KP-II approximation for a scalar FPU system on a 2D square lattice
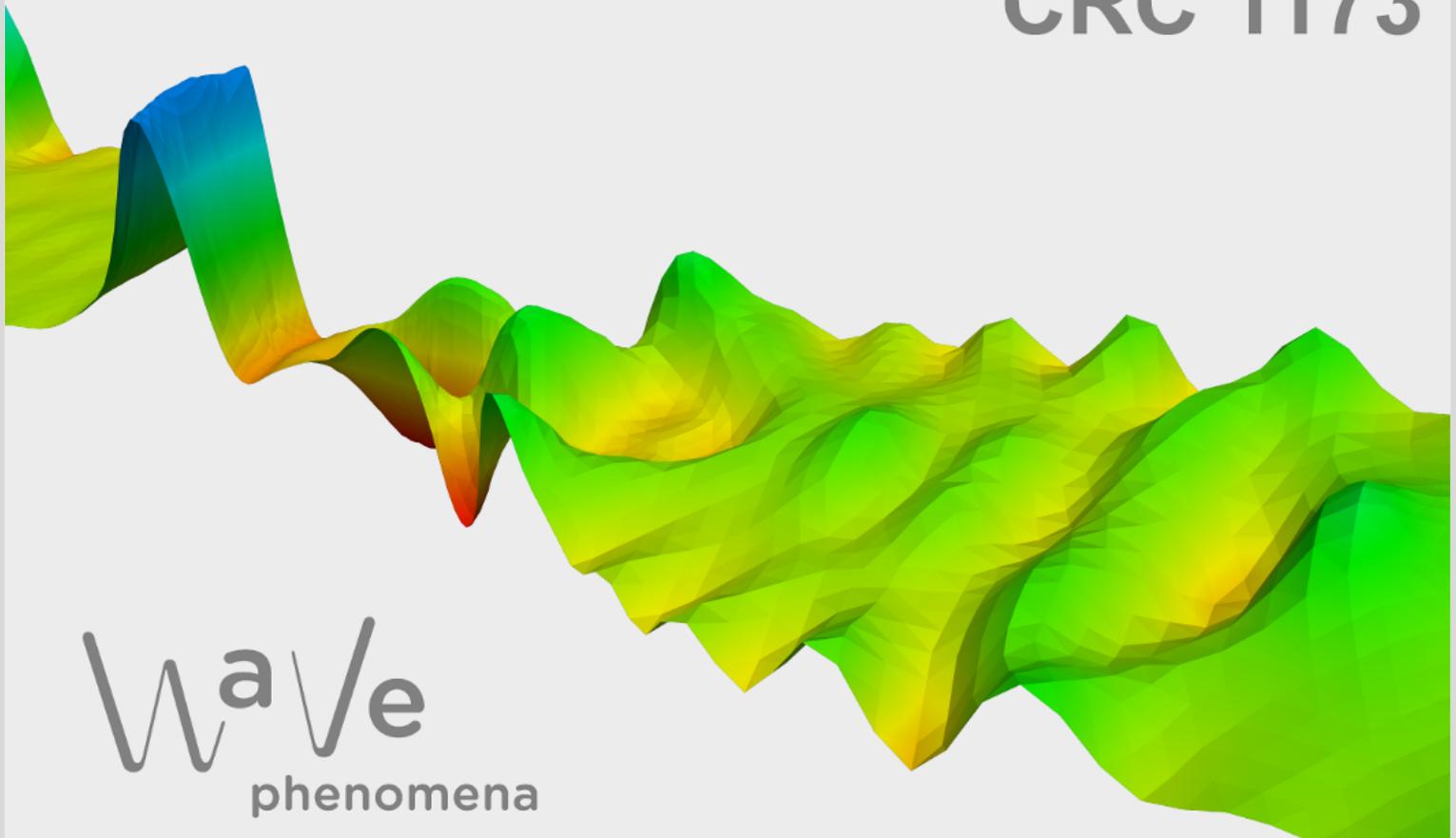
Dmitry E. Pelinovsky, Guido Schneider

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173

Participating universities

Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

TU
WIEN

Funded by

DFG

KP-II APPROXIMATION FOR A SCALAR FPU SYSTEM ON A 2D SQUARE LATTICE

DMITRY E.PELINOVSKY AND GUIDO SCHNEIDER

ABSTRACT. We consider a scalar Fermi–Pasta–Ulam (FPU) system on a square 2D lattice. The Kadomtsev–Petviashvili (KP-II) equation can be derived by means of multiple scale expansions to describe unidirectional long waves of small amplitude with slowly varying transverse modulations. We show that the KP-II approximation makes correct predictions about the dynamics of the original scalar FPU system. An existing approximation result is extended to an arbitrary direction of wave propagation. The main novelty of this work is the use of Fourier transform in the analysis of the FPU system in strain variables.

1. INTRODUCTION

We consider a scalar Fermi–Pasta–Ulam (FPU) system on a square 2D lattice. The equations of motion are given by

$$
\begin{aligned}
\partial_t^2 q_{m,n} = {} & W'(q_{m+1,n} - q_{m,n}) - W'(q_{m,n} - q_{m-1,n}) \\
& + W'(q_{m,n+1} - q_{m,n}) - W'(q_{m,n} - q_{m,n-1}), \quad (m,n) \in \mathbb{Z}^2,
\end{aligned}
\tag{1}
$$

where the scalar variable $q_{m,n}(t)$ describes a vertical displacement in z-direction of a particle with unit mass located at the (m,n)-th site in the (x,y) plane. The interaction potential between a particle and its four neighborhs is described by W. Figure 1 shows the mass–spring system on a square 2D lattice.
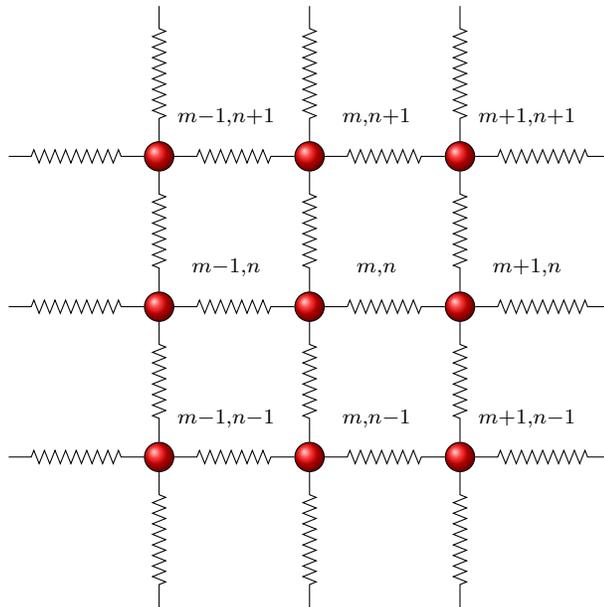

FIGURE 1. A mass–spring system arranged in a 2D square lattice. The masses are fixed at the (m,n) sites with the vertical displacements given by $q_{m,n}(t)$.

The total conserved energy of the FPU system (1) is given by

$$\mathcal{H}(q) = \sum_{(m,n)\in\mathbb{Z}^2} \frac{1}{2}(\partial_t q_{m,n})^2 + W(q_{m+1,n} - q_{m,n}) + W(q_{m,n+1} - q_{m,n}). \qquad (2)$$

For notational simplicity, we assume $W(u) = \frac{1}{2}u^2 - \frac{1}{3}u^3$ so that $W'(u) = u - u^2$.

It is well-known [8] that small-amplitude long-scale waves of the FPU system in one spatial dimension are described by the Korteweg–de Vries (KdV) equation. The KdV equation was first justified with bounds on the approximation error on the unbounded domain in [22] and on the periodic domain in [1, 17]. Applications of these methods for other generalized KdV equations can be found in [5, 11, 13, 15]. Properties of solitary waves in the FPU system in one spatial dimension were recently reviewed in [25].

We are interested to justify the Kadomtsev–Petviashvili (KP-II) approximation which describes unidirectional long waves of small amplitude with slowly varying transverse modulations. The formal derivation of the KP-II equation was reported in [6] but the rigorous justification was considered to be an open problem for some time. Two rigorous results were obtained only very recently. The KP-II equation was justified in the periodic domain among other integrable normal forms [10]. By using a more general setting of the vector FPU systems with particles moving in the (x, y) directions, the KP-II equation was justified in the unbounded 2D square lattice for the propagation along the axes (as well as for the diagonal propagation in the (x, y)-plane under additional constraints on the parameters of the lattice) [14].

The KP-II approximation is different from the KdV equation derived in the vector FPU systems with geometric nonlinearities, where small-amplitude supersonic longitudinal solitary waves may propagate along the horizontal direction [7] and along arbitrary directions [4]. Similarly, the liinearized KdV equation was derived for the linear propagation of rings in two-dimensional lattices [23], where the diffraction properties were neglected.

The purpose of this paper is to improve the justification result obtained in [14] so that it could apply to all directions of propagation in the (x, y) plane and without additional restrictions on parameters of the lattice. We take the normalized potential W and consider the scalar FPU lattice for simplicity, although extensions to the vector case with more complicated potentials W are relatively straightforward. The main novelty of our approach compared to [14] is working in Fourier space. Additionally, we have to construct a higher-order approximation involving the linearized KP-II equation in order to handle arbitrary directions of the wave propagation.

The paper is organized as follows. In Section 2 we introduce the strain variables for which the KP-II equations can be derived. By introducing the strain variables the original system is doubled and an additional compatibility condition has to be satisfied by the two strain variables. The new variables are transformed in Fourier space.

In Section 3 we derive the KP-II equation for the wave propagation along the x-direction. This simple case was considered in [14] and is used here to highlight analysis from the more complicated case of propagation along an arbitrary direction. We use smooth solutions of the KP-II equation to construct a suitable approximation of the FPU system satisfying the compatibility condition. The residual terms from the FPU system with the leading-order approximation are estimated, after which the main approximation result is formulated.

We prove the approximation result in Section 4, where we derive equations for the error terms produced by the leading-order approximation and handle these terms by using energy estimates and Gronwall's inequality. Many terms are miraculously given by the time derivative of the energy quantity due to the energy conservation (2).

Section 5 extends the approximation result for the wave propagation along an arbitrary direction. To keep the residual terms at the same small order, we have to use a higher order approximation leading to the system of the KP-II equation and the linearized KP-II equation as approximation equations. The requirements on smooth solutions of the KP-II and linearized KP-II equations in Sobolev spaces of higher regularity are stated as assumptions for the approximation result.

Section 6 discusses the spatial configurations for which these requirements can be satisfied, e.g. for transversely independent solutions and for periodic solutions, and formulates an open problem for existence of smooth decaying solutions in the unbounded domain.

2. Strain variables and the compatibility condition

For further work we introduce the following strain variables

$$u_{m,n} = q_{m+1,n} - q_{m,n}, \qquad v_{m,n} = q_{m,n+1} - q_{m,n}, \qquad w_{m,n} = \partial_t q_{m,n}. \tag{3}$$

The scalar FPU system (1) can be rewritten in the strain variables as the following system

$$\begin{cases} \partial_t u_{m,n} = w_{m+1,n} - w_{m,n}, \\ \partial_t v_{m,n} = w_{m,n+1} - w_{m,n}, \\ \partial_t w_{m,n} = W'(u_{m,n}) - W'(u_{m-1,n}) + W'(v_{m,n}) - W'(v_{m,n-1}). \end{cases} \tag{4}$$

Alternatively, the component $w_{m,n}$ can be eliminated and the FPU system (4) can be closed as two scalar equations

$$\begin{cases} \partial_t^2 u_{m,n} = W'(u_{m+1,n}) - 2W'(u_{m,n}) + W'(u_{m-1,n}) \\ \qquad + W'(v_{m+1,n}) - W'(v_{m+1,n-1}) - W'(v_{m,n}) + W'(v_{m,n-1}), \\ \partial_t^2 v_{m,n} = W'(v_{m,n+1}) - 2W'(v_{m,n}) + W'(v_{m,n-1}) \\ \qquad + W'(u_{m,n+1}) - W'(u_{m-1,n+1}) - W'(u_{m,n}) + W'(u_{m-1,n}), \end{cases} \tag{5}$$

where the u- and v-variables satisfy a certain compatibility condition due to their relation (3) to the q-variable.

In order to specify the compatibility condition and to develop the justification analysis of the KP-II approximation, we will work in Fourier space, similar to analysis in [19] of the FPU system on 1D lattice. Therefore, we define

$$\widehat{u}(k,l,t) = \frac{1}{2\pi} \sum_{(m,n) \in \mathbb{Z}^2} u_{m,n} e^{ikm+iln}, \qquad u_{m,n} = \frac{1}{2\pi} \iint_{\mathbb{T}^2} \widehat{u}(k,l,t) e^{-ikm-iln} dk dl,$$

and similarly for $v_{m,n}$ and $w_{m,n}$, where $\mathbb{T} := [-\pi, \pi)$ equipped with periodic boundary conditions. The first two equations of the system (4) shows that

$$\partial_t \widehat{u} = (e^{-ik} - 1)\widehat{w}, \quad \partial_t \widehat{v} = (e^{-il} - 1)\widehat{w},$$

which implies that the following compatibility condition is invariant with respect to the time evolution of the FPU system (4):

$$(e^{-ik} - 1)\widehat{v}(k, l, t) = (e^{-il} - 1)\widehat{u}(k, l, t), \quad t \geq 0. \tag{6}$$

Remark 2.1. In general, an arbitrary constant can be added to (6). We set this constant to 0 for the class of solutions we are interested in.

Next we rewrite system (5) with $W'(u) = u - u^2$ in the convenient Fourier form. To do so, let us first inspect the linearized system

$$\begin{cases} \partial_t^2 u_{m,n} = u_{m+1,n} - 2u_{m,n} + u_{m-1,n} + v_{m+1,n} - v_{m+1,n-1} - v_{m,n} + v_{m,n-1}, \\ \partial_t^2 v_{m,n} = v_{m,n+1} - 2v_{m,n} + v_{m,n-1} + u_{m,n+1} - u_{m-1,n+1} - u_{m,n} + u_{m-1,n}, \end{cases}$$

which is written in Fourier space as

$$\begin{cases} \partial_t^2 \widehat{u} = (e^{-ik} - 2 + e^{ik})\widehat{u} + (e^{-ik} - 1)(1 - e^{il})\widehat{v}, \\ \partial_t^2 \widehat{v} = (e^{-il} - 2 + e^{il})\widehat{v} + (e^{-il} - 1)(1 - e^{ik})\widehat{u}, \end{cases}$$

where we have used

$$e^{-ik} - e^{-ik}e^{il} - 1 + e^{il} = e^{-ik}(1 - e^{il}) - (1 - e^{il}) = (e^{-ik} - 1)(1 - e^{il}).$$

To simplify notation, we define

$$\omega_k^2 := 2 - e^{-ik} - e^{ik}, \qquad \omega_l^2 := 2 - e^{-il} - e^{il}.$$

Eliminating $\widehat{v}$ in the first equation of the linearized system yields the following linear equation

$$\partial_t^2 \widehat{u} + (\omega_k^2 + \omega_l^2)\widehat{u} = 0. \tag{7}$$

Extending exactly the same calculations as for the linearized system for $W'(u) = u - u^2$, we obtain the following nonlinear system in Fourier space given by

$$\begin{cases} \partial_t^2 \widehat{u} = -\omega_k^2(\widehat{u} - \widehat{u} * \widehat{u}) + (e^{-ik} - 1)(1 - e^{il})(\widehat{v} - \widehat{v} * \widehat{v}), \\ \partial_t^2 \widehat{v} = -\omega_l^2(\widehat{v} - \widehat{v} * \widehat{v}) + (e^{-il} - 1)(1 - e^{ik})(\widehat{u} - \widehat{u} * \widehat{u}). \end{cases}$$

By using the compatibility condition (6), we rewrite this system in the form:

$$\begin{cases} \partial_t^2 \widehat{u} = -(\omega_k^2 + \omega_l^2)\widehat{u} + \omega_k^2(\widehat{u} * \widehat{u}) - (e^{-ik} - 1)(1 - e^{il})(\widehat{v} * \widehat{v}), \\ \partial_t^2 \widehat{v} = -(\omega_k^2 + \omega_l^2)\widehat{v} + \omega_l^2(\widehat{v} * \widehat{v}) - (e^{-il} - 1)(1 - e^{ik})(\widehat{u} * \widehat{u}). \end{cases} \tag{8}$$

This system in combination with the compatibility condition (6) is the starting point for the derivation and justification of the KP-II equation.

Remark 2.2. By using the compatibility condition (6), the two equations in system (8) can be reduced to a single equation. However, this single equation contains multipliers which are singular with respect to k and l. Since we have to expand these multipliers with respect to k and l in the long wave limit $(k, l) \to (0, 0)$, it is advantageous to work with system (8) where the multipliers are smooth with respect to k and l.

3. PROPAGATION ALONG THE x-DIRECTION

Here we first derive the KP-II equation for the long modulated waves moving along the x-axis. After deriving the KP-II equation, we discuss how to handle the leading-order approximation in Fourier space and the residual terms of the FPU system. We end this section by formulating the approximation theorem, which will be proven in Section 4.

3.1. **The formal long-wave limit.** The long wave limit in physical space corresponds in Fourier space to an expansion of system (8) at the wave vector $(k, l) = (0, 0)$. Expansions

$$\omega_k^2 = k^2 - \frac{1}{12}k^4 + \mathcal{O}(k^6), \qquad 1 - e^{il} = -il + \mathcal{O}(l^2),$$

allows us to rewrite system (8) in physical space formally as

$$\partial_t^2 u = \partial_x^2 u + \partial_y^2 u + \frac{1}{12}\partial_x^4 u + \frac{1}{12}\partial_y^4 u - \partial_x^2 (u^2) - \partial_x \partial_y (v^2) + \text{h.o.t.}, \tag{9}$$

where h.o.t. stands for the higher-order terms. The compatibility condition (6) corresponds in physical space to

$$\partial_x v + \text{h.o.t.} = \partial_y u + \text{h.o.t..} \tag{10}$$

The leading-order approximation is given in physical space by

$$u_{m,n}(t) = \varepsilon^2 A(X, Y, T), \quad v_{m,n}(t) = \varepsilon^3 \partial_X^{-1} \partial_Y A(X, Y, T), \tag{11}$$

with

$$X = \varepsilon(m - t), \quad Y = \varepsilon^2 n, \quad T = \varepsilon^3 t,$$

where A is a suitable solution to the KP-II equation (12) below for which derivatives of A and $\partial_X^{-1}\partial_Y A$ are controlled in Sobolev spaces of sufficiently high regularity. The compatibility condition (10) rewritten in variables (X, Y, T) is satisfied at the order of $\mathcal{O}(\varepsilon^4)$. Substitution of (11) into (9) rewritten in variables (X, Y, T) results in the following KP-II equation at the order of $\mathcal{O}(\varepsilon^6)$:

$$2\partial_X \partial_T A + \partial_Y^2 A + \frac{1}{12}\partial_X^4 A - \partial_X^2 (A^2) = 0. \tag{12}$$

In what follows, we replace the formal approximation in physical space by the precise approximation in Fourier space.

3.2. **The leading-order approximation in Fourier space.** Our goal is to prove a statement of the following form. Let A be a suitable solution of the KP-II equation (12). Then for $\varepsilon > 0$ sufficiently small, there are solutions of system (5) which remain close to the leading-order approximation (11).

In order to establish such an approximation theorem, we have to estimate the residual terms first, i.e., we have to control the terms which do not cancel after inserting the approximation (11) into system (9) and (10). In general, these estimates can be obtained by expanding the multipliers in Fourier space and by assuming a certain regularity of the solutions of the KP-II equation (12). However, additional difficulties occur as we explain below.

(S1) A fundamental difficulty is the fact that the surface of wave frequencies $\omega := \sqrt{\omega_k^2 + \omega_l^2}$ of the linearized problem (7) forms a cone at the wave vector $(k, l) = (0, 0)$. Hence, a Taylor series expansion of the surface is not possible at the tip of the cone. A consequence of non-smoothness is the occurrence of the term $\partial_X^{-1}\partial_Y^2 A$ in the evolution problem

$$\partial_T A = -\frac{1}{2}\left(\partial_X^{-1}\partial_Y^2 A + \frac{1}{12}\partial_X^3 A - 2A\partial_X A\right), \tag{13}$$

which follows from the KP-II equation (12).

(S2) We partially get rid of the first difficulty (S1) by working with the extended system (8). However, the leading-order approximation for the extended system (8) has to satisfy the compatibility condition (6) from which the term $\partial_X^{-1}\partial_Y A$ appear in the approximation (11).

(S3) By looking at the evolution equation (13), it cannot be expected that the solutions of the KP-II equation are arbitrarily smooth. However, a certain smoothness of solutions is needed for estimating the residual terms.

(S4) Finally, solutions of the FPU system (5) live on $\mathbb{Z}^2$, whereas solutions of the KP-II equation (12) live on $\mathbb{R}^2$. In Fourier space, solutions of system (8) live on $\mathbb{T}^2$, whereas solutions of the Fourier-transformed KP-II equation live on $\mathbb{R}^2$,

$$\widehat{A}(\xi, \eta, T) = \frac{1}{2\pi} \iint_{\mathbb{R}^2} A(X, Y, T) e^{i\xi X + i\eta Y} dX dY$$

and

$$A(X, Y, T) = \frac{1}{2\pi} \iint_{\mathbb{R}^2} \widehat{A}(\xi, \eta, T) e^{-i\xi X - i\eta Y} d\xi d\eta.$$

To deal with the difficulties in (S1)–(S3), we use the following well-posedness result obtained in [14, Lemma 1] based on earlier work [9, 24].

Lemma 3.1. *For any $A_0 \in H^{s+9}(\mathbb{R}^2)$ such that $\partial_X^{-2} \partial_Y^2 A_0 \in H^{s+9}(\mathbb{R}^2)$ and*

$$\partial_X^{-1} \partial_Y^2 (\partial_X^{-2} \partial_Y^2 A_0 - A_0^2) \in H^{s+3}(\mathbb{R}^2)$$

with fixed $s \geq 0$, there exists $\tau_0 > 0$ such that the Cauchy problem (13) admits a unique solution

$$A \in C^0([-\tau_0, \tau_0], H^{s+9}) \cap C^1([-\tau_0, \tau_0], H^{s+6}) \cap C^2([-\tau_0, \tau_0], H^{s+3})$$

such that

$$\partial_X^{-1} \partial_Y A \in C^0([-\tau_0, \tau_0], H^{s+8}) \cap C^1([-\tau_0, \tau_0], H^{s+5}) \cap C^2([-\tau_0, \tau_0], H^{s+2})$$

and

$$\partial_X^{-2} \partial_Y^2 A \in C^0([-\tau_0, \tau_0], H^{s+6}) \cap C^1([-\tau_0, \tau_0], H^{s+3}).$$

To deal with the difficulty in (S4), we can use the following approximation result from [14, Lemma 2], which was obtained based on previous estimates in one dimension in [22, Lemma 3.9] and [5, Lemma 5.1].

Lemma 3.2. *Let $A \in C^0([-\tau_0, \tau_0], H^s(\mathbb{R}^2))$ with $s > 1$ and $a_{m,n}(t) := A(\varepsilon(m - t), \varepsilon^2 n, \varepsilon^3 t)$ for $(m, n) \in \mathbb{Z}^2$. Then there exists a constant $C_s > 0$ such that for all $\varepsilon \in (0, 1]$ we have*

$$\|a(t)\|_{\ell^2(\mathbb{Z}^2)} \leq C_s \varepsilon^{-\frac{3}{2}} \|A(\cdot, \cdot, \varepsilon^3 t)\|_{H^s(\mathbb{R}^2)}, \quad \forall t \in [-\varepsilon^{-3}\tau_0, \varepsilon^{-3}\tau_0].$$

Consequently, in Fourier space, we have

$$\|\widehat{a}(\cdot, \cdot, t)\|_{L^2(\mathbb{T}^2)} \leq C_s \varepsilon^{-\frac{3}{2}} \|\widehat{A}(\cdot, \cdot, \varepsilon^3 t)\|_{L^{2,s}(\mathbb{R}^2)}, \quad \forall t \in [-\varepsilon^{-3}\tau_0, \varepsilon^{-3}\tau_0],$$

where $\|\widehat{A}\|_{L^{2,s}(\mathbb{R}^2)} := \| < \cdot >^s \widehat{A}\|_{L^2(\mathbb{R}^2)}$ with $< x >:= \sqrt{1 + |x|^2}$.

Remark 3.3. The proof of Lemma 3.2 is well-known in the existing literature, cf. the book [21]. In Fourier space we lose ε^{-3} due to the Fourier transform of the scaled variables and gain $\varepsilon^{3/2}$ due to the scaling of the L^2-norm. This coincides with the estimates in physical space, where we lose a factor $\varepsilon^{-3/2}$ due to scaling. The bounds in Fourier and physical space agree to each other since Fourier transform is an isomorphism in L^2 spaces.

Finally, we make precise the leading-order approximation in Fourier space, which we denote as $(\widehat{u}, \widehat{v}) = \varepsilon^2 (\widehat{\psi}_u, \widehat{\psi}_v)$. Let $\widehat{A}$ be the Fourier transform of a smooth solution A of the KP-II equation (12) in Lemma 3.1. Let $\chi_{\mathbb{T}^2}$ be the characterstic function on $\mathbb{R}^2$ such that $\chi_{\mathbb{T}^2}(k, l) = 0$ for $(k, l) \notin \mathbb{T}^2$. Then, $\widehat{\psi}_u$ is defined by

$$\widehat{\psi}_u(k, l, t) = \varepsilon^{-3} e^{ikt} \chi_{\mathbb{T}^2}(k, l) \widehat{A}(\varepsilon^{-1} k, \varepsilon^{-2} l, T), \tag{14}$$

whereas $\widehat{\psi}_v$ is obtained from the compatibility relation

$$(e^{-ik} - 1)\widehat{\psi}_v(k, l, t) = (e^{-il} - 1)\widehat{\psi}_u(k, l, t). \tag{15}$$

Expanding as $(k, l) \to (0, 0)$ yields

$$\widehat{\psi}_v(k, l, t) = k^{-1}l\left[1 + \mathcal{O}(|k| + |l|)\right]\widehat{\psi}_u(k, l, t). \tag{16}$$

It is clear that the inverse Fourier transform of (14) and (16) does not recover the approximation (11) in physical space because of the approximation errors. However, the following lemma controls the difference between the approximations in Fourier and physical spaces. The lemma is based on [22, Lemma 3.8] in one dimension and can be proven similarly to Lemma 3.2 proven in [14].

Lemma 3.4. *Let* $A \in C^0([-\tau_0, \tau_0], H^s(\mathbb{R}^2))$ *with* $s > 1$, $a_{m,n}(t) := A(\varepsilon(m - t), \varepsilon^2 n, \varepsilon^3 t)$ *and* $\widehat{\psi}_u$ *be given by (14). Then, there exists a constant* $C_s > 0$ *such that for all* $\varepsilon \in (0, 1]$ *we have*

$$\|\psi_u(t) - a(t)\|_{\ell^2(\mathbb{Z}^2)} \leq C_s \varepsilon^{s - \frac{3}{2}} \|A(\cdot, \cdot, \varepsilon^3 t)\|_{H^s(\mathbb{R}^2)}, \quad \forall t \in [-\varepsilon^{-3}\tau_0, \varepsilon^{-3}\tau_0].$$

Remark 3.5. In view of Lemma 3.2, the difference between the approximation (11) in physical space and $(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v)$ with ψ_u and ψ_v given by the inverse Fourier transform of the approximation (14) and (16) is small.

3.3. **Estimates for the residual.** The residuals contain the terms which do not cancel after inserting the leading-order approximation (14) and (16) into system (8):

$$\begin{cases} \widehat{\mathrm{Res}}_u(u, v) := -\partial_t^2 \widehat{u}(\ell, t) - (\omega_k^2(k) + \omega_l^2(l))\widehat{u}(\ell, t) + \omega_k^2(k)(\widehat{u} * \widehat{u})(\ell, t) \\ \qquad\qquad -(e^{-ik} - 1)(1 - e^{il})(\widehat{v} * \widehat{v})(\ell, t), \\ \widehat{\mathrm{Res}}_v(u, v) := -\partial_t^2 \widehat{v}(\ell, t) - (\omega_k^2(k) + \omega_l^2(l))\widehat{v}(\ell, t) + \omega_l^2(l)(\widehat{v} * \widehat{v})(\ell, t) \\ \qquad\qquad -(e^{-il} - 1)(1 - e^{ik})(\widehat{u} * \widehat{u})(\ell, t). \end{cases}$$

Remark 3.6. The application of Lemmas 3.2 and 3.4 transfers the pure counting of powers of ε into rigorous estimates. Since this is well documented in the existing literature, cf. [21], we refrain from many details.

In physical space, it follows formally from (9) and (11) that

$$\mathrm{Res}_u(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v) = \mathcal{O}(\varepsilon^8 \partial_T^2 A, \varepsilon^8 \partial_X^6 A, \varepsilon^{10} \partial_Y^4 A, \varepsilon^8 \partial_X^4(A^2), \varepsilon^9 \partial_X \partial_Y (\partial_X^{-1} \partial_Y A)^2)$$

and similarly

$$\mathrm{Res}_v(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v) = \mathcal{O}(\varepsilon^9 \partial_X^{-1} \partial_Y \partial_T^2 A, \varepsilon^9 \partial_X^5 \partial_Y A, \varepsilon^{11} \partial_X^{-1} \partial_Y^5 A, \varepsilon^9 \partial_X^3 \partial_Y (A^2), \varepsilon^{10} \partial_Y^2 (\partial_X^{-1} \partial_Y A)^2)$$

Remark 3.7. The notations are to be understood in the following sense. The term $\mathcal{O}(\varepsilon^8 \partial_X^6 A)$ means that in Fourier space the scaled $\widehat{A}$ is multiplied by a function $f = f(k, l)$ satifying $|f(k, l)| \leq C\varepsilon^8 |k|^6$, where possibly different constants are denoted with the same symbol C if they can be chosen independent of the small perturbation parameter $0 < \varepsilon \ll 1$. This estimate is only relevant for small k and l since our equations in Fourier space are posed on the bounded domain $\mathbb{T}^2$.

The residual terms are controlled by the local well-posedness theory of Lemma 3.1 and by the error bounds of Lemma 3.2 when A is a smooth solution of the KP-II equation (13) for any $s \geq 0$. Since we apply ω_k^{-1} to $\widehat{\mathrm{Res}}_u(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v)$ and ω_l^{-1} to $\widehat{\mathrm{Res}}_v(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v)$, both terms in the physical space yield

$$\mathcal{O}(\varepsilon^7 \partial_X^{-1} \partial_T^2 A, \varepsilon^7 \partial_X^5 A, \varepsilon^9 \partial_X^{-1} \partial_Y^4 A, \varepsilon^7 \partial_X^3 (A^2), \varepsilon^8 \partial_Y (\partial_X^{-1} \partial_Y A)^2)$$

Hence, we lose a factor ε^{-1} due to the long wave character of ψ_u and ψ_v. By taking the L^2-norm we lose another factor $\varepsilon^{-3/2}$ due to the involved scalings and the scaling properties of the L^2-norm. Therefore, it follows from the formal order $\mathcal{O}(\varepsilon^8)$ of truncation that we have in the end

$$\|\omega_k^{-1}\widehat{\mathrm{Res}}_u(\varepsilon^2\psi_u, \varepsilon^2\psi_v)\|_{L^2} + \|\omega_l^{-1}\widehat{\mathrm{Res}}_v(\varepsilon^2\psi_u, \varepsilon^2\psi_v)\|_{L^2} = \mathcal{O}(\varepsilon^{\frac{11}{2}}). \tag{17}$$

This means that on the long $\mathcal{O}(1/\varepsilon^3)$-time scale we can choose the error to scale with a factor ε^β with $\beta = \frac{5}{2}$. The precise count of the residual terms with the same bound (17) was obtained in [14, Lemma 3].

Remark 3.8. The formal order $\mathcal{O}(\varepsilon^8)$ is sufficient in the subsequent energy estimates of Section 4 since many terms can be written as time-derivatives which allows us to include these terms in the chosen energy and allow for $\beta = \frac{5}{2}$. If the corresponding terms could not have been written as time-derivatives, then it would be necessary to obtain residual terms of formal order of $\mathcal{O}(\varepsilon^9)$, for which we would have to construct a higher order approximation to the leading-order approximation $(u, v) = (\varepsilon^2\psi_u, \varepsilon^2\psi_v)$.

Additional residual terms arise from the KP-II equation (12) rewritten in Fourier space and truncated on $\mathbb{T}$. The approximation error obtained from this truncation only appears for the nonlinear (quadratic) term of the KP-II equation. Since $\mathbb{T}^2$ is compact, it suffices to control this error by considering the convolution term for A^2 in Fourier space:

$$\chi_{\mathbb{T}^2}(\widehat{A} * \widehat{A}) - \chi_{\mathbb{T}^2}(\chi_{\mathbb{T}^2}\widehat{A} * \chi_{\mathbb{T}^2}\widehat{A})$$
$$= \chi_{\mathbb{T}^2}((\widehat{A} - \chi_{\mathbb{T}^2}\widehat{A}) * (\widehat{A} - \chi_{\mathbb{T}^2}\widehat{A})) + 2\chi_{\mathbb{T}^2}(\chi_{\mathbb{T}^2}\widehat{A} * (\widehat{A} - \chi_{\mathbb{T}^2}\widehat{A})).$$

Each term in the right-hand side is controlled by the application of Lemma 3.4 and the residual error in Fourier space is smaller than the bound (17) for any $s > 1$.

3.4. **The approximation result.** By using the energy estimates for the approximation error, we will prove in Section 4 the following main result for the horizontal propagation in the two-dimensional square lattice.

Theorem 3.9. *There exist C_0 and $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the following holds. Let $A \in C([0, \tau_0], H^{s+9})$ be a solution of the KP-II equation (12) given by Lemma 3.1 with fixed $s \geq 0$. Then there exist solutions (u, v) of system (5) with*

$$\sup_{t \in [0, \varepsilon^{-3}\tau_0]} \|u(t) - \varepsilon^2\psi_u(t)\|_{\ell^2(\mathbb{Z}^2)} + \|v(t) - \varepsilon^2\psi_v(t)\|_{\ell^2(\mathbb{Z}^2)} \leq C_0\varepsilon^{\frac{5}{2}},$$

where (ψ_u, ψ_v) are given by the inverse Fourier transform of (14) and (16).

Remark 3.10. The proof of the approximation result of Theorem 3.9 is a nontrivial task. The KP-II approximation and the associated solution are of order $\mathcal{O}(\varepsilon^2)$ for $\varepsilon \to 0$. Therefore, a simple application of Gronwall's inequality would only provide the boundedness of the solutions on an $\mathcal{O}(\varepsilon^{-2})$-time scale, but not on the natural $\mathcal{O}(\varepsilon^{-3})$-time scale of the KP approximation. There exist many counterexamples where formally derived amplitude equations make wrong predictions about the dynamics of original systems on the natural time scale of the amplitude equations, cf. [18] and recent results in [2, 12, 20].

4. Energy estimates for the approximation error

The approximation error is defined by

$$\varepsilon^\beta(\widehat{R}_u, \widehat{R}_v) := (\widehat{u}, \widehat{v}) - (\varepsilon^2\widehat{\psi}_u, \varepsilon^2\widehat{\psi}_v) \tag{18}$$

with β being suitably chosen as $\beta = \frac{5}{2}$, the approximation $(\varepsilon^2\widehat{\psi}_u, \varepsilon^2\widehat{\psi}_v)$ satisfying the compatibility condition (15), and the error terms $(\widehat{R}_u, \widehat{R}_v)$ satisfying the compatibility condition

$$(e^{-ik} - 1)\widehat{R}_v(k, l, t) = (e^{-il} - 1)\widehat{R}_u(k, l, t). \tag{19}$$

The error terms satisfy equations of motion given by

$$\begin{aligned}
\partial_t^2 \widehat{R}_u = &-(\omega_k^2 + \omega_l^2)\widehat{R}_u + 2\varepsilon^2\omega_k^2(\widehat{\psi}_u * \widehat{R}_u) + \varepsilon^\beta\omega_k^2(\widehat{R}_u * \widehat{R}_u) \\
&- 2\varepsilon^2(e^{-ik} - 1)(1 - e^{il})(\widehat{\psi}_v * \widehat{R}_v) - \varepsilon^\beta(e^{-ik} - 1)(1 - e^{il})(\widehat{R}_v * \widehat{R}_v) \\
&+ \varepsilon^{-\beta}\widehat{\mathrm{Res}}_u(\varepsilon^2\psi_u, \varepsilon^2\psi_v),
\end{aligned} \tag{20}$$

$$\begin{aligned}
\partial_t^2 \widehat{R}_v = &-(\omega_k^2 + \omega_l^2)\widehat{R}_v + 2\varepsilon^2\omega_l^2(\widehat{\psi}_v * \widehat{R}_v) + \varepsilon^\beta\omega_l^2(\widehat{R}_v * \widehat{R}_v) \\
&- 2\varepsilon^2(e^{-il} - 1)(1 - e^{ik})(\widehat{\psi}_u * \widehat{R}_u) - \varepsilon^\beta(e^{-il} - 1)(1 - e^{ik})(\widehat{R}_u * \widehat{R}_u) \\
&+ \varepsilon^{-\beta}\widehat{\mathrm{Res}}_v(\varepsilon^2\psi_u, \varepsilon^2\psi_v).
\end{aligned} \tag{21}$$

To progress further, we recall the conserved energy (2) of the FPU system (1), which suggests that the energy for the error terms in physical space can be defined in the form

$$E = E_0 + \varepsilon^2 E_1 + \varepsilon^\beta E_2$$

with

$$E_0 = \|R_w\|_{\ell^2}^2 + \|R_u\|_{\ell^2}^2 + \|R_v\|_{\ell^2}^2, \tag{22}$$

$$E_1 = -2 \sum_{(m,n)\in\mathbb{Z}^2} (\psi_u)_{m,n}(R_u^2)_{m,n} + (\psi_v)_{m,n}(R_v^2)_{m,n}, \tag{23}$$

and

$$E_2 = -\frac{2}{3} \sum_{(m,n)\in\mathbb{Z}^2} (R_u^3)_{m,n} + (R_v^3)_{m,n}, \tag{24}$$

where R_w is the error term for the third strain variable in (3) and the total energy (2) is is multiplied by a factor of 2 for convenience. We define the L^2-scalar product $(\cdot, \cdot)$ by

$$(\widehat{f}, \widehat{g}) = \int_{\mathbb{T}^2} \overline{\widehat{f}(\ell)}\widehat{g}(\ell)d\ell,$$

where $\ell := (k, l)$. By using Parseval's equality, we have $\sum_{(m,n)\in\mathbb{Z}^2} f_{m,n}g_{m,n} = (\widehat{f}, \widehat{g})$. By using Fourier transform, the first two linear equations in system (4), and the compatibility relation (19), we can rewrite the leading-order energy in the equivalent form

$$E_0 = \frac{1}{2}\|\omega_k^{-1}\partial_t\widehat{R}_u\|_{L^2}^2 + \frac{1}{2}\|\omega_l^{-1}\partial_t\widehat{R}_v\|_{L^2}^2 + \frac{1}{2}\|\omega_k^{-1}\omega\widehat{R}_u\|_{L^2}^2 + \frac{1}{2}\|\omega_l^{-1}\omega\widehat{R}_v\|_{L^2}^2, \tag{25}$$

where $\omega := \sqrt{\omega_k^2 + \omega_l^2}$. Indeed, the first term in (22) after the Fourier transform is split symmetrically by using the first two linear equations in system (4) as

$$\|\widehat{R}_w\|_{L^2}^2 = \frac{1}{2}\|\omega_k^{-1}\partial_t\widehat{R}_u\|_{L^2}^2 + \frac{1}{2}\|\omega_l^{-1}\partial_t\widehat{R}_v\|_{L^2}^2,$$

whereas the second and third terms in (22) after the Fourier transform are rewritten as

$$\|\widehat{R}_u\|_{L^2}^2 + \|\widehat{R}_v\|_{L^2}^2 = \frac{1}{2}\|\omega_k^{-1}\omega\widehat{R}_u\|_{L^2}^2 + \frac{1}{2}\|\omega_l^{-1}\omega\widehat{R}_v\|_{L^2}^2,$$

since the compatibility relation (19) suggests that

$$\omega^2(\omega_l^2|\widehat{R}_u|^2 + \omega_k^2|\widehat{R}_v|^2) = 2\omega_k^2\omega_l^2(|\widehat{R}_u|^2 + |\widehat{R}_v|^2).$$

Similarly, we rewrite E_1 and E_2 after Fourier transform as

$$E_1 := -2\int \overline{\widehat{R}_u(\ell)}\widehat{\psi}_u(\ell-\ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell - 2\int \overline{\widehat{R}_v(\ell)}\widehat{\psi}_v(\ell-\ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell \qquad (26)$$

and

$$E_2 = -\frac{2}{3}\int \overline{\widehat{R}_u(\ell)}\widehat{R}_u(\ell-\ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell - \frac{2}{3}\int \overline{\widehat{R}_v(\ell)}\widehat{R}_v(\ell-\ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell. \qquad (27)$$

The leading-order energy E_0 in (25) suggests that the energy estimates for the system (20) and (21) are obtained by multiplying the first equation by the weighted time derivative $\omega_k^{-2}\overline{\partial_t\widehat{R}_u}$ and the second equation by the weighted time derivative $\omega_l^{-2}\overline{\partial_t\widehat{R}_v}$, after which we add the two equations and integrate in $\ell = (k,l)$. This procedure gives us the following energy balance equation

$$\frac{1}{2}\partial_t\|\omega_k^{-1}\partial_t\widehat{R}_u\|_{L^2}^2 + \frac{1}{2}\partial_t\|\omega_l^{-1}\partial_t\widehat{R}_v\|_{L^2}^2 = \mathrm{Re}(s_1 + s_2 + \ldots + s_{12}), \qquad (28)$$

with

$$s_1 = -\int \overline{\partial_t\widehat{R}_u(\ell)}\omega_k^{-2}\omega^2\widehat{R}_u(\ell)d\ell,$$

$$s_2 = 2\varepsilon^2\int \overline{\partial_t\widehat{R}_u(\ell)}\widehat{\psi}_u(\ell-\ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell,$$

$$s_3 = \varepsilon^\beta\int \overline{\partial_t\widehat{R}_u(\ell)}\widehat{R}_u(\ell-\ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell,$$

$$s_4 = -2\varepsilon^2\int \overline{\partial_t\widehat{R}_u(\ell)}(e^{-ik}-1)(1-e^{il})\omega_k^{-2}\widehat{\psi}_v(\ell-\ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell,$$

$$s_5 = -\varepsilon^\beta\int \overline{\partial_t\widehat{R}_u(\ell)}(e^{-ik}-1)(1-e^{il})\omega_k^{-2}\widehat{R}_v(\ell-\ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell,$$

$$s_6 = \varepsilon^{-\beta}\int \overline{\partial_t\widehat{R}_u(\ell)}\omega_k^{-2}\widehat{\mathrm{Res}}_u(\varepsilon^2\psi_u,\varepsilon^2\psi_v)(\ell)d\ell,$$

$$s_7 = -\int \overline{\partial_t\widehat{R}_v(\ell)}\omega_l^{-2}\omega^2\widehat{R}_v(\ell)d\ell,$$

$$s_8 = 2\varepsilon^2\int \overline{\partial_t\widehat{R}_v(\ell)}\widehat{\psi}_v(\ell-\ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell,$$

$$s_9 = \varepsilon^\beta\int \overline{\partial_t\widehat{R}_v(\ell)}\widehat{R}_v(\ell-\ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell,$$

$$s_{10} = -2\varepsilon^2\int \overline{\partial_t\widehat{R}_v(\ell)}(e^{-il}-1)(1-e^{ik})\omega_l^{-2}\widehat{\psi}_u(\ell-\ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell,$$

$$s_{11} = -\varepsilon^\beta \int \overline{\partial_t \widehat{R}_v(\ell)}(e^{-il} - 1)(1 - e^{ik})\omega_l^{-2}\widehat{R}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell,$$

$$s_{12} = \varepsilon^{-\beta} \int \overline{\partial_t \widehat{R}_v(\ell)}\omega_l^{-2}\widehat{\mathrm{Res}}_v(\varepsilon^2\psi_u, \varepsilon^2\psi_v)(\ell)d\ell.$$

We can now deal with different terms of the energy balance equation (28) as follows.

i) We rewrite s_1 and s_7 as

$$s_1 = -\frac{1}{2}\partial_t \int \overline{\widehat{R}_u(\ell)}\omega_k^{-2}(k)\omega^2(\ell)\widehat{R}_u(\ell)d\ell = -\frac{1}{2}\partial_t \|\omega_k^{-1}\omega\widehat{R}_u\|_{L^2}^2$$

and

$$s_7 = -\frac{1}{2}\partial_t \int \overline{\widehat{R}_v(\ell)}\omega_l^{-2}(l)\omega^2(\ell)\widehat{R}_v(\ell)d\ell - \frac{1}{2}\partial_t \|\omega_l^{-1}\omega\widehat{R}_v\|_{L^2}^2.$$

These terms give the time derivative of the third and fourth terms in the leading-order energy E_0 in (25).

ii) We rewrite s_2 and s_8 as

$$s_2 = \varepsilon^2\partial_t \int \overline{\widehat{R}_u(\ell)}\widehat{\psi}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell - \varepsilon^2 \int \overline{\widehat{R}_u(\ell)}\partial_t\widehat{\psi}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell,$$

and

$$s_8 = \varepsilon^2\partial_t \int \overline{\widehat{R}_v(\ell)}\widehat{\psi}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell - \varepsilon^2 \int \overline{\widehat{R}_v(\ell)}\partial_t\widehat{\psi}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell,$$

The first terms in s_2 and s_8 define one half of the ε^2 correction E_1 to the leading-order energy E_0 given by (26). The other half will come from the terms s_4 and s_{10}. The second term in s_2 is estimated as follows:

$$\varepsilon^2 \left| \int \overline{\widehat{R}_u(\ell)}\partial_t\widehat{\psi}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell \right| \leq \|\widehat{R}_u\|_{L^2}\|\partial_t\widehat{\psi}_u * \widehat{R}_u\|_{L^2}$$

$$\leq \|\partial_t\widehat{\psi}_u\|_{L^1}\|\widehat{R}_u\|_{L^2}^2,$$

due to the Cauchy-Schwarz and Young's inequality. Since

$$\partial_t\widehat{\psi}_u(k, l, t) = ik\varepsilon^{-3}e^{ikt}\chi_{\mathbb{T}^2}(k, l)\widehat{A}(\varepsilon^{-1}k, \varepsilon^{-2}l, T) + e^{ikt}\chi_{\mathbb{T}^2}(k, l)\partial_T\widehat{A}(\varepsilon^{-1}k, \varepsilon^{-2}l, T),$$

we obtain

$$\|\partial_t\widehat{\psi}_u\|_{L^1} \leq \varepsilon \| |\cdot|\widehat{A}(\cdot, \cdot, T)\|_{L^1} + \varepsilon^3 \|\partial_T\widehat{A}(\cdot, \cdot, T)\|_{L^1}.$$

If A is controlled in Sobolev spaces of high regularity with the help of Lemma 3.1, then

$$\varepsilon^2 \left| \int \overline{\widehat{R}_u(\ell)}\partial_t\widehat{\psi}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell \right| \leq C\varepsilon^3\|\widehat{R}_u\|_{L^2}^2.$$

Similarly, we obtain from (16) that

$$\varepsilon^2 \left| \int \overline{\widehat{R}_v(\ell)}\partial_t\widehat{\psi}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell \right| \leq C\varepsilon^4\|\widehat{R}_v\|_{L^2}^2,$$

where the additional power in ε^4 compared to ε^3 is explained by additional power of ε between ψ_v and ψ_u, cf. (11).

iii) We rewrite s_4 as

$$s_4 = 2\varepsilon^2 \int \overline{(e^{-il} - 1)\omega_k^{-1}\partial_t \widehat{R}_u(\ell)}(e^{-ik} - 1)\omega_k^{-1}\widehat{\psi}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell$$

$$= 2\varepsilon^2 \int \overline{(e^{-ik} - 1)\omega_k^{-1}\partial_t \widehat{R}_v(\ell)}(e^{-ik} - 1)\omega_k^{-1}\widehat{\psi}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell$$

$$= 2\varepsilon^2 \int \overline{\partial_t \widehat{R}_v(\ell)}\widehat{\psi}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell$$

$$= \varepsilon^2\partial_t \int \overline{\widehat{R}_v(\ell)}\widehat{\psi}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell - \varepsilon^2 \int \overline{\widehat{R}_v(\ell)}\partial_t \widehat{\psi}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell,$$

and similarly s_{10} as

$$s_{10} = \varepsilon^2\partial_t \int \overline{\widehat{R}_u(\ell)}\widehat{\psi}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell - \varepsilon^2 \int \overline{\widehat{R}_u(\ell)}\partial_t \widehat{\psi}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell.$$

Since $s_4 + s_{10} = s_2 + s_8$, the first terms in s_4 and s_{10} define the other half of $\varepsilon^2 E_1$, where E_1 is given by (26), where the second terms in s_4 and s_{10} have been estimated in (ii).

iv) We rewrite s_3 and s_9 as

$$s_3 = \frac{1}{3}\varepsilon^\beta\partial_t \int \overline{\widehat{R}_u(\ell)}\widehat{R}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell$$

and

$$s_9 = \frac{1}{3}\varepsilon^\beta\partial_t \int \overline{\widehat{R}_v(\ell)}\widehat{R}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell.$$

These terms define one half of the ε^β correction E_2 to the leading-order energy E_0 given by (27). The other half will come from the terms s_5 and s_{11}.

v) We rewrite s_5 as

$$s_5 = 2\varepsilon^\beta \int \overline{(e^{-il} - 1)\omega_k^{-1}\partial_t \widehat{R}_u(\ell)}(e^{-ik} - 1)\omega_k^{-1}\widehat{R}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell$$

$$= 2\varepsilon^\beta \int \overline{(e^{-ik} - 1)\omega_k^{-1}\partial_t \widehat{R}_v(\ell)}(e^{-ik} - 1)\omega_k^{-1}\widehat{R}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell$$

$$= 2\varepsilon^\beta \int \overline{\partial_t \widehat{R}_v(\ell)}\widehat{R}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell$$

$$= \frac{2}{3}\varepsilon^\beta\partial_t \int \overline{\widehat{R}_v(\ell)}\widehat{R}_v(\ell - \ell_1)\widehat{R}_v(\ell_1)d\ell_1 d\ell,$$

and similarly s_{11} as

$$s_{11} = \frac{2}{3}\varepsilon^\beta\partial_t \int \overline{\widehat{R}_u(\ell)}\widehat{R}_u(\ell - \ell_1)\widehat{R}_u(\ell_1)d\ell_1 d\ell.$$

Since $s_5 + s_{11} = s_3 + s_9$, the corresponding terms define the other half of $\varepsilon^\beta E_2$, where E_2 is given by (27).

vi) The residual terms s_6 and s_{12} are estimated with the Cauchy-Schwarz and Young inequalities as

$$|s_6| = \varepsilon^{-\beta} \left| \int \overline{\omega_k^{-1} \partial_t \widehat{R}_u(\ell)} \omega_k^{-1} \widehat{\text{Res}}_u(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v)(\ell) d\ell \right|$$

$$\leq \varepsilon^{-\beta} \|\omega_k^{-1} \partial_t \widehat{R}_u\|_{L^2} \|\omega_k^{-1} \widehat{\text{Res}}_u(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v)\|_{L^2}$$

$$\leq \varepsilon^3 \|\omega_k^{-1} \partial_t \widehat{R}_u\|_{L^2}^2 + (\varepsilon^{-\beta-\frac{3}{2}} \|\omega_k^{-1} \widehat{\text{Res}}_u(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v)\|_{L^2})^2,$$

and similarly,

$$|s_{12}| \leq \varepsilon^3 \|\omega_l^{-1} \partial_t \widehat{R}_v\|_{L^2}^2 + (\varepsilon^{-\beta-\frac{3}{2}} \|\omega_l^{-1} \widehat{\text{Res}}_v(\varepsilon^2 \psi_u, \varepsilon^2 \psi_v).\|_{L^2})^2$$

By using the estimate (17) and setting $\beta = \frac{5}{2}$, we finally obtain

$$|s_6| + |s_{12}| \leq \varepsilon^3 \left(\|\omega_k^{-1} \partial_t \widehat{R}_u\|_{L^2}^2 + \|\omega_l^{-1} \partial_t \widehat{R}_v\|_{L^2}^2 + C_{\text{res}} \right),$$

for some constant $C_{\text{res}} > 0$ that depends on the solution A of the KP-II equation (12).

Combining all estimates together, we have derived the energy balance equation in the form

$$\frac{d}{dt} E \leq C_0 \varepsilon^3 E_0 + C_{\text{res}} \varepsilon^3,$$

for some constant $C_0 > 0$ that also depends on the solution A of the KP-II equation (12), where $E = E_0 + \varepsilon^2 E_1 + \varepsilon^\beta E_2$ and $\beta = \frac{5}{2}$. The corrections of the leading-order energy E_0 are controlled by

$$|E_1| \leq 2(\|\widehat{\psi}_u\|_{L^1} + \|\widehat{\psi}_v\|_{L^1}) E_0,$$

$$|E_2| \leq \frac{2}{3}(\|\widehat{R}_u\|_{L^1} + \|\widehat{R}_v\|_{L^1}) E_0,$$

where

$$\|\widehat{\psi}_u\|_{L^1} + \|\widehat{\psi}_v\|_{L^1} \leq \sqrt{2\pi} \left(\|\widehat{\psi}_u\|_{L^2} + \|\widehat{\psi}_v\|_{L^2} \right) \leq C_A,$$

$$\|\widehat{R}_u\|_{L^1} + \|\widehat{R}_v\|_{L^1} \leq \sqrt{2\pi} \left(\|\widehat{R}_u\|_{L^2} + \|\widehat{R}_v\|_{L^2} \right) \leq 2\sqrt{2\pi E_0},$$

with $C_A > 0$ that depends on the solution A of the KP-II equation (12). As long as there exists $M < \infty$ such that $E_0 \leq M$, there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the energy $E = E_0 + \varepsilon^2 E_1 + \varepsilon^\beta E_2$ is equivalent to the leading-order energy E_0, e.g.,

$$E_0 \leq 2E \leq 4E_0. \tag{29}$$

Hence, the energy balance equation can be written as

$$\frac{d}{dt} E \leq C \varepsilon^3 E + C_{\text{res}} \varepsilon^3,$$

which yields by using Gronwall's inequality for all $t \in [0, \varepsilon^{-3} \tau_0]$ that

$$E(t) \leq C_{\text{res}} \varepsilon^3 t e^{C\varepsilon^3 t} \leq C_{\text{res}} \tau_0 e^{C\tau_0} =: \frac{M}{2}.$$

In view of the equivalence (29), this verifies that $E_0(t) \leq M$ for all $t \in [0, \varepsilon^{-3} \tau_0]$. Due to the scaling (18) with $\beta = \frac{5}{2}$, this bound completes the proof of Theorem 3.9.

5. Propagation along an arbitrary direction

Here we consider an arbitrary angle of propagation with respect to the square lattice $\mathbb{Z}^2$ and derive the extended KP-II equation as the leading-order approximation. This extended KP-II equation is needed to reduce the size of the residual terms and it can be split into the sum of the KP-II equation for the main term and the linearized KP-II equation for the correction term. The approximation theorem is formulated for the smooth solutions to the KP-II and linearized KP-II equations. Classes of such smooth solutions are discussed in Section 6.

5.1. **The formal long-wave limit.** We take the advantage that the Laplacian is invariant under the rotation in the plane $\mathbb{R}^2$ and define the leading-order approximation in physical space by

$$u_{m,n}(t) = \varepsilon^2 A(X, Y, T), \qquad v_{m,n}(t) = \varepsilon^2 B(X, Y, T), \tag{30}$$

where

$$X = \varepsilon((\cos\phi)m + (\sin\phi)n - t), \quad Y = \varepsilon^2(-(\sin\phi)m + (\cos\phi)n), \quad T = \varepsilon^3 t,$$

and the angle of propagation $\phi \in (0, \frac{\pi}{2})$ determines the direction of propagation $(\cos\phi, \sin\phi)$ with respect to the square lattice $\mathbb{Z}^2$. The long wave limit can be written in the extended form compared to system (9) and (10):

$$\partial_t^2 u = \partial_x^2 u + \partial_y^2 u + \frac{1}{12}\partial_x^4 u + \frac{1}{12}\partial_y^4 u - \partial_x^2(u^2) - \partial_x\partial_y(v^2) - \frac{1}{2}(\partial_x - \partial_y)\partial_x\partial_y(v^2) + \text{h.o.t.} \tag{31}$$

and

$$\partial_x v + \frac{1}{2}\partial_x^2 v + \text{h.o.t.} = \partial_y u + \frac{1}{2}\partial_y^2 u + \text{h.o.t..} \tag{32}$$

We find by the chain rule

$$\partial_t^2 = \varepsilon^2 \partial_X^2 - 2\varepsilon^4 \partial_X\partial_T + \varepsilon^6 \partial_T^2,$$
$$\partial_x^2 = \varepsilon^2(\cos\phi)^2\partial_X^2 - 2\varepsilon^3(\cos\phi)(\sin\phi)\partial_X\partial_Y + \varepsilon^4(\sin\phi)^2\partial_Y^2,$$
$$\partial_y^2 = \varepsilon^2(\sin\phi)^2\partial_X^2 + 2\varepsilon^3(\cos\phi)(\sin\phi)\partial_X\partial_Y + \varepsilon^4(\cos\phi)^2\partial_Y^2,$$
$$\partial_x\partial_y = \varepsilon^2(\cos\phi)(\sin\phi)\partial_X^2 + 2\varepsilon^3((\cos\phi)^2 - (\sin\phi)^2)\partial_X\partial_Y - \varepsilon^4(\cos\phi)(\sin\phi)\partial_Y^2.$$

All terms up to the formal order of $\mathcal{O}(\varepsilon^5)$ cancel out, whereas the KP-II equation appears at the formal order of $\mathcal{O}(\varepsilon^6)$. For the propagation in the x-direction, there are no terms of the formal order of $\mathcal{O}(\varepsilon^7)$. If A satisfies the KP-II equation, the residual terms have the formal order of $\mathcal{O}(\varepsilon^8)$. This is no longer the case for the propagation along an arbitrary direction with $\phi \in (0, \frac{\pi}{2})$.

Substituting (30) into (31) and (32) and removing the terms of the formal order of $\mathcal{O}(\varepsilon^6)$ and $\mathcal{O}(\varepsilon^7)$ yield the extended KP-II equation

$$\begin{aligned}
-2\partial_X\partial_T A = &\frac{1}{12}[(\cos\phi)^4 + (\sin\phi)^4]\partial_X^4 A + \partial_Y^2 A \\
&- (\cos\phi)^2\partial_X^2(A^2) - (\sin\phi)(\cos\phi))\partial_X^2(B^2) \\
&- \frac{1}{3}\varepsilon[(\cos\phi)^2 - (\sin\phi)^2](\cos\phi)(\sin\phi)\partial_X^3\partial_Y A + 2\varepsilon(\cos\phi)(\sin\phi)\partial_X\partial_Y(A^2) \\
&- \varepsilon[(\cos\phi)^2 - (\sin\phi)^2]\partial_X\partial_Y(B^2) \\
&- \frac{1}{2}\varepsilon[\cos\phi - \sin\phi](\cos\phi)(\sin\phi)\partial_X^3(B^2).
\end{aligned} \tag{33}$$

and the relation between the amplitudes A and B:

$$(\cos\phi)\partial_X B - \varepsilon(\sin\phi)\partial_Y B + \frac{1}{2}\varepsilon(\cos\phi)^2\partial_X^2 B = (\sin\phi)\partial_X A + \varepsilon(\cos\phi)\partial_Y A + \frac{1}{2}\varepsilon(\sin\phi)^2\partial_X^2 A.$$

Writing

$$A = A_1 + \varepsilon A_2,$$
$$B = B_1 + \varepsilon B_2,$$

we obtain

$$B_1 = (\tan\phi)A_1,$$
$$B_2 = (\tan\phi)A_2 + \frac{1}{(\cos\phi)^2}\partial_X^{-1}\partial_Y A_1 + \frac{1}{2}(\tan\phi)[\sin\phi - \cos\phi]\partial_X A_1,$$

after which the extended KP-II equation (33) can be split into the KP-II equation for A_1 and the linearized KP-II equation for A_2, which are given by

$$-2\partial_X\partial_T A_1 = \frac{1}{12}[(\cos\phi)^4 + (\sin\phi)^4]\partial_X^4 A_1 + \partial_Y^2 A_1$$
$$- [(\cos\phi)^2 + (\sin\phi)^2(\tan\phi)]\partial_X^2(A_1^2) \tag{34}$$

and

$$-2\partial_X\partial_T A_2 = \frac{1}{12}[(\cos\phi)^4 + (\sin\phi)^4]\partial_X^4 A_2 + \partial_Y^2 A_2$$
$$- 2[(\cos\phi)^2 + (\sin\phi)^2(\tan\phi)]\partial_X^2(A_1 A_2)$$
$$- \frac{1}{3}[(\cos\phi)^2 - (\sin\phi)^2](\cos\phi)(\sin\phi)\partial_X^3\partial_Y A_1$$
$$- 2(\tan\phi)^2\partial_X^2(A_1\partial_X^{-1}\partial_Y A_1)$$
$$+ [(\sin\phi)^2(\tan\phi)^2 - (\sin\phi)^2 + 2(\sin\phi)(\cos\phi)]\partial_X\partial_Y(A_1^2). \tag{35}$$

5.2. **Estimates for the residual.** The leading-order approximation in Fourier space is denoted as before by $(\widehat{u}, \widehat{v}) = (\varepsilon^2\widehat{\psi_u}, \varepsilon^2\widehat{\psi_v})$ with

$$\widehat{\psi_u}(k, l, t) = \varepsilon^{-3}e^{i\varkappa t}\chi_{\mathbb{T}^2}(k, l)\widehat{A}(\varepsilon^{-1}\varkappa, \varepsilon^{-2}\vartheta, T), \tag{36}$$

and

$$\widehat{\psi_v}(k, l, t) = \varepsilon^{-3}e^{i\varkappa t}\chi_{\mathbb{T}^2}(k, l)\widehat{B}(\varepsilon^{-1}\varkappa, \varepsilon^{-2}\vartheta, T), \tag{37}$$

where

$$\varkappa = (\cos\phi)k + (\sin\phi)l, \qquad \vartheta = -(\sin\phi)k + (\cos\phi)l,$$

and

$$\widehat{A} = \widehat{A}_1 + \varepsilon\widehat{A}_2, \qquad \widehat{B} = \widehat{B}_1 + \varepsilon\widehat{B}_2.$$

For $\phi \in (0, \frac{\pi}{2})$, the residual terms are formally given in physical space by

$$\mathrm{Res}_u(\varepsilon^2\psi_u, \varepsilon^2\psi_v) = \mathcal{O}(\varepsilon^8\partial_T^2 A_1, \varepsilon^8\partial_X^6 A_1, \varepsilon^8\partial_X^2\partial_Y^2 A_1, \varepsilon^8\partial_X^4(A_1^2), \varepsilon^8\partial_Y^2(A_1^2), \varepsilon^8\partial_Y^2(B_1^2),$$
$$\varepsilon^9\partial_T^2 A_2, \varepsilon^9\partial_X^6 A_2, \varepsilon^9\partial_X^2\partial_Y^2 A_2, \varepsilon^8\partial_X^2(A_2^2), \varepsilon^8\partial_X\partial_Y(A_1 A_2),$$
$$\varepsilon^8\partial_X^2(B_2^2), \varepsilon^8\partial_X\partial_Y(B_1 B_2), \varepsilon^8\partial_X^3(B_1^2))$$

and

$$\text{Res}_v(\varepsilon^2\psi_u, \varepsilon^2\psi_v) = \mathcal{O}(\varepsilon^8\partial_T^2 B_1, \varepsilon^8\partial_X^6 B_1, \varepsilon^8\partial_X^2\partial_Y^2 B_1, \varepsilon^8\partial_X^4(B_1^2), \varepsilon^8\partial_Y^2(B_1^2), \varepsilon^8\partial_Y^2(A_1^2),$$
$$\varepsilon^9\partial_T^2 B_2, \varepsilon^9\partial_X^6 B_2, \varepsilon^9\partial_X^2\partial_Y^2 B_2, \varepsilon^8\partial_X^2(B_2^2), \varepsilon^8\partial_X\partial_Y(B_1 B_2),$$
$$\varepsilon^8\partial_X^2(A_2^2), \varepsilon^8\partial_X\partial_Y(A_1 A_2), \varepsilon^8\partial_X^3(A_1^2)).$$

The residual terms containing A_1 are controlled by the local well-posedness theory of Lemma 3.1 with A_1 being a smooth solution of the KP-II equation (34). If A_2 enjoys the same properties, the bound (17) is justified, from which the proof of the approximation theorem stated below is analogous to the proof of Theorem 3.9.

Theorem 5.1. *There exist C_0 and $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ the following holds. Let $A_1 \in C([0, \tau_0], H^{s+9})$ be a solution of the KP-II equation (34) given by Lemma 3.1 with fixed $s \geq 0$ and assume that $A_2 \in C([0, \tau_0], H^{s+9})$ is a solution of the linearized KP-II equation (35) with the same properties as for A_1. Then there exist solutions (u, v) of system (5) with*

$$\sup_{t \in [0, \varepsilon^{-3}\tau_0]} \|u(t) - \varepsilon^2\psi_u(t)\|_{\ell^2(\mathbb{Z}^2)} + \|v(t) - \varepsilon^2\psi_v(t)\|_{\ell^2(\mathbb{Z}^2)} \leq C_0\varepsilon^{\frac{5}{2}},$$

where (ψ_u, ψ_v) are given by the inverse Fourier transform of (36) and (37).

6. Discussion

Here we discuss classes of solutions to the KP-II equation (34) and the linearized KP-II equation (35) for which the approximation result of Theorem 5.1 can be obtained. This includes transversely independent solutions, periodic solutions, and decaying solutions in the unbounded domain.

6.1. **Transversely independent solutions.** Let $A_{1,2} = A_{1,2}(X, T)$ be Y-independent. Then, A_1 is a solution of the KdV equation

$$-2\partial_T A_1 = \frac{1}{12}[(\cos\phi)^4 + (\sin\phi)^4]\partial_X^3 A_1 - [(\cos\phi)^2 + (\sin\phi)^2(\tan\phi)]\partial_X(A_1^2) \quad (38)$$

and A_2 is a solution of the linearized KdV equation

$$-2\partial_T A_2 = \frac{1}{12}[(\cos\phi)^4 + (\sin\phi)^4]\partial_X^3 A_2 - 2[(\cos\phi)^2 + (\sin\phi)^2(\tan\phi)]\partial_X(A_1 A_2). \quad (39)$$

Clearly, if $A_2|_{T=0} = 0$, then $A_2(X, T) \equiv 0$. Smooth solutions for A_1 to the KdV equation (38) exist in one-dimensional Sobolev spaces without any additional constraints on the initial data $A_1|_{T=0} \in H^s(\mathbb{R})$. As a result, the approximation result of Theorem 5.1 holds in one-dimensional Sobolev spaces $H^s(\mathbb{R})$ for every $s \geq 6$.

6.2. **Periodic solutions.** Let $A_{1,2}$ be spatially periodic in X and Y. Without loss of generality, we assume that the solutions are 1-periodic in X and Y. Such solutions can be expressed in the Fourier form, e.g.

$$A(X, Y, T) = \sum_{j_1 \in \mathbb{Z}} \sum_{j_2 \in \mathbb{Z}} \widehat{A}_{j_1, j_2}(T) e^{2i\pi(j_1 X + j_2 Y)},$$

with $\widehat{A}_{j_1, j_2}(T) \in \mathbb{C}$. The Fourier transformed KP equation (34) vanishes identically for $j_1 = j_2 = 0$ and so there is no governing equation for $\widehat{A}_{0,0}(T)$. Therefore, we are free to set

$\widehat{A}_{0,0}(T) \equiv 0$. Global well-posedness of the KP-II equation (34) was established in $H^s(\mathbb{T}^2)$ for any $s \geq 0$ in [3], provided that the initial data satisfies

$$\oint A(X, Y, 0)dX = 0, \quad \text{for every } Y, \tag{40}$$

that is, $\widehat{A}_{0,j_2} = 0$ for every $j_2 \in \mathbb{Z}$.

The evolution problem for the inhomogeneous linearized KP equation (35) is also well-defined in Sobolev spaces $H^s(\mathbb{T}^2)$ with the same constraint (40). Thus, antiderivatives in X presents no problem on existence of smooth solutions in Sobolev spaces $H^s(\mathbb{T}^2)$ both for the KP-II equation (34) and the linearized KP-II equation (35).

However, the choice of periodic boundary conditions in the KP equation (34) leads to the problem of having to choose corresponding boundary conditions in the original FPU system (5). The resulting difficulties are illustrated in Figure 2. An irrational propagation direction $(\cos\phi, \sin\phi)$ leads to a quasi-periodic lattice in Fourier space which lies densely in the torus. The treatment of this problem leads to functional analytical difficulties whose solution is outside the scopes of this work. Therefore, in the following we restrict ourselves to the case of rational propagation directions $(\cos\phi, \sin\phi)$.
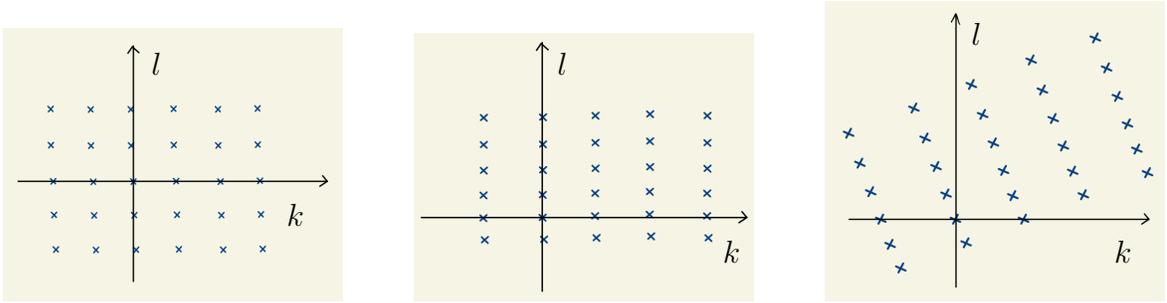

FIGURE 2. The left panel shows the distribution of Fourier modes in the case of periodic boundary conditions for the KP equation (34). The dots are located at integer multiples of 2π. The middle panel shows the resulting distribution of Fourier modes in the original system (5) in the case of wave propagation along the x-axis. The distance of the dots in k-direction is $\mathcal{O}(\varepsilon)$ and in l-direction $\mathcal{O}(\varepsilon^2)$. The right panel shows the corresponding distribution of Fourier modes in the original system (5) in case of wave propagation along an arbitrary direction. For a rational propagation direction $(\cos\phi, \sin\phi)$, we obtain a periodic lattice in Fourier space with finitely many modes on the torus. For an irrational propagation direction $(\cos\phi, \sin\phi)$, we obtain a quasi-periodic lattice that lies densely in the torus due to non-linear interactions.

In case of a propagation along the x-axis spatially periodic solutions for the KP equation (12) of period 1 in X and Y correspond in the original FPU system (5) to

$$u_{m,n} = u_{m+1/\varepsilon,n} = u_{m,n+1/\varepsilon^2},$$

and so we should restrict to values of $\varepsilon > 0$ such that $1/\varepsilon \in \mathbb{N}$ for which $1/\varepsilon^2 \in \mathbb{N}$. Such solutions can be represented by the finite Fourier polynomial

$$u_{m,n} = \sum_{j_1=-1/\varepsilon}^{1/\varepsilon-1} \sum_{j_2=-1/\varepsilon^2}^{1/\varepsilon^2-1} \widehat{u}_{j_1,j_2} e^{2\pi i(j_1\varepsilon m)} e^{2\pi i(j_2\varepsilon^2 n)}$$

In case of a rational direction of propagation $(\cos\phi, \sin\phi)$ with respect to the square lattice $\mathbb{Z}^2$ the finite Fourier polynomial is given by

$$
\begin{aligned}
u_{m,n} &= \sum_{j_1=-1/\varepsilon}^{1/\varepsilon-1} \sum_{j_2=-1/\varepsilon^2}^{1/\varepsilon^2-1} \widehat{u}_{j_1,j_2} e^{2\pi i(j_1\varepsilon((\cos\phi)m+(\sin\phi)n))} e^{2\pi i(j_2\varepsilon^2(-(\sin\phi)m+(\cos\phi)n))} \\
&= \sum_{j_1=-1/\varepsilon}^{1/\varepsilon-1} \sum_{j_1=-1/\varepsilon^2}^{1/\varepsilon^2-1} \widehat{u}_{j_1,j_2} e^{2\pi i(j_1\varepsilon(\cos\phi)-j_2\varepsilon^2(\sin\phi))m} e^{2\pi i(j_1\varepsilon(\sin\phi)+j_2\varepsilon^2(\cos\phi))n}.
\end{aligned}
$$

Since the indices do not reflect the position in Fourier space we introduce

$$
\widehat{u}(2\pi(j_1\varepsilon(\cos\phi)-j_2\varepsilon^2(\sin\phi)), 2\pi(j_1\varepsilon(\sin\phi)+j_2\varepsilon^2(\cos\phi))) = \widehat{u}_{j_1,j_2}
$$

and introduce the lattice

$$
\Sigma_2 = \{(2\pi(j_1\varepsilon(\cos\phi)-j_2\varepsilon^2(\sin\phi)), 2\pi(j_1\varepsilon(\sin\phi)+j_2\varepsilon^2(\cos\phi))) \in [0,2\pi)^2 : j_1, j_2 \in \mathbb{Z}\}
$$

The essential difference to the above calculations is that the space $L^2(\mathbb{T}^2)$ has to be replaced by a sequence space, i.e., instead of $L^2(\mathbb{T}^2)$ for solving the FPU system in Fourier space we consider the space $\ell^2(\Sigma_2)$ which is equipped with the norm

$$
\|\widehat{u}\|_{\ell^2(\Sigma_2)} = \varepsilon^{3/2} \left(\sum_{j_1=-1/\varepsilon}^{1/\varepsilon-1} \sum_{j_2=-1/\varepsilon^2}^{1/\varepsilon^2-1} |\widehat{u}_{j_1,j_2}|^2 \right)^{1/2} = \varepsilon^{3/2} \left(\sum_{\ell \in \Sigma_2} |\widehat{u}(\ell)|^2 \right)^{1/2}
$$

in order to have the same scaling as above. For estimating the leading-order approximation $(\widehat{u}, \widehat{v}) = (\varepsilon^2\widehat{\psi}_u, \varepsilon^2\widehat{\psi}_v)$ in the equations for the error we need similar to above a space $\ell^1(\Sigma_2)$ which is equipped with the norm $\|\widehat{u}\|_{\ell^1(\Sigma_2)} = \varepsilon^3 \sum_{\ell \in \Sigma_2} |\widehat{u}(\ell)|$. By Cauchy–Schwarz inequality, we obtain

$$
\|\widehat{u}\|_{\ell^1(\Sigma_2)} = \varepsilon^3 \sum_{\ell \in \Sigma_2} 1 \cdot |\widehat{u}(\ell)| \leq \varepsilon^3 (\sum_{\ell \in \Sigma_2} 1^2)^{1/2} (\sum_{\ell \in \Sigma_2} |\widehat{u}(\ell)|^2)^{1/2} \leq C\|\widehat{u}\|_{\ell^2(\Sigma_2)}
$$

where we used $\sum_{\ell \in \Sigma_2} 1^2 = \mathcal{O}(\varepsilon^{-3})$ due to the $\mathcal{O}(\varepsilon^{-3})$ many summands. As a result, we have $\|\widehat{\psi}_u\|_{\ell^1(\Sigma_2)} = \mathcal{O}(1)$ and $\|\widehat{\psi}_v\|_{\ell^1(\Sigma_2)} = \mathcal{O}(1)$. With these norms and notations the approximation result of Theorem 5.1 transfers to the periodic Sobolev spaces $H^s(\mathbb{T}^2)$ for every $s \geq 9$.

6.3. **Unbounded domain.** One needs to be careful in analyzing smooth solutions of the KP-II equation (34) and the linearized KP-II equation (35) in Sobolev spaces $H^s(\mathbb{R}^2)$. As was pointed out in [16], if A belongs to $H^s(\mathbb{R}^2)$, then $\partial_X^{-1}\partial_Y A^2$ may not be in $H^s(\mathbb{R}^2)$ as the integral of the positive function cannot decay to zero both as $X \to -\infty$ and $X \to +\infty$. This was also observed in the proof of Lemma 1 in [14], where a constraint was added on the combined quantity $\partial_X^{-1}\partial_Y^2(\partial_X^{-2}\partial_Y^2 A_0 - A_0^2)$ rather than on $\partial_X^{-3}\partial_Y^3 A$ or $\partial_X^{-1}\partial_Y A^2$.

Rewriting the linearized KP-II equation (35) in the evolution form yields

$$
\begin{aligned}
-2\partial_T A_2 = &\frac{1}{12}[(\cos\phi)^4 + (\sin\phi)^4]\partial_X^3 A_2 + \partial_X^{-1}\partial_Y^2 A_2 \\
&- 2[(\cos\phi)^2 + (\sin\phi)^2(\tan\phi)]\partial_X(A_1 A_2) \\
&- \frac{1}{3}[(\cos\phi)^2 - (\sin\phi)^2](\cos\phi)(\sin\phi)\partial_X^2\partial_Y A_1 \\
&- 2(\tan\phi)^2\partial_X(A_1\partial_X^{-1}\partial_Y A_1) \\
&+ [(\sin\phi)^2(\tan\phi)^2 - (\sin\phi)^2 + 2(\sin\phi)(\cos\phi)]\partial_Y(A_1^2).
\end{aligned}
\tag{41}
$$

In the evolution form, the right-hand side of the linearized KP-II equation contains terms $\mathcal{O}(\partial_X^2\partial_Y A_1, \partial_X(A_1\partial_X^{-1}\partial_Y A_1), \partial_Y(A_1^2))$, which are controlled by Lemma 3.1 in Sobolev norms. Therefore, by Duhamel's principle, we have $A_2 \in C([0,\tau_0], H^{s+6}) \cap C^1([0,\tau_0], H^{s+3})$. However, for the justification analysis, we need to estimate $\partial_X^{-1}\partial_T^2 A_2$ in Sobolev norms.

For $D_2 := \partial_X^{-1}\partial_Y A_2$, we can obtain from (41)

$$
\begin{aligned}
-2\partial_T D_2 = &\frac{1}{12}[(\cos\phi)^4 + (\sin\phi)^4]\partial_X^3 D_2 + \partial_X^{-1}\partial_Y^2 D_2 \\
&- 2[(\cos\phi)^2 + (\sin\phi)^2(\tan\phi)]\partial_Y(A_1 A_2) \\
&- \frac{1}{3}[(\cos\phi)^2 - (\sin\phi)^2](\cos\phi)(\sin\phi)\partial_X\partial_Y^2 A_1 \\
&- 2(\tan\phi)^2\partial_Y(A_1\partial_X^{-1}\partial_Y A_1) \\
&+ [(\sin\phi)^2(\tan\phi)^2 - (\sin\phi)^2 + 2(\sin\phi)(\cos\phi)]\partial_X^{-1}\partial_Y^2(A_1^2).
\end{aligned}
$$

By Duhamel's principle, we obtain $\partial_X^{-1}\partial_Y A_2 \in C^0([0,\tau_0], H^{s+6}) \cap C^1([0,\tau_0], H^{s+3})$ if the initial data satisfy the constraint

$$
\partial_X^{-1}\partial_Y^2\left[\partial_X^{-1}\partial_Y A_2|_{T=0} + [(\sin\phi)^2(\tan\phi)^2 - (\sin\phi)^2 + 2(\sin\phi)(\cos\phi)]A_1^2|_{T=0}\right] \in H^{s+6}.
$$

Since we need to control $\partial_X^{-1}\partial_T^2 A_2$, we need to extend this method to $\partial_X^{-1}\partial_Y A_2 \in C^2([0,\tau_0], H^s)$ or equivalently to $\partial_X^{-2}\partial_Y^2 A_2 \in C^0([0,\tau_0], H^{s+3}) \cap C^1([0,\tau_0], H^s)$. However, this is out of reach at the present time because of overdetermined set of constraints on the initial data $A_1|_{T=0}$ and $A_2|_{T=0}$. Further work on extending the well-posedness results for the KP-II equation (35) is needed to satisfy the requirements of Theorem 5.1 in $H^s(\mathbb{R}^2)$ for every $s \geq 9$.

References

[1] D. Bambusi and A. Ponno. "On metastability in FPU", Comm. Math. Phys. **264** (2006), 539–561.

[2] S. Baumstark, G. Schneider, K. Schratz, and D. Zimmermann, "Effective slow dynamics models for a class of dispersive systems", J. Dyn. Differ. Eqs. **32** (2020) 1867–1899.

[3] J. Bourgain, "On the Cauchy problem for the Kadomstev–Petviashvili equation," Geom. Funct. Anal. **3** (1993), 315–341.

[4] F. Chen and M. Herrmann, "KdV-like solitary waves in two-dimensional FPU-lattices", Discrete Contin. Dyn. Syst. **38** (2018), 2305–2332.

[5] E. Dumas and D.E. Pelinovsky, "Justification of the log-KdV equation in granular chains: the case of precompression", SIAM J. Math. Anal. **46** (2014), 4075–4103.

[6] D.B. Duncan, J.C. Eilbeck, C.H. Walshaw, and V.E. Zakharov, "Solitary waves on a strongly anisotropic KP lattice," Phys. Lett. A **158** (1991), 107–111.

[7] G. Friesecke and K. Matthies, "Geometric solitary waves in a 2D mass-spring lattice," Discr. Cont. Dynam. Syst. B **3** (2003), 105–114.

[8] G. Gallavotti, "Introduction to FPU", in *The Fermi-Pasta-Ulam Problem. A status report*, G. Gallavotti (Editor), Lecture Notes in Physics **728** (Springer, Berlin, Heidelberg, 2008), pp.1–19

[9] T. Gallay and G. Schneider, "KP description of unidirectional long waves. The model case," Proc. R. Soc. Edinburgh A **131** (2001), 885–898.

[10] M. Gallone and S. Pasquali, "Metastability phenomena in two-dimensional rectangular lattices with nearest-neighbour interaction", Nonlinearity **34** (2021), 4983–5044.

[11] M. Gallone, A. Ponno, and B. Rink, "Korteweg–de Vries and Fermi-Pasta-Ulam-Tsingou: asymptotic integrability of quasi unidirectional waves", J. Phys. A: Math. Theor. **54** (2021) 305701.

[12] T. Haas and G. Schneider, "Failure of the n-wave interaction approximation without imposing periodic boundary conditions", Zeitschrift für Angewandte Mathematik und Mechanik **100** (2020), e201900230.

[13] M. Herrmann and A. Mikikits-Leitner, "KdV waves in atomic chains with nonlocal interactions", Discrete Contin. Dyn. Syst. **36** (2016), 2047–2067.

[14] N. Hristov and D. E. Pelinovsky, "Justification of the KP-II approximation in dynamics fo two-dimensional FPU systems", ZAMP (2022), in print.

[15] A. Khan and D. Pelinovsky, "Long-time stability of small FPU solitary waves", Discr. Cont. Dynam. Syst. A **37** (2017), 2065–2075.

[16] L. Molinet, J.-C. Saut, and N. Tzvetkov, "Global well-posedness for the KP-I equation", Math. Ann. **324** (2002) 255–275

[17] A. Ponno and D. Bambusi, "Korteweg–de Vries equation and energy sharing in Fermi–Pasta–Ulam", Chaos **15** (2005), 015107.

[18] G. Schneider, "Validity and limitation of the Newell–Whitehead equation", Math. Nachr. **176** (1995) 249–263.

[19] G. Schneider, "Bounds for the nonlinear Schrödinger approximation of the Fermi-Pasta-Ulam system", Appl. Anal. **89** (2010) 1523–1539.

[20] G. Schneider, D. Ali Sunny, and D. Zimmermann, "The NLS approximation makes wrong predictions for the water wave problem in case of small surface tension and spatially periodic boundary conditions", J. Dyn. Differ. Eqs **27** (2015), 1077–1099.

[21] G. Schneider and H. Uecker, Nonlinear PDEs. A dynamical systems approach, volume 182 of Grad. Stud. Math (AMS, Providence, RI, 2017).

[22] G. Schneider and C. E. Wayne, "Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model," in *International conference on differential equations Equadiff–99, Berlin, Germany August 1-7, 1999* (World Scientific Publishing Company, Singapore, 2000), pp. 390–404.

[23] B. Schweizer and F. Theil, "Lattice dynamics on large time scales and dispersive effective equations", SIAM J. Appl. Math. **78** (2018), 3060–3086.

[24] S. Ukai, "Local solutions of the Kadomtsev–Petviashvili equation," J. Fac. Sci. Univ. Tokyo **36** (1989), 193–209.

[25] A. Vainchtein, "Solitary waves in FPU-type lattices", Physica D **434** (2022) 133252 (22 pages)

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCMASTER UNIVERSITY, HAMILTON, ONTARIO, CANADA, L8S 4K1

Email address: dmpeli@math.mcmaster.ca

INSTITUT FÜR ANALYSIS, DYNAMIK UND MODELLIERUNG, UNIVERSITÄT STUTTGART, 70569 STUTTGART, GERMANY

Email address: guido.schneider@mathematik.uni-stuttgart.de