_2|\leqslant|K_1|}{\leqslant} 2\pi^2\, e^{-2\rho|K|}\|(\xi,\eta)\|_{\mathcal{H}^{\rho,0}}^2. \tag{208}$$

In order to estimate $\mathcal{E}_\kappa$ for large K_1 and K_2, it is convenient to divide the frequency-space in different regions and bound the terms supported in each region separately. Unlike the one-dimensional case, only few terms will turn out to be exponentially small with respect to μ, so the introduction of different regions in the frequency space will help us estimating most of the terms in an efficient way. Let us define

$$
\mathscr{L}_{\mu,\delta,\rho} := \left\{ L = (L_1, L_2) \in \mathbb{Z}^2 : \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}, |K_1| \right.
$$
$$
\left. + |K_2| > \frac{(2+\delta)|\log\mu|}{\rho} \right\}.
\tag{209}
$$

Using the explicit expression (150) we obtain

$$
\frac{\mathcal{E}_\kappa}{\mu^4} = \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2| \leqslant |K_1+L_1|}} \left(|\hat{q}_{K+L}|^2 + \omega_k^2 \left| \frac{\hat{p}_{K+L}}{\mu} \right|^2 \right)
$$

$$
+ \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2| > |K_1+L_1|}} \left(|\hat{q}_{K+L}|^2 + \omega_k^2 \left| \frac{\hat{p}_{K+L}}{\mu} \right|^2 \right)
$$

$$
\overset{(206),(208),(77)}{\leqslant} \pi^2 \|(\xi,\eta)\|_{\mathcal{H}^{\rho,0}}^2 2 \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2| \leqslant |K_1+L_1|}} \mathrm{e}^{-2\rho|K+L|}
$$

$$
+ \pi^2 \|(\xi,\eta)\|_{\mathcal{H}^{\rho,0}}^2 \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2| > |K_1+L_1| \\ K_1+L_1 \neq 0}} \mathrm{e}^{-2\rho|K+L|}
$$

$$
\times \left(1 + \mu^{2(\sigma-1)} \frac{(K_2+L_2)^2}{(K_1+L_1)^2} \right) + \pi^2 \|(\xi,\eta)\|_{\mathcal{H}^{\rho,0}}^2
$$

$$
\times \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2| > |K_1+L_1| \\ K_1+L_1 = 0}} \mathrm{e}^{-2\rho|K_2+L_2|}.
$$

Now,

$$
\sum_{L \in \mathscr{L}_{\mu,\delta,\rho}} \mathrm{e}^{-2\rho|K+L|} \leqslant \mathrm{e}^{-2\rho|K|} + \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1=0, L_2 \neq 0}} \mathrm{e}^{-2\rho|K+L|}
$$

$$
+ \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1 \neq 0, L_2 = 0}} \mathrm{e}^{-2\rho|K+L|} + \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1, L_2 \neq 0}} \mathrm{e}^{-2\rho|K+L|}. \tag{210}
$$

We now estimate the last sum in (210); we point out that for $L_1, L_2 \neq 0$ we have $|L| \geqslant \frac{2}{\mu} + \frac{2}{\mu^\sigma}$, hence $2|K| \leqslant |L|$.

Therefore, for any k such that $\kappa(k) = (\mu K_1,\ \mu^\sigma K_2)$ and $|K_1| + |K_2| \geqslant \frac{(2+\delta)|\log \mu|}{\rho}$

$$
\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1, L_2 \neq 0}} \mathrm{e}^{-2\rho|K+L|} \leqslant \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1, L_2 \neq 0}} \mathrm{e}^{-2\rho||K|-|L||} \leqslant \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1, L_2 \neq 0}} \mathrm{e}^{2\rho|K|} \, \mathrm{e}^{-2\rho|L|}
$$

$$
\leqslant \mathrm{e}^{2\rho|K|} 2\pi \int_{2|K|}^{+\infty} R \, \mathrm{e}^{-2\rho R} \, \mathrm{d}R
$$

$$
= 2\pi \, \mathrm{e}^{2\rho|K|} \left(-\frac{1}{2} \right) \frac{\mathrm{d}}{\mathrm{d}\rho} \left[\int_{2|K|}^{+\infty} \mathrm{e}^{-2\rho R} \, \mathrm{d}R \right]
$$

$$
= -\pi \, \mathrm{e}^{2\rho|K|} \frac{\mathrm{d}}{\mathrm{d}\rho} \left(\frac{\mathrm{e}^{-4\rho|K|}}{2\rho} \right) = \frac{\pi}{2\rho} \left(\frac{1}{\rho} + 4 \right) \mathrm{e}^{-2\rho|K|}. \tag{211}
$$

Next we estimate the second sum in (210); we have

$$\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1 \neq 0, L_2 = 0}} e^{-2\rho|K+L|} \leqslant e^{-2\rho(|K_1|+|K_2|)} \sum_{\ell \in \mathbb{Z}\setminus\{0\}} e^{-4\rho|\ell|/\mu}, \tag{212}$$

which is exponentially small with respect to μ. Similarly,

$$\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1 = 0, L_2 \neq 0}} e^{-2\rho|K+L|} \leqslant e^{-2\rho(|K_1|+|K_2|)} \sum_{\ell \in \mathbb{Z}\setminus\{0\}} e^{-4\rho|\ell|/\mu^\sigma}. \tag{213}$$

Then,

$$\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2|>|K_1+L_1| \\ K_1+L_1 \neq 0}} e^{-2\rho|K+L|} \frac{(K_2+L_2)^2}{(K_1+L_1)^2}$$

$$\leqslant e^{-2\rho|K|}\left(\frac{K_2}{K_1}\right)^2 + \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2|>|K_1+L_1| \\ K_1+L_1 \neq 0 \\ L_1 \neq 0, L_2 = 0}} e^{-2\rho|K+L|} \frac{(K_2+L_2)^2}{(K_1+L_1)^2}$$

$$+ \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2|>|K_1+L_1| \\ K_1+L_1 \neq 0 \\ L_1 = 0, L_2 \neq 0}} e^{-2\rho|K+L|} \frac{(K_2+L_2)^2}{(K_1+L_1)^2}$$

$$+ \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2|>|K_1+L_1| \\ K_1+L_1 \neq 0 \\ L_1, L_2 \neq 0}} e^{-2\rho|K+L|} \frac{(K_2+L_2)^2}{(K_1+L_1)^2}. \tag{214}$$

First we estimate the last term in (214): we have that $|L+K| \geqslant |K|$, hence

$$\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ |K_2+L_2|>|K_1+L_1| \\ K_1+L_1 \neq 0 \\ L_1, L_2 \neq 0}} e^{-2\rho|K+L|} \frac{(K_2+L_2)^2}{(K_1+L_1)^2}$$

$$= \int_{|K|}^{+\infty} \int_0^{\pi/4} e^{-2\rho\xi} \xi \tan^2\phi \, \mathrm{d}\phi \, \mathrm{d}\xi$$

$$= \left(1 - \frac{\pi}{4}\right) e^{-2\rho|K|} \frac{1 + 2\rho|K|}{4\rho^2}$$

$$\leqslant \left(1 - \frac{\pi}{4}\right) \mu^4 e^{-2\rho\left[|K| - \frac{2|\log\mu|}{\rho} - \frac{1}{2\rho}\log(2\rho|K|)\right]}$$

$$\overset{\delta < 1 - 1/e}{\leqslant} \left(1 - \frac{\pi}{4}\right) \mu^4 e^{-2\rho\left[\delta|K| - \frac{2|\log\mu|}{\rho}\right]}$$

$$= \left(1 - \frac{\pi}{4}\right) \mu^8 e^{-2\rho\delta|K|}. \tag{215}$$

Now we bound the other two nontrivial terms in (214); on the one hand, we notice that

$$
\sum_{\substack{L\in\mathscr{L}_{\mu,\delta,\rho}\\ |K_2+L_2|>|K_1+L_1|\\ K_1+L_1\neq 0\\ L_1\neq 0,L_2=0}} \mathrm{e}^{-2\rho|K+L|}\frac{(K_2+L_2)^2}{(K_1+L_1)^2}
$$

$$
= K_2^2 \sum_{\substack{L\in\mathscr{L}_{\mu,\delta,\rho}\\ |K_2+L_2|>|K_1+L_1|\\ K_1+L_1\neq 0\\ L_1\neq 0,L_2=0}} \mathrm{e}^{-2\rho|K+L|}\frac{1}{(K_1+L_1)^2}
$$

$$
\leqslant K_2^2\, \mathrm{e}^{-2\rho|K|} \sum_{\ell\in\mathbb{Z}\setminus\{0\}} \mathrm{e}^{-4\rho|\ell|/\mu}
$$

$$
\leqslant 2K_2^2\, \mathrm{e}^{-2\rho|K|} \int_1^{+\infty} \mathrm{e}^{-4\rho|\ell|/\mu}\, \mathrm{d}\ell, \tag{216}
$$

where the last integral is exponentially small with respect to μ, while on the other hand

$$
\sum_{\substack{L\in\mathscr{L}_{\mu,\delta,\rho}\\ |K_2+L_2|>|K_1+L_1|\\ K_1+L_1\neq 0\\ L_1\neq 0,L_2=0}} \mathrm{e}^{-2\rho|K+L|}\frac{(K_2+L_2)^2}{(K_1+L_1)^2}
$$

$$
\leqslant \frac{1}{K_1^2} \sum_{\substack{L\in\mathscr{L}_{\mu,\delta,\rho}\\ |K_2+L_2|>|K_1+L_1|\\ K_1+L_1\neq 0\\ L_1=0,L_2\neq 0}} \mathrm{e}^{-2\rho|K+L|}(K_2+L_2)^2
$$

$$
\leqslant \frac{1}{K_1^2}\, \mathrm{e}^{-2\rho|K|} \sum_{\ell\in\mathbb{Z}\setminus\{0\}} \mathrm{e}^{-4\rho|\ell|/\mu^\sigma}\left(K_1^2 + 2K_1\frac{\ell}{\mu^\sigma} + \frac{\ell^2}{\mu^{2\sigma}}\right)
$$

$$
\leqslant \frac{2\,\mathrm{e}^{-2\rho|K|}}{K_1^2} \int_1^{+\infty} \mathrm{e}^{-4\rho|\ell|/\mu^\sigma}\left(K_1^2 + 2K_1\frac{\ell}{\mu^\sigma} + \frac{\ell^2}{\mu^{2\sigma}}\right)\mathrm{d}\ell, \tag{217}
$$

and the last integral is again exponentially small with respect to μ.

On the other hand, for any k such that $\kappa(k)=(\mu K_1,\ \mu^\sigma K_2)$ and $|K_1|+|K_2|\leqslant \frac{(2+\delta)|\log\mu|}{\rho}$

$$
\left|\frac{\mathcal{E}_\kappa}{\mu^4} - \frac{|\hat{\xi}_K|^2 + |\hat{\eta}_K|^2}{2}\right|
$$

$$
\leqslant \left|\frac{\omega_k^2 - \mu^2\pi^2 K_1^2}{2\mu^2}\right| |\hat{p}_K|^2 + \frac{1}{2}\sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\}\\ \mu L_1,\mu^\sigma L_2\in 2\mathbb{Z}}}\left(|\hat{q}_{K+L}|^2 + \omega_k^2\left|\frac{\hat{p}_{K+L}}{\mu}\right|^2\right),
$$

$$
\overset{(206)}{\leqslant} (\mu^2\pi^4 K_1^4 + \pi^2\mu^{2(\sigma-1)}K_2^2)|\hat{p}_K|^2
$$

$$
+ \frac{1}{2}\sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\}\\ \mu L_1,\mu^\sigma L_2\in 2\mathbb{Z}}}\left(|\hat{q}_{K+L}|^2 + \pi^2\left[(K_1+L_1)^2\right.\right.
$$

$$+ \mu^{2(\sigma-1)}(K_2 + L_2)^2 \big] |\hat{p}_{K+L}|^2 \Big) ,$$

$$\leqslant \left(\pi^4 \mu^2 K_1^4 + \pi^2 \mu^{2(\sigma-1)} \frac{9|\log \mu|^2}{\rho^2} \right) |\hat{p}_K|^2$$

$$+ \frac{1}{2} \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \left(|\hat{q}_{K+L}|^2 + \pi^2 \big[(K_1 + L_1)^2 \right.$$

$$\left. + \mu^{2(\sigma-1)}(K_2 + L_2)^2 \big] |\hat{p}_{K+L}|^2 \right) ,$$

$$\leqslant \left(\pi^4 \mu^2 + \pi^2 \mu^{2(\sigma-1)} \right) \frac{9|\log \mu|^2}{\rho^2} 2 \|(\xi, \eta)\|_{\mathcal{H}^{\rho,0}}^2 \tag{218}$$

$$+ \frac{\pi^2}{2} \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} e^{2\rho|K+L|}(|\hat{\xi}_{K+L}|^2 + |\hat{\eta}_{K+L}|^2)$$

$$\times \left(1 + 2\mu^{2(\sigma-1)} \frac{K_2^2 + L_2^2}{(K_1 + L_1)^2} \right) e^{-2\rho|K+L|}, \tag{219}$$

and we can conclude by estimating (218) by exploiting the fact that $|\log \mu| \leqslant \mu^{-1/4}$, while we can estimate (219) by

$$\frac{\pi^2}{2} \|(\xi, \eta)\|_{\mathcal{H}^{\rho,0}}^2 \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \left(1 + 2\mu^{2(\sigma-1)} \frac{K_2^2 + L_2^2}{(K_1 + L_1)^2} \right) e^{-2\rho|K+L|}$$

$$\leqslant \frac{\pi^2}{2} \|(\xi, \eta)\|_{\mathcal{H}^{\rho,0}}^2 \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \left(1 + 2\mu^{2(\sigma-1)} K_2^2 + 2\mu^{2(\sigma-1)} L_2^2 \right)$$

$$\times e^{-2\rho|K+L|}$$

$$\leqslant \frac{\pi^2}{2} \|(\xi, \eta)\|_{\mathcal{H}^{\rho,0}}^2 \left[(1 + 2\mu^{2(\sigma-1)} K_2^2) 2\pi \int_{2/\mu}^{+\infty} e^{-2\rho\ell} \ell \, d\ell \right.$$

$$+ 4\pi \int_{2/\mu}^{+\infty} e^{-2\rho\ell} \ell^3 \, d\ell \bigg]$$

$$= \frac{\pi^2}{2} \|(\xi, \eta)\|_{\mathcal{H}^{\rho,0}}^2$$

$$\times \left[2\pi \left(1 + 2\mu^{2(\sigma-1)} \frac{9|\log \mu|^2}{\rho^2} \right) e^{-4\rho/\mu} \frac{\mu + 4\rho}{4\mu\rho^2} \right.$$

$$+ 4\pi e^{-4\rho/\mu} \frac{3\mu^3 + 12\rho\mu^2 + 24\rho^2\mu + 32\rho^3}{8\mu^3\rho^4} \bigg]. \tag{220}$$

$\square$

Proof of proposition 6.2. Proposition 6.2 is obtained as a corollary of proposition 6.5 by setting $\sigma = 2$. $\square$

Appendix C. Proof of propositions 6.6 and 6.3

Proof of proposition 6.6. The argument follows along the lines of appendix C in [BP06]. Exploiting the canonical transformation found in theorem 3.3, we also define

$$\zeta_a := (\xi_a, \eta_a) = \mathcal{T}_{\mu^2}(\widetilde{\xi_a}, \widetilde{\eta_a}) = \tilde{\zeta}_a + \psi_a(\tilde{\zeta}_a), \tag{221}$$

where $\psi_a(\tilde{\zeta}_a) := (\psi_\xi(\tilde{\zeta}_a), \psi_\eta(\tilde{\zeta}_a))$; by (46) we have

$$\sup_{\zeta \in \mathcal{B}_{\rho,n}(R)} \|\psi_a(\zeta)\|_{\mathcal{H}^{\rho,n}} \leqslant C'_n \mu^2 R. \tag{222}$$

For convenience we define

$$q_a(\tau, y) := \frac{1}{\sqrt{2}} \left[\xi_a(\mu^2 \tau, y_1 - \tau, y_2) + \eta_a(\mu^2 \tau, y_1 + \tau, y_2) \right] \tag{223}$$

$$\partial_{y_1} p_a(\tau, y) := \frac{1}{\sqrt{2}} \left[\xi_a(\mu^2 \tau, y_1 - \tau, y_2) - \eta_a(\mu^2 \tau, y_1 + \tau, y_2) \right]. \tag{224}$$

We observe that the pair (q_a, p_a) satisfies

$$\mu^2 (q_a)_t = -\Delta_1 \mu p_a + \mu^{\min(2\sigma,6)} \mathcal{R}_q, \mu (p_a)_t$$
$$= -\mu^2 q_a - \mu^4 \alpha \overline{\pi_0} q_a^2 + \mu^{\min(2\sigma-1,5)} \mathcal{R}_p,$$

where the operator Δ_1 acts on the variable x, $\overline{\pi_0}$ is the projector on the space of the functions with zero average (defined in section 3), and the remainders are functions of the rescaled variables τ and y which satisfy

$$\sup_{\mathcal{B}_{\rho,n}(R)} \|\mathcal{R}_q\|_{\ell^2_{\rho,0}} \leqslant C, \qquad \sup_{\mathcal{B}_{\rho,n}(R)} \|\mathcal{R}_p\|_{\ell^2_{\rho,1}} \leqslant C.$$

We now restrict the space variables to integer values; keeping in mind that q_a and p_a are periodic, we assume that $j \in \mathbb{Z}^2_{N_1, N_1^\sigma}$.

For a finite sequence $Q = (Q_j)_{j \in \mathbb{Z}^2_{N_1, N_1^\sigma}}$ we define the norm

$$\|Q\|^2_{\ell^2_{N_1, N_1^\sigma}} := \sum_{j \in \mathbb{Z}^2_{N_1, N_1^\sigma}} |Q_j|^2. \tag{225}$$

Now we consider the discrete model (5): we rewrite in the following form,

$$\dot{Q}_j = -(\Delta_1 P)_j \tag{226}$$

$$\dot{P}_j = -Q_j - \alpha \overline{\pi_0} Q_j^2 \tag{227}$$

and we want to show that there exist two sequences $E = (E_j)_{j \in \mathbb{Z}^2_{N_1, N_1^\sigma}}$ and $F = (F_j)_{j \in \mathbb{Z}^2_{N_1, N_1^\sigma}}$ such that

$$Q = \mu^2 q_a + \mu^{2+\gamma} E, P = \mu p_a + \mu^{2+\gamma} F$$

fulfills (226) and (227), where $\gamma > 0$ is a parameter we will fix later in the proof. Therefore, we have that

$$\dot{E} = -\Delta_1 F - \mu^{\min(2\sigma,6)-2-\gamma} \mathcal{R}_q \tag{228}$$

$$\dot{F} = -E - \alpha\overline{\pi_0}(\mu^2 2q_a E + \mu^{2+\gamma}E^2) - \mu^{\min(2\sigma-1,5)-2-\gamma}\mathcal{R}_p, \tag{229}$$

where we impose initial conditions on (E, F) such that $(\tilde{q}, \tilde{p})$ has initial conditions corresponding to the ones of the true initial datum,

$$\mu^2 q_a(0, \mu j_1, \mu^\sigma j_2) + \mu^{2+\gamma}E_{0,j} = Q_{0,j},$$

$$\mu p_a(0, \mu j_1, \mu^\sigma j_2) + \mu^{2+\gamma}F_{0,j} = P_{0,j}.$$

We now define the operator ∂_i, $i = 1, 2$, by $(\partial_i f)_j := f_j - f_{j-e_i}$ for each $f \in \ell^2_{N_1, N_1^\sigma}$.

- Claim 1: let $\sigma > 2$ and $\gamma > 0$, we have

$$\|E_0\|_{\ell^2_{N_1, N_1^\sigma}} \leqslant C'\mu^{(3-2\gamma-\sigma)/2}, \quad \|\partial_1 F_0\|_{\ell^2_{N_1, N_1^\sigma}}$$

$$\leqslant C'\mu^{(3-2\gamma-\sigma)/2}, \quad \|\partial_2 F_0\|_{\ell^2_{N_1, N_1^\sigma}}$$

$$\leqslant C'\mu^{(1-2\gamma+\sigma)/2}.$$

To prove claim 1 we observe that

$$E_0 = \mu^2 \frac{\xi_a + \eta_a - (\widetilde{\xi}_a + \widetilde{\eta}_a)}{\sqrt{2}\mu^{2+\gamma}} = \mu^{-\gamma}\frac{\psi_\xi + \psi_\eta}{\sqrt{2}},$$

$$F_0 = \mu\frac{\partial_{y_1}^{-1}[\xi_a - \eta_a - (\widetilde{\xi}_a - \widetilde{\eta}_a)]}{\sqrt{2}\mu^{2+\gamma}} = \mu^{-1-\gamma}\frac{\partial_{y_1}^{-1}(\psi_\xi - \psi_\eta)}{\sqrt{2}},$$

from which we can deduce

$$\|E_0\|^2_{\ell^2_{N_1, N_1^\sigma}} \leqslant \sum_{j\in\mathbb{Z}^2_{N_1, N_1^\sigma}} |E_{0,j}|^2 \leqslant C4N_1^{\sigma+1}(\mu^{2-\gamma})^2 = C\mu^{3-2\gamma-\sigma},$$

$$\|\partial_1 F_0\|^2_{\ell^2_{N_1, N_1^\sigma}} \leqslant \sum_{j\in\mathbb{Z}^2_{N_1, N_1^\sigma}} |\partial_1 F_{0,j}|^2 \leqslant C4N_1^{\sigma+1}(\mu^{2-\gamma})^2 \leqslant C\mu^{3-2\gamma-\sigma},$$

$$\|\partial_2 F_0\|^2_{\ell^2_{N_1, N_1^\sigma}} \leqslant \sum_{j\in\mathbb{Z}^2_{N_1, N_1^\sigma}} |\partial_2 F_{0,j}|^2 \leqslant C4N_1^{\sigma+1}(\mu^{1+\sigma-\gamma})^2 = C\mu^{1-2\gamma+\sigma}$$

and this leads to the thesis.

- Claim 2: fix $n \geqslant 0$, $T_0 > 0$ and $K_* > 0$, then for any $\mu < \mu_s$ and for any $\sigma > 2$ and $\gamma > 0$ such that $\sigma + 2\gamma < \min(4\sigma - 5, 7)$ we have

$$\|E\|^2_{\ell^2_{N_1, N_1^\sigma}} + \|\partial_1 F\|^2_{\ell^2_{N_1, N_1^\sigma}} + \|\partial_2 F\|^2_{\ell^2_{N_1, N_1^\sigma}} \leqslant K_*, \quad \forall\, |t| < \frac{T_0}{\mu^3}. \tag{230}$$

To prove the claim, we define

$$\mathcal{G}(E, F) := \sum_{j\in\mathbb{Z}^2_{N_1, N_1^\sigma}} \left(\frac{E_j^2 + F_j(-\Delta_1 F)_j}{2} + \frac{2\mu^2\alpha q_{a,j}E_j^2}{2} \right), \tag{231}$$

and we remark that, using the boundedness of $q_{a,j}$,

$$\frac{1}{2}\mathcal{G}(E,F) \leqslant \|E\|^2_{\ell^2_{N_1,N_1^\sigma}} + \|\partial_1 F_0\|^2_{\ell^2_{N_1,N_1^\sigma}} + \|\partial_2 F_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant 4\mathcal{G}(E,F).$$

Now we compute the time derivative of $\mathcal{G}$. Exploiting (228) and (229)

$$\dot{\mathcal{G}} = \sum_j E_j \left[-(\Delta_1 F)_j - \mu^{\min(2\sigma-2,4)-\gamma}(\mathcal{R}_q)_j \right] \tag{232}$$

$$+ \sum_j (-\Delta_1 F)_j \left[-E_j - \alpha(\mu^2 2q_{a,j}E_j + \mu^{2+\gamma}E_j^2) \right.$$

$$\left. - \mu^{\min(2\sigma-3,3)-\gamma}(\mathcal{R}_p)_j \right] \tag{233}$$

$$+ \sum_j 2\mu^2 \alpha q_{a,j} E_j \left[-(\Delta_1 F)_j - \mu^{\min(2\sigma-2,4)-\gamma}(\mathcal{R}_q)_j \right] \tag{234}$$

$$+ \sum_j \mu^2 \alpha E_j^2 \mu \frac{\partial q_{a,j}}{\partial \tau} \tag{235}$$

$$= \sum_j -E_j \mu^{\min(2\sigma-2,4)-\gamma}(\mathcal{R}_q)_j + \sum_j (-\Delta_1 F)_j \left[-\alpha\mu^{2+\gamma}E_j^2 \right.$$

$$\left. - \mu^{\min(2\sigma-3,3)-\gamma}(\mathcal{R}_p)_j \right] \tag{236}$$

$$- \sum_j 2\mu^2 \alpha q_{a,j} E_j \mu^{\min(2\sigma-2,4)-\gamma}(\mathcal{R}_q)_j + \sum_j \mu^2 \alpha E_j^2 \mu \frac{\partial q_{a,j}}{\partial \tau} \tag{237}$$

In order to estimate (236) and (237), we notice that

$$\sup_j |(\Delta_1 F)_j| \leqslant 2 \sup_j |(\partial_1 F)_j| + |(\partial_2 F)_j| \leqslant 4\sqrt{\mathcal{G}},$$

$$\|\mathcal{R}_q\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant \sum_j |(\mathcal{R}_q)_j|^2 \leqslant 4N_1^{\sigma+1} \sup_y |\mathcal{R}_q(y)|^2 \leqslant C\mu^{-1-\sigma},$$

and that $|(\partial_i \mathcal{R}_p)_j| \leqslant \mu \sup_y \left| \frac{\partial \mathcal{R}_p}{\partial y}(y) \right|$, which implies

$$\|\partial_i \mathcal{R}_p\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant C\mu^{1-\sigma}.$$

Now, the first sum in (236) is estimated by $C\mathcal{G}^{1/2}\mu^{(2\min(2\sigma-2,4)-1-2\gamma-\sigma)/2}$; the second sum in (236) can be bounded by

$$C(\mu^{2+\gamma}\mathcal{G}^{3/2} + \mu^{(2\min(2\sigma-3,3)+1-2\gamma-\sigma)/2}\mathcal{G}^{1/2}).$$

Recalling that $q_{a,j}$ is bounded, the first sum in (237) can be bounded by $C\mathcal{G}^{1/2}\mu^{(2\min(2\sigma,6)-1-2\gamma-\sigma)/2}$, while the second one is estimated by $C\mu^3\mathcal{G}$. Hence, as long as $\mathcal{G} < 2K_*$ we have

$$\left| \dot{\mathcal{G}} \right| \leqslant C \left| \left(\mu^{(\min(4\sigma-5,7)-2\gamma-\sigma)/2} + \mu^{(\min(4\sigma-5,7)-2\gamma-\sigma)/2} \right. \right.$$

$$\left. \left. + \mu^{(\min(4\sigma-1,11)-2\gamma-\sigma)/2} \right) \mathcal{G}^{1/2} + \mu^{2+\gamma} \mathcal{G}^{3/2} + \mu^3 \mathcal{G} \right|$$

$$\leqslant C(\mu^{2+\gamma}\sqrt{2}K_*^{1/2} + \mu^3)\mathcal{G} + C \left(2\mu^{(\min(4\sigma-5,7)-2\gamma-\sigma)/2} \right.$$

$$\left. + \mu^{(\min(4\sigma-1,11)-2\gamma-\sigma)/2} \right) \sqrt{2}K_*^{1/2}, \tag{238}$$

$$\overset{\gamma \geqslant 1}{\leqslant} C\mu^3 2\sqrt{2}K_*^{1/2}\mathcal{G} + C3\mu^{(\min(4\sigma-5,7)-2\gamma-\sigma)/2}\sqrt{2}K_*^{1/2}, \tag{239}$$

and by applying Gronwall's lemma we get

$$\mathcal{G}(t) \leqslant \mathcal{G}(0)e^{C2\sqrt{2}K_*^{1/2}\mu^3 t} + e^{C2\sqrt{2}K_*^{1/2}\mu^3 t}C2\sqrt{2}K_*^{1/2}\mu^3 t$$

$$\times \ C3\mu^{(\min(4\sigma-5,7)-2\gamma-\sigma)/2}\sqrt{2}K_*^{1/2}, \tag{240}$$

from which we can deduce the thesis. $\qquad\square$

Proof of proposition 6.3. Proposition 6.3 can be obtained from proposition 6.6 by setting $\sigma = 2$.
$\qquad\square$

Appendix D. Proof of proposition 6.9

We argue as in the proof of proposition 6.5.

First we remark that for all k such that $\kappa(k) = (\mu K_1, \ \mu^\sigma K_2)$ we have

$$|\omega_k^2| \overset{(24)}{=} 1 + 4 \left[\sin^2 \left(\frac{k_1 \pi}{2N_1 + 1} \right) + \sin^2 \left(\frac{k_2 \pi}{2N_2 + 1} \right) \right]$$

$$= 1 + 4 \left[\sin^2 \left(\frac{\mu K_1 \pi}{2} \right) + \sin^2 \left(\frac{\mu^\sigma K_2 \pi}{2} \right) \right]$$

$$\leqslant 1 + \pi^2(\mu^2 K_1^2 + \mu^{2\sigma} K_2^2) \leqslant \pi^2(1 + \mu^2 K_1^2 + \mu^{2\sigma} K_2^2), \tag{241}$$

hence

$$\frac{|\hat{p}_K|^2 + \pi^2(1 + \mu^2 K_1^2 + \mu^{2\sigma} K_2^2)|\hat{q}_K|^2}{2}$$

$$\leqslant \pi^2 \, e^{-2\rho|K|} \frac{|\hat{p}_K|^2 + (1 + \mu^2 K_1^2 + \mu^{2\sigma} K_2^2)|\hat{q}_K|^2}{2} \, e^{2\rho|K|}$$

$$\leqslant \pi^2 \, e^{-2\rho|K|} \left(1 + \mu^2 K_1^2 + \mu^{2\sigma} K_2^2 \right) \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2. \tag{242}$$

Hence, analogously to appendix B, it is convenient to define $\mathscr{L}_{\mu,\delta,\rho}$ as in (209) and by (165) we obtain that for all k such that $\kappa(k) = (\mu K_1, \ \mu^\sigma K_2)$ and $|K_1| + |K_2| > \frac{(2+\delta)|\log \mu|}{\rho}$

$$\frac{\mathcal{E}_\kappa}{\mu^2} \leqslant \sum_{L \in \mathscr{L}_{\mu,\delta,\rho}} \left(|\hat{p}_{K+L}|^2 + \omega_k^2 |\hat{q}_{K+L}|^2 \right)$$

$$\overset{(241),(242)}{\leqslant} \pi^2 \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 2 \sum_{L \in \mathscr{L}_{\mu,\delta,\rho}} \mathrm{e}^{-2\rho|K+L|} \left[1 + \mu^2 (K_1 + L_1)^2 \right.$$

$$\left. + \mu^{2\sigma} (K_2 + L_2)^2 \right], \tag{243}$$

$$= \sum_{L \in \mathscr{L}_{\mu,\delta,\rho}} \mathrm{e}^{-2\rho|K+L|} + \mu^2 \sum_{L \in \mathscr{L}_{\mu,\delta,\rho}} \mathrm{e}^{-2\rho|K+L|} (K_1 + L_1)^2$$

$$+ \mu^{2\sigma} \sum_{L \in \mathscr{L}_{\mu,\delta,\rho}} \mathrm{e}^{-2\rho|K+L|} (K_2 + L_2)^2. \tag{244}$$

Now,

$$\sum_{L \in \mathscr{L}_{\mu,\delta,\rho}} \mathrm{e}^{-2\rho|K+L|} \leqslant \mathrm{e}^{-2\rho|K|} + \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1=0, L_2 \neq 0}} \mathrm{e}^{-2\rho|K+L|} + \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1 \neq 0, L_2=0}} \mathrm{e}^{-2\rho|K+L|}$$

$$+ \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1, L_2 \neq 0}} \mathrm{e}^{-2\rho|K+L|}, \tag{245}$$

and we can estimate the above terms as for (210) in proposition 6.5; indeed, by (211)–(213) we have that the last sum in (245) is bounded by

$$\mathrm{e}^{-2\rho|K|} + \pi \left(\frac{1}{2\rho^2} + 2|K| \right) \mathrm{e}^{-2\rho|K|} + \mathrm{e}^{-2\rho(|K_1|+|K_2|)} \sum_{\ell \in \mathbb{Z} \setminus \{0\}} \mathrm{e}^{-4\rho|\ell|/\mu}$$

$$+ \mathrm{e}^{-2\rho(|K_1|+|K_2|)} \sum_{\ell \in \mathbb{Z} \setminus \{0\}} \mathrm{e}^{-4\rho|\ell|/\mu^\sigma}. \tag{246}$$

Next, we have

$$\sum_{L \in \mathscr{L}_{\mu,\delta,\rho}} \mathrm{e}^{-2\rho|K+L|} (K_1 + L_1)^2 \leqslant \mathrm{e}^{-2\rho|K|} K_1^2 + \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1 \neq 0, L_2=0}} \mathrm{e}^{-2\rho|K+L|}$$

$$\times (K_1 + L_1)^2 + \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1=0, L_2 \neq 0}} \mathrm{e}^{-2\rho|K+L|} K_1^2$$

$$+ \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1, L_2 \neq 0}} \mathrm{e}^{-2\rho|K+L|} (K_1 + L_1)^2. \tag{247}$$

First we estimate the last term in (247): we have that $|L + K| \geqslant |K|$, hence

$$\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1,L_2 \neq 0}} e^{-2\rho|K+L|} (K_1 + L_1)^2 = \int_{|K|}^{+\infty} \int_0^{2\pi} e^{-2\rho\xi} \xi \cos^2 \phi \, d\phi \, d\xi$$

$$= \pi e^{-2\rho|K|} \frac{1 + 2\rho|K|}{4\rho^2}$$

$$\leqslant \pi \mu^4 e^{-2\rho \left[|K| - \frac{2|\log \mu|}{\rho} - \frac{1}{2\rho} \log(2\rho|K|) \right]}$$

$$\overset{\delta < 1 - 1/e}{\leqslant} \pi \mu^4 e^{-2\rho \left[\delta|K| - \frac{2|\log \mu|}{\rho} \right]}. \tag{248}$$

Now we bound the other two nontrivial terms in (247); on the one hand, we notice that

$$\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1=0, L_2 \neq 0}} e^{-2\rho|K+L|}(K_1 + L_1)^2 \leqslant 2 \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1=0, L_2 \neq 0}} e^{-2\rho|K+L|} K_1^2$$

$$+ 2 \sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1=0, L_2 \neq 0}} e^{-2\rho|K+L|} L_1^2, \tag{249}$$

where the first sum can be bounded as the second term in (245), while

$$\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1=0, L_2 \neq 0}} e^{-2\rho|K+L|} L_1^2 \leqslant e^{-2\rho|K|} \sum_{\ell \in \mathbb{Z} \setminus \{0\}} e^{-4\rho|\ell|/\mu} \frac{\ell^2}{\mu^2}$$

$$\leqslant 2 e^{-2\rho|K|} \int_1^{+\infty} e^{-4\rho|\ell|/\mu} \frac{\ell^2}{\mu^2} \, d\ell, \tag{250}$$

where the last integral is exponentially small with respect to μ. Similarly,

$$\sum_{\substack{L \in \mathscr{L}_{\mu,\delta,\rho} \\ L_1=0, L_2 \neq 0}} e^{-2\rho|K+L|} L_2^2 \leqslant e^{-2\rho|K|} \sum_{\ell \in \mathbb{Z} \setminus \{0\}} e^{-4\rho|\ell|/\mu^\sigma} \frac{\ell^2}{\mu^{2\sigma}}$$

$$\leqslant 2 e^{-2\rho|K|} \int_1^{+\infty} e^{-4\rho|\ell|/\mu^\sigma} \frac{\ell^2}{\mu^{2\sigma}} \, d\ell, \tag{251}$$

where the last integral is exponentially small with respect to μ.

On the other hand, for any k such that $\kappa(k) = (\mu K_1, \ \mu^\sigma K_2)$ and $|K_1| + |K_2| \leqslant \frac{(2+\delta)|\log \mu|}{\rho}$

$$\left| \frac{\mathcal{E}_\kappa}{\mu^2} - \frac{|\hat{\psi}_K|^2}{2} \right| \leqslant |\omega_k^2 - 1| \, |\hat{q}_K|^2 + \frac{1}{2} \sum_{\substack{L=(L_1,L_2) \in \mathbb{Z}^2 \setminus \{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \left(|\hat{p}_{K+L}|^2 + \omega_k^2 |\hat{q}_{K+L}|^2 \right)$$

$$\overset{(241)}{\leqslant} (\mu^2 \pi^2 K_1^2 + \pi^2 \mu^{2\sigma} K_2^2) |\hat{p}_K|^2$$

$$+ \frac{1}{2} \sum_{\substack{L=(L_1,L_2) \in \mathbb{Z}^2 \setminus \{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \left(|\hat{p}_{K+L}|^2 + |\hat{q}_{K+L}|^2 \right)$$

$$+ \pi^2 [\mu^2 (K_1 + L_1)^2 + \mu^{2\sigma} (K_2 + L_2)^2] |\hat{q}_{K+L}|^2 \big),$$

$$\leqslant \left(\pi^2 \, \mu^2 K_1^2 + \pi^2 \mu^{2\sigma} \, K_2^2\right) |\hat{p}_K|^2 \; + \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2$$

$$\times \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \mathrm{e}^{-2\rho|K+L|} \left[1 + \pi^2\mu^2(K_1 + L_1)^2\right.$$

$$\left. + \pi^2\mu^{2\sigma}(K_2 + L_2)^2\right]$$

$$\leqslant \pi^2\mu^2 \left(1 + \mu^{2(\sigma-1)}\right) \frac{9|\log \mu|^2}{\rho^2} \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 \qquad (252)$$

$$+ \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \mathrm{e}^{-2\rho|K+L|} \qquad (253)$$

$$+ \pi^2\mu^2 \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \mathrm{e}^{-2\rho|K+L|}(K_1 + L_1)^2$$

$$(254)$$

$$+ \pi^2\mu^{2\sigma} \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} \mathrm{e}^{-2\rho|K+L|}(K_2 + L_2)^2$$

$$(255)$$

and we can conclude by estimating (252) by exploiting the fact that $|\log \mu| \leqslant \mu^{-1/4}$, while we can bound (253) and (254) by

$$\frac{\pi^2}{2} \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} [1 + \mu^2(K_1 + L_1)^2]\mathrm{e}^{-2\rho|K+L|}$$

$$\leqslant \frac{\pi^2}{2}\|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 \sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\} \\ \mu L_1, \mu^\sigma L_2 \in 2\mathbb{Z}}} (1 + 2\mu^2 K_1^2 + 2\mu^2 L_1^2)\mathrm{e}^{-2\rho|K+L|}$$

$$\leqslant \frac{\pi^2}{2}\|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 \left[(1 + 2\mu^2 K_1^2)2\pi \int_{2/\mu}^{+\infty} \mathrm{e}^{-2\rho\ell} \, \ell \, \mathrm{d}\ell \right.$$

$$+ 4\pi \, \mu^2 \int_{2/\mu}^{+\infty} \mathrm{e}^{-2\rho\ell}\ell^3 \, \mathrm{d}\ell\Bigg]$$

$$= \frac{\pi^2}{2} \|(\psi, \bar{\psi})\|_{\mathcal{H}^{\rho,0}}^2 \left[2\pi \left(1 + 2\mu^2 \frac{9|\log \mu|^2}{\rho^2}\right) \mathrm{e}^{-4\rho/\mu} \frac{\mu + 4\rho}{4\mu\rho^2}\right.$$

$$+ 4\pi\mu^2 \mathrm{e}^{-4\rho/\mu} \frac{3\mu^3 + 12\rho\mu^2 + 24\rho^2 \, \mu + 32\rho^3}{8\mu^3 \rho^4}\Bigg] , \qquad (256)$$

and we can estimate (255) by

$$\frac{\pi^2}{2}\|(\psi,\bar{\psi})\|^2_{\mathcal{H}^{\rho,0}}\mu^{2(\sigma-1)}\sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\}\\\mu L_1,\mu^\sigma L_2\in 2\mathbb{Z}}}(K_2+L_2)^2\,\mathrm{e}^{-2\rho|K+L|}$$

$$\leqslant\frac{\pi^2}{2}\|(\psi,\bar{\psi})\|^2_{\mathcal{H}^{\rho,0}}\mu^{2(\sigma-1)}\sum_{\substack{L=(L_1,L_2)\in\mathbb{Z}^2\setminus\{0\}\\\mu L_1,\mu^\sigma L_2\in 2\mathbb{Z}}}(2K_2^2+2L_2^2)\mathrm{e}^{-2\rho|K+L|}$$

$$\leqslant\frac{\pi^2}{2}\|(\psi,\bar{\psi})\|^2_{\mathcal{H}^{\rho,0}}\mu^{2(\sigma-1)}\left[2K_1^2 2\pi\int_{2/\mu^\sigma}^{+\infty}\mathrm{e}^{-2\rho\ell}\ell\mathrm{d}\ell\right.$$

$$\left.+4\pi\int_{2/\mu^\sigma}^{+\infty}\mathrm{e}^{-2\rho\ell}\ell^3\,\mathrm{d}\ell\right]$$

$$=\frac{\pi^2}{2}\|(\psi,\bar{\psi})\|^2_{\mathcal{H}^{\rho,0}}\mu^{2(\sigma-1)}\left[4\pi\frac{9|\log\mu|^2}{\rho^2}\mathrm{e}^{-4\rho/\mu^\sigma}\frac{\mu^\sigma+4\rho}{4\mu^\sigma\rho^2}\right.$$

$$\left.+4\pi\mathrm{e}^{-4\rho/\mu^\sigma}\frac{3\mu^{3\sigma}+12\rho\mu^{2\sigma}+24\rho^2\mu^\sigma+32\rho^3}{8\mu^{3\sigma}\rho^4}\right]. \tag{257}$$

Appendix E. Proof of proposition 6.10

The argument follows along the lines of appendix C in [BP06].

Exploiting the canonical transformation found in theorem 3.3, we also define

$$\zeta_a := (\psi_a,\bar{\psi}_a) = \mathcal{T}_{\mu^2}(\widetilde{\psi_a},\widetilde{\bar{\psi}_a}) = \tilde{\zeta}_a + \phi_a(\tilde{\zeta}_a), \tag{258}$$

where $\phi_a(\tilde{\zeta}_a) := (\phi_\xi(\tilde{\zeta}_a),\phi_\eta(\tilde{\zeta}_a))$; by (46) we have

$$\sup_{\zeta\in\mathcal{B}_{\rho,n}(R)}\|\phi_a(\zeta)\|_{\mathcal{H}^{\rho,n}}\leqslant C_n'\mu^2 R. \tag{259}$$

For convenience we define

$$q_a(\tau,y):=\frac{1}{\sqrt{2}}\left[\mathrm{e}^{\mathrm{i}\tau}\widetilde{\psi_a}(\tau,y_1,y_2)+\mathrm{e}^{-\mathrm{i}\tau}\widetilde{\bar{\psi}_a}(\tau,y_1,y_2)\right] \tag{260}$$

$$p_a(\tau,y):=\frac{1}{\sqrt{2}\mathrm{i}}\left[\mathrm{e}^{\mathrm{i}\tau}\widetilde{\psi_a}(\tau,y_1,y_2)-\mathrm{e}^{-\mathrm{i}\tau}\widetilde{\bar{\psi}_a}(\tau,y_1,y_2)\right]. \tag{261}$$

We observe that the pair (q_a,p_a) satisfies

$$\mu(q_a)_t = \mu p_a + \mu^{\min(2\sigma+1,5)}\mathcal{R}_q,$$

$$\mu(p_a)_t = -\mu q_a + \mu\Delta_1 q_a - \mu^3\beta\overline{\pi_0}q_a^3 + \mu^{\min(2\sigma+1,5)}\mathcal{R}_p,$$

where the operator Δ_1 acts on the variable x, $\overline{\pi_0}$ is the projector on the space of the functions with zero average, and the remainders are functions of the rescaled variables τ and y which satisfy

$$\sup_{\mathcal{B}_{\rho,n}(R)}\|\mathcal{R}_q\|_{\ell^2_{\rho,0}}\leqslant C,\qquad\sup_{\mathcal{B}_{\rho,n}(R)}\|\mathcal{R}_p\|_{\ell^2_{\rho,1}}\leqslant C.$$

We now restrict the space variables to integer values; keeping in mind that q_a and p_a are periodic, we assume that $j \in \mathbb{Z}^2_{N_1,N_1^\sigma}$.

For a finite sequence $Q = (Q_j)_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}}$ we use the norm $\|Q\|^2_{\ell^2_{N_1,N_1^\sigma}}$ defined in (225).

Now we consider the discrete model (5): we rewrite in the following form,

$$\dot{Q}_j = P_j \tag{262}$$

$$\dot{P}_j = -Q_j + (\Delta_1 Q)_j - \beta \,\overline{\pi_0} Q_j^3 \tag{263}$$

and we want to show that there exist two sequences $E = (E_j)_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}}$ and $F = (F_j)_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}}$ such that

$$Q = \mu q_a + \mu^{1+\gamma} E, P = \mu p_a + \mu^{1+\gamma} F$$

fulfills (262) and (263), where $\gamma > 0$ is a parameter we will fix later in the proof. Therefore, we have that

$$\dot{E} = F - \mu^{\min(2\sigma+1,5)-1-\gamma} \mathcal{R}_q \tag{264}$$

$$\dot{F} = -E + \Delta_1 E - \beta\overline{\pi_0}(3\mu^2 q_a^2 E + 3\mu^{2+\gamma} q_a E^2 + \mu^{2+2\gamma} E^3)$$
$$\quad - \mu^{\min(2\sigma+1,5)-1-\gamma} \mathcal{R}_p, \tag{265}$$

where we impose initial conditions on (E, F) such that $(\tilde{q}, \tilde{p})$ has initial conditions corresponding to the ones of the true initial datum,

$$\mu q_a(0, \mu j_1, \mu^\sigma j_2) + \mu^{1+\gamma} E_{0,j} = Q_{0,j},$$

$$\mu p_a(0, \mu j_1, \mu^\sigma j_2) + \mu^{1+\gamma} F_{0,j} = P_{0,j}.$$

We now define the operator ∂_i, $i = 1, 2$, by $(\partial_i f)_j := f_j - f_{j-e_i}$ for each $f \in \ell^2_{N_1,N_1^\sigma}$.

- Claim 1: let $\sigma > 1$ and $\gamma > 0$, we have

$$\|E_0\|_{\ell^2_{N_1,N_1^\sigma}} \leqslant C'\mu^{(3-2\gamma-\sigma)/2},$$

$$\|F_0\|_{\ell^2_{N_1,N_1^\sigma}} \leqslant C'\mu^{(3-2\gamma-\sigma)/2},$$

$$\|\partial_1 E_0\|_{\ell^2_{N_1,N_1^\sigma}} \leqslant C'\mu^{(5-2\gamma-\sigma)/2},$$

$$\|\partial_2 E_0\|_{\ell^2_{N_1,N_1^\sigma}} \leqslant C'\mu^{(3-2\gamma+\sigma)/2},$$

$$\|\partial_1 F_0\|_{\ell^2_{N_1,N_1^\sigma}} \leqslant C'\mu^{(5-2\gamma-\sigma)/2},$$

$$\|\partial_2 F_0\|_{\ell^2_{N_1,N_1^\sigma}} \leqslant C'\mu^{(3-2\gamma+\sigma)/2}.$$

To prove claim 1 we observe that

$$E_0 = \mu \frac{\psi_a + \bar{\psi}_a - (\widetilde{\psi}_a + \widetilde{\bar{\psi}}_a)}{\sqrt{2}\mu^{1+\gamma}} = \mu^{-\gamma} \frac{\phi_\xi + \phi_\eta}{\sqrt{2}},$$

$$F_0 = \mu \frac{\psi_a - \bar{\psi}_a - (\widetilde{\psi}_a - \bar{\widetilde{\psi}}_a)}{\sqrt{2}\mathrm{i}\,\mu^{1+\gamma}} = \mu^{-\gamma} \frac{\phi_\xi - \phi_\eta}{\sqrt{2}\mathrm{i}},$$

from which we can deduce

$$\|E_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant \sum_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}} |E_{0,j}|^2 \leqslant C\,4N_1^{\sigma+1} \left(\mu^{2-\gamma}\right)^2 = C\,\mu^{3-2\gamma-\sigma},$$

$$\|F_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant \sum_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}} |F_{0,j}|^2 \leqslant C\,4N_1^{\sigma+1} \left(\mu^{2-\gamma}\right)^2 = C\,\mu^{3-2\gamma-\sigma},$$

$$\|\partial_1 E_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant \sum_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}} |\partial_1 E_{0,j}|^2 \leqslant C\,4N_1^{\sigma+1} \left(\mu^{2+1-\gamma}\right)^2 \leqslant C\,\mu^{5-2\gamma-\sigma},$$

$$\|\partial_2 E_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant \sum_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}} |\partial_2 E_{0,j}|^2 \leqslant C\,4N_1^{\sigma+1} \left(\mu^{2+\sigma-\gamma}\right)^2 = C\,\mu^{3-2\gamma+\sigma},$$

$$\|\partial_1 F_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant \sum_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}} |\partial_1 F_{0,j}|^2 \leqslant C\,4N_1^{\sigma+1} \left(\mu^{2+1-\gamma}\right)^2 \leqslant C\,\mu^{5-2\gamma-\sigma},$$

$$\|\partial_2 F_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant \sum_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}} |\partial_2 F_{0,j}|^2 \leqslant C\,4N_1^{\sigma+1} \left(\mu^{2+\sigma-\gamma}\right)^2 = C\,\mu^{3-2\gamma+\sigma},$$

and this leads to the thesis.

- Claim 2: fix $n \geqslant 0$, $T_0 > 0$ and $K_* > 0$, then for any $\mu < \mu_s$ and for any $\sigma > 1$ and $\gamma > 0$ such that $\sigma + 2\gamma < \min(4\sigma - 1, 7)$ we have

$$\|E\|^2_{\ell^2_{N_1,N_1^\sigma}} + \|F\|^2_{\ell^2_{N_1,N_1^\sigma}} + \|\partial_1 E_0\|^2_{\ell^2_{N_1,N_1^\sigma}} + \|\partial_2 E_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant K_*, \quad \forall\, |t| < \frac{T_0}{\mu^2}. \quad (266)$$

To prove the claim, we define

$$\mathcal{G}(E,F) := \sum_{j \in \mathbb{Z}^2_{N_1,N_1^\sigma}} \left(\frac{F_j^2 + E_j^2 + E_j(-\Delta_1 E)_j}{2} \right.$$
$$\left. + \frac{3\mu^2 \beta q_a^2 E_j^2 + 3\mu^{2+\gamma}\beta q_a E_j^3}{2} \right), \quad (267)$$

and we remark that

$$\frac{1}{2}\mathcal{G}(E,F) \leqslant \|E\|^2_{\ell^2_{N_1,N_1^\sigma}} + \|\partial_1 F_0\|^2_{\ell^2_{N_1,N_1^\sigma}} + \|\partial_2 F_0\|^2_{\ell^2_{N_1,N_1^\sigma}} \leqslant 2\mathcal{G}(E,F).$$

Now we compute the time derivative of $\mathcal{G}$. Exploiting (228) and (229) and using a procedure analogous to appendix C, we finally obtain as long as $\mathcal{G} < 2K_*$

$$\left|\dot{\mathcal{G}}\right| \leqslant C \left[\mu^{2+\gamma}K_*^{1/2} + \mu^{2+2\gamma}K_* + \mu^3 + \mu^{2+\gamma}K_*^{1/2} + \mu^{\min(2\sigma+2,6)-(1+\sigma)/2}\right.$$
$$\left. + \mu^{2+\gamma}K_*^{1/2}\right]\mathcal{G} \tag{268}$$

$$+ C \left[2\mu^{\min(2\sigma,4)-\gamma-(1+\sigma)/2} + \mu^{\min(2\sigma,4)-\gamma+(1-\sigma)/2} + \mu^{\min(2\sigma+2,6)-\gamma-(1+\sigma)/2}\right]K_*^{1/2} \tag{269}$$

$$\overset{\sigma+2\gamma<\min(4\sigma-1,7)}{\leqslant} C\mu^2(1+K_*^{1/2})\mathcal{G} + C\mu^{(\min(4\sigma-1,7)-2\gamma-\sigma)/2}K_*^{1/2} \tag{270}$$

and by applying Gronwall's lemma we get

$$\mathcal{G}(t) \leqslant \mathcal{G}(0)e^{C(1+K_*^{1/2})\mu^2 t} + e^{C(1+K_*^{1/2})\mu^2 t}C(1+K_*^{1/2})\mu^2 t$$
$$\times C\mu^{(\min(4\sigma-1,7)-2\gamma-\sigma)/2}K_*^{1/2}, \tag{271}$$

from which we can deduce the thesis.

ORCID iDs

M Gallone ⓘ https://orcid.org/0000-0003-3999-8982
S Pasquali ⓘ https://orcid.org/0000-0003-2712-2000

References

[Bam99] Bambusi D 1999 Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations *Math. Z.* **230** 345–87
[Bam05] Bambusi D 2005 Galerkin averaging method and Poincaré normal form for some quasilinear PDEs *Ann. della Scuola Norm. Super. Pisa - Cl. Sci.* **4** 669–702
[BCMM15] Bambusi D, Carati A, Maiocchi A and Maspero A 2015 Some analytic results on the FPU paradox *Hamiltonian Partial Differential Equations and Applications* (New York: Springer) pp 235–54
[BCP02] Bambusi D, Carati A and Ponno A 2002 The nonlinear Schrödinger equation as a resonant normal form *Discrete Continuous Dyn. Syst.* B **2** 109–28
[BPP10] Bambusi D, Paleari S and Penati T 2010 Existence and continuous approximation of small amplitude breathers in 1D and 2D Klein–Gordon lattices *Appl. Anal.* **89** 1313–34
[BP06] Bambusi D and Ponno A 2006 On metastability in FPU *Commun. Math. Phys.* **264** 539–61
[Bam08] Bambusi D and Ponno A 2008 *Resonance, Metastability and Blow up in FPU* (Berlin: Springer) pp 191–205
[Ben05] Benettin G 2005 Time scale for energy equipartition in a two-dimensional FPU model *Chaos* **15** 015108
[BG08] Benettin G and Gradenigo G 2008 A study of the Fermi–Pasta–Ulam problem in dimension two *Chaos* **18** 013112
[BVT80] Benettin G, Vecchio G L and Tenenbaum A 1980 Stochastic transition in two-dimensional Lennard-Jones systems *Phys. Rev.* A **22** 1709
[BW06] Butt I A and Wattis J A D 2006 Discrete breathers in a two-dimensional Fermi–Pasta–Ulam lattice *J. Phys. A: Math. Gen.* **39** 4955
[BW07] Butt I A and Wattis J A D 2007 Discrete breathers in a two-dimensional hexagonal Fermi–Pasta–Ulam lattice *J. Phys. A: Math. Theor.* **40** 1239
[CLL83] Chen H H, Lee Y C and Lin J-E 1983 On a new hierarchy of symmetries for the Kadomtsev–Petviashvili equation *Physica* D **9** 439–45
[CL87] Chen H H and Lin J E 1987 On the infinite hierarchies of symmetries and constants of motion for the Kadomtsev–Petviashvili equation *Physica* D **26** 171–80

[FPU95]   Fermi E, Pasta J and Ulam S 1995 *Studies of Non Linear Problems Los Alamos Rpt. LA-1940.* Los Alamos Scientific Lab.

          Fermi E, Pasta J and Ulam S 1965 *Collected Papers of Enrico Fermi* vol 2 (Chicago: University of Chicago Press)

[Gal07]   Gallavotti G 2007 *The Fermi–Pasta–Ulam Problem: A Status Report* vol 728 (Berlin: Springer)

[GKK14]   Grébert B, Kappeler T and Kappeler T 2014 *The Defocusing NLS Equation and its Normal Form* (*EMS Series of Lectures in Mathematics*) (Zürich: European Mathematical Society)

[GKM98]   Grébert B, Kappeler T and Mityagin B 1998 Gap estimates of the spectrum of the Zakharov–Shabat system *Appl. Math. Lett.* **11** 95–7

[SW00]   Guido S and Wayne C E 2000 Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi–Pasta–Ulam model *Equadiff* vol 99 (Singapore: World Scientific) pp 390–404

[Bou93]   Jean B 1993 On the Cauchy problem for the Kadomstev–Petviashvili equation *Geom. Funct. Anal.* **3** 315–41

[KP03]   Kappeler T and Pöschel J 2003 *Kdv & KAM* vol 45 (Berlin: Springer)

[KP08]   Kappeler T and Pöschel J 2008 On the well-posedness of the periodic KdV equation in high regularity classes *Hamiltonian Dynamical Systems and Applications* (Dordrecht: Springer) pp 431–41

[KST08]   Kappeler T, Schaad B and Topalov P 2008 mKdV and its Birkhoff coordinates *Physica D* **237** 1655–62

[LLPR07] Lichtenberg A J, Livi R, Pettini M and Ruffo S 2007 Dynamics of oscillator chains *The Fermi–Pasta–Ulam Problem* (Berlin: Springer) pp 21–121

[LC82]   Lin J-E and Chen H H 1982 Constraints and conserved quantities of the Kadomtsev–Petviashvili equations *Phys. Lett.* A **89** 163–7

[Mol14]   Molnar J-C 2014 New estimates of the nonlinear Fourier transform for the defocusing NLS equation *Int. Math. Res. Not.* **2015** 8309–52

[Pas19]   Pasquali S 2019 Dynamics of the nonlinear Klein–Gordon equation in the nonrelativistic limit *Ann. Mat. Pura Appl.* **198** 903–72

[Pös11]   Pöschel J 2011 Hill's potentials in weighted Sobolev spaces and their spectral gaps *Math. Ann.* **349** 433–58

[Ros03]   Rosenau P 2003 Hamiltonian dynamics of dense chains and lattices: or how to correct the continuum *Phys. Lett.* A **311** 39–52

[SJ87]   Schwarz M Jr 1987 Periodic solutions of Kadomtsev–Petviashvili *Adv. Math.* **66** 217–33

[Wat94]   Wattis J A D 1994 Solitary waves on a two-dimensional lattice *Phys. Scr.* **50** 238

[YWSC09] Xiang Y, Wattis J A D, Susanto H and Cummings L J 2009 Discrete breathers in a two-dimensional spring-mass lattice *J. Phys. A: Math. Theor.* **42** 355207