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) [FPU95], 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 outgoing 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 [Gal07]. For a more recent account on analytic results on the ‘FPU paradox’ we refer to [BCMM15].

In particular, we mention the papers [BP06] and [Bam08], in which the authors used the techniques of canonical perturbation theory for PDEs in order to show that the FPU $\alpha$ model (respectively, $\beta$ 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 $\mu \sim 1/N \ll 1$ 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 $\beta$ model), then the dynamics of the KdV (resp. mKdV) equations approximates the solutions of the FPU model up to a time of order $O(\mu^{-3})$. 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 $\epsilon$ (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 [BVT80] [Ben05] [BG08] studied numerically a two-dimensional FPU lattice with triangular cells and different boundary conditions in order to estimate the equipartition time-scale, and 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 [Wat94] [BW06] [BW07] [YWSC09] [WJ14] [BPP10].

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, under some suitable assumptions on the ratio between the sides of the lattice and on the type of small-amplitude solution we want to describe, we obtain for a 2D Electrical Transmission lattice (ETL) either a system of two uncoupled KdV equations or a system of two uncoupled KP-II equations 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. 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 genuinely two-dimensional regime is described by a system of two-dimensional integrable PDEs.

Some comments are in order:

i. denoting by \( \mu \ll 1 \) the wave-number of the Fourier mode initially excited, we have that the time-scale of validity of our result is of order \( O(\mu^{-1}) \) for the 2D ETL lattice, and of order \( O(\mu^{-2}) \) for the 2D Klein-Gordon lattice;

ii. the ansatz about the small amplitude solutions gives a relation between the specific energy of the system \( \epsilon \) and the wave-number \( \mu \sim \frac{1}{N} \) 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;

iii. our result can be easily generalized to higher-dimensional lattices (see Remark 2.8 and Remark 2.9), such as the physical case of three-dimensional rectangular lattices with cubic cells.

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. 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 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 two-dimensional KG lattice by the dynamics of the one-dimensional nonlinear Schrödinger 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 [BCP02] [Bam05] [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 [SW00] [BP06] for the one-dimensional case).

2. Main Results

2.1. The Electrical transmission lattice. We describe a lossless periodic two-dimensional electrical transmission lattice (ETL), given by a rectangular configuration of repeating units, each made up of two linear inductors and a nonlinear capacitor; in the non-periodic setting, the model has been studied in [BW06]. We define lattice nodes by the locations of capacitors. We denote

\[
Z_{N_1,N_2}^2 := \{(j_1, j_2) : j_1, j_2 \in \mathbb{Z}, |j_1| \leq N_1, |j_2| \leq N_2\};
\]

we also write \( e_1 := (1, 0) \), \( e_2 := (0, 1) \) and \( Z_N^2 := \mathbb{Z}^2_{N,N} \).

The variable \( V_j(t), j \in Z_{N_1,N_2}^2 \), denotes the voltage across the \( j \)-th capacitor, \( Q_j(t) \) denotes the charge stored on the \( j \)-th capacitor and \( I_j(t) \) denotes the current through the \( j \)-th inductor along direction \( e_1 \).

To derive the equations for the voltage \( V_j \) and the charge \( Q_j \) in the lattice one can proceed as follows. Considering a section of the lattice and applying Faraday’s law and Lenz’s law, the difference in shunt voltage at site \( j \) and site \( j + e_1 \) is given by

\[
V_{j+e_1} - V_j = -L \frac{dI_j}{dt},
\]
where $L$ is the inductance, which we assume to be constant. Assuming the capacitance $C$ to be an analytic function of the voltage $V$ we can expand it in Taylor series, obtaining for small voltages
\begin{equation}
C_j(V) \sim C_0(1 + 2aV_j + 3bV_j^2),
\end{equation}
where $C_0 := C_1(0)$, $a$ and $b$ are real constants determined by the physical realisation of the network. Using standard relations between electrical quantities we finally obtain a closed equation for the charge
\begin{equation}
\frac{d^2Q_j}{dt^2} = \frac{1}{LC_0}(\Delta_1((Q + \alpha Q^2 + \beta Q^3))_j,
\end{equation}
\begin{equation}
(\Delta_1 Q)_j := (Q_j+d_j - 2Q_j + Q_{j-1}) + (Q_{j+1} - 2Q_j + Q_{j-2}).
\end{equation}
where $\alpha, \beta$ are real parameters related to $a$ and $b$. Up to a rescaling of time, we can set $LC_0 = 1$ without loss of generality. The Hamiltonian associated to (4) is given by
\begin{equation}
H(Q, P) = \sum_{j \in \mathbb{Z}_{N_1}^2, N_2} -\frac{1}{2} P_j (\Delta_1 P)_j + (F(Q))_j,
\end{equation}
\begin{equation}
(F(Q))_j = \frac{Q_j^2}{2} + \alpha \frac{Q_j^3}{3} + \beta \frac{Q_j^4}{4}.
\end{equation}
We refer to (6) as $\alpha + \beta$ model (respectively, $\beta$ model) if $\alpha \neq 0$ (respectively $\alpha = 0$). With the above Hamiltonian formulation the equations of motion associated to (6) are given by
\begin{equation}
\begin{cases}
\dot{Q}_j = - (\Delta_1 P)_j \\
\dot{P}_j = - (F'(Q))_j
\end{cases}.
\end{equation}
(8)
We also introduce the Fourier coefficients of $Q$ via the following standard relation,
\begin{equation}
Q_j := \frac{1}{\sqrt{(2N_1+1)(2N_2+1)}} \sum_{k \in \mathbb{Z}_{N_1}^2, N_2} \hat{Q}_k e^{j \frac{\pi}{2N_1+1} (k_1+1)j_1 + j \frac{\pi}{2N_2+1} j_2}, \quad j \in \mathbb{Z}_{N_1}^2, N_2,
\end{equation}
and similarly for $P_j$. We denote by
\begin{equation}
E_k := \frac{\omega_k^2 |\hat{P}_k|^2 + |\hat{Q}_k|^2}{2},
\end{equation}
\begin{equation}
\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),
\end{equation}
the energy and the square of the frequency of the mode at site $k = (k_1, k_2) \in \mathbb{Z}_{N_1}^2, N_2$. 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
\begin{equation}
\mathbb{Z}_{N_1, N_2, +}^2 := \{(k_1, k_2) \in \mathbb{Z}_{N_1}^2, N_2 : k_1, k_2 \geq 0\}.
\end{equation}
It is also convenient to introduce the following specific quantities,
\begin{equation}
\kappa := \kappa(k) = \frac{k_1}{N_1 + \frac{1}{2}}, \frac{k_2}{N_2 + \frac{1}{2}},
\end{equation}
\begin{equation}
E_\kappa := \frac{E_k}{(N_1 + \frac{1}{2})(N_2 + \frac{1}{2})},
\end{equation}
where (13) is the specific energy of the normal mode with index $\kappa$.
We want to study the behaviour of small amplitude solutions of (8), with initial data in which only one low-frequency Fourier mode is excited.
We assume $N_1 \leq N_2$, and we introduce the quantities
\begin{align}
\mu &:= \frac{2}{2N_1 + 1}, \\
\sigma &:= \log_{N_1 + \frac{1}{2}} \left( N_2 + \frac{1}{2} \right),
\end{align}
which play the role of parameters in our construction.

We study the $\alpha + \beta$ model of (8) in the following regimes:
\begin{itemize}
  \item [(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 $2 < \sigma < 5$.
  \item [(KP)] the weakly transverse regime, where the effective dynamics is a described by a system of two uncoupled Kadomtsev-Petviashvili (KP) equation. This corresponds to taking $\mu \ll 1$ and $\sigma = 2$.
\end{itemize}

From now on, we denote by $\kappa_0 := \left( \frac{1}{1 + \frac{1}{N_1 + \frac{1}{2} \sigma}}, \frac{1}{1 + \frac{1}{N_1 + \frac{1}{2} \sigma}} \right) = (\mu, \mu^\gamma)$.

**Theorem 2.2.** Consider (8) with $\alpha \neq 0$, $2 < \sigma < 5$.

Fix $1 \leq \gamma \leq \frac{\sigma}{2}$ and two positive constants $C_0$ and $T_0$, then there exist positive constants $\mu_0$, $C_1$ and $C_2$ (depending only on $\gamma$, $C_0$ and on $T_0$) such that the following holds. Consider an initial datum with
\begin{align}
\mathcal{E}_n(0) &= C_0 \mu^4, \quad \mathcal{E}_n(0) = 0, \quad \forall \kappa = (\kappa_1, \kappa_2) \neq \kappa_0, \\
\end{align}
and assume that $\mu < \mu_0$. Then there exists $\rho > 0$ such that along the corresponding solution one has
\begin{align}
\mathcal{E}_n(t) &\leq C_1 \mu^4 e^{-\rho(\kappa_1, \kappa_2, \mu^\gamma)} + C_2 \mu^{4+\gamma}, \quad |t| \leq \frac{T_0}{\mu^\gamma},
\end{align}
for all $\kappa$. Moreover, for any $n_2$ with $0 \leq n_2 \leq N_2$ there exists a sequence of almost-periodic functions $(\mathcal{F}_n)_{n=(n_1, n_2) \in \mathbb{Z}_{N_1 + \frac{1}{2}} \times \mathbb{N}_{n_2}^+}$ such that, if we denote
\begin{align}
\mathcal{F}_n &= \mu^4 \mathcal{F}_n, \quad \mathcal{F}_n = 0 \quad \forall \kappa \neq n\kappa_0 \\
\end{align}
then
\begin{align}
|\mathcal{E}_n(t) - \mathcal{F}_n(t)| &\leq C_2 \mu^{4+\gamma}, \quad |t| \leq \frac{T_0}{\mu^\gamma},
\end{align}
for all $\kappa$.\[\] **Theorem 2.2.** Consider (8) with $\alpha \neq 0$, $\sigma = 2$.

Fix $1 \leq \gamma < \frac{\sigma}{4}$ and two positive constants $C_0$ and $T_0$, then there exist positive constants $\mu_0$, $C_1$ and $C_2$ (depending only on $\gamma$, $C_0$ and on $T_0$) such that the following holds. Consider an initial datum with
\begin{align}
\mathcal{E}_n(0) &= C_0 \mu^4, \quad \mathcal{E}_n(0) = 0 \quad \forall \kappa = (\kappa_1, \kappa_2) \neq \kappa_0, \\
\end{align}
and assume that $\mu < \mu_0$. Then there exists $\rho > 0$ such that along the corresponding solution one has
\begin{align}
\mathcal{E}_n(t) &\leq C_1 \mu^4 e^{-\rho(\kappa_1, \kappa_2, \mu^\gamma)} + C_2 \mu^{4+\gamma}, \quad |t| \leq \frac{T_0}{\mu^\gamma},
\end{align}
for all $\kappa$.

**Remark 2.3.** In Theorem 2.2 we do not mention the existence of a sequence of almost-periodic functions approximating the specific energies of the modes. 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.

2.2. The 2D Klein-Gordon lattice. Among the lattices that have received a great amount of attention, we mention the class of Klein-Gordon (KG) lattices, which combine the nearest-neighbour potential with an on-site one. The Hamiltonian of the system with $2N + 1$ particles in the one-dimensional case is
\begin{align}
H(r, s) &= \sum_{j=-N}^{N} \frac{s^2}{2} + \frac{(r_{j+1} - r_j)^2}{2} + U(r_j), \\
U(x) &= m^2 \frac{x^2}{2} + \beta \frac{x^{2p+2}}{2p + 2}, \quad m > 0, \quad p \geq 1.
\end{align}
We now pass to two-dimensional KG lattices: the scalar model

\begin{equation}
H(Q, P) = \sum_{j \in \mathbb{Z}^2_{N_1, N_2}} \frac{P_j^2}{2} + \frac{1}{2} \sum_{j, k \in \mathbb{Z}^2_{N_1, N_2} \setminus \{j = k\}} \frac{(Q_j - Q_k)^2}{2} + \sum_{j \in \mathbb{Z}^2_{N_1, N_2}} U(Q_j),
\end{equation}

\begin{equation}
U(x) = m^2 \frac{x^2}{2} + \beta \frac{x^{2p+2}}{2p+2}, \quad m > 0, \quad \beta > 0, \quad p \geq 1,
\end{equation}

can be used to describe rigid rotating molecules in the lattice plane (\(Q\) being the angle of rotation), where each molecule interacts with its neighbors and with the periodic substrate potential \(U\); alternatively, \(Q\) can represent the transverse motion of a planar lattice \([Ros03]\).

Using the operator \(\Delta_1\) introduced in (5), the Hamiltonian (24) can be rewritten as

\begin{equation}
H(Q, P) = \sum_{j \in \mathbb{Z}^2_{N_1, N_2}} \frac{P_j^2}{2} + \frac{1}{2} \sum_{j, k \in \mathbb{Z}^2_{N_1, N_2}} Q_j (-\Delta(Q)_k) + \sum_{j \in \mathbb{Z}^2_{N_1, N_2}} U(Q_j),
\end{equation}

the associated equations of motion are

\begin{equation}
\dot{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}.
\end{equation}

If we take \(p = 1\), we obtain a generalization of the one-dimensional \(\phi^4\) model.

We also introduce the Fourier coefficients of \(Q\) via the following relation,

\begin{equation}
Q_j := \frac{1}{(2N_1 + 1)(2N_2 + 1)} \sum_{k \in \mathbb{Z}^2_{N_1+N_2}} \hat{Q}_k e^{i \frac{k \cdot j}{2N_1+1}}, \quad j \in \mathbb{Z}^2_{N_1, N_2},
\end{equation}

and similarly for \(P_j\), and we denote by

\begin{equation}
E_k := \frac{|P_k|^2 + \omega_k^2 |Q_k|^2}{2},
\end{equation}

\begin{equation}
\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),
\end{equation}

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

In the rest of the paper we will assume that \(m = 1\).

We study the two-dimensional KG lattice (24) in the following regime:

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

**Theorem 2.4.** Consider (24) with \(\beta > 0, 1 < \sigma < 7\).

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

\begin{equation}
\mathcal{E}_\nu(0) = C_0 \nu^2, \quad \mathcal{E}_\nu(0) = 0, \quad \forall \nu \in (\kappa_1, \kappa_2) \neq \kappa_0,
\end{equation}

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

\begin{equation}
\mathcal{E}_\nu(t) \leq C_1 \mu^2 e^{-\rho |(\kappa_1, \mu, \kappa_2)|} + C_2 \mu^{2+\gamma}, \quad |t| \leq \frac{T_0}{\mu^2},
\end{equation}

for all \(\kappa\). Moreover, for any \(n_2\) with \(0 \leq n_2 \leq N_2\) 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

\begin{equation}
\mathcal{F}_{\kappa_0} = \mu^2 F_n, \quad \mathcal{F}_\nu = 0 \quad \forall \nu \neq \nu_0,
\end{equation}

then

\begin{equation}
|\mathcal{E}_\nu(t) - \mathcal{F}_\nu(t)| \leq C_2 \mu^{2+\gamma}, \quad |t| \leq \frac{T_0}{\mu^2}.
\end{equation}
2.3. Further remarks.

Remark 2.5. The specific choice of the direction of longitudinal propagation in the regimes that we have considered is not relevant.

Remark 2.6. Using the definition of \( \sigma \) and \( \mu \) in (15), (14) we can read Theorems 2.1, 2.2 using, as parameter, the total number of sites in the lattice \( N \). The statement should read as follows:

Consider (8) with \( \alpha \neq 0 \) and \( 2 \leq \sigma < 5 \). Fix \( 1 \leq \gamma \leq \frac{2}{\mu} \) and two positive constants \( C_0 \) and \( T_0 \), then there exists positive constants \( N, C_1 \) and \( C_2 \) (depending only on \( \gamma, C_0 \) and \( T_0 \)) such that if we consider an initial datum with

\[
\mathcal{E}_{\kappa}(0) = \frac{C_0}{N^{\frac{1}{\mu+\sigma}}}, \quad \mathcal{E}_{\kappa}(0) = 0 \quad \forall \kappa \neq \kappa_0
\]

with \( N > N_0 \). There exists \( \rho > 0 \) such that along the corresponding solution one has

\[
\mathcal{E}_{\kappa}(t) \leq C_1 \frac{1}{N^{\frac{1}{\mu+\sigma}}} e^{-\rho \left| (\kappa_1, \ldots, \kappa_d) \right|} + C_2 \frac{1}{N^{\frac{1}{\mu+\sigma}}}, \quad |t| \leq T_0 N^{\frac{1}{\mu+\sigma}}.
\]

for all \( \kappa \).

Remark 2.7. We point out that the time of validity of Theorem 2.4 for the KG lattice is of order \( O(\mu^{-2}) \), which is different from the time of validity of Theorem 2.1 and Theorem 2.2 for the FPU lattice. In the one-dimensional case it has been observed that, for a fixed value of specific energy \( \epsilon \) and for long-wavelength modes initially excited, the \( \phi^4 \) model reached equipartition faster than the FPU \( \beta \) model (see [LLPR07], sec. 2.1.8).

Remark 2.8. Theorem 2.1 and Theorem 2.2 can be generalized to higher dimensional lattices. Indeed, let \( d \leq 4 \), define

\[
\mathbb{Z}_{N_1, \ldots, N_d} := \{(j_1, \ldots, j_d) : j_1, \ldots, j_d \in \mathbb{Z}, |j_1| \leq N_1, \ldots, |j_d| \leq N_d\},
\]

and consider the d-dimensional ETL

\[
H(Q, P) = \sum_{j \in \mathbb{Z}_{N_1, \ldots, N_d}} -\frac{1}{2} P_j (\Delta_1 P)_j + (F(Q))_j,
\]

\[
(F(Q))_j = \frac{Q_j^2}{2} + \frac{Q_j^3}{3} + \beta \frac{Q_j^4}{4}, \quad j \in \mathbb{Z}_{N_1, \ldots, N_d}.
\]

We assume \( N_1 \leq N_2, \ldots, N_d \), and we introduce the quantities

\[
\mu := \frac{2}{2N_1 + 1},
\]

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

Then we can describe the following regimes:

(KdV-d) the \( \alpha + \beta \) model, in the very weakly transverse regime with \( \mu \ll 1 \) and \( 2 < \sigma_1, \ldots, \sigma_{d-1} < 5 \);

(KP-d) the \( \alpha + \beta \) model, in the weakly transverse regime with \( \mu \ll 1 \) and \( \sigma_1 = 2, 2 < \sigma_2, \ldots, \sigma_{d-1} < 5 \).

Moreover, in order to obtain Theorem 2.1 and Theorem 2.2 we will have to assume that

\[
2\gamma + \sum_{i=1}^{d-1} \sigma_i < 7.
\]

which, together with the fact that \( \sigma_i > 2 \) for all \( i = 1, \ldots, d-1 \), is consistent with the assumption \( d \leq 4 \).
Remark 2.9. Theorem 2.4 can be generalized to higher dimensional lattices. Indeed, let \( d \leq 6 \), define \( \mathbb{Z}_{N_1, \ldots, N_d}^d \) as in (37) and consider the \( d \)-dimensional NLKG lattice

\[
H(Q, P) = \sum_{j \in \mathbb{Z}_{N_1, \ldots, N_d}^d} \frac{P_j^2}{2} + \frac{1}{2} \sum_{j, k \in \mathbb{Z}_{N_1, \ldots, N_d}^d} \frac{(Q_j - Q_k)^2}{2} + \sum_{j \in \mathbb{Z}_{N_1, \ldots, N_d}^d} U(Q_j),
\]

(43)

\[
U(x) = m^2 x^2 + \frac{\beta}{2} x^{2p+2}, \quad m > 0, \quad \beta > 0, \quad p \geq 1.
\]

(44)

We assume \( N_1 \leq N_2, \ldots, N_{d-1} \), and we introduce the quantities \( \mu \) and \( \sigma_i \) (\( 1 \leq i \leq d-1 \)) as in (40) and (41).

Then we can describe the following regime:

(1DNLS-d) the model (43) with \( m = 1 \) and \( p = 1 \) in the very weakly transverse regime, with \( \mu \ll 1 \), \( 1 < \sigma_1, \ldots, \sigma_{d-1} < 7 \);

Moreover, in order to obtain Theorem 2.4 we will have to assume that

\[
2\gamma + \sum_{i=1}^{d-1} \sigma_i < 7.
\]

(45)

which, together with the fact that \( \sigma_i > 1 \) for all \( i = 1, \ldots, d-1 \), is consistent with the assumption \( d \leq 6 \).

Remark 2.10. There are other interesting regimes for (8) and (27) especially for their relation with the modified KdV equation and two-dimensional Non-Linear Schrödinger equation respectively. These will be discussed in Remark 4.7 and Remark 4.12 respectively.

3. Galerkin Averaging

3.1. An Averaging Theorem. Following Pas19 (see also BP06 and Bam05) we use a Galerkin averaging method in order to approximate the solutions of the continuous approximation of the lattice with the solutions of the system in normal form.

To this end we first have to introduce a topology in the phase space. This is conveniently done in terms of Fourier coefficients.

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

\[
(v, w)_{\rho, s} := \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \rho_n |n|^{2s} e^{i\rho_n n \cdot y},
\]

(46)

and such that

\[
\|v\|_{\rho, s}^2 := (v, v)_{\rho, s} = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} |v_n|^2 |n|^{2s}
\]

(47)

is finite. We will denote by \( \ell^2 \) the space \( \ell^2_{0,0} \).

We will identify a 2-periodic function \( v \) with the sequence of its Fourier coefficients \( \{\hat{v}_n\}_n \),

\[
v(y) = \frac{1}{2} \sum_{n \in \mathbb{Z}^d} \hat{v}_n e^{i\pi n \cdot y},
\]

and we will say that \( v \in \ell^2_{\rho, s} \) if the sequence of its Fourier coefficients belong to \( \ell^2_{\rho, s} \).

Now fix \( \rho \geq 0 \) and \( s \geq 1 \), 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:

\[
\Omega_1 := \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \Omega_2 := \begin{pmatrix} -\partial_{x_1} & 0 \\ 0 & \partial_{x_1} \end{pmatrix}.
\]

(48)

Observe that \( \Omega_1 : \mathcal{H}^{\rho, s} \to \mathcal{H}^{\rho, s+\gamma-1} \) (\( \gamma = 1, 2 \)) is a well-defined operator. Moreover, \( \Omega_2 \) is well-defined on the space of functions with zero-average with respect to the \( x_1 \)-variable, i.e. on those functions \( \zeta(x_1, x_2) \) such that for every \( x_2 \) we have \( \int_{-1}^{1} \zeta(x_1, x_2) \, dx_1 = 0 \).
If we fix $\gamma \in \{1, 2\}$, $s$ and $U_s \subset \ell^2_{p,s}$ open, we define the gradient of $K \in C^\infty(U_s, \mathbb{R})$ with respect to $\xi \in \ell^2_{p,s}$ as the unique function s.t.

$$\langle \nabla \xi K, h \rangle = d_\xi K h, \quad \forall h \in \ell^2_{p,s}.$$  

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

$$X_H(\xi) = \Omega^{-1}_\xi \nabla H(\xi).$$

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

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

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

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

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

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

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

**Lemma 3.2.** The projection operators defined in (50) and (51) satisfy the following properties for any $\zeta \in \mathcal{H}^{p,s}$:

1. for any $j \geq 0$

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

2. for any $M \geq 0$

$$\|\Pi_M \zeta\|_{\mathcal{H}^{p,s}} \leq \|\zeta\|_{\mathcal{H}^{p,s}};$$

3. the following equality holds

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

where $|\zeta|$, for $\zeta \in \mathcal{H}^{p,s}$ is the element $|\zeta| \in \mathcal{H}^{p,s}$ whose $n$-th element is

$$|\zeta|_n := \langle (\zeta_n, |\eta_n|) \rangle$$

and $(\zeta^n)_n := (\zeta^n, \eta^n)$.

Now we consider a Hamiltonian system of the form

$$(53) \quad H = h_0 + \delta F,$$

where we assume that

(PER) $h_0$ generates a linear periodic flow $\Phi^T_{h_0}$ with period $T$,

$$\Phi^T_{h_0} = \Phi^{0,T}_{h_0} \quad \forall \tau,$$

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

(INV) for any $s \geq 1$, $\Phi^T_{h_0}$ leaves invariant the space $\Pi_j \mathcal{H}^{p,s}$ for any $j \geq 0$. Furthermore, for any $j \geq 0$

$$\pi_j \circ \Phi^T_{h_0} = \Phi^T_{h_0} \circ \pi_j.$$
Next, we assume that the vector field of $F$ admits an asymptotic expansion in $\delta$ of the form

$$F \sim \sum_{j \geq 1} \delta^{j-1} F_j,$$

and that the following property is satisfied

(HVF) There exists $R^* > 0$ such that for any $j \geq 1$

- $X_{F_j}$ is analytic from $B_{p,s+2j+1}(R^*)$ to $\mathcal{H}^{p,s}$.

Moreover, for any $r \geq 1$ we have that

- $X_{F - \sum_{j=1}^{r} \delta^{j-1} F_j}$ is analytic from $B_{p,s+2(r+1)+1}(R^*)$ to $\mathcal{H}^{p,s}$.

The main result of this section is the following theorem.

**Theorem 3.3.** Fix $R > 0$, $s_1 > 1$. Consider (53), and assume (PER), (INV) and (HVF). Then $\exists s_0 > 0$ with the following properties: for any $s \geq s_1$ there exists $\delta_* \ll 1$ such that for any $\delta < \delta_*$ there exists $T_0 : B_{p,s}(R/2) \to B_{p,s}(R)$ analytic canonical transformation such that

$$H_1 := H \circ T_0 = h_0 + \delta Z_1 + \delta^2 R^{(1)},$$

where $Z_1$ is in normal form, namely

$$\{Z_1, h_0\} = 0,$$

and there exists a positive constant $C'_s$ such that

$$\sup_{B_{p,s+2}(R/2)} \|X_{Z_1}\|_{\mathcal{H}^{p,s}} \leq C'_s,$$

$$\sup_{B_{p,s+2}(R/2)} \|X_{Z_1^{(1)}}\|_{\mathcal{H}^{p,s}} \leq C'_s,$$

$$\sup_{B_{p,s}(R/2)} \|T_0 - \text{id}\|_{\mathcal{H}^{p,s}} \leq C'_s \delta.$$

In particular,

$$Z_1(\zeta) = \langle F_1 \rangle(\zeta),$$

where $\langle F_1 \rangle(\zeta) := \int_0^T F_1 \circ \Phi^\tau_{h_0}(\zeta) d\tau$.

**Remark 3.4.** By using the same arguments of [Bam05] and [Pas19] one can prove a more general version of Theorem 3.3, in which the Hamiltonian is put in normal form up to order $r$, for any $r \geq 1$. In this latter case, both $\delta_*$ and $s_0$ will also depend on $r$.

### 3.2. Proof of the Averaging Theorem

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

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

Now we set $H = H_{1,M} + R_{1,M} + R_1$, where

$$H_{1,M} := h_0 + \delta F_{1,M},$$

$$F_{1,M} := F_1 \circ \Pi_M,$$

and

$$R_{1,M} := h_0 + \delta F_1 - H_{1,M},$$

$$R_1 := \delta (F - F_1).$$

The system described by the Hamiltonian (61) 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 $a \leq Kb$. We exploit the following intermediate results:
Lemma 3.5. For any $s \geq s_1$ there exists $R > 0$ such that $\forall \, \sigma > 0, M > 0$

\[
\sup_{B_{\rho,s+\gamma+\sigma+2}(R)} \|X_{\mathcal{R}_{1,M}}(\zeta)\|_{H^{\rho,s}} \lesssim \frac{\delta}{(M+1)^{\sigma}},
\]

\[
\sup_{B_{\rho,s+\gamma+4}(R)} \|X_{\mathcal{R}_{1}}(\zeta)\|_{H^{\rho,s}} \lesssim \delta^2.
\]

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

We first notice that $\|id - \Pi_M\|_{H^{\rho,s} \rightarrow H^{\rho,s}} = (M+1)^{-\sigma}$: indeed, using (52) we obtain

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

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

Inequality (65) follows from

\[
\sup_{\zeta \in B_{\rho,s+\gamma+2+\sigma}(R)} \|X_{\mathcal{R}_{1,M}}(\zeta)\|_{H^{\rho,s}} \lesssim \|dX_{\mathcal{R}_1} \|_{L^\infty(B_{\rho,s+2+\gamma}(R),H^{\rho,s})}\|id - \Pi_M\|_{L^\infty(B_{\rho,s+2+\gamma}(R),B_{\rho,s+2+\gamma}(R))} \lesssim \delta (M+1)^{-\sigma},
\]

while estimate (66) is an immediate consequence of (HVF). \qed

Lemma 3.6. For any $s \geq s_1$

\[
\sup_{B_{\rho,s}(R^*)} \|X_{F_{1,M}}(\zeta)\|_{H^{\rho,s}} \leq K_{1,s}^{(F)} M^{2+\gamma},
\]

where

\[
K_{1,s}^{(F)} := \sup_{B_{\rho,s}(R^*)} \|X_{F_1}(\zeta)\|_{H^{\rho,s-2-\gamma}} < +\infty.
\]

Proof. Using (52) we have

\[
\sup_{(\zeta) \in B_{\rho,s}(R)} \left\| \sum_{h \leq M} h X_{F_{1,M}}(\zeta) \right\|_{H^{\rho,s}} = \sup_{(\zeta) \in B_{\rho,s}(R)} \left\| \left[ \sum_{h \leq M} |h^s X_{F_{1,M}}(\zeta)|^2 \right]^{1/2} \right\|_{H^{\rho,0}} \leq M^{2+\gamma} \sup_{(\zeta) \in B_{\rho,s}(R)} \left\| X_{F_{1,M}}(\zeta) \right\|_{H^{\rho,s-2-\gamma}} = K_{1,s}^{(F)} M^{2+\gamma},
\]

where the last quantity is finite for $R \leq R^*$ by property (HVF). \qed

To normalize (61) we need a slight 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 3.7. Let $s \geq s_1 + 2 + \gamma, R > 0$, and consider the system (61). Assume that $\delta < \frac{1}{M}$, and that

\[
12 T K_{1,s}^{(F)} M^{2+\gamma} \delta < R
\]

where

\[
K_{1,s}^{(F)} := \sup_{(\zeta) \in B_{\rho,s}(R)} \|X_{F_1}(\zeta)\|_{H^{\rho,s-2-\gamma}}.
\]
Then there exists an analytic canonical transformation \( T_{h,M}^{(0)} : \mathcal{B}_{h,s}(R/2) \rightarrow \mathcal{B}_{h,s}(R) \) such that
\[
\sup_{\mathcal{B}_{h,s}(R/2)} \| T_{h,M}^{(0)}(C) - C \|_{H^{\rho,s}} \leq 2T K_{1,s}^{(F)} M^{2+\gamma} \delta,
\]
and that puts (61) in normal form up to a small remainder,
\[
H_{1,M} \circ T_{h,M}^{(0)} = h_0 + \delta Z_{h,M}^{(1)} + \delta^2 R_{h,M}^{(1)},
\]
with \( Z_{h,M}^{(1)} \) in normal form, namely \( \{ h_{0,M}, Z_{h,M}^{(1)} \} = 0, \) and
\[
\sup_{\mathcal{B}_{h,s}(R/2)} \| X_{h,M}^{(1)}(\cdot) \|_{H^{\rho,s}} \leq K_{1,s}^{(F)} M^{2+\gamma}
\]
\[
\sup_{\mathcal{B}_{h,s}(R/2)} \| X_{h,M}^{(1)}(\cdot) \|_{H^{\rho,s}} \leq 15K_{1,s}^{(F)} M^{2+\gamma}
\]

Now we conclude with the proof of the Theorem 3.3.

Proof. If we define \( \delta_s := \min\left\{ \frac{1}{12T K_{1,s}^{(F)} M^{2+\gamma}}, \frac{B}{12T K_{1,s}^{(F)} M^{2+\gamma}} \right\} \) and we choose
\[
\delta_0 = \sigma + 2 + \gamma,
\]
\[
\sigma \geq 2,
\]
then the transformation \( T_{S} := T_{h,M}^{(0)} \) defined by Lemma 3.7 satisfies (56) because of (72).

Next, Eq. (57) follows from Lemma 3.7, Eq. (58) follows from (73) and (74), while (59) is precisely (71). Finally, (60) can be deduced by applying Lemma A.6 to \( G = F_1. \)

4. Applications to two-dimensional lattices

4.1. The KdV regime for the ETL lattice. We want to study the behaviour of small amplitude solutions of (8), 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 \( j \in \mathbb{Z}_{N_1,N_2}; \)

(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, \( \int_{-(N_1+\frac{1}{2}),N_1+\frac{1}{2} \times -(N_2+\frac{1}{2}),N_2+\frac{1}{2}} Q(t,j) \) \( \forall t; \)

(A4) \( \tilde{Q} \) fulfills
\[
\bar{Q} = \Delta_1(Q + \alpha Q^2 + \beta Q^4),
\]
\[
\Delta_1 := 4 \sinh^2 \left( \frac{\partial_{x_1}}{2} \right) + 4 \sinh^2 \left( \frac{\partial_{x_2}}{2} \right).
\]

It is easy to verify that (75) is Hamiltonian with Hamiltonian function
\[
H(Q,P) = \int_{\left[ \frac{1}{2}, \frac{1}{2} \right] \times \left[ \frac{1}{2}, \frac{1}{2} \right]} \frac{-P \Delta_1 P + Q^2}{2} + \alpha Q^3 + \beta Q^4 \ dx,
\]
where \( P \) is a periodic function which has zero average and is canonically conjugated to \( Q. \)

First we consider (75), with \( \alpha \neq 0, \) and we look for small amplitude solutions of the form
\[
Q(t,x) = \mu^2 q(\mu t, \mu x_1, \mu^\sigma x_2),
\]
where \( q : \mathbb{R} \times \mathbb{T}^2 \rightarrow \mathbb{R} \) is a periodic function and \( \mu, \sigma \) are defined in (14)-(15). We introduce the rescaled variables \( \tau = \mu t, y_1 = \mu x_1, y_2 = \mu^\sigma x_2, \) and we denote
\[
I := [-1,1]^2.
\]
Plugging (78) into (75), we get

\begin{align}
q_{\tau \tau} &= \frac{\Delta_{\mu,y_1,\sigma}}{\mu^2} (q + \mu^2 \alpha q^2), \\
\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),
\end{align}

which is a Hamiltonian PDE corresponding to the Hamiltonian functional

\begin{align}
K_1(q, p) &= \int_I -p \frac{\Delta_{\mu,y_1,\sigma} p}{2\mu^2} + \frac{q^2}{2} + \alpha \mu^2 \frac{q^3}{3} dy,
\end{align}

and \(p\) is the variable canonically conjugated to \(q\).

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

\begin{align}
\frac{\Delta_{\mu,y_1,\sigma}}{\mu^2} &\sim \partial_y^2 + \mu^2(\sigma - 1) \partial_{y_2}^2 + \sum_{m \geq 1} c_m \left( \mu^{2m} \partial_{y_1}^{2(m+1)} + \mu^{2[(m+1)\sigma - 1]} \partial_{y_2}^{2(m+1)} \right), \\
c_m &:= \frac{2}{(2m)!},
\end{align}

which, up to terms of order \(O(\mu^4)\), reads

\begin{align}
\frac{\Delta_{\mu,y_1,\sigma}}{\mu^2} &\sim \partial_y^2 + \frac{\mu^2}{12} \partial_{y_1}^4 + O(\mu^4),
\end{align}

(recall that \(\sigma > 2\)). Therefore the Hamiltonian (82) admits the following asymptotic expansion

\begin{align}
K_1(q, p) &\sim \hat{h}_0(q, p) + \mu^2 \hat{F}_1(q, p) + \mu^4 \hat{R}(q, p), \\
\hat{h}_0(q, p) &= \int_I -p \left( \mu \frac{\partial_y^2 p}{2} + q^2 \right) dy, \\
\hat{F}_1(q, p) &= \int_I -p \frac{\partial_y^4 p}{24} + \alpha \frac{q^3}{3} dy.
\end{align}

Note that the nonlinearity of degree 4 does not affect the Hamiltonian up to order \(O(\mu^4)\). Following the approach of [BP06], we can introduce the following non-canonical change of coordinates

\begin{align}
\xi &:= \frac{1}{\sqrt{2}} (q + \partial_y p), \\
\eta &:= \frac{1}{\sqrt{2}} (q - \partial_y p).
\end{align}

Since the previous transformation is not canonical, the Poisson tensor in these new coordinates is

\begin{align}
J &= \partial_{y_1} \left( \begin{array}{cc}
-1 & 0 \\
0 & 1
\end{array} \right),
\end{align}

and Hamilton equations associated to a Hamiltonian \(K_1\) are

\begin{align}
\partial_y \xi &= -\partial_{y_1} \frac{\delta K_1}{\delta \xi}, \\
\partial_y \eta &= \partial_{y_1} \frac{\delta K_1}{\delta \eta}.
\end{align}

**Remark 4.1.** The explicit expression of the Poisson tensor (91) let us compute straightforwardly Casimir invariants associated to \(J\), which are

\begin{align}
C(\xi, \eta) &= A + B \int_{-1}^{1} \xi(\tau, y_1, y_2) dy_1 + C \int_{-1}^{1} \eta(\tau, y_1, y_2) dy_1,
\end{align}

where \(A, B, C\) are arbitrary real constants.
However, by recalling (93) which is true due to periodic boundary conditions.

\[ \int_{-1}^{1} \xi(\tau, y_1, y_2) - \eta(\tau, y_1, y_2) \, dy_1 = 0 \quad \forall \tau \in \mathbb{R}, \ |y_2| \leq 1. \]

(94)

Thus, Lemma is proved. \( \square \)

Moreover, if we denote by \( K_1(\xi, \eta) \) the Fourier series of \( \frac{1}{2} \int_{-1}^{1} \mathcal{A}(\xi, \eta) \, dy \),

\[ K_1(\xi, \eta) = \int_{-1}^{1} \frac{(\partial_{\eta_1} \xi)^2 + (\partial_{\eta_1} \eta)^2}{48} \, dy + \frac{a}{3 \cdot 2^{3/2}} (|\xi|^3 + |\eta|^3), \]

Now we apply the averaging Theorem 3.3 to the Hamiltonian (95), with \( \delta = \mu^2 \): observe that the equations of motion of \( h_0 \) have the following simple form:

\[ \xi_t = -\partial_{\eta_1} \xi, \quad \eta_t = \partial_{\eta_1} \eta. \]

Proposition 4.2. The average of \( F_1 \) in (95) with respect to the flow of \( h_0 \) in (96) is given by

\[ \langle F_1 \rangle(\xi, \eta) = -\int_{-1}^{1} (\partial_{\eta_1} \xi)^2 + (\partial_{\eta_1} \eta)^2\, dy + \frac{a}{3 \cdot 2^{3/2}} (|\xi|^3 + |\eta|^3), \]

where we denote by \( [f^+] \) the average \( \int_{-1}^{1} f^+ \, dy \).

The proof of this proposition is a straightforward application of the following two lemmas.

Lemma 4.3. Given two functions \( u, v \in L^2([-1, 1]) \)

\[ \int_{-1}^{1} dy \int_{-1}^{1} ds \, u(y \pm s) v(y \mp s) = \int_{-1}^{1} u(y) \, dy \int_{-1}^{1} v(y) \, dy. \]

Proof. Denoting with \( \{\hat{u}_k\}_k \) and \( \{\hat{v}_k\}_k \) the Fourier series of \( u \) and \( v \) respectively and using Plancherel theorem one obtains

\[ \int_{-1}^{1} dy \int_{-1}^{1} ds \, u(y \pm s) v(y \mp s) = \frac{1}{2} \int_{-1}^{1} dy \int_{-1}^{1} ds \sum_{k, k' \in \mathbb{Z}} \hat{u}_k \hat{v}_{k'} e^{i \pi k (y + s)} e^{-i \pi k' (y - s)} = \hat{u}_0 \hat{v}_0 \]

and thus Lemma is proved. \( \square \)

Lemma 4.4. Given a function \( u \in L^1([-1, 1]) \) then

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

Proof. The thesis follows by a simple change of coordinates \( x := y \pm s. \) \( \square \)

Proof of Proposition 4.2. For the computation of \( \langle F_1 \rangle(\xi, \eta) \) one can exchange the order of the integrations and apply Lemma 4.3 and 4.4. \( \square \)

Corollary 4.5. The equations of motion associated to \( h_0(\xi, \eta) + \mu^2 \langle F_1 \rangle(\xi, \eta) \) are given by

\[ \begin{cases} 
\xi_t = -\partial_{\eta_1} \xi - \frac{\omega^2}{24} \partial_{\eta_1} \xi - \frac{\mu^2}{2 \cdot 2^{3/2}} \partial_{\eta_1} (\xi^3) \\
\eta_t = \partial_{\eta_1} \eta + \frac{\omega^2}{24} \partial_{\eta_1} \eta + \frac{\mu^2}{2 \cdot 2^{3/2}} \partial_{\eta_1} (\eta^3). 
\end{cases} \]
Remark 4.7. One can also study the regime, corresponding Hamiltonian, leads to the system (106) which, combined with the fact that

\[ H(Q, P) = \int_{\tau}^{\frac{1}{2}P_{\Delta}P + \frac{Q^2}{2} + \alpha \frac{Q^3}{3} + \beta \frac{Q^4}{4} \, dx, \]

and makes the ansatz (78) about the solution, one gets the rescaled Hamiltonian

\[ K_1(q, p) = \int_{\mu, \sigma}^{P_{\Delta}P + \frac{Q^2}{2} + \alpha \frac{Q^3}{3} + \beta \frac{Q^4}{4} \, dy, \]

\[ \Delta_{\mu, \gamma, \sigma} := 4 \sinh^2 \left( \frac{\mu \partial_{q}}{\mu \partial_{\gamma}} \right) + 4 \sinh^2 \left( \frac{\sigma \partial_{q}}{\sigma \partial_{\gamma}} \right), \]

\[ I_{\mu, \sigma} := [-1, 1] \times [-\mu^{-1}, \mu^{-1}], \]

which, combined with the fact that

\[ \int_{-1}^{1} \xi(\tau, y_1, y_2) = 0 \quad \forall \tau \in \mathbb{R}, \quad |y_2| \leq \mu^{-1}, \]

leads to the system (100) of two uncoupled KdV equations in translating frames with respect to the \( y_1 \)-direction.

Remark 4.6. If one considers a square lattice, namely

(101) \[ H(Q, P) = \sum_{j \in \mathbb{Z}} \frac{1}{2} P_{\Delta}P + (F(Q)), \]

with \( F(Q) \) as in (7), with its continuous approximation

(102) \[ H(Q, P) = \int_{\tau}^{\frac{1}{2}P_{\Delta}P + \frac{Q^2}{2} + \alpha \frac{Q^3}{3} + \beta \frac{Q^4}{4} \, dx, \]

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 (79); plugging (107) into (75), we get

(108) \[ q_{\tau} = \frac{\Delta_{\mu, \gamma, \sigma}}{\mu^2} (q + \mu^2 \beta q^3), \]

where \( \Delta_{\mu, \gamma, \sigma} \) is the operator introduced in (81). Eq. (108) is a Hamiltonian PDE with the following corresponding Hamiltonian,

(109) \[ K_2(q, p) = \int_{I}^{P_{\Delta}P + \frac{Q^2}{2} + \alpha \frac{Q^3}{3} + \beta \frac{Q^4}{4} \, dy, \]

where \( p \) is the variable canonically conjugated to \( q \).

Recalling that (93) holds true, we exploit again the non-canonical change of coordinates (89)-(90) and the Poisson tensor (91), obtaining that

(110) \[ K_2(\xi, \eta) = h_0(\xi, \eta) + \mu^2 F_1(\xi, \eta) + \mu^4 R(\xi, \eta), \]

where \( h_0 \) is the same as in (96), while

(111) \[ F_1(\xi, \eta) = \int_{\tau}^{\frac{[\partial_{\xi}(\xi - \eta)]^2}{48} + \beta (\xi + \eta)^4 \, dy. \]

Applying Theorem 3.3 to the Hamiltonian (110) with \( \delta = \mu^2 \), we get that the equations of motion associated to \( h_0(\xi, \eta) + \mu^2 (F_1) \) \( (\xi, \eta) \) are given by

(112) \[ \begin{cases} \xi_{\tau} = - (1 + \frac{4}{3}[\xi^2]) \partial_{\xi} \xi - \frac{\mu^2}{3} \partial_{\eta} \xi - \frac{\mu^2}{2} \partial_{\eta} (\xi^3) \\ \eta_{\tau} = (1 + \frac{4}{3}[\xi^2]) \partial_{\eta} \eta + \frac{\mu^2}{3} \partial_{\xi} \eta + \frac{\mu^2}{2} \partial_{\xi} (\eta^3) \end{cases}. \]
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].

4.2. The KP regime for the ETL lattice. For this regime we consider (75), with \( \alpha \neq 0 \), and we look for small amplitude solutions of the form

\[
Q(t, x) = \mu^2 q(\mu t, \mu x_1, \mu^2 x_2),
\]

with \( \mu \) as in (14). We introduce the rescaled variables \( \tau = \mu t, y_1 = \mu x_1, y_2 = \mu^2 x_2 \).

Plugging (113) into (75), leads to

\[
q_{\tau} = \frac{\Delta_{\mu,y_1}}{\mu^2} (q + \mu^2 \alpha q^3),
\]

(114)

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

(115)

which is a Hamiltonian PDE corresponding to the Hamiltonian functional,

\[
K_3(q, p) = \int_I - \frac{p \Delta_{\mu,y_1} p}{2\mu^2} + q^2 + \alpha \mu^2 \frac{q^4}{3} + \beta \mu^4 \frac{q^4}{4} dy,
\]

(116)

where \( I \) is as in (79), and \( p \) is the variable canonically conjugated to \( q \).

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

\[
\frac{\Delta_{\mu,y_1}}{\mu^2} \sim \delta_{y_1}^2 + \mu^2 \delta_{y_2}^2 + \frac{\mu^2}{12} \delta_{y_1}^4 + O(\mu^4),
\]

(117)

Therefore the Hamiltonian (116) admits the following asymptotic expansion

\[
K_3(q, p) \sim h_0(q, p) + \mu^2 F_1(q, p) + \mu^4 \mathcal{R}(q, p),
\]

(118)

\[
h_0(q, p) = \int_I - \frac{p \partial_{y_1}^2 p}{2} \partial_{y_1}^2 + \frac{p \partial_{y_2}^2 p}{2} + \alpha \frac{q^3}{3} dy,
\]

(119)

\[
F_1(q, p) = \int_I - \frac{p \partial_{y_1}^4 p}{24} - \frac{p \partial_{y_2}^4 p}{2} + \alpha \frac{q^3}{3} dy.
\]

(120)

By exploiting again the non-canonical change of coordinates \( (q, p) \mapsto (\xi, \eta) \) introduced in (89)-(90) and the Poisson tensor (91), and

\[
\int_{-1}^{1} \xi(\tau, y_1, y_2) - \eta(\tau, y_1, y_2) dy_1 = 0 \quad \forall \tau \in \mathbb{R}, \ |y_2| \leq 1,
\]

(121)

we obtain

\[
K_3(\xi, \eta) \sim h_0(\xi, \eta) + \mu^2 F_1(\xi, \eta) + \mu^4 \mathcal{R}(\xi, \eta),
\]

(122)

\[
h_0(\xi, \eta) = \int_I \frac{\xi^2 + \eta^2}{2} dy,
\]

(123)

\[
F_1(\xi, \eta) = \int_I - \frac{\partial_{y_1}^2 (\xi - \eta)}{48} + \frac{\partial_{y_2}^2 (\xi - \eta)}{4} + \alpha \frac{4}{3} \frac{\xi^3}{2} \frac{\eta^3}{2} dy,
\]

(124)

where (124) is well defined because of (121).

Now we apply the averaging Theorem 3.3 to the Hamiltonian (122), with \( \delta = \mu^2 \).

**Proposition 4.8.** The average of \( F_1 \) in (122) with respect to the flow of \( h_0 \) in (122) is given by

\[
\langle F_1 \rangle (\xi, \eta) = \int_I - \frac{\partial_{y_1}}{48} (\xi - \eta)^2 + \frac{\partial_{y_2}^2 (\xi - \eta)}{4} dy + \frac{\alpha}{3} \frac{\xi^3}{2} \frac{\eta^3}{2} (\xi^3 + \eta^3)
\]

(125)

where we denote by \( [f] \) the average \( \int_I f(y) dy \).
Corollary 4.9. The equations of motion associated to \( h_0(\xi, \eta) + \mu^2 \langle F_1 \rangle (\xi, \eta) \) are given by

\[
\begin{align*}
\xi_{tt} & = -\partial_{\xi_t} \xi - \frac{\mu^2}{24} \partial_{\xi_t}^3 \xi - \frac{\mu^2}{2} \partial_{\eta_{1t}}^2 \partial_{\eta_{2t}} \xi - \frac{\mu^2}{2} \partial_{\eta_{1t}} \partial_{\eta_{2t}} \xi, \\
\eta_{tt} & = \partial_{\xi_t} \eta + \frac{\mu^2}{2} \partial_{\eta_{1t}} \partial_{\eta_{2t}} \eta + \frac{\mu^2}{2} \partial_{\eta_{1t}}^2 \eta + \frac{\mu^2}{2} \partial_{\eta_{1t}} \partial_{\eta_{2t}} \eta.
\end{align*}
\]

More explicitly, we observe that (126) is a system of two uncoupled KP equations on a two-dimensional torus in translating frames.

4.3. The one-dimensional NLS regime for the KG Lattice. We want to study small amplitude solutions of (27), with initial data in which only one low-frequency Fourier mode is excited.

Starting from the Hamiltonian (24), where \( p \) is a periodic function and is canonically conjugated to \( q \), we introduce the rescaled variable \( y = \mu^t t, \mu x_1, \mu^s x_2 \).

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

\[
\psi = \frac{1}{\sqrt{2}} (q - ip),
\]

and the equation of motion associated to \( h_0 + F_1 \) is given by the following cubic one-dimensional nonlinear Klein-Gordon (NLKG) equation.

\[
q_{tt} = -(q - \mu^2 \partial_q q) - \mu^2 \beta q^3.
\]

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

\[
\psi = \frac{1}{\sqrt{2}} (q - ip),
\]

and the equation of motion associated to \( h_0 + F_1 \) is given by the following cubic one-dimensional nonlinear Klein-Gordon (NLKG) equation.

\[
q_{tt} = -(q - \mu^2 \partial_q q) - \mu^2 \beta q^3.
\]
therefore the inverse change of coordinates is given by
\begin{align}
q &= \frac{1}{\sqrt{2}}(\psi + \bar{\psi}), \\
p &= \frac{1}{\sqrt{2}}(\psi - \bar{\psi}),
\end{align}
while the symplectic form is given by $-id\psi \wedge d\bar{\psi}$. With this change of variables the Hamiltonian takes the form
\begin{align}
K_4(\psi, \bar{\psi}) &\sim \hbar(\psi, \bar{\psi}) + \mu^2 F_1(\psi, \bar{\psi}) + \mu^{2(\sigma - 1)} R(\psi, \bar{\psi}), \\
\hbar(\psi, \bar{\psi}) &= \int \psi \bar{\psi} dy,
\end{align}
\begin{align}
F_1(\psi, \bar{\psi}) &= \int -\frac{(\psi + \bar{\psi})(-\partial_y^2(\psi + \bar{\psi}))}{4} + \frac{\beta(\psi + \bar{\psi})^4}{16} dy.
\end{align}
Now we apply the averaging Theorem 3.3 to the Hamiltonian (139), with $\delta = \mu^2$. Observe that $\hbar$ generates a periodic flow,
\begin{align}
-i\partial_t \psi &= \psi; \\
\psi(t, y) &= e^{it} \psi_0(y).
\end{align}

**Proposition 4.10.** The average of $F_1$ in (139) with respect to the flow of $\hbar$ (133) is given by
\begin{equation}
\langle F_1 \rangle(\psi, \bar{\psi}) = \int \psi \left(-\frac{\partial_y^2}{2}\right) dy + \frac{3\beta}{8} \int |\psi|^4 dy.
\end{equation}

**Corollary 4.11.** The equations of motion associated to $\hbar(\psi, \bar{\psi}) + \mu^2 \langle F_1 \rangle(\psi, \bar{\psi})$ are given by a cubic one-dimensional nonlinear Schrödinger equation for each fixed value of $y_2$,
\begin{align}
-i\psi_t &= -\mu^2 \partial_y^2 \psi + \mu \frac{3\beta}{4} |\psi|^2 \psi.
\end{align}

**Remark 4.12.** Let us consider the Hamiltonian (24) in the following regime,(2-D NLS) the scalar model (24) with $m = 1, p = 1$ and
\begin{equation}
Q(t, x) = \mu q(\mu^2 t, \mu x),
\end{equation}
where $\mu < 1$ and $\sigma = 1$.

If we introduce the rescaled variable $y = \mu x$ and we define $I$ as in (79), we have that the Hamiltonian takes the following form (we denote by $p$ the variable canonically conjugated to $q$)
\begin{align}
K_5(q, p) &= \int \frac{p^2}{2} + q^2 - \frac{q \Delta_\mu q}{2} + \beta \mu^4 q^4 dy, \\
\Delta_\mu := 4 \sinh^2 \left( \frac{\mu \partial_y}{2} \right) + 4 \sinh^2 \left( \frac{\partial_y}{\mu} \right).
\end{align}

By expanding the operator $\Delta_\mu$ and by exploiting the change of variable (136), we get
\begin{align}
K_5(\psi, \bar{\psi}) &\sim \hbar(\psi, \bar{\psi}) + \mu^2 F_1(\psi, \bar{\psi}) + \mu^4 R(\psi, \bar{\psi}), \\
\hbar(\psi, \bar{\psi}) &= \int \psi \bar{\psi} dy,
\end{align}
\begin{align}
F_1(\psi, \bar{\psi}) &= \int -\frac{(\psi + \bar{\psi})(-\partial_y^2(\psi + \bar{\psi}))}{4} + \frac{\beta(\psi + \bar{\psi})^4}{16} dy.
\end{align}
By applying Theorem 3.3 to the Hamiltonian (139), with $\delta = \mu^2$, we obtain that the equation of motion associated to $\hbar(\psi, \bar{\psi}) + \mu^2 \langle F_1 \rangle(\psi, \bar{\psi})$ is given by the cubic nonlinear Schrödinger (NLS) equation
\begin{equation}
-i\psi_t = -\mu^2 \Delta \psi + \mu^2 \frac{3\beta}{4} |\psi|^2 \psi.
\end{equation}

The local well-posedness of the NLS equation (151) in the Sobolev space $H^s(\mathbb{T}^2), s > 0$, has been discussed by Bourgain in [Bou93a]; along with the conservation laws, this implies the global existence in
the defocusing case ($\beta > 0$), and the global existence for small solutions in the focusing case ($\beta < 0$).

The long time dynamics of the NLS equation has also been studied, in relation with the transfer of energy among Fourier modes and with the growth of Sobolev norms [CKS+10] [CF12] [Han14] [GK15] [GHP16].

5. Dynamics of the normal form equation

5.1. 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 [KP03].

Consider the KdV equation

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

Through the Lax pair formulation of the evolution problem (152) one gets that the periodic eigenvalues ($\lambda_n$)$_{n \in \mathbb{N}}$ of the Sturm-Liouville operator

$$L_\xi := -\partial^2_y + 6\sqrt{2} \xi (\tau, y_1)$$

are conserved quantities under the evolution of the KdV equation (152). Moreover, if we define the gaps of the spectrum $\gamma_m := \lambda_{2m} - \lambda_{2m-1} (m \geq 1)$, it is well known that the squared spectral gaps ($\gamma^2_m$)$_{m \geq 1}$ form a complete set of constants of motion for (152).

The following relation between the sequence of the spectral gaps and the regularity of the corresponding solution to the KdV equation holds (see Theorem 9, Theorem 10 and Theorem 11 in [KP08]; see also [Pôs11])

**Theorem 5.1.** Assume that $\xi \in L^2$, then $\xi \in \ell^2_{\rho,s}$ if and only if its spectral gaps satisfy

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

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

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

conversely, if (154) 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 [KP03])

**Theorem 5.2.** There exists a diffeomorphism $\Omega : L^2 \to \ell^2_{0,1/2} \times \ell^2_{0,1/2}$ such that:

- $\Omega$ is bijective, bianalytic and canonical;
- for each $s \geq 0$, the restriction of $\Omega$ 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 (152) depends only on the action $I_m := \frac{x^2_m + y^2_m}{2} (m \geq 1)$.

The dynamics of the KdV equation (152) 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.
5.2. The KP equation. In this section we recall some known facts on the dynamics of the KP equation on the two-dimensional torus

\begin{equation}
\xi_t = -\frac{1}{24} \partial_{\psi^1}^3 \xi - \frac{1}{2} \partial_{\psi^1}^{-1} \partial_{\psi^2}^2 \xi - \frac{\alpha}{2\sqrt{2}} \partial_{\psi^1}(\xi^2), \quad \alpha = \pm 1, \quad y \in \mathbb{T}^2 := \mathbb{R}^2/(2\pi \mathbb{Z})^2.
\end{equation}

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 \[LC82\] \[CLL83\] \[CL87\]. It is customary to refer to (155) 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 \[Bou93b\]. 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 Theorem 8.10 and Theorem 8.12 in \[Bou93b\]).

**Theorem 5.3.** Consider (155) with \(\alpha = 1\).

Let \(\rho \geq 0\) and \(s \geq 0\), and assume that the initial datum \(\xi(0, \cdot, \cdot) = \xi_0 \in \ell^2_{\rho,s}\). Then (155) is globally well-posed in \(\ell^2_{\rho,s}\). Moreover, the \(\ell^2\)-norm of the solution is conserved,

\begin{equation}
||\xi(t)||_{\ell^2} = ||\xi_0||_{\ell^2},
\end{equation}

while

\begin{equation}
||\xi(t)||_{\ell^2_{0,s}} \leq e^{C||t||} ||\xi_0||_{\ell^2_{0,s}},
\end{equation}

where \(C\) depends on \(s\).

**Remark 5.4.** As pointed out by Bourgain in Sec. 10.2 of \[Bou93b\], a global well-posedness result for sufficiently smooth solution of the KP-I equation (namely, (155) 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.

For the KP equation the construction of action-angle/Birkhoff coordinates is still an open problem.

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

\begin{equation}
i\psi_t = -\partial_{\psi^1}^2 \psi + 2|\psi|^2 \psi, \quad y_1 \in \mathbb{T} := \mathbb{R}/(2\pi \mathbb{Z}).
\end{equation}

Eq. (158) 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

\begin{equation}
\{F,G\} := -i \int_{\mathbb{T}} \left( \partial_{\phi_1} F \partial_{\phi_2} G - \partial_{\phi_2} F \partial_{\phi_1} G \right) d y_1,
\end{equation}

\begin{equation}
H_{NLS}(\phi_1, \phi_2) := \int_{\mathbb{T}} \partial_{\phi_1} \phi_1 \partial_{\phi_2} \phi_2 + \phi_1^2 \phi_2^2 d y_1.
\end{equation}

The defocusing NLS equation (158) is obtained by restricting (160) to the invariant subspace of states of real type,

\begin{equation}
\mathcal{H}^{\rho,s} := \{ \phi \in \mathcal{H}^{\rho,s} : \phi_2 = \bar{\phi}_1 \}.
\end{equation}

The above Hamiltonian (160) is well-defined on \(\mathcal{H}^{\rho,s}\) with \(s \geq 1\) and \(\rho \geq 0\), while the initial value problem for the NLS equation (158) 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 (158) 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

\begin{equation}
L(\phi) = \begin{pmatrix}
i & 0 \\
0 & -i \end{pmatrix} \partial_{y_1} + \begin{pmatrix} 0 & \phi_1 \\
\phi_2 & 0 \end{pmatrix}.
\end{equation}
where we call $\phi$ 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
\[
\cdots < \lambda_{-1} < \lambda_0^+ < \lambda_0^- < \lambda_1^- < \lambda_1^+ < \cdots ,
\]
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 (158).

Grébert, Kappeler and Mityagin proved the following relation between the sequence of the squared spectral gaps and the regularity of the corresponding potential (see Theorem in [GKK98]).

**Theorem 5.5.** Let $\rho \geq 0$ and $s > 0$, then for any bounded subset $B \subset L^2_s \times L^2_s$ there exists $n_0 \geq 1$ and $M \geq 1$ such that for any $|k| \geq n_0$ and any $(\phi_1, \phi_2) \in B$, the following estimate holds
\[
\sum_{|k| \geq n_0} (1 + |k|)^{2s} e^{2\rho|k|} |\gamma_m| \leq M.
\]

Moreover, Grébert and Kappeler constructed the following global Birkhoff coordinates (see Theorem 20.1 - Theorem 20.3 in [GKK14]):

**Theorem 5.6.** There exists a diffeomorphism $\Omega : L^2 \to \mathcal{H}^{0,0}_r$ such that:
- $\Omega$ is biaalytic and canonical;
- for each $s \geq 0$, the restriction of $\Omega$ to $\mathcal{H}^{0,s}_r$, namely the map
  \[ \Omega : \mathcal{H}^{0,s}_r \to \mathcal{H}^{0,s}_r \]
is again biaalytic and canonical;
- the coordinates $(x, y) \in \mathcal{H}^{0,1}_r$ are Birkhoff coordinates for the NLS equation, namely they form a set of canonically conjugated coordinates in which the Hamiltonian of the NLS equation (158) 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 (158) 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 (8) and (27), and we estimate the difference with respect to the true solutions with corresponding initial data.

The approach is the same for all the regimes (78), (113) and (129). First, we have to point out a relation between the energy of normal mode $E_k$ (defined in (10) for (8), and in (10) for (27)), $k \in \mathbb{Z}_{2N+1}$, and the Fourier coefficients of the solutions of the normal form equations. Then we have to prove that the approximate solutions approximate the energy of the true normal mode $E_k$ up to the time-scale in which the continuous approximation is valid, and finally we can deduce the result about the dynamics of the lattice.

**6.1. The KdV regime.** Let $I = [-1, 1]^2$ be as in (79), we define the Fourier coefficients of the function $q : I \to \mathbb{R}$ by
\[
\hat{q}(j) := \frac{1}{2} \int_I q(y_1, y_2) e^{-i\pi (j_1 y_1 + j_2 y_2)} dy_1 dy_2 ,
\]
and similarly for the Fourier coefficients of the function $p$.

**Lemma 6.1.** Consider the lattice (6) in the regime (KdV) and with interpolating function (78). Then for a state corresponding to $(q, p)$ one has
\[
\mathcal{E}_n = \frac{\mu^4}{2} \sum_{L = (L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu^2 L_2 \in \mathbb{Z}} |\hat{q}_{K + L}|^2 + \omega_k^2 \left| \frac{\hat{p}_{K + L}}{\mu} \right|^2 , \quad \forall k : k(k) = (\mu K_1, \mu^2 K_2)
\]
(where the $\omega_k$ are defined as in (11) and the $\mathcal{E}_n$ in (13)), and $\mathcal{E}_n = 0$ otherwise.
Proof. First we introduce a \((2N_1 + 1)(2N_2 + 1)\)-periodic interpolating function for \(Q_j\), namely a smooth function \(Q : (t,x) \mapsto Q(t,x)\) such that
\[
Q_j(t) = Q(t,j), \quad \forall t,j,
\]
\[
Q(t,x_1,x_2 + 2N_2 + 1) = Q(t,x), \quad \forall t,x,
\]
\[
Q(t,x_1 + 2N_1 + 1,x_2) = Q(t,x), \quad \forall t,x,
\]
and similarly for \(P_j\). We denote by
\[
(167)
\]
\[
\hat{Q}(j) := \frac{1}{(2N_1 + 1)^{1/2}(2N_2 + 1)^{1/2}} \left[ \int_{-\cdot (N_1 + \frac{1}{2}), (N_1 + \frac{1}{2})}^{\cdot (N_1 + \frac{1}{2}), (N_1 + \frac{1}{2})} Q(x)e^{-i\cdot (2N_1 + 1)^{1/2}(2N_2 + 1)^{1/2}} dx, 
\]
so that by the interpolation property we obtain
\[
(168)
\]
\[
\hat{Q}_k = \sum_{h \in \mathbb{Z}^2} \hat{Q}(k_1 + (2N_1 + 1)h_1, k_2 + (2N_2 + 1)h_2).
\]
The relation between \(\hat{Q}(k)\) and \(q_k\) can be deduced from (78),
\[
(169)
\]
and similarly
\[
(170)
\]
By using (10), (13) and (168)-(170) we have
\[
(169) , (169) \Rightarrow (16)
\]
for all \(k\) such that \(\kappa(k) = (\mu K_1, \mu^n K_2)\), and this leads to (166).
Proposition 6.2. Fix \( \rho > 0 \) and \( 0 < \delta \ll 1 \). Consider the normal form system (100), and define the Fourier coefficients of \( (\xi, \eta) \) through the following formula

\[
\xi(y) = \frac{1}{2} \sum_{h \in \mathbb{Z}^2} \xi_h e^{ih \cdot y_1},
\]

\[
\eta(y) = \frac{1}{2} \sum_{h \in \mathbb{Z}^2} \eta_h e^{ih \cdot y_1},
\]

Consider \( (\xi, \eta) \in \mathcal{H}^{\rho,0} \), and denote by \( E_\kappa \) the specific energy of the normal mode with index \( \kappa \) as defined in (12)-(13). Then for any positive \( \mu \) sufficiently small

\[
|E_\kappa - \mu \frac{|\xi_\kappa|^2 + |\eta_\kappa|^2}{2}| \leq C \mu^{4+\frac{4}{\rho}} \| (\xi, \eta) \|_{\mathcal{H}^{\rho,0}}^2
\]

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

\[
|E_\kappa| \leq C \mu^8 \| (\xi, \eta) \|_{\mathcal{H}^{\rho,0}}^2
\]

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

We defer the proof of the above Proposition to Appendix B.

Now, consider the following system of uncoupled KdV equations

\[
\partial_\tau \xi = -\frac{1}{24} \partial^3 \xi - \frac{\alpha}{2 \sqrt{2}} \partial_\eta (\xi^2),
\]

\[
\partial_\tau \eta = \frac{1}{24} \partial^3 \eta + \frac{\alpha}{2 \sqrt{2}} \partial_\eta (\eta^2),
\]

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

We consider the approximate solutions \((Q_\alpha, P_\alpha)\) of the FPU model (75)

\[
Q_\alpha(\tau, y) := \frac{\mu}{\sqrt{2}} \left[ \tilde{\xi}(\mu^2 \tau, y_1 - \tau, y_2) + \tilde{\eta}(\mu^2 \tau, y_1 + \tau, y_2) \right],
\]

\[
\partial_\eta P_\alpha(\tau, y) := \frac{\mu}{\sqrt{2}} \left[ \tilde{\xi}(\mu^2 \tau, y_1 - \tau, y_2) - \tilde{\eta}(\mu^2 \tau, y_1 + \tau, y_2) \right].
\]

We need to compare the difference between the approximate solution (177)-(178) and the true solution of (8). Let consider an initial datum \((Q_0, P_0)\) with corresponding Fourier coefficients \((Q_{0,k}, P_{0,k})\) given by (9), where

\[
Q_{0,k} \neq 0 \quad \text{only if} \quad \kappa(k) = (\mu K_1, \mu^n K_2).
\]

We also assume that there exist \( C, \rho > 0 \) such that

\[
\frac{|Q_{0,k}|^2 + \omega_1^2 |P_{0,k}|^2}{(2N_1 + 1)(2N_2 + 1)} \leq C \mu^{-2\rho(|\kappa(k)|/\mu^n, 2)}.
\]

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

\[
Q_0(y) = \frac{1}{(2N_1 + 1)(2N_2 + 1)} \sum_{K: |\mu^2| |K_1|^2 + \mu^{2n} |K_2|^2 |^{1/2} = |\kappa(k)| \leq 1} Q_{0,k} e^{i\kappa(y_1 + \mu^n K_2 y_2)},
\]

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

Proposition 6.3. Consider (8) with \( \sigma > 2 \) and \( \gamma \geq 1 \) such that \( \sigma + 2\gamma < 7 \). Let us assume that the initial datum satisfies (179)-(180), and denote by \((Q(t), P(t))\) the corresponding solution. Consider the approximate solution \((\tilde{\xi}(t, x), \tilde{\eta}(t, x))\) with the corresponding initial datum. Assume that \((\tilde{\xi}, \tilde{\eta}) \in \mathcal{H}^{\rho,n}\) for some \( \rho > 0 \) and for some \( n \geq 1 \) for all times, and fix \( T_0 > 0 \) and \( 0 < \delta \ll 1 \).
Then there exists \( \mu_0 = \mu_0(T_0, \| (\xi_\alpha(0), \eta_\alpha(0)) \|_{\mathcal{H}^\alpha}) \) such that, if \( \mu < \mu_0 \), we have that there exists \( C > 0 \) such that

\[
\sup_j |Q_j(t) - Q_a(t, j)| + |P_j(t) - P_a(t, j)| \leq C \mu^\gamma, \quad |t| \leq \frac{T_0}{\mu^2},
\]

where \((Q_a, P_a)\) are given by (177)-(178). Moreover, \( C > 0 \) and we want to show that there exist two sequences \( j \), \( k \) such that

\[
\| E_\alpha - \mu \frac{\| \xi K \|^2 + \| \eta K \|^2}{2} \| \leq C \mu^{4+\gamma}
\]

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

\[
\| E_\alpha \| \leq \mu^{4+\gamma}
\]

for all \( k \) such that \( \kappa(k) = (\mu K_1, \mu^2 K_2) \) and \( |K_1| + |K_2| > \frac{(2+\delta)\log \mu}{C} \), and \( E_\alpha = 0 \) otherwise.

Proof. The argument follows along the lines of Appendix C in [BP06].

Exploiting the canonical transformation found in Theorem 3.3, we also define

\[
\zeta_\alpha := (\xi_\alpha, \eta_\alpha) = T_{x^2} (\tilde{\xi}_\alpha, \tilde{\eta}_\alpha) = \tilde{\zeta} + \psi_\alpha(\tilde{\zeta}_\alpha),
\]

where \( \psi_\alpha(\tilde{\zeta}_\alpha) := (\psi_\tau(\tilde{\zeta}_\alpha), \psi_\eta(\tilde{\zeta}_\alpha)) \); by (59) we have

\[
\sup_{\zeta \in \mathcal{H}_{\mu, \alpha}} \| \psi_\alpha(\zeta) \|_{\mathcal{H}_{\mu, \alpha}} \leq C_\alpha \mu^2 R.
\]

For convenience we define

\[
q_\alpha(\tau, y) := \frac{1}{\sqrt{2}} \left[ \xi_\alpha(\mu^2 \tau, y_1 - \tau, y_2) + \eta_\alpha(\mu^2 \tau, y_1 + \tau, y_2) \right],
\]

\[
\partial_{\xi_\alpha} p_\alpha(\tau, y) := \frac{1}{\sqrt{2}} \left[ \xi_\alpha(\mu^2 \tau, y_1 - \tau, y_2) - \eta_\alpha(\mu^2 \tau, y_1 + \tau, y_2) \right],
\]

We observe that the pair \( (q_\alpha, p_\alpha) \) satisfies

\[
\mu^2 (q_\alpha)_\tau = -\Delta_1 \mu p_\alpha + \mu^6 R_q
\]

\[
\mu (p_\alpha)_\tau = -\mu^2 q_\alpha - \mu^4 \alpha \pi_0 q_\alpha^2 + \mu^5 R_p,
\]

where the operator \( \Delta_1 \) acts on the variable \( x \), \( \pi_0 \) is the projector on the space of the functions with zero average, and the remainders are functions of the rescaled variables \( \tau \) and \( y \) which satisfy

\[
\sup_{\mathcal{H}_{\mu, \alpha}} \| R_q \|_{L^2_{\mu, \alpha}} \leq C,
\]

\[
\sup_{\mathcal{H}_{\mu, \alpha}} \| R_p \|_{L^2_{\mu, \alpha}} \leq C.
\]

We now restrict the space variables to integer values; keeping in mind that \( q_\alpha \) and \( p_\alpha \) are periodic, we assume that \( j \in \mathbb{Z}^2_{N, N^*} \).

For a finite sequence \( Q = (Q_j)_{j \in \mathbb{Z}^2_{N, N^*}} \) we define the norm

\[
\| Q \|^2_{N, N^*} := \sum_{j \in \mathbb{Z}^2_{N, N^*}} |Q_j|^2.
\]

Now we consider the discrete model (8): we rewrite in the following form,

\[
Q_j = -(\Delta_1 P)_j
\]

\[
\hat{P}_j = -Q_j - \alpha \pi_0 Q_j^2
\]

and we want to show that there exist two sequences \( E = (E_j)_{j \in \mathbb{Z}^2_{N, N^*}} \) and \( F = (F_j)_{j \in \mathbb{Z}^2_{N, N^*}} \) such that

\[
Q = \mu^2 q_\alpha + \mu^{2+\gamma} E, \quad P = \mu p_\alpha + \mu^{2+\gamma} F
\]
fulfills (191)-(192), where $\gamma > 0$ is a parameter we will fix later in the proof. Therefore, we have that

\begin{align}
\dot{E} &= -\Delta_1 F - \mu^6 e^{-2-\gamma} R_q, \\
\dot{F} &= -E - \alpha \pi_0 (\mu^2 \alpha q_{0} E + \mu^2 \gamma E^2) - \mu^2 e^{-2-\gamma} R_q,
\end{align}

where we impose initial conditions on $(E, F)$ such that $Q_0$ has initial conditions corresponding to the ones of the true initial datum,

\[
\mu^2 q_0(0, \mu j_1, \mu j_2) + \mu^{2+\gamma} E_{0,j} = Q_{0,j},
\]

\[
\mu p_0(0, \mu j_1, \mu j_2) + \mu^{2+\gamma} F_{0,j} = F_{0,j}.
\]

We now define the operator $\partial_i$, $i = 1, 2$, by $(\partial_i f)_j := f_j - f_{j \pm e_i}$ for each $f \in L^2_{E,N}$. 

- **Claim 1:** Let $\sigma > 2$ and $\gamma > 0$, we have

\[
\|E_0\|_{L^2_{E,N}} \leq C\mu^{(3-2\gamma-\sigma)/2},
\]

\[
\|\partial_1 F_0\|_{L^2_{E,N}} \leq C\mu^{(3-2\gamma-\sigma)/2},
\]

\[
\|\partial_2 F_0\|_{L^2_{E,N}} \leq C\mu^{(1-2\gamma+\sigma)/2}.
\]

To prove Claim 1 we observe that

\[
E_0 = \mu \frac{j \bar{\xi}_a + \bar{\eta}_a - (\bar{\xi}_a + \bar{\eta}_a)}{\sqrt{2\mu^{2+\gamma}}} = \mu^{-\gamma} \psi \xi + \psi \eta,
\]

\[
F_0 = \mu \frac{-\alpha q_{\bar{E},\bar{N}}}{\sqrt{2\mu^{2+\gamma}}} = \mu^{-1-\gamma} \alpha q_{\bar{E},\bar{N}} (\psi \xi + \psi \eta),
\]

from which we can deduce

\[
\|E_0\|_{L^2_{E,N}}^2 \leq \sum_{j \in \mathbb{Z}^2_{E,N}} |E_{0,j}|^2 \leq C 4N^{\sigma+1} (\mu^{2-\gamma})^2 = C \mu^{3-2\gamma-\sigma},
\]

\[
\|\partial_1 F_0\|_{L^2_{E,N}}^2 \leq \sum_{j \in \mathbb{Z}^2_{E,N}} |\partial_1 F_{0,j}|^2 \leq C 4N^{\sigma+1} (\mu^{2-\gamma})^2 \leq C \mu^{3-2\gamma-\sigma},
\]

\[
\|\partial_2 F_0\|_{L^2_{E,N}}^2 \leq \sum_{j \in \mathbb{Z}^2_{E,N}} |\partial_2 F_{0,j}|^2 \leq C 4N^{\sigma+1} (\mu^{1+\gamma-\sigma})^2 = C \mu^{1-2\gamma+\sigma}
\]

and this leads to the thesis.

- **Claim 2:** Fix $n \geq 1$, $T_0 > 0$ and $K_0 > 0$, then for any $\mu < \mu_0$ and for any $\sigma > 2$ and $\gamma \geq 1$ such that $\sigma + 2\gamma < 7$ we have

\[
\|E\|_{L^2_{E,N}}^2 + \|\partial_1 F\|_{L^2_{E,N}}^2 + \|\partial_2 F\|_{L^2_{E,N}}^2 \leq K_0, \quad |t| < \frac{T_0}{\mu^3}.
\]

To prove the claim, we define

\[
F(E, F) := \sum_{j \in \mathbb{Z}^2_{E,N}} \frac{E_j^2 + F_j(-\Delta_1 F)_j + 2\mu^2 \alpha q_{a,j} E_j^2}{2},
\]

and we remark that, using the boundedness of $q_{a,j}$,

\[
\frac{1}{2} F(E, F) \leq \|E\|_{L^2_{E,N}}^2 + \|\partial_1 F\|_{L^2_{E,N}}^2 + \|\partial_2 F\|_{L^2_{E,N}}^2 \leq 4F(E, F).
\]
Now we compute the time derivative of $F$. Exploiting (193)-(194)

\begin{align}
\hat{F} &= \sum_j E_j \left[- (\Delta_1 F)_j - \mu^{4-\gamma} (\mathcal{R}_q)_j\right] \\
+ \sum_j \left[- (\Delta_1 F)_j - E_j - \alpha (\mu^2 2q_{a,j} E_j + \mu^{2+\gamma} E_j^2) - \mu^{3-\gamma} (\mathcal{R}_p)_j\right] \\
+ \sum_j 2 \mu^2 \alpha q_{a,j} E_j \left[- (\Delta_1 F)_j - \mu^{4-\gamma} (\mathcal{R}_q)_j\right] \\
+ \sum_j \mu^2 \alpha E_j \mu^2 \frac{\partial q_{a,j}}{\partial \tau} \\
= \sum_j -E_j \mu^{4-\gamma} (\mathcal{R}_q)_j + \sum_j \left[- \alpha \mu^{2+\gamma} E_j^2 - \mu^{3-\gamma} (\mathcal{R}_p)_j\right] \\
- \sum_j 2 \mu^2 \alpha q_{a,j} E_j \mu^{4-\gamma} (\mathcal{R}_q)_j + \sum_j \mu^2 \alpha E_j \mu^2 \frac{\partial q_{a,j}}{\partial \tau}
\end{align}

In order to estimate (201)-(202), we notice that

\[
\sup_j \| (\Delta_1 F)_j \| \leq 2 \sup_j \| (\partial_1 F)_j \| + \| (\partial_2 F)_j \| \leq 4 \sqrt{\mathcal{F}},
\]

\[
\| R_q \|_{L^2_{\mathbf{N},\mathbf{\sigma}}}^2 \leq \sum_j \| (R_q)_j \|^2 \leq 4 N^{\sigma+1} \sup_y \| R_q(y) \|^2 \leq C \mu^{1-\sigma},
\]

and that \( \| (\partial_\tau R_p)_j \| \leq \mu \sup_y \left| \frac{\partial_\tau R_p}{\partial \tau}(y) \right| \), which implies

\[
\| \partial_\tau R_p \|_{L^2_{\mathbf{N},\mathbf{\sigma}}}^2 \leq C \mu^{1-\sigma}.
\]

Now, the first sum in (201) is estimated by \( C \mathcal{F}^{1/2} \mu^{(7-2\gamma-\sigma)/2} \); the second sum in (201) can be bounded by

\[
C (\mu^{2+\gamma} \mathcal{F}^{3/2} + \mu^{(7-2\gamma-\sigma)/2} \mathcal{F}^{1/2}).
\]

Recalling that \( q_{a,j} \) is bounded, the first sum in (202) can be bounded by \( C \mathcal{F}^{1/2} \mu^{(11-2\gamma-\sigma)/2} \), while the second one is estimated by \( C \mu^3 \mathcal{F} \). Hence, as long as \( \mathcal{F} < 2K_s \), we have

\begin{align}
\left| \hat{F} \right| &\leq C \left[ \mathcal{F}^{1/2} \mu^{(7-2\gamma-\sigma)/2} + \mu^{2+\gamma} \mathcal{F}^{3/2} + \mu^{(7-2\gamma-\sigma)/2} \mathcal{F}^{1/2} + \mathcal{F}^{1/2} \mu^{(11-2\gamma-\sigma)/2} + \mu^3 \mathcal{F} \right] \\
&\leq C (\mu^{2+\gamma} \sqrt{2} K_s^{1/2} + \mu^3) \mathcal{F} + C (2 \mu^{(7-2\gamma-\sigma)/2} + \mu^{(11-2\gamma-\sigma)/2}) \sqrt{2} K_s^{1/2},
\end{align}

\[
\gamma \geq 1 \leq C \mu^3 2 \sqrt{2} K_s^{1/2} \mathcal{F} + C 3 \mu^{(7-2\gamma-\sigma)/2} \sqrt{2} K_s^{1/2},
\]

and by applying Gronwall’s lemma we get

\[
\mathcal{F}(t) \leq \mathcal{F}(0) e^{C 2 \sqrt{2} K_s^{1/2} \mu^3 t} + e^{C 2 \sqrt{2} K_s^{1/2} \mu^3 t} C 2 \sqrt{2} K_s^{1/2} \mathcal{F}^{1/2} + C 3 \mu^{(7-2\gamma-\sigma)/2} \sqrt{2} K_s^{1/2},
\]

from which we can deduce the thesis.

\[ \square \]

**Proof of Theorem 2.1.** First we prove (17).

We consider an initial datum as in (16); when passing to the continuous approximation (75), this initial datum corresponds to an initial data \((\xi_0, \eta_0) \in H^{n-\alpha} \) for some \( \rho_0 > 0 \) and \( \alpha \geq 1 \). By Theorem 5.1 the corresponding sequence of gaps belongs to \( H^{n-\alpha} \), and that the solution \((\xi(t), \eta(t)) \) is analytic in a complex strip of width \( \rho(t) \). Taking the minimum of such quantities one gets the coefficient \( \rho \) appearing in the statement of Theorem 2.1. Applying Proposition 6.3, we can deduce the corresponding result for the discrete model (8) and the specific quantities (13).

Next, we prove (19). In order to do so, we exploit the Birkhoff coordinates \((x, y)\) introduced in Theorem 5.2; indeed, by rewriting the normal form system (100) in Birkhoff coordinates we get that
every solution is almost-periodic in time. Now, let us introduce the quantities
\[ E^{(1)}_K := \left| \hat{\xi}_K \right|^2, \]
\[ E^{(2)}_K := \left| \hat{\eta}_K \right|^2, \]
then \( \tau \mapsto E^{(1)}_K(x(\tau), y(\tau)) \) and \( \tau \mapsto E^{(2)}_K(x(\tau), y(\tau)) \) are almost-periodic. If we set \( E_K := \frac{1}{4} \left( E^{(1)}_K + E^{(2)}_K \right) \), we can exploit (182) of Proposition 6.3 to translate the results in terms of the specific quantities \( E_n \), and we get the thesis.

6.2. The KP regime. Similarly to Lemma 6.1, Proposition 6.2 we can prove the following results

**Lemma 6.4.** Consider the lattice (6) in the regime (KP) and with interpolating function (113). Then for a state corresponding to \( (q, p) \) one has
\[
E_n = \frac{\mu^4}{2} \sum_{L = (L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \not\in 2\mathbb{Z}} \left| \hat{\eta}_{K+L} \right|^2 + \omega^2 \left| \frac{\partial K+L}{\mu} \right|^2, \quad \forall k : \kappa(k) = (\mu K_1, \mu^2 K_2)
\]
(where the \( \omega_k \) are defined as in (11)), and \( E_n = 0 \) otherwise.

**Proposition 6.5.** Fix \( \mu > 0 \) and \( 0 < \delta \ll 1 \). Consider the normal form system (126), and define the Fourier coefficients of \( (\xi, \eta) \) through the following formula
\[
\xi(y) = \frac{1}{2} \sum_{h \in \mathbb{Z}^2} \hat{\xi}_h e^{ih \cdot y},
\]
\[
\eta(y) = \frac{1}{2} \sum_{h \in \mathbb{Z}^2} \hat{\eta}_h e^{ih \cdot y},
\]
Consider \( (\xi, \eta) \in H^{\alpha, 0} \), and denote by \( E_n \) the specific energy of the normal mode with index \( \kappa \) as defined in (12)-(13). Then for any positive \( \mu \) sufficiently small
\[
\left| E_n - \mu^4 \frac{\left| \hat{\xi}_K \right|^2 + \left| \hat{\eta}_K \right|^2}{2} \right| \leq C \mu^{1+\frac{\alpha}{p}} \| (\xi, \eta) \|_{H^{\alpha, 0}}^2
\]
for all \( k \) such that \( \kappa(k) = (\mu K_1, \mu^2 K_2) \) and \( |K_1| + |K_2| \leq \frac{(2+\delta)}{\log \mu} \). Moreover,
\[
|E_n| \leq C \mu^8 \| (\xi, \eta) \|_{H^{\alpha, 0}}^2
\]
for all \( k \) such that \( \kappa(k) = (\mu K_1, \mu^2 K_2) \) and \( |K_1^2 + K_2^2|^{1/2} \geq \frac{(2+\delta)}{\log \mu} \), and \( E_n = 0 \) otherwise.

Now, consider the following systems of uncoupled KP equations
\[
\xi_\tau = \frac{1}{24} \partial^3_{y_1} \xi - \frac{1}{2} \partial_{y_1}^{-1} \partial^2_{y_2} \xi - \frac{\alpha}{2 \sqrt{2}} \partial_{y_1} \xi^2,
\]
\[
\eta_\tau = \frac{1}{2} \partial^3_{y_1} \eta + \frac{1}{24} \partial^2_{y_2} \eta + \frac{\alpha}{2 \sqrt{2}} \partial_{y_1} \eta^2.
\]
and consider a solution \( (\tau, y) \mapsto (\hat{\xi}_n(\tau, y), \hat{\eta}_n(\tau, y)) \) such that it belongs to \( H^{\alpha, n} \), for some \( n \geq 1 \).

We consider the approximate solutions \( (\bar{Q}_n, \bar{P}_n) \) of the FPU model (75)
\[
\bar{Q}_n(\tau, y) := \mu^2 \sqrt{2} \left[ \hat{\xi}_n(\mu^2 \tau, y_1 - \tau, y_2) + \hat{\eta}_n(\mu^2 \tau, y_1 + \tau, y_2) \right]
\]
\[
\partial_{y_1} \bar{P}_n(\tau, y) := \mu \sqrt{2} \left[ \hat{\xi}_n(\mu^2 \tau, y_1 - \tau, y_2) - \hat{\eta}_n(\mu^2 \tau, y_1 + \tau, y_2) \right].
\]
We need to compare the difference between the approximate solution (213)-(214) and the true solution of (8). Let us consider an initial datum \( (\bar{Q}_0, \bar{P}_0) \) with corresponding Fourier coefficients \( \bar{Q}_{0, k}, \bar{P}_{0, k} \) given by (9), where
\[
Q_{0, k} \neq 0 \quad \text{only if } \kappa(k) = (\mu K_1, \mu^2 K_2).
We also assume that there exist $C, \rho > 0$ such that

$$
|\hat{Q}_{0,k}|^2 + \omega_k^2|\hat{P}_{0,k}|^2 \leq C e^{-2\rho|\kappa_{1,k}/\mu,\kappa_{2,k}/\mu^2|}.
$$

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

$$
Q_0(y) = \frac{1}{(2N_1 + 1)(2N_2 + 1)} \sum_{K} Q_{0,k} e^{i\pi(\mu K_1 y_1 + \mu^2 K_2 y_2)},
$$

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

Arguing as for Proposition 6.3, we obtain

**Proposition 6.6.** Consider (8) with $\sigma = 2$, and fix $1 \leq \gamma \leq \frac{3}{2}$. Let us assume that the initial datum for (8) satisfying (215)-(216), and denote by $(Q(t), P(t))$ the corresponding solution. Consider the approximate solution $(\xi_0, \eta_0)$ with the corresponding initial datum. Assume that $(\xi_0, \eta_0) \in H^{\alpha,n}$ for some $\rho > 0$ and for some $n \geq 1$ for all times, and fix $T_0 > 0$ and $0 < \delta \ll 1$.

Then there exists $\mu_0 = \mu_0(T_0, \| (\xi_0(0), \eta_0(0)) \|_{H^{\alpha,n}})$ such that, if $\mu < \mu_0$, we have that there exists $C > 0$ such that

$$
sup_j |Q_j(t) - Q_0(t, j)| + |P_j(t) - P_0(t, j)| \leq C \mu^\gamma, \quad |t| \leq \frac{T_0}{\mu^3},
$$

where $(Q_0, P_0)$ are given by (177)-(178). Moreover,

$$
|E_n - \mu^2 \frac{\hat{\xi}_K^2 + \hat{\eta}_K^2}{2} \leq C \mu^{4+\gamma}
$$

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

$$
|E_n| \leq \mu^{4+\gamma}
$$

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 $E_n = 0$ otherwise.

**Proof of Theorem 2.2.** First we prove (21).

We consider an initial datum as in (20); when passing to the continuous approximation (75), this initial datum corresponds to an initial data $(\xi_0, \eta_0) \in H^{\alpha,n}$ for some $\rho_0 > 0$ and $n \geq 1$. By Theorem 5.3 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 $\rho$ appearing in the statement of Theorem 2.2. Applying Proposition 6.6, we can deduce the corresponding result for the discrete model (8) and the specific quantities (13). \qed

6.3. The one-dimensional NLS regime. Let $\beta > 0$ and let $I$ be as in (79), we define the Fourier coefficients of the function $q : I \to \mathbb{R}$ by

$$
\hat{q}(j) := \frac{1}{2} \int_{I} q(y_1, y_2) e^{-i\pi(j_1 y_1 + j_2 y_2)} dy_1 dy_2,
$$

and similarly for the Fourier coefficients of the function $p$.

**Lemma 6.7.** Consider the lattice (24) in the regime (1D NLS) and with interpolating function (129). Then for a state corresponding to $(q, p)$ one has

$$
E_n = \frac{\mu^2}{2} \sum_{L \in (L_1, L_2) \in \mathbb{Z}^2; \mu |L_1, \mu^\sigma |L_2 \in \mathbb{Z}} |\hat{p}_{k+L}|^2 + \omega_k^2 |\hat{q}_{k+L}|^2, \quad \forall k : \kappa(k) = (\mu K_1, \mu^\sigma K_2)
$$

(where the $\omega_k$ are defined as in (30)), and $E_n = 0$ otherwise.
Proof. First we introduce a \((2N_1 + 1)(2N_2 + 1)\)-periodic interpolating function for \(Q_j\), namely a smooth function \(Q : (t, x) \mapsto Q(t, x)\) such that

\[
Q_j(t) = Q(t, j), \quad \forall t, j,
\]

and similarly for \(P_j\). We denote by

\[
Q_j(t) = Q(t, j) = \frac{1}{(2N_1 + 1)^{1/2}(2N_2 + 1)^{1/2}} \sum_{k \in \mathbb{Z}^2} \hat{Q}(k) e^{i(2N_1 + 1)^{1/2}x(2N_2 + 1)^{1/2}t} dx,
\]

so that by the interpolation property we obtain

\[
Q_j(t) = Q(t, j) = \frac{1}{(2N_1 + 1)^{1/2}(2N_2 + 1)^{1/2}} \sum_{k \in \mathbb{Z}^2} \hat{Q}(k) e^{i(2N_1 + 1)^{1/2}x(2N_2 + 1)^{1/2}t} dx,
\]

hence

\[
\hat{Q}_k = \sum_{h \in \mathbb{Z}^2} \hat{Q}(k + (2N_1 + 1)h_1, k_2 + (2N_2 + 1)h_2).
\]

The relation between \(\hat{Q}(k)\) and \(\hat{q}_k\) can be deduced from (129),

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

\[
Q_k = \frac{1}{2} \mu^{(\sigma + 1)/2} \int \left[ \frac{-1}{\mu^\sigma} \right] \left[ \frac{-1}{\mu^\sigma} \right] \mu q(x_1, x_2) e^{-i\pi (k_1 x_1 + k_2 x_2)} dx_1 dx_2
\]

\[
= \frac{1}{2} \mu^{(\sigma + 1)/2} \int \left[ \frac{-1}{\mu^\sigma} \right] \left[ \frac{-1}{\mu^\sigma} \right] \mu q(x_1, x_2) e^{-i\pi (k_1 x_1 + k_2 x_2)} dx_1 dx_2
\]

\[
= \frac{1}{2} \mu^{(1 - \sigma)/2} q_k,
\]

and similarly

\[
\hat{P}_k = \mu^{(1 - \sigma)/2} \hat{p}_k.
\]

By using (29), (13) and (223)-(225) we have

\[
E_\alpha = \mu^{\alpha + 1} \sum_{L = (L_1, L_2) \in \mathbb{Z}^2, \mu L_1 = \mu^\sigma L_2} |\hat{P}_{K + L}|^2 + \omega^2 |\hat{Q}_{K + L}|^2
\]

\[
= \mu^{\alpha + 1} \mu^{1 - \sigma} \sum_{L = (L_1, L_2) \in \mathbb{Z}^2, \mu L_1 = \mu^\sigma L_2} |\hat{p}_{K + L}|^2 + \omega^2 |\hat{q}_{K + L}|^2
\]

for all \(k\) such that \(\kappa(k) = (\mu K_1, \mu^\sigma K_2)\), and this leads to (221).

\[
\text{Proposition 6.8. Fix} \ \rho > 0 \ \text{and} \ 0 < \delta < 1. \ \text{Consider the normal form equation (144), and define the Fourier coefficients of} \ (\psi, \psi) \ \text{through the following formula}
\]

\[
\psi(y) = \frac{1}{2} \sum_{h \in \mathbb{Z}^2} \hat{\psi}_h e^{ih \cdot y}.
\]
Consider $(\psi, \tilde{\psi}) \in H^{\sigma, 0}$, and denote by $E_\kappa$ the specific energy of the normal mode with index $\kappa$ as defined in (12)-(13). Then for any positive $\mu$ sufficiently small
\begin{equation}
|E_\kappa - \mu^2 \frac{|\tilde{\psi}_K|^2}{2}| \leq C \mu^{2+\frac{\gamma}{2}} \|(\psi, \tilde{\psi})\|_{H^{\rho, 0}}^2
\end{equation}
for all $k$ such that $\kappa(k) = (\mu K_1, \mu^n K_2)$ and $|K_1| + |K_2| \leq \frac{(2+\delta) \log \mu}{\rho}$. Moreover,
\begin{equation}
|E_\kappa| \leq C \mu^\delta \|(\psi, \tilde{\psi})\|_{H^{\rho, 1}}^2
\end{equation}
for all $k$ such that $\kappa(k) = (\mu K_1, \mu^n K_2)$ and $|K_1| + |K_2| > \frac{(2+\delta) \log \mu}{\rho}$, and $E_\kappa = 0$ otherwise.

We defer the proof of the above Proposition to Appendix C.

Now, consider the normal form equation, namely the following cubic defocusing one-dimensional NLS
\begin{equation}
-i\psi_t = -\partial_y^2 \psi + \frac{3\gamma}{4} |\psi|^2 \psi.
\end{equation}
and consider a solution $(\tilde{\psi}_n, \tilde{\psi}_a)$ such that it belongs to $H^{\rho, n}$, for some $n > 0$.

We consider the approximate solutions $(Q_\alpha, P_\alpha)$ of the KG lattice (24) (in the following $\mu = \mu^2 t$)
\begin{align}
Q_\alpha(\tau, y) &:= \frac{\mu}{\sqrt{2}} \left[ e^{i\tau} \tilde{\psi}_\alpha(\tau, y_1, y_2) + e^{-i\tau} \tilde{\psi}_\alpha(\tau, y_1, y_2) \right] \\
P_\alpha(\tau, y) &:= \frac{\mu}{\sqrt{2}} \left[ e^{i\tau} \tilde{\psi}_a(\tau, y_1, y_2) + e^{-i\tau} \tilde{\psi}_a(\tau, y_1, y_2) \right]
\end{align}

We need to compare the difference between the approximate solution (177)-(178) and the true solution of (24). Let consider an initial datum $(Q_0, P_0)$ with corresponding Fourier coefficients $(\tilde{Q}_{0,k}, \tilde{P}_{0,k})$ given by (9), where
\begin{equation}
\tilde{Q}_{0,k} \neq 0 \quad \text{only if} \quad \kappa(k) = (\mu K_1, \mu^n K_2).
\end{equation}
We also assume that there exist $C, \rho > 0$ such that
\begin{equation}
|\tilde{P}_{0,k}|^2 + |\tilde{Q}_{0,k}|^2 \leq C e^{-2\rho(|\kappa(k)/\mu|^{2/n})}.
\end{equation}
Moreover, we define an interpolating function for the initial datum $(Q_0, P_0)$ by
\begin{equation}
Q_0(y) = \frac{1}{(2N_1 + 1)(2N_2 + 1)} \sum_{\kappa, |\kappa| = |\kappa(k)| < 1} \tilde{Q}_{0,k} e^{i\kappa(y_1 y_2)} + \sum_{\kappa, |\kappa| > 1} \tilde{Q}_{0,k} e^{i\kappa(y_1 y_2)},
\end{equation}
and similarly for $y \mapsto P_0(y)$.

**Proposition 6.9.** Consider (24) with $\sigma > 1$ and $\gamma > 0$ such that $\sigma + 2\gamma < 7$. Let us assume that the initial datum satisfies (233)-(234), and denote by $(Q(t), P(t))$ the corresponding solution. Consider the approximate solution $(\tilde{\psi}_n(t, x), \tilde{\psi}_a(t, x))$ with the corresponding initial datum. Assume that $(\tilde{\psi}_n, \tilde{\psi}_a) \in H^{\rho, n}$ for some $\rho > 0$ and for some $n \geq 0$ for all times, and fix $T_0 > 0$ and $0 < \delta < 1$.

Then there exists $\mu_0 = \mu_0(T_0, \|(\tilde{\psi}_n(0), \tilde{\psi}_a(0))\|_{H^{\rho, n}})$ such that, if $\mu < \mu_0$, we have that there exists $C > 0$ such that
\begin{equation}
\sup_j |Q_j(t) - Q_n(t, j)| + |P_j(t) - P_n(t, j)| \leq C \mu^7, \quad |t| \leq \frac{T_0}{\mu^2},
\end{equation}
where $(Q_\alpha, P_\alpha)$ are given by (230)-(231). Moreover,
\begin{equation}
|E_\kappa - \mu^2 \frac{|\tilde{\psi}_K|^2}{2}| \leq C \mu^{2+\gamma}
\end{equation}
for all $k$ such that $\kappa(k) = (\mu K_1, \mu^n K_2)$ and $|K_1| + |K_2| \leq \frac{(2+\delta) \log \mu}{\rho}$. Moreover,
\begin{equation}
|E_\kappa| \leq C \mu^{2+\gamma}
\end{equation}
for all $k$ such that $\kappa(k) = (\mu K_1, \mu^n K_2)$ and $|K_1| + |K_2| > \frac{(2 + 3\delta)\log p}{p}$, and $E_n = 0$ otherwise.

**Proof.** The argument follows along the lines of Appendix C in [BP06].

Exploiting the canonical transformation found in Theorem 3.3, we also define
\begin{equation}
\zeta_a := (\psi_a, \tilde{\psi}_a) = T_{\omega_a}^* (\psi_a, \tilde{\psi}_a) = \zeta_a + \phi_a (\tilde{\zeta}_a),
\end{equation}
where $\phi_a (\tilde{\zeta}_a) := (\phi_1 (\tilde{\zeta}_a), \phi_2 (\tilde{\zeta}_a))$; by (59) we have
\begin{equation}
\sup_{\zeta \in B_{p_n(R)}} \| \phi_a (\zeta) \|_{\mathcal{H} p_n} \leq C_n \mu^2 R.
\end{equation}

For convenience we define
\begin{align}
q_a (\tau, y) & := \frac{1}{\sqrt{2}} \left[ e^{i\tau} \bar{\psi}_a (\tau, y_1, y_2) + e^{-i\tau} \tilde{\psi}_a (\tau, y_1, y_2) \right] \\
p_a (\tau, y) & := \frac{1}{\sqrt{2}} \left[ e^{i\tau} \bar{\psi}_a (\tau, y_1, y_2) - e^{-i\tau} \tilde{\psi}_a (\tau, y_1, y_2) \right],
\end{align}

We observe that the pair $(q_a, p_a)$ satisfies
\begin{align}
\mu (q_a) & = \mu p_a + \mu^5 R_q \\
\mu (p_a) & = -\mu q_a + \mu \Delta_1 q_a - \mu^3 \beta \pi_0 q_a^3 + \mu^5 R_p,
\end{align}
where the operator $\Delta_1$ acts on the variable $x$, $\pi_0$ is the projector on the space of the functions with zero average, and the remainders are functions of the rescaled variables $\tau$ and $y$ which satisfy
\begin{align}
\sup_{B_{p_n(R)}} \| R_q \|_{p_0}^2 & \leq C, \\
\sup_{B_{p_n(R)}} \| R_p \|_{p_1}^2 & \leq C.
\end{align}

We now restrict the space variables to integer values; keeping in mind that $q_a$ and $p_a$ are periodic, we assume that $j \in \mathbb{Z}_{N, N^\gamma}^2$.

For a finite sequence $Q = (Q_j)_{j \in \mathbb{Z}_{N, N^\gamma}^2}$ we define the norm
\begin{equation}
\| Q \|^2_{\mathcal{H}_{N, N^\gamma}} := \sum_{j \in \mathbb{Z}_{N, N^\gamma}^2} |Q_j|^2.
\end{equation}

Now we consider the discrete model (8): we rewrite in the following form,
\begin{align}
\hat{Q}_j & = P_j \\
\hat{P}_j & = -Q_j + (\Delta_1 Q)_j - \beta \pi_0 Q_j^3
\end{align}
and we want to show that there exist two sequences $E = (E_j)_{j \in \mathbb{Z}_{N, N^\gamma}^2}$ and $F = (F_j)_{j \in \mathbb{Z}_{N, N^\gamma}^2}$ such that
\begin{equation}
Q = \mu q_a + \mu^{1+\gamma} E, \quad P = \mu p_a + \mu^{1+\gamma} F
\end{equation}
fulfills (245)-(246), where $\gamma > 0$ is a parameter we will fix later in the proof. Therefore, we have that
\begin{align}
\hat{E} & = F - \mu^{5-1-\gamma} R_q \\
\hat{F} & = -E + \Delta_1 E - \beta \pi_0 (3 \mu^3 \gamma - 1 - \gamma) q_a^3 E + 3 \mu^{1+2+2\gamma-1-\gamma} q_a E^2 + \mu^{3+3\gamma-1-\gamma} E^3 - \mu^{5-1-\gamma} R_p,
\end{align}
where we impose initial conditions on $(E, F)$ such that $(\bar{q}, \bar{p})$ has initial conditions corresponding to the ones of the true initial datum,
\begin{align}
\mu q_a (0, \mu j_1, \mu^n j_2) + \mu^{1+\gamma} E_{0,j} & = Q_{0,j}, \\
\mu p_a (0, \mu j_1, \mu^n j_2) + \mu^{1+\gamma} F_{0,j} & = P_{0,j}.
\end{align}

We now define the operator $\partial_i$, $i = 1, 2$, by $(\partial_i f)_j := f_j - f_{j-i}$ for each $f \in \mathcal{H}_{N, N^\gamma}$. 

• Claim 1: Let $\sigma > 1$ and $\gamma > 0$, we have

$$
\|E_0\|_{\ell_2^{2,N,\sigma}}^2 \leq C \mu^{(3-2\gamma-\sigma)/2},
$$
$$
\|F_0\|_{\ell_2^{2,N,\sigma}}^2 \leq C \mu^{(3-2\gamma-\sigma)/2},
$$
$$
\|\partial_t E_0\|_{\ell_2^{2,N,\sigma}}^2 \leq C \mu^{(5-2\gamma-\sigma)/2},
$$
$$
\|\partial_2 E_0\|_{\ell_2^{2,N,\sigma}}^2 \leq C \mu^{(3-2\gamma+\sigma)/2},
$$
$$
\|\partial_t F_0\|_{\ell_2^{2,N,\sigma}}^2 \leq C \mu^{(5-2\gamma-\sigma)/2},
$$
$$
\|\partial_2 F_0\|_{\ell_2^{2,N,\sigma}}^2 \leq C \mu^{(3-2\gamma+\sigma)/2}.
$$

To prove Claim 1 we observe that

$$
E_0 = \mu \frac{\psi_a + \tilde{\psi_a} - (\tilde{\psi_a} + \tilde{\psi_a})}{\sqrt{2\mu^{1+\gamma}}} = \mu^{-\gamma} \phi_x + \phi_y,
$$
$$
F_0 = \mu \frac{\psi_a - \psi_a - (\tilde{\psi_a} - \tilde{\psi_a})}{\sqrt{2\mu^{1+\gamma}}} = \mu^{-\gamma} \phi_x - \phi_y,
$$

from which we can deduce

$$
\|E_0\|_{\ell_2^{2,N,\sigma}}^2 \leq \sum_{j \in \mathbb{Z}_N^{2,\sigma}} |E_{0,j}|^2 \leq C 4 N^{\sigma+1} (\mu^{2-\gamma})^2 = C \mu^{3-2\gamma-\sigma},
$$
$$
\|F_0\|_{\ell_2^{2,N,\sigma}}^2 \leq \sum_{j \in \mathbb{Z}_N^{2,\sigma}} |F_{0,j}|^2 \leq C 4 N^{\sigma+1} (\mu^{2-\gamma})^2 = C \mu^{3-2\gamma-\sigma},
$$
$$
\|\partial_t E_0\|_{\ell_2^{2,N,\sigma}}^2 \leq \sum_{j \in \mathbb{Z}_N^{2,\sigma}} |\partial_t E_{0,j}|^2 \leq C 4 N^{\sigma+1} (\mu^{2+\gamma-\sigma})^2 \leq C \mu^{5-2\gamma-\sigma},
$$
$$
\|\partial_2 E_0\|_{\ell_2^{2,N,\sigma}}^2 \leq \sum_{j \in \mathbb{Z}_N^{2,\sigma}} |\partial_2 E_{0,j}|^2 \leq C 4 N^{\sigma+1} (\mu^{2+\gamma-\sigma})^2 = C \mu^{3-2\gamma+\sigma},
$$
$$
\|\partial_t F_0\|_{\ell_2^{2,N,\sigma}}^2 \leq \sum_{j \in \mathbb{Z}_N^{2,\sigma}} |\partial_t F_{0,j}|^2 \leq C 4 N^{\sigma+1} (\mu^{2+\gamma-\sigma})^2 \leq C \mu^{5-2\gamma-\sigma},
$$
$$
\|\partial_2 F_0\|_{\ell_2^{2,N,\sigma}}^2 \leq \sum_{j \in \mathbb{Z}_N^{2,\sigma}} |\partial_2 F_{0,j}|^2 \leq C 4 N^{\sigma+1} (\mu^{2+\gamma-\sigma})^2 = C \mu^{3-2\gamma+\sigma},
$$

and this leads to the thesis.

• Claim 2: Fix $n \geq 0$, $T_0 > 0$ and $K_\ast > 0$, then for any $\mu < \mu_\ast$ and for any $\sigma > 1$ and $\gamma > 0$ such that $\sigma + 2\gamma < 7$ we have

$$
\|E\|_{\ell_2^{2,N,\sigma}}^2 + \|F\|_{\ell_2^{2,N,\sigma}}^2 + \|\partial_t E_0\|_{\ell_2^{2,N,\sigma}}^2 + \|\partial_2 E_0\|_{\ell_2^{2,N,\sigma}}^2 \leq K_\ast, \ |t| < \frac{T_0}{\mu^2}.
$$

To prove the claim, we define

$$
F(E, F) := \sum_{j \in \mathbb{Z}_N^{2,\sigma}} \frac{E_j^2 + E_j^2 + E_j(-\Delta)E_j}{2} + \frac{3\mu^2 \beta q_0 E_j^2 + 3\mu^{2+\gamma} \beta q_0 E_j^2}{2},
$$

and we remark that

$$
\frac{1}{2} F(E, F) \leq \|E\|_{\ell_2^{2,N,\sigma}}^2 + \|\partial_t F_0\|_{\ell_2^{2,N,\sigma}}^2 + \|\partial_2 F_0\|_{\ell_2^{2,N,\sigma}}^2 \leq 2 F(E, F).
$$
Now we compute the time derivative of $\mathcal{F}$. Exploiting (193)-(194)

\begin{align}
\dot{\mathcal{F}} &= \sum_j F_j \left[ -E_j + (\Delta_1 E_j) - \beta \pi_0 (3\mu^2 q_s^2 E_j + 3\mu^{2+\gamma} q_s E_j^2 + \mu^{2+2\gamma} E_j^3) - \mu^{4-\gamma} (R_p)_j \right] \\
&+ \sum_j (E_j - (\Delta_1 E)_j) \left[ F_j - \mu^{4-\gamma} (R_q)_j \right] \\
&+ \sum_j 3\mu^2 \beta q_s^2 E_j \left[ F_j - \mu^{4-\gamma} (R_q)_j \right] + 3\mu^2 \beta E_j^2 q_s \mu \frac{\partial q_s}{\partial \tau} \\
&+ \sum_j \frac{9}{7} \mu^{2+\gamma} \beta E_j^2 \left[ F_j - \mu^{4-\gamma} (R_q)_j \right] + \frac{3}{7} \mu^{2+\gamma} \beta E_j^3 \mu \frac{\partial q_s}{\partial \tau} \\
&= \sum_j F_j \left[ -\beta \pi_0 (3\mu^{2+\gamma} q_s E_j^2 + \mu^{2+2\gamma} E_j^3) - \mu^{4-\gamma} (R_p)_j \right] \\
&+ \sum_j E_j \left[ -\mu^{4-\gamma} (R_q)_j \right] - (\Delta_1 E)_j \left[ -\mu^{4-\gamma} (R_q)_j \right] \\
&+ \sum_j 3\mu^2 \beta q_s^2 E_j \left[ -\mu^{4-\gamma} (R_q)_j \right] + 3\mu^2 \beta E_j^2 q_s \mu \frac{\partial q_s}{\partial \tau} \\
&+ \sum_j \frac{9}{7} \mu^{2+\gamma} \beta E_j^2 \left[ F_j - \mu^{4-\gamma} (R_q)_j \right] + \frac{3}{7} \mu^{2+\gamma} \beta E_j^3 \mu \frac{\partial q_s}{\partial \tau}
\end{align}

In order to estimate (255)-(258), we notice that

$$
\sup_j \| (\Delta_1 E)_j \| \leq 2 \sup_j \| (\partial_1 E)_j \| + \| (\partial_2 E)_j \| \leq 4 \sqrt{\mathcal{F}},
$$

$$
\| R_q \|^2 \leq \sum_j \| (R_q)_j \|^2 \leq 4N^{\sigma+1} \sup_y |R_q(y)|^2 \leq C \mu^{-1-\sigma},
$$

$$
\| R_p \|^2 \leq \sum_j \| (R_q)_j \|^2 \leq C \mu^{-1-\sigma},
$$

and that $\| (\partial_\tau R_q)_j \| \leq \mu \sup_y \frac{|R_q|}{\| R_q \|} (y)$, which implies

$$
\| (\partial_\tau R_q)_j \|^2 \leq C \mu^{-1-\sigma}.
$$

Now, we can estimate (255) by

\begin{align}
C \left( \mu^{2+\gamma} \mathcal{F}^{3/2} + \mu^{2+2\gamma} \mathcal{F}^{2} + \mu^{4-\gamma} \mu^{-1+\sigma} // \mathcal{F}^{1/2} \right)
\end{align}

Then, (256) can be bounded by

\begin{align}
C \left( \mu^{4-\gamma} \mu^{-1+\sigma} // \mathcal{F}^{1/2} + \mu^{4-\gamma} \mu^{-1+\sigma} // \mathcal{F}^{1/2} \right)
\end{align}

next, we can estimate (257) by

\begin{align}
C \left( \mu^{6-\gamma} \mu^{-1+\sigma} // \mathcal{F}^{1/2} + \mu^3 \mathcal{F} \right)
\end{align}

while (258) can be bounded by

\begin{align}
C \left( \mu^{2+\gamma} \mathcal{F}^{3/2} + \mu^{6-1+\sigma} // \mathcal{F} + \mu^{2+\gamma} \mathcal{F}^{3/2} \right)
\end{align}

Hence, as long as $\mathcal{F} < 2K_\ast$, we have

\begin{align}
\left| \dot{\mathcal{F}} \right| \leq C \left[ \mu^{2+\gamma} K_\ast^{1/2} + \mu^{2+2\gamma} K_\ast + \mu^3 + \mu^{2+\gamma} K_\ast^{1/2} + \mu^{6-1+\sigma} / \mathcal{F} + \mu^{2+\gamma} K_\ast^{1/2} \right] \mathcal{F}
\end{align}

\begin{align}
+ C \left[ \mu^{4-\gamma} \mu^{-1+\sigma} // \mathcal{F} + \mu^{4-\gamma} \mu^{-1+\sigma} // \mathcal{F} + \mu^{6-1+\sigma} // \mathcal{F} \right] K_\ast^{1/2}
\end{align}

\begin{align}
\leq C \mu^2 (1 + K_\ast^{1/2}) \mathcal{F} + C \mu^{1+\gamma} K_\ast^{1/2}
\end{align}
and by applying Gronwall’s lemma we get
\begin{equation}
\mathcal{F}(t) \leq \mathcal{F}(0)e^{C(1+K^1_{\rho,s})\mu^2 t} + e^{C(1+K^1_{\rho,s})\mu^2 t} C (1 + K^1_{\rho,s}) \mu^2 t C \mu^{(7-2\gamma-\sigma)/2} K^1_{\rho,s},
\end{equation}
from which we can deduce the thesis. \hfill \Box

**Proof of Theorem 2.4.** First we prove (32).

We consider an initial datum as in (31); when passing to the continuous approximation (128), this initial datum corresponds to an initial datum \((\psi_0, \tilde{\psi}_0) \in \mathcal{H}^{\rho,s}. \) By Theorem 5.5 the corresponding sequence of gaps belongs to \(\mathcal{H}^{\rho,s}. \) and that the solution \((\psi(\tau), \tilde{\psi}(\tau))\) is analytic in a complex strip of width \(\mu(\tau).\) Taking the minimum of such quantities one gets the coefficient \(\rho\) appearing in the statement of Theorem 2.4. Applying Proposition 6.9, we can deduce the corresponding result for the discrete model (27) and the specific quantities (13).

Next, we prove (34). In order to do so, we exploit the Birkhoff coordinates \((x, y)\) introduced in Theorem 5.6; indeed, by rewriting the normal form system (144) in Birkhoff coordinates we get that every solution is almost-periodic in time. Now, let us introduce the quantity
\[ E_K := \frac{1}{2} \left\| \psi_K \right\|^2, \]
then \(\tau \mapsto E_K(x(\tau), y(\tau))\) is almost-periodic. Hence we can exploit (236) of Proposition 6.9 to translate the results in terms of the specific quantities \(E_n, \) and we get the thesis. \hfill \Box

**Appendix A. Proof of Lemma 3.7**

This appendix is devoted to the proof of the Lemma 3.7, which is a key step in order to normalize the system (61). This result is an adaptation of Theorem 4.4 in [Bam99] and its proof is based on the method of Lie transforms, briefly recalled in the following. Throughout this Section, we consider \(s \geq s_1\) and \(\rho \geq 0\) to be fixed quantities.

Given an auxiliary function \(\chi\) analytic on \(\mathcal{H}^{\rho,s}, \) we consider the auxiliary differential equation
\begin{equation}
\dot{\zeta} = X_\chi(\zeta)
\end{equation}
and denote by \(\Phi^t_\chi\) its flow at time \(t.\)

**Lemma A.1.** Let \(\chi\) and its vector field be analytic in \(B_{\rho,s}(R).\) Fix \(\delta < R, \) and assume that
\[ \sup_{B_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq \delta. \]
Then, if we consider the time-\(t\) flow \(\Phi^t_\chi\) of \(X_\chi\) we have that for \(|t| \leq 1\)
\[ \sup_{B_{\rho,s}(R-\delta)} \|\Phi^t_\chi(\zeta) - \zeta\|_{\mathcal{H}^{\rho,s}} \leq \sup_{B_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}. \]

**Definition A.2.** The map \(\Phi^1_\chi := \Phi^1_\chi\) is called the Lie transform generated by \(\chi.\)

Given \(G\) analytic on \(\mathcal{H}^{\rho,s}, \) let us consider the differential equation
\begin{equation}
\dot{\tilde{\zeta}} = X_G(\zeta),
\end{equation}
where by \(X_G\) we denote the vector field of \(G.\) Now define
\[ \Phi^t_\chi G(\zeta) := G \circ \Phi^t_\chi(\zeta). \]
By exploiting the fact that \(\Phi^t_\chi\) is a canonical transformation, we have that in the new variable \(\tilde{\zeta}\) defined by \(\zeta = \Phi^t_\chi(\zeta)\) equation (268) is equivalent to
\begin{equation}
\dot{\tilde{\zeta}} = X_{\Phi^t_\chi G}(\tilde{\zeta}).
\end{equation}
Using the relation
\begin{equation}
\frac{d}{dt} \Phi^t_\chi G = \Phi^t_\chi [\chi, G],
\end{equation}
and the Poisson bracket formalism \( \{ G_1, G_2 \}(\zeta) := dG_1(\zeta)[X_{G_2}(\zeta)] \) we formally get

\[
\Phi^\ell G = \sum_{\ell=0}^\infty G_\ell,
\]

(271)

\( G_0 := G \),

\( G_\ell := \frac{1}{\ell!} \{ \chi, G_{\ell-1} \}, \ \ell \geq 1. \)

In order to estimate the vector field of the terms appearing in (271), we exploit the following results

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

\[
\sup_{B_{\rho,s}(R-d)} \| X_{\{ \chi, G \}}(\zeta) \|_{H^{\rho,s}} \leq \frac{2}{d} \left( \sup_{B_{\rho,s}(R)} \| X_\chi(\zeta) \|_{H^{\rho,s}} \right) \left( \sup_{B_{\rho,s}(R)} \| X_G(\zeta) \|_{H^{\rho,s}} \right).
\]

(272)

Proof. Observe that

\[
\| X_{\{ \chi, G \}}(\zeta) \|_{H^{\rho,s}} = \| dX_\chi(\zeta) X_G(\zeta) - dX_G(\zeta) X_\chi(\zeta) \|_{H^{\rho,s}} \\
\leq \| dX_\chi(\zeta) X_G(\zeta) \|_{H^{\rho,s}} + \| dX_G(\zeta) X_\chi(\zeta) \|_{H^{\rho,s}},
\]

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

\[
\sup_{B_{\rho,s}(R-d)} \| dX_\chi(\zeta) X_G(\zeta) \|_{H^{\rho,s}} \leq \frac{1}{d} \sup_{B_{\rho,s}(R)} \| X_\chi(\zeta) \|_{H^{\rho,s}},
\]

we finally get

\[
\sup_{B_{\rho,s}(R-d)} \| dX_\chi(\zeta) X_G(\zeta) \|_{H^{\rho,s}} \leq \frac{1}{d} \left( \sup_{B_{\rho,s}(R)} \| X_\chi(\zeta) \|_{H^{\rho,s}} \right) \left( \sup_{B_{\rho,s}(R)} \| X_G(\zeta) \|_{H^{\rho,s}} \right).
\]

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

**Lemma A.4.** Let \( R > 0 \), and assume that \( \chi, G \) are analytic on \( B_{\rho,s}(R) \) as well as their vector fields. Let \( \ell \geq 1 \), and consider \( G_\ell \) as defined in (271); for any \( d \in (0, R) \), \( G_\ell \) is analytic on \( B_{\rho,s}(R - d) \) as well as it vector field, and

\[
\sup_{B_{\rho,s}(R-d)} \| X_{G_\ell}(\zeta) \|_{H^{\rho,s}} \leq \frac{2\ell}{d} \left( \sup_{B_{\rho,s}(R)} \| X_\chi(\zeta) \|_{H^{\rho,s}} \right)^\ell \left( \sup_{B_{\rho,s}(R)} \| X_G(\zeta) \|_{H^{\rho,s}} \right).
\]

(273)

Proof. Fix \( \ell \), and denote \( \delta := d/\ell \). We look for a sequence \( C_\ell^{(m)} \) such that

\[
\sup_{B_{\rho,s}(R-\delta m)} \| X_{G_m}(\zeta) \|_{H^{\rho,s}} \leq C_\ell^{(m)} \quad \forall m \leq \ell.
\]

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

\[
C_0^{(\ell)} := \sup_{B_{\rho,s}(R)} \| X_G(\zeta) \|_{H^{\rho,s}},
\]

\[
C_m^{(\ell)} = \left( \frac{2}{\delta m} \right)^\ell \sup_{B_{\rho,s}(R)} \| X_\chi(\zeta) \|_{H^{\rho,s}}
\]

\[
= \left( \frac{2}{\delta m} \right)^\ell \sup_{B_{\rho,s}(R)} \| X_\chi(\zeta) \|_{H^{\rho,s}}.
\]

One has

\[
C_\ell^{(\ell)} = \frac{1}{\ell!} \left( \frac{2}{d} \right)^\ell \left( \sup_{B_{\rho,s}(R)} \| X_\chi(\zeta) \|_{H^{\rho,s}} \right)^\ell \left( \sup_{B_{\rho,s}(R)} \| X_G(\zeta) \|_{H^{\rho,s}} \right),
\]

and by using the inequality \( \ell^\ell < \ell e^\ell \) one obtains the estimate (273). \( \square \)
Lemma A.5. Let $\chi$ and $F$ be analytic on $B_{\rho,s}(R)$ as well as their vector fields. Fix $d \in (0, R)$, and assume also that

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

Then for $|t| \leq 1$

$$\sup_{B_{\rho,s}(R-d)} \|X_\chi(t\Phi^*_\chi - F)(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq \frac{3}{d} \sup_{B_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{B_{\rho,s}(R)} \|X_F(\zeta)\|_{\mathcal{H}^{\rho,s}}. \tag{274}$$

Proof. Since the bound on the norm of $X_\chi$ implies that $\Phi^*_\chi(\zeta) \in B_{\rho,s}(R)$ when $\zeta \in B_{\rho,s}(R - d/3)$, using Cauchy inequality and Lemma A.1

$$\sup_{B_{\rho,s}(R-d)} \|d\Phi^*_\chi(t\Phi^*_\chi(\zeta)) - id\|_{\mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}} \leq \sup_{B_{\rho,s}(R-2d/3)} \|d\Phi^*_\chi(t\Phi^*_\chi(\zeta)) - id\|_{\mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}} \leq \frac{3}{d} \sup_{B_{\rho,s}(R-d/3)} \|\Phi^*_\chi(t\Phi^*_\chi(\zeta)) - \zeta\|_{\mathcal{H}^{\rho,s}} \leq \frac{3}{d} \sup_{B_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}}.$$

Since $\Phi^*_\chi$ is a canonical transformation, a direct computation shows

$$\Omega^{-1} d(F \circ \Phi^*_\chi)(\zeta) = (d\Phi^*_\chi(F^*_\chi(\zeta)) - id)\Omega^{-1} dF^*_\chi + \Omega^{-1} dF^*_\chi(\zeta))$$

whence

$$\sup_{B_{\rho,s}(R-d)} \|X_\chi(t\Phi^*_\chi - F)(\zeta)\|_{\mathcal{H}^{\rho,s}} = \sup_{B_{\rho,s}(R-d)} \|\Omega^{-1} d(F^*_\chi(\zeta)) - F(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq \sup_{B_{\rho,s}(R-d)} \|(d\Phi^*_\chi(F^*_\chi(\zeta)) - id)\Omega^{-1} dF^*_\chi + \Omega^{-1} dF^*_\chi(\zeta)) - F(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq \sup_{B_{\rho,s}(R-d)} \|d\Phi^*_\chi(t\Phi^*_\chi(\zeta)) - id\|_{\mathcal{H}^{\rho,s} \to \mathcal{H}^{\rho,s}} \sup_{B_{\rho,s}(R-d)} \|X_F(\Phi^*_\chi(\zeta))\|_{\mathcal{H}^{\rho,s}} + \sup_{B_{\rho,s}(R-d)} \|X_F(\Phi^*_\chi(\zeta)) - X_F(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq \frac{3}{d} \sup_{B_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{B_{\rho,s}(R)} \|X_F(\zeta)\|_{\mathcal{H}^{\rho,s}} + \sup_{B_{\rho,s}(R-d)} \int_0^t \|X_\chi, X_F(\Phi^*_\chi(\zeta))ds\|_{\mathcal{H}^{\rho,s}}.$$

To estimate the last term we use Cauchy inequality

$$\sup_{B_{\rho,s}(R-d)} \int_0^t \|X_\chi, X_F(\Phi^*_\chi(\zeta))ds\|_{\mathcal{H}^{\rho,s}} \leq 2 \sup_{B_{\rho,s}(R-2d/3)} \|X_\chi, X_F(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq \frac{6}{2d} \sup_{B_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{B_{\rho,s}(R)} \|X_F(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq \frac{6}{d} \sup_{B_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{B_{\rho,s}(R)} \|X_F(\zeta)\|_{\mathcal{H}^{\rho,s}}.$$

Then the thesis follows. □
Lemma A.6. Assume that $G$ is analytic on $B_{p,s}(R)$ as well as its vector field, and that $h_0$ satisfies (PER). Then there exists $\chi$ analytic on $B_{p,s}(R)$ and $Z$ analytic on $B_{p,s}(R)$ with $Z$ in normal form, namely $\{h_0, Z\} = 0$, such that

\begin{equation}
\{\chi, h_0\} + G = Z.
\end{equation}

Such $Z$ and $\chi$ are given explicitly by

\begin{equation}
Z(\zeta) = \frac{1}{T} \int_0^T G(\Phi^t_{h_0}(\zeta)) \, dt,
\end{equation}

\begin{equation}
\chi(\zeta) = \frac{1}{T} \int_0^T \frac{t}{\zeta} \left[ Z(\Phi^t_{h_0}(\zeta)) - G(\Phi^t_{h_0}(\zeta)) \right] \, dt.
\end{equation}

Furthermore, we have that the vector fields of $\chi$ and $Z$ are analytic on $B_{p,s}(R)$, and satisfy

\begin{equation}
\sup_{B_{p,s}(R)} \|X_\zeta(\zeta)\|_{H^{p,s}} \leq \sup_{B_{p,s}(R)} \|X_G(\zeta)\|_{H^{p,s}},
\end{equation}

\begin{equation}
\sup_{B_{p,s}(R)} \|X_\chi(\zeta)\|_{H^{p,s}} \leq 2T \sup_{B_{p,s}(R)} \|X_G(\zeta)\|_{H^{p,s}}.
\end{equation}

Proof. We check directly that the solution of (275) is (277). Indeed,

\begin{equation*}
\{\chi, h_0\}(\zeta) = \frac{d}{d\xi|_{\xi=0}} \chi(\Phi^0_{h_0}(\zeta))
= \frac{1}{T} \int_0^T \frac{d}{d\xi|_{\xi=0}} \left[ Z(\Phi^t_{h_0}(\zeta)) - G(\Phi^t_{h_0}(\zeta)) \right] \, dt
= \frac{1}{T} \int_0^T \frac{d}{d\xi|_{\xi=0}} \left[ Z(\Phi^t_{h_0}(\zeta)) - G(\Phi^t_{h_0}(\zeta)) \right] \, dt
= \frac{1}{T} \left[ tZ(\Phi^t_{h_0}(\zeta)) - tG(\Phi^t_{h_0}(\zeta)) \right]_{t=0} - \frac{1}{T} \int_0^T \left[ Z(\Phi^t_{h_0}(\zeta)) - G(\Phi^t_{h_0}(\zeta)) \right] \, dt
= Z(\zeta) - G(\zeta).
\end{equation*}

In the last step we used the explicit expression of $Z$ provided in (276). Finally, the first estimate in (278) follows from the explicit expression of $Z$ in (276) while for the second estimate we write explicitly the vector field $X_\chi$:

\begin{equation*}
X_\chi(\zeta) = \frac{1}{T} \int_0^T t \, d\Phi^t_{h_0}(\Phi^t_{h_0}(\zeta)) \circ X_Z(\Phi^t_{h_0}(\zeta)) \, dt.
\end{equation*}

Assumption (PER) ensures that $\Phi^t_{h_0}$ as well as its derivatives and the inverses are uniformly bounded as operators from $H^{p,s}$ into itself. Moreover, for any $t \in \mathbb{R}$, the map $\zeta \mapsto \Phi^t_{h_0}(\zeta)$ is a diffeomorphism of $B_{p,s}(R)$ into itself. Thus

\begin{equation*}
\sup_{B_{p,s}(R)} \|X_\zeta(\zeta)\|_{H^{p,s}} \leq T \sup_{t \in [0,T]} \sup_{\zeta \in H^{p,s}} \left( \|(D\Phi^t_{h_0}(\zeta))^{-1}\|_{H^{p,s} \rightarrow H^{p,s}} \right) \sup_{B_{p,s}} \left( \|X_Z(\zeta)\|_{H^{p,s}} + \|X_G(\zeta)\|_{H^{p,s}} \right)
\leq 2T \sup_{t \in [0,T]} \sup_{\zeta \in H^{p,s}} \left( \|(D\Phi^t_{h_0}(\zeta))^{-1}\|_{H^{p,s} \rightarrow H^{p,s}} \right) \sup_{B_{p,s}} \|X_G(\zeta)\|_{H^{p,s}}
\end{equation*}

where in the last step we used the first inequality in (278). Since by assumption (PER) $\Phi^t_{h_0}$ is an isometry, $\sup_{t \in [0,T]} \sup_{\zeta \in H^{p,s}} \left( \|(D\Phi^t_{h_0}(\zeta))^{-1}\|_{H^{p,s} \rightarrow H^{p,s}} \right) = 1$ and the thesis follows. \hfill \Box

Lemma A.7. Assume that $G$ and its vector fields are analytic on $B_{p,s}(R)$, and that $h_0$ satisfies (PER). Let $\chi$ and its vector field be analytic on $B_{p,s}(R)$, and assume that $\chi$ solves (275). For any $\ell \geq 1$ denote by $h_{0,\ell}$ the functions defined recursively as in (271) from $h_0$. Then for any $d \in (0, R)$ one has that $h_{0,\ell}$ and its vector field are analytic on $B_{p,s}(R - d)$, and

\begin{equation}
\sup_{B_{p,s}(R - d)} \|X_{h_{0,\ell}}(\zeta)\|_{H^{p,s}} \leq 2 \sup_{B_{p,s}(R)} \|X_G(\zeta)\|_{H^{p,s}} \left( \frac{d}{d} \sup_{B_{p,s}(R)} \|X_\chi(\zeta)\|_{H^{p,s}} \right)^\ell.
\end{equation}
Proof. By using (275) one gets that \( h_{0,1} = Z - G \) is analytic on \( B_{\rho,s}(R) \). Then by exploiting (274) one gets the result.

**Lemma A.8.** Assume that \( G \) and its vector field are analytic on \( B_{\rho,s}(R) \), and that \( h_0 \) satisfies (PER). Let \( \chi \) be the solution of (275), denote by \( \Phi^\chi \) the flow of the Hamiltonian vector field associated to \( \chi \) and by \( \Phi_\chi \) the corresponding time-one map. Moreover, denote by

\[
F(\zeta) := h_0(\Phi_\chi(\zeta)) - h_0(\zeta) - \{\chi, h_0\}(\zeta).
\]

Let \( d < R \), and assume that

\[
\sup_{\mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq d/3.
\]

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

\[
\sup_{\mathcal{B}_{\rho,s}(R-d)} \|X_F(\zeta)\|_{\mathcal{H}^{\rho,s}} \leq \frac{18}{d} \sup_{\mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{\mathcal{B}_{\rho,s}(R)} \|X_G(\zeta)\|_{\mathcal{H}^{\rho,s}}.
\]

**Proof.** Since

\[
h_0(\Phi_\chi(\zeta)) - h_0(\zeta) = \int_0^1 \{\chi, h_0\} \circ \Phi^\chi(\zeta) \, dt
\]

(275)\[
= \int_0^1 Z(\Phi^\chi(\zeta)) - G(\Phi^\chi(\zeta)) \, dt,
\]

if we define \( F(\zeta) := Z(\zeta) - G(\zeta) \), we get

\[
F(\zeta) = \int_0^1 F(\Phi^\chi(\zeta)) - F(\zeta) \, dt.
\]

Now, we have

\[
\sup_{\mathcal{B}_{\rho,s}(R-d)} \|X_F(\zeta)\|_{\mathcal{H}^{\rho,s}}
\]

\[
= \sup_{\mathcal{B}_{\rho,s}(R-d)} \|\Omega^{-1} d \left( \int_0^1 F(\Phi^\chi(\zeta)) - F(\zeta) \, dt \right) \|_{\mathcal{H}^{\rho,s}}
\]

\[
\leq \sup_{\mathcal{B}_{\rho,s}(R-d)} \left\| \int_0^1 (d\Phi^{-1}(\Phi^\chi(\zeta)) - id)\Omega^{-1} dF(\Phi^\chi(\zeta)) + \Omega^{-1} d(F(\Phi^\chi(\zeta)) - F(\zeta)) \right\|_{\mathcal{H}^{\rho,s}}
\]

\[
\leq \sup_{\mathcal{B}_{\rho,s}(R-d)} \| \int_0^1 (d\Phi^{-1}(\Phi^\chi(\zeta)) - id)\Omega^{-1} dF(\Phi^\chi(\zeta)) \|_{\mathcal{H}^{\rho,s}}
\]

\[+ \sup_{\mathcal{B}_{\rho,s}(R-d)} \| \int_0^1 X_F(\Phi^\chi(\zeta)) - X_F(\zeta) \|_{\mathcal{H}^{\rho,s}} \]

and by dominated convergence we can bound the last quantity by

\[
\sup_{\mathcal{B}_{\rho,s}(R-d)} \sup_{t \in [0,1]} \|d\Phi^{-1}(\Phi^\chi(\zeta)) - id\|_{\mathcal{H}^{\rho,s} \rightarrow \mathcal{H}^{\rho,s}} \sup_{\mathcal{B}_{\rho,s}(R-d)} \|X_F(\Phi^\chi(\zeta))\|_{\mathcal{H}^{\rho,s}}
\]

\[+ \sup_{\mathcal{B}_{\rho,s}(R-d)} \sup_{t \in [0,1]} \|X_F(\Phi^\chi(\zeta)) - X_F(\zeta)\|_{\mathcal{H}^{\rho,s}} \]

\[
\leq \frac{2}{d} \sup_{\mathcal{B}_{\rho,s}(R)} \|X_\chi(\zeta)\|_{\mathcal{H}^{\rho,s}} \sup_{\mathcal{B}_{\rho,s}(R)} \|X_F(\zeta)\|_{\mathcal{H}^{\rho,s}} + \sup_{t \in [0,1]} \sup_{\mathcal{B}_{\rho,s}(R-d)} \| \int_0^t \|X_\chi, X_F\|\Phi^\chi(\zeta) ds \|_{\mathcal{H}^{\rho,s}}.
\]
where we can estimate the last term by Cauchy inequality
\[
\sup_{B_{p,s}(R-d)} \| \int_0^t [X_{\chi}, X_F](\Phi^s_\xi(\zeta)) \, ds \|_{H^{p,s}} \leq 2 \sup_{B_{p,s}(R-2d/3)} \| [X_{\chi}, X_F](\zeta) \|_{H^{p,s}} \\
\leq \frac{6}{2d} \sup_{B_{p,s}(R)} \| X_\chi(\zeta) \|_{H^{p,s}} \sup_{B_{p,s}(R)} \| X_F(\zeta) \|_{H^{p,s}} \\
\leq \frac{6}{d} \sup_{B_{p,s}(R)} \| X_\chi(\zeta) \|_{H^{p,s}} \sup_{B_{p,s}(R)} \| X_F(\zeta) \|_{H^{p,s}}.
\]
By the above computations and (278) we obtain
\[
\sup_{B_{p,s}(R-d)} \| X_F(\zeta) \|_{H^{p,s}} \leq \frac{9}{d} \sup_{B_{p,s}(R)} \| X_\chi(\zeta) \|_{H^{p,s}} \sup_{B_{p,s}(R)} \| X_F(\zeta) \|_{H^{p,s}} \leq \frac{18}{d} \sup_{B_{p,s}(R)} \| X_\chi(\zeta) \|_{H^{p,s}} \sup_{B_{p,s}(R)} \| X_F(\zeta) \|_{H^{p,s}}.
\]

**Lemma A.9.** Let \( s \geq s_1 \gg 1, R > 0, m \geq 0, \) and consider the Hamiltonian
\[
H^{(m)}(\zeta) = h_0(\zeta) + \delta Z^{(m)}(\zeta) + \delta^{m+1} F^{(m)}(\zeta).
\]
Assume that \( h_0 \) satisfies (PER) and (INV), and that
\[
\sup_{B_{p,s}(R)} \| X_F(0)(\zeta) \|_{H^{p,s}} \leq F.
\]
Fix \( d < \frac{1}{m+1} \), and set \( R_m := R - md \) (\( m \geq 1 \)).
Assume also that \( Z^{(m)} \) is analytic on \( B_{p,s}(R_m) \), and that
\[
\sup_{B_{p,s}(R_m)} \| X_{Z^{(0)}}(\zeta) \|_{H^{p,s}} = 0,
\]
(282)
\[
\sup_{B_{p,s}(R_m)} \| X_{Z^{(m)}}(\zeta) \|_{H^{p,s}} \leq F \sum_{i=0}^{m-1} \delta^i K_0, \quad m \geq 1,
\]
(283)
\[
\sup_{B_{p,s}(R_m)} \| X_{F^{(m)}}(\zeta) \|_{H^{p,s}} \leq F K_0^m, \quad m \geq 1,
\]
with \( K_0 \geq 15 \) and \( d > 3 \delta F \).
Then, if \( \delta K_0 < 1/2 \) there exists a canonical transformation \( T^{(m)}_\delta(\zeta) \) analytic on \( B_{p,s}(R_{m+1}) \) such that
\[
\sup_{B_{p,s}(R_{m+1})} \| T^{(m)}_\delta(\zeta) - \zeta \|_{H^{p,s}} \leq 2 \delta^{m+1} K_0^m F,
\]
(284)
\[
H^{(m+1)} := H^{(m)} \circ T^{(m)}_\delta(\zeta) \text{ has the form (281) and satisfies (283) with } m \text{ replaced by } m + 1.
\]

**Proof.** The key point of the proof is to look for \( T^{(m)}_\delta(\zeta) \) as the time-one map of the Hamiltonian vector field of an analytic function \( \delta^{m+1} \chi_m \). Hence, consider the differential equation
\[
\dot{\zeta} = X_{\delta^{m+1} \chi_m}(\zeta).
\]
By standard theory we have that, if \( X_{\delta^{m+1} \chi_m} \|_{B_{p,s}(R_m)} \) is small enough (e.g. \( X_{\delta^{m+1} \chi_m} \|_{B_{p,s}(R_m)} \leq \frac{1}{m+1} \)) and \( \zeta_0 \in B_{p,s}(R_{m+1}) \), then the solution of (285) exists for \( |t| \leq 1. \)
Therefore we can define $T_{m+1} : B_{p,s}(R_{m+1}) \to B_{p,s}(R_m)$, and in particular the corresponding time-one map $T_{m+1}^2 := T_{m+1}$, which is an analytic canonical transformation, $\delta^{m+1}$-close to the identity. We have

\[
(T_{m+1}^2)^* \left( h_0 + \delta Z^{(m)} + \delta^{m+1} F^{(m)} \right) = h_0 + \delta Z^{(m)} + \delta^{m+1} \left[ \{\chi_m, h_0\} + F^{(m)} \right] +
\]

\[
\left( h_0 \circ T_{m+1}^2 - h_0 - \delta^{m+1} \{\chi_m, h_0\} \right) + \delta \left( Z^{(m)} \circ T_{m+1}^2 - Z^{(m)} \right)
\]

\[
\delta^{m+1} \left( F^{(m)} \circ T_{m+1}^2 - F^{(m)} \right).
\]

(286)

(287)

It is easy to see that the first two terms are already normalized, that the term in the second line is the non-normalized part of order $m+1$ that can be normalized through the choice of a suitable $\chi_m$, and that (286)-(287) contain all the terms of order higher than $m+1$.

In order to normalize the terms in the second line we solve the homological equation

\[
\{\chi_m, h_0\} + F^{(m)} = Z_{m+1},
\]

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

\[
Z_{m+1}(\zeta) = \frac{1}{T} \int_0^T F^{(m)}(\Phi_{h_0}^t(\zeta)) \, dt,
\]

\[
\chi_m(\zeta) = \frac{1}{T} \int_0^T \frac{1}{2} T_0^T t F^{(m)}(\Phi_{h_0}^t(\zeta)) - Z_{m+1}(\Phi_{h_0}^t(\zeta)) \, dt.
\]

The explicit expression of $X_{\chi_m}$ can be computed following the argument of Lemma A.6. Using this explicit expression, the analyticity of the flow $\Phi_{h_0}$ ensured by (PER) and (287) one has

\[
\sup_{B_{p,s}(R_m)} \|X_{\chi_m}(\zeta)\|_{H_p \times s} \leq 2T \sup_{B_{p,s}(R_m)} \|X_{\Phi_{h_0}^t}(\zeta)\|_{H_p \times s} \leq 2T K_0^m F.
\]

(\*)

Straightforwardly, from the explicit expression of $Z_{m+1}(\zeta)$ and (283) one has

\[
\sup_{B_{p,s}(R_m)} \|X_{Z_{m+1}}\|_{H_p \times s} \leq K_0^m F.
\]

Now define $Z^{(m+1)} := Z^{(m)} + \delta^m Z_{m+1}$ and notice that as a consequence of the latter estimate and (282) we have

\[
\sup_{B_{p,s}(R_{m+1})} \|X_{Z^{(m+1)}}(\zeta)\| \leq \sup_{B_{p,s}(R_{m+1})} \|X_{Z^{(m)}}(\zeta)\|_{H_p \times s} + \sup_{B_{p,s}(R_{m+1})} \|X_{\chi_m Z_{m+1}}(\zeta)\|_{H_p \times s}
\]

\[
\leq F \left( \sum_{j=0}^{m-1} \delta^j K_0^j + \delta^m K_0^m \right)
\]

Defining now $T_{m+1}^2(\zeta) := \Phi_{\Phi_{h_0}^{m+1} \chi_m}^1(\zeta)$ we can apply Lemma A.1 and (\*) to obtain

\[
\sup_{B_{p,s}(R_{m+1})} \|T_{m+1}^{(m)}(\zeta) - \zeta\|_{H_p \times s} = \sup_{B_{p,s}(R_{m+1})} \|\Phi_{\Phi_{h_0}^{m+1} \chi_m}(\zeta) - \zeta\|_{H_p \times s}
\]

\[
\leq \sup_{B_{p,s}(R_m)} \|X_{\Phi_{h_0}^{m+1} \chi_m}(\zeta)\|_{H_p \times s} \leq 2T \delta^{m+1} K_0^m F.
\]

Let us set now $\delta^{m+2} F^{(m+1)} := (286) + (287)$. Using Lemma A.5 one can estimate separately the three pieces. We notice that $\sup_{B_{p,s}(R_m)} \|X_{\Phi_{h_0}^{m+1} \chi_m}\|_{H_p \times s} \leq 2T \delta^{m+1} K_0^m F$ and since $\delta K_0 < \frac{1}{2}$ we have

\[
\sup_{B_{p,s}(R_m)} \|X_{\Phi_{h_0}^{m+1} \chi_m}\|_{H_p \times s} < T \delta F < \frac{1}{2} \delta^{m+1} \delta K_0^m.
\]

We can thus apply Lemma A.5 and Lemma A.8 to
If \( m \geq 1 \), we exploit the smallness condition \( \delta K_0 < \frac{1}{2} \) to get \( \sum_{i=0}^{m-1} \delta K_0^i < 2 \) and
\[
\sup_{B_{\rho,s}(R_{m+1})} \| X_{z_{m+2}f(m+1)} \|_{H^{\rho,s}} \leq \delta^{m+2} \left( 6F + \frac{9F}{2m} + \frac{6F}{2m} \right) \leq 15 \delta^{m+2}F.
\]

By means of these inequalities, with the additional information \( \sup_{B_{\rho,s}(R_{m+1})} \| f_{z_{m+2}f(m+1)} \|_{H^{\rho,s}} \leq \frac{(m+1)d}{2} \) and the hypotheses (282) and (283), we can estimate
\[
\sup_{B_{\rho,s}(R_{m+1})} \| X_{z_{m+2}f(m+1)} \|_{H^{\rho,s}} \leq \delta^{m+2} \sup_{B_{\rho,s}(R_{m+1})} \| X_{z_{m+2}f(m)} \|_{H^{\rho,s}} + \delta^{m+2} \sup_{B_{\rho,s}(R_{m+1})} \| X_{f_{z_{m+2}f(m+1)}} \|_{H^{\rho,s}}
\]
\[
+ \delta^{m+2} \sum_{i=0}^{m-1} \delta K_0^i + 9 \delta^{m+2} F K_0^m + 6 \delta^{m+2} F K_0^m
\]
\[
= \delta^{m+2} \left( 9F \sum_{i=0}^{m-1} \delta K_0^i + 9F \delta^{m+2} F K_0^m + 6 \delta^{m+2} F K_0^m \right)
\]

If \( m = 0 \) the first term is not present and then
\[
\sup_{B_{\rho,s}(R_{1})} \| X_{z_{2f(1)}} \|_{H^{\rho,s}} \leq \delta^2 (9F + 6F) + \delta^2 (9F + 6F).
\]

If \( m \geq 1 \) we exploit the smallness condition \( \delta K_0 < \frac{1}{2} \) to get \( \sum_{i=0}^{m-1} \delta K_0^i < 2 \) and
\[
\sup_{B_{\rho,s}(R_{m+1})} \| X_{z_{m+2}f(m+1)} \|_{H^{\rho,s}} \leq \delta^{m+2} \left( 6F + \frac{9F}{2m} + \frac{6F}{2m} \right) \leq 15 \delta^{m+2} F.
\]

\[\]
APPENDIX B. PROOF OF Proposition 6.2

In order to prove Proposition 6.2 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

$$\| \frac{\omega_m^2}{\mu^2} \|^{(11)} = \frac{4}{\mu^2} \left[ \sin^2 \left( \frac{K_1 \pi}{2N + 1} \right) + \sin^2 \left( \frac{K_2 \pi}{2N + 1} \right) \right]$$

$$\leq \pi^2 (K_1^2 + \mu^{2(\sigma-1)} K_2^2);$$

moreover, for $K_1 \neq 0$

$$\frac{|\hat{q}_K|^2 + \pi^2 (K_1^2 + \mu^{2(\sigma-1)} K_2^2)}{2} |\hat{p}_K|^2 \leq \pi^2 e^{-2\rho |K|} \left( \frac{|\hat{q}_K|^2 + (K_1^2 + \mu^{2(\sigma-1)} K_2^2)}{2} |\hat{p}_K|^2 \right) e^{2\rho |K|}$$

while for $|K_2| \leq |K_1|$

$$\frac{|\hat{q}_K|^2 + \pi^2 (K_1^2 + \mu^{2(\sigma-1)} K_2^2)}{2} |\hat{p}_K|^2 \leq 2 \pi^2 e^{-2\rho |K|} \| (\xi, \eta) \|^2_{H^{1,0}}.$$
Now,
\[
\sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|} 
\leq e^{-2\rho |K|} + \sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|} + \sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|}
\]
\[
= e^{-2\rho |K|} + \sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|} + \sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|}
\]
\[
= e^{-2\rho |K|} + \sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|} + \sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|}
\]
\[
+ \sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|}.
\]

We now estimate the last sum in (297); we point out that for $L_1, L_2 \neq 0$ we have
\[
|L| \geq \frac{2}{\mu} + \frac{2}{\mu^2},
\]
hence
\[
2|K| \leq |L|.
\]
Therefore, for any $k$ such that $\kappa(k) = (\mu K_1, \mu^2 K_2)$ and $|K_1| + |K_2| \geq \frac{(2+\delta) \log \mu}{\rho}$
\[
\sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|} \leq \sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-2\rho |K+L|} \leq e^{2\rho |K|} \int_{2|K|}^{+\infty} e^{-2\rho R} dR
\]
\[
= 2\pi e^{2\rho |K|} \left( \frac{1}{2} \right) d \frac{d}{d\rho} \left[ \int_{2|K|}^{+\infty} e^{-2\rho R} dR \right]
\]
\[
= -\pi e^{2\rho |K|} \left( \frac{e^{-4\rho |K|}}{2\rho} \right) - \pi e^{2\rho |K|} \left( \frac{1}{2\rho} \right) - 2|K| e^{-4\rho |K|} - 2|K| e^{-4\rho |K|}
\]
\[
= \frac{\pi}{2\rho} \left( \frac{1}{\rho} + 4 \right) e^{-2\rho |K|}
\]
\[
= \frac{\pi}{2\rho} \left( \frac{1}{\rho} + 4 \right) e^{-2\rho |K|}
\]
Next we estimate the second sum in (297); we have
\[
\sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-4\rho |l|/\mu} \leq e^{-2\rho |(K_1+K_2)|} \sum_{\ell \in \mathbb{Z}/\{0\}} e^{-4\rho |\ell|/\mu},
\]
which is exponentially small with respect to $\mu$. Similarly,
\[
\sum_{L=(l_1, l_2) \in \mathbb{Z}^2 \text{ such that } |l_1| > (2+\delta) \log \mu} e^{-4\rho |l|/\mu^2} \leq e^{-2\rho |(K_1+K_2)|} \sum_{\ell \in \mathbb{Z}/\{0\}} e^{-4\rho |\ell|/\mu^2},
\]
Then,

\[
\sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu^2 L_2 \in 2\mathbb{Z}, \left|K_1 + K_2\right| > \frac{(2+\delta)^{1/2}}{\log(2\rho|K|)}} e^{-2\rho|K + L|} \left(\frac{K_2 + L_2}{K_1 + L_1}\right)^2 \leq e^{-2\rho|K|} \left(\frac{K_2}{K_1}\right)^2
\]

\[
+ \sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu^2 L_2 \in 2\mathbb{Z}, \left|K_1 + K_2\right| > \frac{(2+\delta)^{1/2}}{\log(2\rho|K|)}} \left|K_1 + L_1\right| \neq 0 L_1 \neq 0, L_2 \neq 0 \left(\frac{K_2 + L_2}{K_1 + L_1}\right)^2 + \sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu^2 L_2 \in 2\mathbb{Z}, \left|K_1 + K_2\right| > \frac{(2+\delta)^{1/2}}{\log(2\rho|K|)}} \left|K_1 + L_1\right| \neq 0 L_1 \neq 0, L_2 \neq 0 \left(\frac{K_2 + L_2}{K_1 + L_1}\right)^2
\]

(302)

First we estimate the last term in (302): we have that \(|L + K| \geq |K|\), hence

\[
\sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu^2 L_2 \in 2\mathbb{Z}, \left|K_1 + K_2\right| > \frac{(2+\delta)^{1/2}}{\log(2\rho|K|)}} \left|K_1 + L_1\right| \neq 0 L_1 \neq 0, L_2 \neq 0 e^{-2\rho|K + L|} \left(\frac{K_2 + L_2}{K_1 + L_1}\right)^2
\]

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

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

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

(303)

(304)

Now we bound the other two nontrivial terms in (302); on the one hand, we notice that

\[
\sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu^2 L_2 \in 2\mathbb{Z}, \left|K_1 + K_2\right| > \frac{(2+\delta)^{1/2}}{\log(2\rho|K|)}} \left|K_1 + L_1\right| \neq 0 L_1 \neq 0, L_2 = 0 e^{-2\rho|K + L|} L_2^2
\]

(305)
vanishes, while on the other hand
\[ \sum_{L=(L_1,L_2) \in \mathbb{Z}^2, \mu L_1, \mu^* L_2 \in \mathbb{Z}} e^{-2\rho |K+L|^2} \leq e^{-2\rho |K|} \sum_{\ell \in \mathbb{Z}\setminus\{0\}} e^{-4\rho |\ell|/\mu^*} \frac{\ell^2}{\mu^{2*}} \]
where the last integral is exponentially small with respect to \( \mu \).

On the other hand, for any \( k \) such that \( \kappa(k) = (\mu K_1, \mu^* K_2) \) and \( |K_1| + |K_2| \leq \frac{(2+\delta)|\log \mu|}{\rho} \)
\[ \left( \frac{\pi^4}{2} \frac{\xi^2 + \hat{\eta}^2}{\rho^2} \right) \leq 2e^{-2\rho |K|} \int_1^{+\infty} e^{-4\rho |\ell|/\mu^*} \frac{\ell^2}{\mu^{2*}} d\ell, \]
and we can conclude by estimating (307) by exploiting the fact that \( |\log \mu| \leq \mu^{-2/5} \), while we can estimate (308) by
\[ \frac{\pi^2}{2} \| (\xi, \eta) \|_{H^{1,0}}^2 \sum_{L=(L_1,L_2) \in \mathbb{Z}^2, \mu L_1, \mu^* L_2 \in \mathbb{Z}} \left( 1 + 2\mu^2(s-1) \frac{K_1^2 + L_2^2}{(K_1 + L_1)^2} \right) e^{-2\rho |K+L|^2} \]
\[ \leq \frac{\pi^2}{2} \| (\xi, \eta) \|_{H^{1,0}}^2 \sum_{L=(L_1,L_2) \in \mathbb{Z}^2, \mu L_1, \mu^* L_2 \in \mathbb{Z}} \left( 1 + 2\mu^2(s-1) \frac{K_2^2 + L_1^2}{(K_2 + L_2)^2} \right) e^{-2\rho |K+L|^2} \]
\[ \leq \frac{\pi^2}{2} \| (\xi, \eta) \|_{H^{1,0}}^2 \left( 1 + 2\mu^2(s-1) \frac{K_2^2}{(K_1 + L_1)^2} \right) 2\pi \int_0^{+\infty} e^{-2\rho |\ell|} d\ell + 4\pi \int_0^{+\infty} e^{-2\rho |\ell|} \ell^2 d\ell \]
\[ = \frac{\pi^2}{2} \| (\xi, \eta) \|_{H^{1,0}}^2 \times \left[ 2\pi \left( 1 + 2\mu^2(s-1) \frac{9|\log \mu|}{\rho^2} \right) e^{-4\rho/\mu + 4\rho/\mu^* + 8\mu^3(\delta^3 + 3\mu^2 + 2\mu^2(2^2 + 32\mu^3)/8) \mu^*} \right]. \]
Hence, by (221) we obtain that for all $k$ such that $\kappa(k) = (\mu K_1, \mu^2 K_2)$ we have
\[
|\omega_k|^2 = 1 + 4 \left( \sin^2 \left( \frac{k_1 \pi}{2N+1} \right) + \sin^2 \left( \frac{k_2 \pi}{2N+1} \right) \right)
\]
\[
= 1 + 4 \left| \sin \left( \frac{\mu K_1 \pi}{2} \right) + \sin \left( \frac{\mu^2 K_2 \pi}{2} \right) \right|^2
\]
\[
\leq 1 + \pi^2 (\mu^2 K_1^2 + \mu^2 K_2^2),
\]
hence
\[
\frac{|\hat{p}_K|^2 + \pi^2 (1 + \mu^2 K_1^2 + \mu^2 K_2^2)|\hat{q}_K|^2}{2} \leq \pi^2 e^{-2\rho|K|} \frac{|\hat{p}_K|^2 + (1 + \mu^2 K_1^2 + \mu^2 K_2^2)|\hat{q}_K|^2}{2}
\]
\[
\leq \pi^2 e^{-2\rho|K|} (1 + \mu^2 K_1^2 + \mu^2 K_2^2) \| \psi, \bar{\psi} \|_{H^{n,o}}^2.
\]
Hence, by (221) we obtain that for all $k$ such that $\kappa(k) = (\mu K_1, \mu^2 K_2)$ and $|K_1| + |K_2| > \frac{(2+\delta) \log \eta}{\rho}$

\[
E_\rho 
\]

\[
\leq \sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|} \left( |\hat{p}_K+L|^2 + \omega_k^2 |\hat{q}_K+L|^2 \right)
\]

\[
\leq \pi^2 \| \psi, \bar{\psi} \|_{H^{n,o}}^2 2 \sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|} \left[ (1 + \mu^2 (K_1 + L_1)^2 + \mu^2 (K_2 + L_2)^2) \right],
\]

where the sum in (312) can be rewritten as follows,

\[
\sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|} \left[ (1 + \mu^2 (K_1 + L_1)^2 + \mu^2 (K_2 + L_2)^2) \right]
\]

\[
+ \mu^2 \sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|} (K_1 + L_1)^2
\]

\[
+ \mu^2 \sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|} (K_2 + L_2)^2.
\]

Now,

\[
\sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|}
\]

\[
\leq e^{-2\rho|K|} + \sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|} + \sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|}
\]

\[
+ \sum_{L=(L_1,L_2) \in \mathbb{Z}^2 : \rho L_1, \rho L_2 \in \mathbb{Z}} e^{-2\rho|K+L|}.
\]
and we can estimate the above terms as for (297) in Proposition 6.2; indeed, by (299), (300) and (301) we have that (316) is bounded by
\[
e^{-2\rho |K|} + \pi \left( \frac{1}{2\rho^2} + 2|K| \right) e^{-2\rho |K|} + e^{-2\rho (|K_1|+|K_2|)} \sum_{\ell \in \mathbb{Z}\setminus\{0\}} e^{-4\rho |\ell|/\mu}
\]
\[
+ e^{-2\rho (|K_1|+|K_2|)} \sum_{\ell \in \mathbb{Z}\setminus\{0\}} e^{-4\rho |\ell|/\mu''}.
\]
(317)

Now we estimate (314). We have
\[
\sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in 2\mathbb{Z}, |K_1|+|K_2| \geq 2 \frac{\log \mu}{\rho}} e^{-2\rho |K| K_1^2}
\]
\[
\leq e^{-2\rho |K| K_1^2} \sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in 2\mathbb{Z}, |K_1|+|K_2| \geq 2 \frac{\log \mu}{\rho}} e^{-2\rho |K+L| (K_1 + L_1)^2} + \sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in 2\mathbb{Z}, |K_1|+|K_2| \geq 2 \frac{\log \mu}{\rho}} e^{-2\rho |K+L| K_1^2}
\]
\[
+ \sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in 2\mathbb{Z}, |K_1|+|K_2| \geq 2 \frac{\log \mu}{\rho}} e^{-2\rho |K+L| (K_1 + L_1)^2}.
\]
(318)

First we estimate the last term in (318): we have that $|L+K| \geq |K|$, hence
\[
\sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in 2\mathbb{Z}, |K_1|+|K_2| \geq 2 \frac{\log \mu}{\rho}} e^{-2\rho |K+L| (K_1 + L_1)^2}
\]
\[
= \int_{|K|}^{+\infty} \int_{0}^{e^{-2\rho |K| \log \mu/\rho}} e^{-2\rho \xi \cos^2 \phi} \phi \, d\phi \, d\xi
\]
\[
= \pi e^{-2\rho |K|} \frac{1+2\rho |K|}{4\rho^2}
\]
\[
\leq \pi \mu^4 e^{-2\rho (|K| - \frac{2\log \mu}{\rho} - \frac{1}{2\rho} \log(2\rho |K|))}
\]
\[
\leq 2^{1-1/\rho} \pi \mu^4 e^{-2\rho (|K| - \frac{2\log \mu}{\rho})}
\]
(319)

Now we bound the other two nontrivial terms in (318); on the one hand, we notice that
\[
\sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in 2\mathbb{Z}, |K_1|+|K_2| \geq 2 \frac{\log \mu}{\rho}} e^{-2\rho |K+L| (K_1 + L_1)^2}
\]
\[
\leq 2 \sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in 2\mathbb{Z}, |K_1|+|K_2| \geq 2 \frac{\log \mu}{\rho}} e^{-2\rho |K+L| K_1^2}
\]
\[
+ 2 \sum_{L=(L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in 2\mathbb{Z}, |K_1|+|K_2| \geq 2 \frac{\log \mu}{\rho}} e^{-2\rho |K+L| L_1^2},
\]
(320)
where the first sum can be bounded as the second term in (316), while

\[ \sum_{L = (L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in \mathbb{Z}, |K_1| + |K_2| > \frac{2}{\log \mu}} e^{-2 \rho |K + L|} |L|^2 \leq e^{-2 \rho |K|} \sum_{\ell \in \mathbb{Z} \setminus \{0\}} e^{-4 \rho |\ell|/\mu} \frac{\ell^2}{\mu^2} \]

\[(321) \leq 2e^{-2 \rho |K|} \int_1^{+\infty} e^{-4 \rho |\ell|/\mu} \frac{\ell^2}{\mu^2} d\ell, \]

where the last integral is exponentially small with respect to \( \mu \).

Similarly,

\[ \sum_{L = (L_1, L_2) \in \mathbb{Z}^2, \mu L_1, \mu L_2 \in \mathbb{Z}, |K_1| + |K_2| > \frac{2}{\log \mu}} e^{-2 \rho |K + L|} |L|^2 \leq e^{-2 \rho |K|} \sum_{\ell \in \mathbb{Z} \setminus \{0\}} e^{-4 \rho |\ell|/\mu^\sigma} \frac{\ell^2}{\mu^{2\sigma}} \]

\[(322) \leq 2e^{-2 \rho |K|} \int_1^{+\infty} e^{-4 \rho |\ell|/\mu^\sigma} \frac{\ell^2}{\mu^{2\sigma}} d\ell, \]

where the last integral is exponentially small with respect to \( \mu \).

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

\[ \left| \frac{\xi_k}{\mu^2} - \frac{\hat{\varphi}_K}{2} \right| \leq |\omega_2|^2 - 1 |\hat{q}_K|^2 + 1 \left| \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} |\hat{p}_{K+L}|^2 + \omega_2^2 |\hat{q}_{K+L}|^2 \right| \]

\[(310) \leq (\mu^2 \tau^2 K_1^2 + \tau^2 \mu^{2\sigma} K_2^2) |\hat{p}_K|^2 + 1 \left| \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} |\hat{p}_{K+L}|^2 + |\hat{q}_{K+L}|^2 + \tau^2 |\mu^2 (K_1 + L_1)^2 + \mu^{2\sigma} (K_2 + L_2)^2||\hat{q}_{K+L}|^2, \]

\[ \leq (\tau^2 \mu^2 K_1^2 + \tau^2 \mu^{2\sigma} K_2^2) |\hat{p}_K|^2 + \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} e^{-2 \rho |K+L|} [1 + \tau^2 \mu^2 (K_1 + L_1)^2 + \tau^2 \mu^{2\sigma} (K_2 + L_2)^2] \]

\[(323) \leq \tau^2 \mu^2 \left( 1 + \mu^{2(\ell - 1)} \right) \frac{9 \log |\mu|^2}{\rho^2} \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} e^{-2 \rho |K+L|} \left| (\psi, \tilde{\psi}) \right|_{\mathcal{H}^p,0}^2 \]

\[ \left(324 \right) + \left| (\psi, \tilde{\psi}) \right|_{\mathcal{H}^p,0}^2 \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} e^{-2 \rho |K+L|} \left| (\psi, \tilde{\psi}) \right|_{\mathcal{H}^p,0}^2 \]

\[ \left(325 \right) + \tau^2 \mu^{2\sigma} \left| (\psi, \tilde{\psi}) \right|_{\mathcal{H}^p,0}^2 \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} e^{-2 \rho |K+L|} (K_1 + L_1)^2 \]

\[ \left(326 \right) + \tau^2 \mu^{2\sigma} \left| (\psi, \tilde{\psi}) \right|_{\mathcal{H}^p,0}^2 \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} e^{-2 \rho |K+L|} (K_2 + L_2)^2 \]
and we can conclude by estimating (323) by exploiting the fact that $|\log \mu| \leq \mu^{-2/5}$, while we can bound (324)-(325) by

$$\frac{\pi^2}{2} \left\| (\psi, \tilde{\psi}) \right\|_{H^{s_p,0}}^2 \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} \left(1 + 2 \mu^2 (K_1 + L_1)^2 \right) e^{-2\mu^2 |K + L|}$$

$$\frac{\pi^2}{2} \left\| (\psi, \tilde{\psi}) \right\|_{H^{s_p,0}}^2 \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} \left(1 + 2 \mu^2 K_1^2 + 2 \mu^2 L_1^2 \right) e^{-2\mu^2 |K + L|}$$

$$\leq \frac{\pi^2}{2} \left\| (\psi, \tilde{\psi}) \right\|_{H^{s_p,0}}^2 \left(1 + 2 \mu^2 K_1^2 \right) 2\pi \int_{2/\mu}^{+\infty} e^{-2\mu^2 \ell \rho} d\ell + 4\pi \mu^2 \int_{2/\mu}^{+\infty} e^{-2\mu^2 \ell^3} d\ell \right]$$

$$= \frac{\pi^2}{2} \left\| (\psi, \tilde{\psi}) \right\|_{H^{s_p,0}}^2 \times$$

$$\left(2\pi \left(1 + 2 \mu^2 \frac{9}{\rho^2} |\log \mu|^2 \right) e^{-4\rho/\mu} / 4\mu^2 + 4\pi \mu^2 e^{-4\rho/\mu} \frac{3\mu^3 + 12\rho \mu^2 + 24\rho^2 \mu + 32\rho^3}{8\mu^3 \rho^3} \right)\right.$$

and we can estimate (326) by

$$\frac{\pi^2}{2} \left\| (\psi, \tilde{\psi}) \right\|_{H^{s_p,0}}^2 \mu^{2(s-1)} \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} \left(K_1^2 + L_1^2 \right) e^{-2\mu^2 |K + L|}$$

$$\frac{\pi^2}{2} \left\| (\psi, \tilde{\psi}) \right\|_{H^{s_p,0}}^2 \mu^{2(s-1)} \sum_{L = (L_1, L_2) \in \mathbb{Z}^2 \setminus \{0\}} (2K_1^2 + 2L_1^2) e^{-2\mu^2 |K + L|}$$

$$\leq \frac{\pi^2}{2} \left\| (\psi, \tilde{\psi}) \right\|_{H^{s_p,0}}^2 \mu^{2(s-1)} \left[2K_1^2 \int_{2/\mu}^{+\infty} e^{-2\mu^2 \ell \rho} d\ell + 4\pi \int_{2/\mu}^{+\infty} e^{-2\mu^2 \ell^3} d\ell \right]$$

$$= \frac{\pi^2}{2} \left\| (\psi, \tilde{\psi}) \right\|_{H^{s_p,0}}^2 \times$$

$$\left[2\pi \left(1 + 2 \mu^2 \frac{9}{\rho^2} |\log \mu|^2 \right) \frac{\mu^3 + 4\rho}{4\mu^2 \rho^2} + 4\pi e^{-4\rho/\mu} \frac{3\mu^3 + 12\rho \mu^2 + 24\rho^2 \mu + 32\rho^3}{8\mu^3 \rho^3} \right)\right.$$

Acknowledgements. The authors would like to thank Dario Bambusi, Alberto Maspero, Tiziano Penati and Antonio Ponno for useful comments and suggestions. S.P. acknowledges financial support from the Spanish "Ministerio de Ciencia, Innovación y Universidades", through the María de Maeztu Programme for Units of Excellence (2015-2019) and the Barcelona Graduate School of Mathematics, and partial support by the Spanish MINECO-FEDER Grant MTM2015-65715-P. M.G. thanks for the kind hospitality of the Department of Mathematics at "Universitat Politécnica de Catalunya" where part of this work was carried out. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 757802).

References

[Bam99] Dario Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. *Mathematische Zeitschrift*, 230(2):345–387, 1999.

[Bam05] Dario Bambusi. Galerkin averaging method and Poincaré normal form for some quasilinear PDEs. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 4(4):669–702, 2005.

[Bam08] Bambusi, Dario and Ponno, Antonio. Resonance, Metastability and Blow up in FPU. Springer Berlin Heidelberg, 2008.

[BCCM15] Dario Bambusi, Andrea Carati, Alberto Maiocchi, and Alberto Maspero. Some analytic results on the FPU paradox. In *Hamiltonian partial differential equations and applications*, pages 235–254. Springer, 2015.

[BCP02] Dario Bambusi, Andrea Carati, and Antonio Ponno. The nonlinear Schrödinger equation as a resonant normal form. *Discrete and Continuous Dynamical Systems Series B*, 2(1):109–128, 2002.

[Ben05] Giancarlo Benettin. Time scale for energy equipartition in a two-dimensional fpu model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 15(1):015108, 2005.
