BOSTON UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

Dissertation

A BU THESIS LATEX TEMPLATE

by

TREVOR NORTON

B.S., Virginia Polytechnic Institute and State University, 2015
M.S., Virginia Polytechnic Institute and State University, 2018

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2022

Approved by

First Reader

First M. Last, PhD
Professor of Electrical and Computer Engineering

Second Reader

First M. Last
Associate Professor of . . .

Third Reader

First M. Last
Assistant Professor of . . .

Facilis descensus Averni;
Noctes atque dies patet atri janua Ditis;
Sed revocare gradum, superasque evadere ad auras,
Hoc opus, hic labor est. Virgil (from Don's thesis!)

Acknowledgments

[This is where the acknowledgments go...]

A BU THESIS LATEX TEMPLATE

TREVOR NORTON

Boston University, College of Arts and Sciences, 2022

Major Professors: First M. Last, PhD
Professor of Electrical and Computer Engineering
Secondary appointment

First M. Last, PhD
Professor of computer Science

ABSTRACT

[This is where the text for the abstract will go]

Contents

List of Abbreviations

FPUT	....	Fermi-Pasta-Ulam-Tsingou
mKdV	....	modified Korteweg-De Vries

Chapter 1

Long-Time approximations of small-amplitude, long-wavelength FPUT solutions

1.1 Introduction

As shown in earlier work, there exists a wave solution of the FPUT lattice whose profile is well approximated by that of the kink solution to the (defocusing) mKdV. This approximation holds globally in time, but is restricted to one special solution of the FPUT. We are now interested in studying more general solutions of the FPUT which can be approximated by solutions of the mKdV for long (but finite) time. The equations of motion on the lattice are given by

$$\ddot{x}_n = V'(x_{n+1} - x_n) - V'(x_n - x_{n-1}), \quad n \in \mathbb{Z}. \tag{1.1}$$

where V is the interaction potential between neighboring particles and $\dot{}$ denotes the derivative with respect to the time $t \in \mathbb{R}$. Equation (1.1) can be rewritten in the strain variables $u_n := x_{n+1} - x_n$ as follows

$$\ddot{u}_n = V'(u_{n+1}) - 2V'(u_n) + V'(u_{n-1}), \quad n \in \mathbb{Z} \tag{1.2}$$

The moving wave solution in eq. (1.1) corresponds to a kink solution in eq. (1.2).

For the case where V is of the form $V(u) = \frac{1}{2}u^2 + \frac{\epsilon^2}{p+1}u^{p+1}$ for $p \geq 2$, the generalized

KdV equation given by

$$2\partial_T W + \frac{1}{12}\partial_X^3 W + \partial_X(W^p) = 0, \quad X \in \mathbb{R} \tag{1.3}$$

serves as a modulation equation for solutions of eq. (1.2) (Bambusi and Ponno, 2006; Friesecke and Pego, 1999). That is, for a local solution $W \in C([-\tau_0, \tau_0], H^s(\mathbb{R}))$ of eq. (1.3) there exist positive constants ϵ_0 and C_0 such that, for all $\epsilon \in (0, \epsilon_0)$, when initial data $(u_{\text{in}}, \dot{u}_{\text{in}}) \in \ell^2(\mathbb{R})$ satisfy

$$\|u_{\text{in}} - W(\epsilon\cdot, 0)\|_{\ell^2} + \|\dot{u}_{\text{in}} + \epsilon\partial_X W(\epsilon\cdot, 0)\|_{\ell^2} \leq \epsilon^{3/2}, \tag{1.4}$$

the unique solution to eq. (1.2) with initial data $(u_{\text{in}}, \dot{u}_{\text{in}})$ belongs to $C^1([-\tau_0\epsilon^{-3}, \tau_0\epsilon^{-3}]; \ell^2(\mathbb{Z}))$ and satisfies

$$\|u(t) - W(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2(\mathbb{Z})} + \|\dot{u}(t) + \epsilon\partial_X W(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2(\mathbb{Z})} \leq C_0\epsilon^{3/2},$$
$$t \in [-\tau_0\epsilon^{-3}, \tau_0\epsilon^{-3}]. \tag{1.5}$$

Furthermore, the approximation can also be extended to include counter-propagating solutions of the KdV in the case where $p = 2$ (Schneider and Wayne, 2000; Hong et al., 2021).

The KdV approximation was extended to longer time scales on the order of $\epsilon^{-3}|\log(\epsilon)|$ by Khan and Pelinovsky in order to deduce the nonlinear metastability of small FPUT solitary waves from the orbital stability of the corresponding KdV solitary waves (Khan and Pelinovsky, 2017).

We consider the FPUT with potential

$$V(u) = \frac{1}{2}u^2 - \frac{1}{24}u^4. \tag{1.6}$$

This potential differs from those studied in (Khan and Pelinovsky, 2017) in that it admits kink solutions. Numerical experiments show that these kink solutions play an important role in the FPUT recurrence for lattices with potential give in eq. (1.6) (Pace et al., 2019). We will introduce an ansatz that solutions of the FPUT with this potential can be well-approximated by counter-propagating solutions of mKdV equations.

The technique of the proof follows from ideas in (Schneider and Wayne, 2000; Khan and Pelinovsky, 2017) and is roughly sketched out as follows. First the system is rewritten into a Hamiltonian system on a Hilbert space, H:

$$\dot{X}(t) = J\mathcal{H}'(X) \tag{1.7}$$

where $J : H \to H$ is a skew symmetric operator and $\mathcal{H}$ is the Hamiltonian such that $\mathcal{H}'(X) = LX + N(X)$ with $L := \mathcal{H}'(0)$. We introduce some ansatz $\tilde{X}_\epsilon$ which is an approximate solution to eq. (1.7) in the sense that

$$\mathrm{Res}(t) := J[L\tilde{X}_\epsilon(t) + N(\tilde{X}_\epsilon(t))] - \dot{\tilde{X}}_\epsilon(t) \tag{1.8}$$

has norm of order ϵ^α for $\alpha > 0$ for all time t. The approximate solution will be "small-amplitude" in the sense that $\|\tilde{X}_\epsilon\| = \mathcal{O}(\epsilon^k)$ for $k > 0$. Then we can write the evolution equation for the $R(t) = X(t) - \tilde{X}_\epsilon(t)$ as

$$\dot{R}(t) = J[L + N'(\tilde{X}_\epsilon(t))]R(t) + \mathrm{Res}(t) + \mathcal{N}(\tilde{X}_\epsilon, R) \tag{1.9}$$

with $\mathcal{N}(X_\epsilon, R) := J[N(\tilde{X}_\epsilon + R) - N(\tilde{X}_\epsilon) - N'(\tilde{X}_\epsilon)R]$. The goal is then to show that $R(t)$ remains small for long periods of time so that the approximation $X \approx \tilde{X}_\epsilon$ is valid for that time. The standard way to prove this is to find a suitable energy function to control the norm of R with. If $L + N'(\tilde{X}_\epsilon(t))$ is self-adjoint, then eq. (1.9) is up to first order a linear, non-autonomous, Hamiltonian system with Hamiltonian $\mathcal{H}_1(R,t) = \frac{1}{2}\langle (L + N'(\tilde{X}_\epsilon))R, R\rangle$. Therefore, $\mathcal{E}(t) := \mathcal{H}_1(R(t), t)$ serves as a natural choice of energy function for eq. (1.9). Hence, if one shows that $\|R\|^2 \lesssim \mathcal{E}(t)$ and that $\|\mathcal{N}(\tilde{X}_\epsilon, R)\| \lesssim \epsilon^{k+2}\mathcal{E}(t)$, then can show that $\mathcal{S}(t) = \mathcal{E}(t)^{1/2}$ satisfies

$$|\dot{\mathcal{S}}(t)| \lesssim \epsilon^\alpha + \epsilon^{k+2}\mathcal{S}(t). \tag{1.10}$$

Intuitively, one would expect $\mathcal{S}(t)$ to grow like $\mathcal{S}(t) \sim \epsilon^\alpha t + e^{\epsilon^{k+2}t}\mathcal{S}(0)$. Taking $\mathcal{S}(0) = \epsilon^\gamma$ for $\gamma \geq 1$ and assuming $\alpha > 2(k+2)$, we have $\mathcal{S}(t) \sim \epsilon^\gamma$ for $|t| \lesssim \epsilon^{-(k+2)}$. One can further the time where the approximation holds by relaxing how big $\mathcal{S}(t)$ can get. Taking $r > 0$ small, one can show that $\mathcal{S}(t) \sim \epsilon^{\gamma - r}$ for $|t| \lesssim r\epsilon^{-(k+2)}|\log(\epsilon)|$.

1.2 Counter-Propagating Waves Ansatz

We make the assumption that solutions of eq. (1.2) can be expressed as a sum of two counter-propagating small-amplitude waves, i.e.,

$$u_n(t) \approx \epsilon f(\epsilon(n+t), \epsilon^3 t) + \epsilon g(\epsilon(n - ct), \epsilon^3 t) + \epsilon^3 \phi(\epsilon n, \epsilon t) \tag{1.11}$$

where we allow f to have a fixed non-zero limits, $f_{\pm\infty}$, at positive and negative infinity and ϕ captures the interaction effects between f and g. The wave speed of g is given by

$$c = c(\epsilon, f_\infty) = 1 - \frac{\epsilon^2 f_\infty^2}{4}. \tag{1.12}$$

Plugging in the ansatz in eq. (1.11) back into eq. (1.2) and grouping terms of the same order ϵ together gives

$$\begin{aligned}
&\epsilon^3\Big(\partial_1^2 f(\cdot,\epsilon^3 t)+\partial_1^2 g(\cdot,\epsilon^3 t)\Big)\\
&+\epsilon^5\Big(2\partial_1\partial_2 f(\cdot,\epsilon^3 t)-2\partial_1\partial_2 g(\cdot,\epsilon^3 t)-\frac{f_\infty^2}{2}\partial_1^2 g+\partial_2^2\phi(\epsilon x,\epsilon t)\Big)\\
&+\mathcal{O}(\epsilon^6)\\
&\qquad=\quad \epsilon^3\Big(\partial_1^2 f(\cdot,\epsilon^3 t)+\partial_1^2 g(\cdot,\epsilon^3 t)\Big)\\
&\qquad\quad+\epsilon^5\Big(\partial_1^2\phi(\epsilon x,\epsilon t)\\
&\qquad\qquad-\frac{1}{6}\partial_1^2\big[f^3(\cdot,\epsilon^3 t)+3f^2(\cdot,\epsilon^3)g(\cdot,\epsilon^3 t)+3f(\cdot,\epsilon t)g^2(\cdot,\epsilon^3 t)+g^3(\cdot,\epsilon^3 t)\big]\\
&\qquad\qquad+\frac{1}{12}\partial_1^4 f(\cdot,\epsilon^3 t)+\frac{1}{12}\partial_1^4 g(\cdot,\epsilon^3 t)\Big)\\
&\qquad\quad+\mathcal{O}(\epsilon^6).
\end{aligned} \tag{1.13}$$

Clearly the equation will hold up to order ϵ^3. For the order ϵ^5 terms, the equation will again hold if f, g, and ϕ satisfy

$$2\partial_2 f=-\frac{1}{6}\partial_1(f^3)+\frac{1}{12}\partial_1^3 f \tag{1.14}$$

and

$$-2\partial_2 g=-\frac{1}{6}\partial_1(g^3+3f_\infty g^2)+\frac{1}{12}\partial_1^3 g, \tag{1.15}$$

and

$$\begin{aligned}
\partial_2^2\phi(\xi,\tau)=\partial_1^2\phi(\xi,\tau)-\frac{1}{6}\partial_1^2\big[&3(f^2(\xi+\tau,\epsilon^2\tau)-f_\infty^2)g(\xi-c\tau,\epsilon^2\tau)\\
&+3(f(\xi+\tau,\epsilon^2\tau)-f_\infty)g^2(\xi-c\tau,\epsilon^2\tau)\big]
\end{aligned} \tag{1.16}$$

$$\phi(\xi,0)=\partial_1\phi(\xi,0)=0.$$

Note that eq. (1.14) is the defocusing mKdV equation and eq. (1.15) is a type of generalized KdV equation. This formal calculation shows that the mKdV can serve as a modulation equation. That is, for ϵ sufficiently small, one would expect the ansatz in eq. (1.11) to hold for time on the order of ϵ^{-3}. We make precise this notion, but we must first make decisions for the function spaces in which the functions f, g, and ϕ must live.

A natural choice of function space for g is a Sobolev space like $H^k(\mathbb{R})$. However, for f, we want to allow the possibility of the function approaching a non-zero limit at positive and negative infinity while also having sufficient regularity.

Definition 1. *For $k \in \mathbb{N}$, let $\mathcal{X}^k(\mathbb{R})$ be the Banach space*

$$\mathcal{X}^k(\mathbb{R}) := \{f \in L^\infty(\mathbb{R}) \mid f' \in H^{k-1}(\mathbb{R})\} \tag{1.17}$$

with norm

$$\|f\|_{\mathcal{X}^k(\mathbb{R})} := \|f\|_{L^\infty(\mathbb{R})} + \|f'\|_{H^{k-1}(\mathbb{R})}. \tag{1.18}$$

Then $\mathcal{X}^k$ is the set of L^∞ functions which are k times weakly differentiable and whose derivatives are in L^2. That this is a Banach space follows from the Banach space isomorphism

$$\mathcal{X}^k(\mathbb{R}) \cong L^\infty(\mathbb{R}) \cap \dot{H}^1(R) \cap \dot{H}^k(\mathbb{R}), \tag{1.19}$$

where $\dot{H}^k(\mathbb{R})$ denotes the homogeneous Sobolev spaces. For convenience, we let $\mathcal{X}^0(\mathbb{R})$ denote $L^\infty(\mathbb{R})$

Note that eq. (1.14) has kink solutions of the form

$$f(X,T) = -\sqrt{12v}\tanh\left(\sqrt{12v}(X - vT)\right). \tag{1.20}$$

In particular, setting $v = 24$ we get the approximate solution on the lattice given by

$$-\frac{\epsilon}{\sqrt{2}}\tanh\left(\frac{\epsilon}{\sqrt{2}}\left(n + \left(1 - \frac{\epsilon^2}{24}\right)t\right)\right), \tag{1.21}$$

which seems to agree with the kink solution on the lattice for long periods of time (i.e. it should hold formally for t of order $\mathcal{O}(\epsilon^{-4})$). The space $\mathcal{X}^k$ allows for f to be these kink solutions and thus allows us to study the kink solution of the lattice found previously.

We also have the following inequalities for products of functions in $\mathcal{X}^k$ and H^k that will be useful.

Lemma 1. *For non-negative integers k, there is a $C > 0$ such that*

$$\|fg\|_{H^k} \leq C\|f\|_{\mathcal{X}^k}\|g\|_{H^k} \tag{1.22}$$

for any $f \in \mathcal{X}^k(\mathbb{R})$ and $g \in H^k(\mathbb{R})$.

Lemma 2. *For non-negative integers k, there is a $C > 0$ such that*

$$\|fg\|_{\mathcal{X}^k} \leq C\|f\|_{\mathcal{X}^k}\|h\|_{\mathcal{X}^k} \tag{1.23}$$

for any $f, g \in \mathcal{X}^k(\mathbb{R})$.

See appendix A for proofs.

However, for our main result, we require that ϕ, the term which captures the interaction effects, remains uniformly bounded for all time. Intuitively, if f and g are localized, the inhomogeneous term in eq. (1.16) will quickly go to zero, and the equation governing ϕ eq. (1.16) will approach the homogeneous wave equation, for which Sobolev norms remain uniformly bounded. Since the two functions are localized and counter-propagating, their product will quickly decay in time as the two wave profiles move in opposite directions. Thus we require that f and g quickly decay to their respective limits at infinity. This is enforced by assuming the functions belong to appropriate weighted Banach spaces.

A suitable choice of space for g is the weighted Sobolev spaces $H^k_n(\mathbb{R})$. Here, H^k_n for $k, n \in \mathbb{N} \cup \{0\}$

$$H^k_n(\mathbb{R}) := \{g \in H^k(\mathbb{R}) \mid g\langle\cdot\rangle^n \in H^k\} \tag{1.24}$$

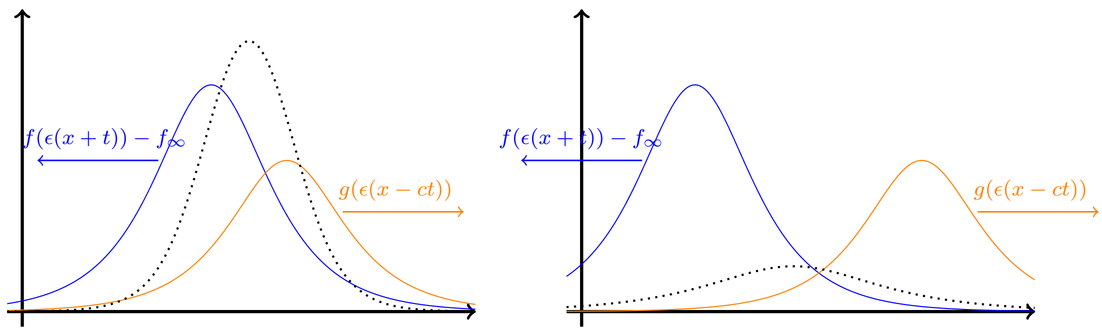

Figure 1·1: The function $f(\epsilon(x+t)) - f_\infty$ (shown in blue) moves to the left while $g(\epsilon(x-ct))$ (shown in orange) moves to the right. Since they are localized, the product (shown by the dotted line) will quickly decay in time.

where $\langle x\rangle = \sqrt{1+x^2}$. The norm on this space is

$$\|g\|_{H^k_n(\mathbb{R})} := \|g\langle\cdot\rangle^n\|_{H^k(\mathbb{R})}. \tag{1.25}$$

This space has the useful property that if $g \in H^k_n$, then its Fourier transform, $\hat{g}$, is in H^n_k and

$$c\|\hat{g}\|_{H^n_k} \le \|g\|_{H^k_n} \le C\|\hat{g}\|_{H^n_k} \tag{1.26}$$

for $c, C > 0$ independent of g.

We want an analogous space for f, but allowing for non-zero limits at infinity. Let $\langle\cdot\rangle_+ : \mathbb{R} \to \mathbb{R}$ be a smooth function such that

$$\langle x\rangle_+ = \begin{cases} \langle x\rangle, & x > 1 \\ 1, & x < 0 \end{cases} \tag{1.27}$$

and $\langle\cdot\rangle_+$ continued smoothly between 0 and 1 such that it is always greater than or equal to 1. Thus $\langle\cdot\rangle_+$ is a function that only acts like $\langle\cdot\rangle$ for numbers greater than 1. The function $\langle\cdot\rangle_-$ is similarly defined but for numbers less than -1.

Definition 2. *Define $\mathcal{X}^k_{n^+}(\mathbb{R})$ to be the Banach space of functions where*

$$\mathcal{X}^k_{n^+}(\mathbb{R}) := \{f \in \mathcal{X}^k(\mathbb{R}) \mid \lim_{x\to\infty} f(x) = f_\infty \text{ and } (f - f_\infty)\langle\cdot\rangle^n_+ \in \mathcal{X}^k(\mathbb{R})\} \tag{1.28}$$

with norm given by

$$\|f\|_{\mathcal{X}^k_{n^+}(\mathbb{R})} := |f_\infty| + \|(f - f_\infty)\langle\cdot\rangle^n_+\|_{\mathcal{X}^k(\mathbb{R})} \tag{1.29}$$

Similarly,

$$\mathcal{X}^k_{n^-}(\mathbb{R}) := \{f \in \mathcal{X}^k(\mathbb{R}) \mid \lim_{x\to-\infty} f(x) = f_{-\infty} \text{ and } (f - f_{-\infty})\langle\cdot\rangle^n_- \in \mathcal{X}^k(\mathbb{R})\} \tag{1.30}$$

and

$$\|f\|_{\mathcal{X}^k_{n^-}(\mathbb{R})} := |f_{-\infty}| + \|(f - f_{-\infty})\langle\cdot\rangle^n_-\|_{\mathcal{X}^k(\mathbb{R})} \tag{1.31}$$

Define $\mathcal{X}^k_n(\mathbb{R})$ to be the intersection of these Banach spaces. That is,

$$\mathcal{X}^k_n(\mathbb{R}) := \mathcal{X}^k_{n^+}(\mathbb{R}) \cap \mathcal{X}^k_{n^-}(\mathbb{R}), \quad \|f\|_{\mathcal{X}^k_n(\mathbb{R})} := \|f\|_{\mathcal{X}^k_{n^+}(\mathbb{R})} + \|f\|_{\mathcal{X}^k_{n^-}(\mathbb{R})}. \tag{1.32}$$

That $\mathcal{X}^k_{n^\pm}$ are Banach spaces follows from the fact that there exists a linear isomorphism between the Banach space $\mathbb{R} \times \mathcal{X}^k$ and these spaces, which is given by

$$(\alpha, f) \mapsto \alpha + f\langle\cdot\rangle^{-n}_\pm. \tag{1.33}$$

One can show that the kink solutions as specified in eq. (1.20) lie in $\mathcal{X}^k_n$ for all $k, n \geq 0$; the derivatives are smooth and decay exponentially to zero, and the kink solutions approach the limits $\mp\sqrt{12v}$ exponentially fast. These spaces also contain bounded rational functions. For instance, the function

$$\frac{1}{x^2+1}$$

is in $\mathcal{X}^k_2(\mathbb{R})$ since it approaches its limit at infinity (which in this case is 1) at a rate of $\mathcal{O}(1/x^2)$, and its derivatives are in $H^0_2(\mathbb{R})$.

The definitions above are used to prove that ϕ remains bounded for all time. The idea behind the proof is similar to that of (Schneider and Wayne, 2000, Lemma 3.1). The following lemma will be useful in showing the decay in products of $f - f_\infty$ and g.

Lemma 3. *For each $k \geq 0$ and $c > 0$, there exists $C > 0$ depending only on k such that*

$$\left\| \frac{1}{\langle \cdot + \tau \rangle_+^2 \langle \cdot - c\tau \rangle^2} \right\|_{C^k} \leq C \sup_{x \in \mathbb{R}} \frac{1}{\langle x + \tau \rangle_+^2 \langle x - c\tau \rangle^2}. \tag{1.34}$$

Furthermore,

$$\int_0^\infty \sup_{x \in \mathbb{R}} \frac{1}{\langle x + \tau \rangle_+^2 \langle x - c\tau \rangle^2} \, d\tau < \infty. \tag{1.35}$$

See appendix A for proof.

We are now ready to prove that ϕ (and its time derivative) remain uniformly bounded in time.

Proposition 1. *Fix $T_0 > 0$ and suppose that $f \in C([-T_0, T_0], \mathcal{X}_2^{k+1}(\mathbb{R}))$ and $g \in C([-T_0, T_0], H_2^{k+1}(\mathbb{R}))$, with $k > 2$ an integer. Also, suppose that $f(X,T) \to f_\infty$ as $X \to \infty$ for any $T \in [-T_0, T_0]$. Then there exists a constant $C > 0$ such that*

$$\sup_{t \in [-\epsilon^{-3}T_0, \epsilon^{-3}T_0]} \|\phi(\cdot, \epsilon t)\|_{H^k} \leq C \left(\sup_{t \in [-\epsilon^{-3}T_0, \epsilon^{-3}T_0]} \left\{ \|f(\cdot, \epsilon^3 t)\|_{\mathcal{X}_2^{k+1}}, \|g(\cdot, \epsilon^3 t)\|_{H_2^{k+1}} \right\} \right)^3 \tag{1.36}$$

and

$$\sup_{t \in [-\epsilon^{-3}T_0, \epsilon^{-3}T_0]} \|\psi(\cdot, \epsilon t)\|_{H^{k-1}} \leq C \left(\sup_{t \in [-\epsilon^{-3}T_0, \epsilon^{-3}T_0]} \left\{ \|f(\cdot, \epsilon^3 t)\|_{\mathcal{X}_2^{k+1}}, \|g(\cdot, \epsilon^3 t)\|_{H_2^{k+1}} \right\} \right)^3, \tag{1.37}$$

where $\psi = \partial_2 \phi$.

Proof. Set $\partial_2 \phi = \psi$. Taking the Fourier transform $\mathcal{F}$ on both sides of eq. (1.16) and writing the ODE as a first order system, we get that

$$\begin{aligned} \partial_2 \begin{bmatrix} \hat{\phi}(k,\tau) \\ \hat{\psi}(k,\tau) \end{bmatrix} &= \begin{bmatrix} \hat{\psi}(k,\tau) \\ -k^2 \hat{\phi}(k,\tau) \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \frac{1}{2} k^2 \mathcal{F}[(f^2(\cdot + \tau), \epsilon^2 \tau) - f_\infty^2) g(\cdot - c\tau, \epsilon^2 \tau) + (f(\cdot + \tau, \epsilon^2 \tau) - f_\infty) g^2(\cdot - c\tau, \epsilon^2 \tau)](k) \end{bmatrix}. \end{aligned} \tag{1.38}$$

The semigroup generated by the linear part can be computed explicitly. Putting the solution into variation of constants form with initial conditions set to zero gives

$$\hat{\phi}(k,T) = \frac{1}{2}\int_0^T k\sin(k(T-\tau))\times$$
$$\mathcal{F}[(f^2(\cdot+\tau),\epsilon^2\tau)-f_\infty^2)g(\cdot-c\tau,\epsilon^2\tau)+(f(\cdot+\tau,\epsilon^2\tau)-f_\infty)g^2(\cdot-c\tau,\epsilon^2\tau)](k)\,d\tau \tag{1.39}$$

and

$$\hat{\psi}(k,T) = \frac{1}{2}\int_0^T k^2\cos(k(T-\tau))\times$$
$$\mathcal{F}[(f^2(\cdot+\tau,\epsilon^2\tau)-f_\infty^2)g(\cdot-c\tau,\epsilon^2\tau)+(f(\cdot+\tau,\epsilon^2\tau)-f_\infty)g^2(\cdot-c\tau,\epsilon^2\tau)](k)\,d\tau \tag{1.40}$$

Hence we can get that

$$\begin{aligned}
&\|\phi(\cdot,T)\|_{H^k} \\
&\quad\le C\|\hat{\phi}(\cdot,T)\|_{H^0_k} \\
&\quad\le C\int_0^T \|\partial_1((f^2(\cdot+\tau)-f_\infty^2)g(\cdot-c\tau))\|_{H^k}+\|\partial_1((f(\cdot+\tau)-f_\infty)g^2(\cdot-c\tau))\|_{H^k}\,\mathrm{d}\tau \\
&\quad\le C\int_0^T \|f(\cdot+\tau)\partial_1 f(\cdot+\tau)g(\cdot-c\tau)\|_{H^k}+\|(f^2(\cdot+\tau)-f_\infty^2)\partial_1 g(\cdot-c\tau)\|_{H^k} \\
&\qquad+\|\partial_1 f(\cdot+\tau)g^2(\cdot-c\tau)\|_{H^k}+\|(f(\cdot+\tau)-f_\infty)\partial_1 g(\cdot-c\tau)\|_{H^k}\,\mathrm{d}\tau \\
&\quad\le C\int_0^T \sup_{x\in\mathbb{R}}\frac{1}{\langle x+\tau\rangle_+^2\langle x-c\tau\rangle^2}\times\left(\|f\|^2_{\mathcal{X}_2^{k+1}}\|g\|_{H_2^{k+1}}+\|f\|_{\mathcal{X}_2^{k+1}}\|g\|^2_{H_2^{k+1}}\right)\mathrm{d}\tau,
\end{aligned} \tag{1.41}$$

whence eq. (1.36) follows. The proof for eq. (1.37) is analogous. □

1.3 Setup of Lattice Equations

The scalar second-order differential equation eq. (1.2) with potential V given by eq. (1.6) can be rewritten as the following first-order system:

$$\begin{cases} \dot{u}_n = q_{n+1} - q_n, \\ \dot{q}_n = u_n - u_{n-1} - \dfrac{1}{6}(u_n^3 - u_{n-1}^3), \end{cases} \qquad n \in \mathbb{Z}. \tag{1.42}$$

Recall that $u_n = x_{n+1} - x_n$, so we have that u_n physically represents the displacement between two neighbors on the lattice and q_n is equal to

$$q_n(t) = \sum_{k=-\infty}^{n-1} \dot{u}_k(t) = \sum_{k=-\infty}^{n-1} [\dot{x}_{k+1}(t) - \dot{x}_k(t)] = \dot{x}_n(t) \tag{1.43}$$

and so represents the velocity at a lattice point (assuming that $\dot{x}_k(t) \to 0$ as $k \to -\infty$). Note that we have the flexibility to add or subtract a constant from q without changing the dynamics on u (a fact that we use later to adjust the approximation and guarantee the error terms are in $\ell^2(\mathbb{Z})$). Writing the equations for the FPUT lattice in the form given by eq. (1.42) also puts the system into a Hamiltonian framework (when $u, q \in \ell^2(\mathbb{Z})$). Here the equations are of the form

$$\dot{U} = J\mathcal{H}'(U) \tag{1.44}$$

where $U = (u, q)$, J is the skew-symmetric operator given by

$$J = \begin{bmatrix} 0 & e^{\partial} - 1 \\ 1 - e^{-\partial} & 0 \end{bmatrix} \tag{1.45}$$

and $\mathcal{H}(U) = \sum_{n\in\mathbb{Z}} \frac{1}{2}q_n^2 + V(u_n)$.

We will now introduce the traveling wave ansatz for the system in eq. (1.42), but we first must assume certain regularity and decay of f and g.

Assumption 1. *Let f and g be solutions of eqs.* (1.14) *and* (1.15), *respectively.*

Assume that

$$f \in C_b(\mathbb{R}, \mathcal{X}_2^6(\mathbb{R})) \quad \textit{and} \quad g \in C_b(\mathbb{R}, H_2^6(\mathbb{R})).$$

Furthermore, assume that f has fixed limits in its spatial variable at $\pm\infty$ given by $f_{\pm\infty}$.

The traveling wave ansatz for u_n and q_n is then given by

$$u_n(t) = \epsilon f(\epsilon(n+t), \epsilon^3 t) + \epsilon g(\epsilon(n-ct), \epsilon^3 t) + \epsilon^3 \phi(\epsilon n, \epsilon t) + \mathcal{U}_n(t) \tag{1.46}$$

and

$$q_n(t) = \epsilon F(\epsilon(n+t), \epsilon^3 t) + \epsilon G(\epsilon(n-ct), \epsilon^3 t) + \epsilon^3 \Phi(\epsilon n, \epsilon t) - \epsilon F_{-\infty} + \mathcal{Q}_n(t). \tag{1.47}$$

The wave speed c is again given by eq. (1.12).

The form that the ansatz takes for $u_n(t)$ is clear. For $q_n(t)$ we need to define F, G, and Φ (where $F_{-\infty}$ is a constant to specified shortly thereafter). One would expect

$$\begin{aligned} q_n(t) &= \sum_{k=-\infty}^{n-1} \dot{u}_n(t) \\ &\approx \sum_{k=-\infty}^{n-1} [\epsilon^2 \partial_1 f(\epsilon(k+t)) + \epsilon^4 \partial_2 f(\epsilon(k+t)) \\ &\quad + \epsilon^2 c \partial_1 g(\epsilon(k-ct)) + \epsilon^4 \partial_2 g(\epsilon(k-ct)) \\ &\quad + \epsilon^4 \partial_2 \phi(\epsilon k)]. \end{aligned} \tag{1.48}$$

However, the final summation does not have a simple closed form, and so would be difficult to use. Instead, the summation is approximated with simpler terms up to an appropriate order of ϵ. We choose F, G, and Φ so that

$$\begin{aligned}
\epsilon F(\epsilon(n+1+t)) - \epsilon F(\epsilon(n+t)) &= \epsilon^2 \partial_1 f(\epsilon(n+t)) + \epsilon^4 \partial_2 f(\epsilon(n+1)) + \mathcal{O}(\epsilon^6) \\
\epsilon G(\epsilon(n+1-ct)) - \epsilon G(\epsilon(n-ct)) &= \epsilon^2 c \partial_1 g(\epsilon(n-ct)) + \epsilon^4 \partial_2 g(\epsilon(n-ct)) + \mathcal{O}(\epsilon^6) \\
\epsilon^3 \Phi(\epsilon(n+1)) - \epsilon^3 \Phi(\epsilon(n)) &= \epsilon^4 \partial_2 \phi(\epsilon n) + \mathcal{O}(\epsilon^6).
\end{aligned} \tag{1.49}$$

After this choice, the summation of the terms on the left has a simpler and explicit representation. Thus, following some calculations, we get the following:

$$F := f - \frac{\epsilon}{2}\partial_1 f + \frac{\epsilon^2}{8}\partial_1^2 f - \frac{\epsilon^2}{12} f^3 - \frac{\epsilon^3}{48}\partial_1^3 f + \frac{\epsilon^3}{8} f^2 \partial_1 f \tag{1.50}$$

$$\begin{aligned}
G := &-g + \frac{\epsilon}{2}\partial_1 g + \frac{\epsilon^2 f_\infty^2}{4} g + \frac{\epsilon^2}{12}(g^3 + 3 f_\infty g^2) - \frac{\epsilon^2}{8}\partial_1^2 g + \frac{\epsilon^3}{48}\partial_1^3 g \\
&- \frac{\epsilon^3}{24}\partial_1 (g^3 + 3 f_\infty g^2) - \frac{\epsilon^3 f_\infty^2}{8}\partial_1 g
\end{aligned} \tag{1.51}$$

$$\Phi := \partial_1^{-1}\psi - \frac{\epsilon}{2}\psi. \tag{1.52}$$

Here $\psi = \partial_2 \phi$ and ∂_1^{-1} is defined as a Fourier multiplier. That $\partial_1^{-1}\psi$ is well-defined and in $H^5(\mathbb{R})$ follows from eq. (1.40). Namely, we have that

$$\begin{aligned}
&\mathcal{F}[\partial_1^{-1}\psi(\cdot, T)](k) = (ik)^{-1}\hat{\psi}(k, T) \\
&\quad = \frac{-i}{2}\int_0^T k\cos(k(T-\tau)) \times \\
&\quad \mathcal{F}[(f^2(\cdot+\tau, \epsilon^2\tau) - f_\infty^2) g(\cdot - c\tau, \epsilon^2\tau) + (f(\cdot+\tau, \epsilon^2\tau) - f_\infty) g^2(\cdot - c\tau, \epsilon^2\tau)](k)\, d\tau
\end{aligned} \tag{1.53}$$

and (following the same calculations in eq. (1.41))

$$\|\partial_1^{-1}\psi(\cdot, T)\|_{H^5} \le C \int_0^T \sup_{x\in\mathbb{R}} \frac{1}{\langle x+\tau\rangle_+^2 \langle x - c\tau\rangle^2} \times \left(\|f\|_{\mathcal{X}_2^6}^2 \|g\|_{H_2^6} + \|f\|_{\mathcal{X}_2^6} \|g\|_{H_2^6}^2 \right) \mathrm{d}\tau. \tag{1.54}$$

Assumption 1 implies that F has fixed limits in its spatial variable at $\pm\infty$ given by

$F_{\pm\infty} = f_{\pm\infty} - \frac{\epsilon^2}{12} f_{\pm\infty}^3$.

We want $\mathcal{U}(t)$ and $\mathcal{Q}(t)$ to be elements of $\ell^2(\mathbb{Z})$ (at least locally in time). However, to satisfy $\mathcal{Q}(0) \in \ell^2(\mathbb{Z})$ and $\dot{u}_n(0) = q_{n+1}(0) - q_n(0)$, a compatibility condition must hold.

Assumption 2. *Assume that*

$$\sum_{n=-\infty}^{\infty} \dot{u}_n(0) = \epsilon F_{+\infty} - \epsilon F_{-\infty}.$$

Note that if this did not hold, then $\mathcal{Q}_n(0) \not\to 0$ as $n \to \infty$ and $\mathcal{Q}(0) \notin \ell^2(\mathbb{Z})$. That $\mathcal{Q}(0)_n \to 0$ as $n \to -\infty$ follows directly from the ansatz. The introduction of the constant $\epsilon F_{-\infty}$ in eq. (1.47) does not affect the dynamics of q in eq. (1.42)

An equivalent set of equations to eq. (1.42) are given by

$$\begin{cases} \dot{\mathcal{U}}_n(t) = \mathcal{Q}_{n+1}(t) - \mathcal{Q}_n(t) + \mathrm{Res}_n^{(1)}(t) \\ \dot{\mathcal{Q}}_n(t) = \mathcal{U}_n(t) - \mathcal{U}_{n-1}(t) \\ \qquad - \frac{1}{2}(\epsilon f(\epsilon(n+t)) + \epsilon g(\epsilon(n-ct)) + \epsilon^3 \phi(\epsilon n))^2 \mathcal{U}_n(t) & n \in \mathbb{Z}, \\ \qquad + \frac{1}{2}(\epsilon f(\epsilon(n-1+t)) + \epsilon g(\epsilon(n-1-ct)) + \epsilon^3 \phi(\epsilon(n-1)))^2 \mathcal{U}_{n-1}(t) \\ \qquad + \mathrm{Res}_n^{(2)}(t) + \mathcal{B}_n(\epsilon f + \epsilon g + \epsilon^3 \phi, \mathcal{U}) \end{cases} \tag{1.55}$$

where

$$\begin{aligned} \mathrm{Res}_n^{(1)}(t) =& \epsilon F(\epsilon(n+1+t)) - \epsilon F(\epsilon(n+t)) \\ & + \epsilon G(\epsilon(n+1-ct) - \epsilon G(\epsilon(n-ct) + \epsilon^3 \Phi(\epsilon(n+1)) - \epsilon^3 \Phi(\epsilon n) \\ & - \epsilon^2 \partial_1 f(\epsilon(n+t)) - \epsilon^4 \partial_2 f(\epsilon(n+t)) \\ & + \epsilon^2 c \partial_1 g(\epsilon(n-ct)) - \epsilon^4 \partial_2 g(\epsilon(n-ct)) - \epsilon^4 \partial_2 \phi(\epsilon n), \end{aligned} \tag{1.56}$$

$$\begin{aligned}
\mathrm{Res}_n^{(2)}(t) =& \epsilon f(\epsilon(n+t)) - \epsilon f(\epsilon(n-1+t)) \\
&+ \epsilon g(\epsilon(n-ct)) - \epsilon g(\epsilon(n-1-ct) + \epsilon^3\phi(\epsilon n) - \epsilon^3\phi(\epsilon(n-1)) \\
&- \epsilon^2\partial_1 F(\epsilon(n+t)) - \epsilon^4\partial_2 F(\epsilon(n+t)) \\
&+ \epsilon^2 c\partial_1 G(\epsilon(n-ct)) - \epsilon^4\partial_2 G(\epsilon(n-ct)) - \epsilon^4\partial_2\Phi(\epsilon n) \\
&- \frac{1}{6}\Big((\epsilon f(\epsilon(n+t)) + \epsilon g(\epsilon(n-ct)) + \epsilon^3\phi(\epsilon n))^3 \\
&\qquad - (\epsilon f(\epsilon(n-1+t)) + \epsilon g(\epsilon(n-1-ct)) + \epsilon^3\phi(\epsilon(n-1)))^3\Big),
\end{aligned} \tag{1.57}$$

and

$$\begin{aligned}
&\mathcal{B}_n(\epsilon f + \epsilon g + \epsilon^3\phi, \mathcal{U}) \\
&\quad = -\frac{1}{6}\Big(3(\epsilon f(\epsilon(n+t) + \epsilon g(\epsilon(n-ct)) + \epsilon^3\phi(\epsilon n))\mathcal{U}_n^2(t) \\
&\qquad - 3(\epsilon f(\epsilon(n-1+t) + \epsilon g(\epsilon(n-1-ct)) + \epsilon^3\phi(\epsilon(n-1)))\mathcal{U}_{n-1}^2(t) \\
&\qquad + \mathcal{U}_n^3(t) - \mathcal{U}_{n-1}^3(t)\Big).
\end{aligned} \tag{1.58}$$

The terms $\mathcal{U}$ and $\mathcal{Q}$ control the error associated with the ansatz in eqs. (1.46) and (1.47). Thus if these terms remain small in the $\ell^2(\mathbb{Z})$ norm, then the traveling wave ansatz will remain valid. In particular, if one has that $\|\mathcal{U}\|_\ell^2 \leq C\epsilon^{5/2}$, then the ansatz $\epsilon f + \epsilon g$ is valid up to order $\epsilon^{5/2}$ (since ϕ is uniformly bounded in norm and is thus $\mathcal{O}(1)$). Similarly, if $\mathcal{Q}$ is of order $\epsilon^{5/2}$, then one can show that $\dot{u}_n(t)$ is approximated by $\epsilon^2\partial_1 f + \epsilon^2\partial_1 g$ up to order $\epsilon^{5/2}$. Hence, controlling the norms of $\mathcal{U}$ and $\mathcal{Q}$ is sufficient in proving the approximation holds.

1.4 Preparatory Estimates

To control the dynamics of $\mathcal{U}$ and $\mathcal{Q}$, we need estimates of the residuals and the nonlinearity. We will frequently need to bound the $\ell^2(\mathbb{Z})$ of a term by the $H^1(\mathbb{R})$ norm of a function. To this end the following lemma proved in (Dumas and Pelinovsky,

2014) is useful.

Lemma 4. *There exists $C > 0$ such that for all $X \in H^1(\mathbb{R})$ and $\epsilon \in (0, 1)$,*

$$\|x\|_{\ell^2} \leq C\epsilon^{-1/2}\|X\|_{H^1},$$

where $x_n := X(\epsilon n)$, $n \in \mathbb{Z}$.

Lemma 5. *Let f and g be solutions of eqs.* (1.14) *and* (1.15)*, respectively, such that $f \in C([-\tau_0, \tau_0], \mathcal{X}_2^6)$ and $g \in C([-\tau_0, \tau_0], H_2^6)$. Let $\tau_0 > 0$ be fixed and $\delta > 0$ be as*

$$\delta := \max\left\{\sup_{\tau\in[-\tau_0,\tau_0]} \|f(\cdot,\tau)\|_{\mathcal{X}_2^6}, \sup_{\tau\in[-\tau_0,\tau_0]} \|g(\cdot,\tau)\|_{H_2^6}\right\} \tag{1.59}$$

Then there exists a δ-independent constant $C > 0$ such that the residual and nonlinear terms satisfy

$$\|\mathrm{Res}^{(1)}(t)\|_{\ell^2} + \|\mathrm{Res}^{(2)}(t)\|_{\ell^2} \leq C\epsilon^{11/2}(\delta + \delta^5) \tag{1.60}$$

and

$$\|\mathcal{B}_n(\epsilon f + \epsilon g + \epsilon^3\phi, \mathcal{U})\|_{\ell^2} \leq C\epsilon[(\delta + \epsilon^2\delta^3)\|\mathcal{U}\|_{\ell^2}^2 + \|\mathcal{U}\|_{\ell^2}^3] \tag{1.61}$$

for every $t \in [-\epsilon^{-3}\tau_0, \epsilon^{-3}\tau_0]$ and $\epsilon \in (0, 1)$.

Proof. We first focus on bounding $\mathrm{Res}^{(1)}(t)$. Looking at the terms in $\mathrm{Res}^{(1)}(t)$ involving f and F and using Taylor expansions and eq. (1.14), we get the following:

$$
\begin{array}{l}
\epsilon F(\cdot+\epsilon) - \epsilon F - \epsilon^2 \partial_1 f - \epsilon^4 \partial_2 f = \\
\qquad \begin{array}{lllll}
\epsilon^2 \partial_1 f + & \frac{\epsilon^3}{2}\partial_1^2 f + & \frac{\epsilon^4}{6}\partial_1^3 f & +\frac{\epsilon^5}{24}\partial_1^4 f & \\
 & -\frac{\epsilon^3}{2}\partial_1^2 f & -\frac{\epsilon^4}{4}\partial_1^3 f & -\frac{\epsilon^5}{12}\partial_1^4 f & \\
 & & +\frac{\epsilon^4}{8}\partial_1^3 f & +\frac{\epsilon^5}{16}\partial_1^4 f & \\
 & & -\frac{\epsilon^4}{12}\partial_1(f^3) & -\frac{\epsilon^5}{24}\partial_1^2(f^3) & \\
 & & & -\frac{\epsilon^5}{48}\partial_1^4 f & \\
 & & & +\frac{\epsilon^5}{24}\partial_1^2(f^3) & \\
-\epsilon^2 \partial_1 f & & & & \\
 & & +\frac{\epsilon^4}{12}\partial_1(f^3) & & \\
 & & -\frac{\epsilon^4}{24}\partial^3 f & & +I_{f,1}(n,t),
\end{array}
\end{array}
\tag{1.62}
$$

where $I_{f,1}$ contains the integral remainder terms:

$$
\begin{aligned}
I_{f,1}(n,t) := {} & \frac{\epsilon^6}{24}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(1-s)^4\,ds - \frac{\epsilon^6}{12}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(1-s)^3\,ds \\
& + \frac{\epsilon^6}{16}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(1-s)^2\,ds - \frac{\epsilon^6}{24}\int_0^1 \partial_1^3(f^3)(\epsilon(n+t+s))(1-s)^2\,ds \\
& - \frac{\epsilon^6}{48}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(1-s)\,ds + \frac{\epsilon^6}{24}\int_0^1 \partial_1^3(f^3)(\epsilon(n+t+s))(1-s)\,ds.
\end{aligned}
\tag{1.63}
$$

Note that all the terms in eq. (1.62) cancel except $I_{f,1}$, and so we are only left with terms of order ϵ^6. Applying lemma 4 (and lemmas 1 and 2 when needed) to the terms in eq. (1.63) gives that the ℓ^2 norm on the left-hand side of eq. (1.62) can be bounded by

$$C(\epsilon^{11/2}(\delta+\delta^3))$$

for some choice of constant $C > 0$.

Doing the same Taylor expansion for the g and G gives

$$
\begin{array}{lllll}
\multicolumn{5}{l}{\epsilon G(\cdot+\epsilon) - \epsilon G + \epsilon^2 c\partial_1 g - \epsilon^4 \partial_2 g =} \\
-\epsilon^2\partial_1 g & -\frac{\epsilon^3}{2}\partial_1^2 g & -\frac{\epsilon^4}{6}\partial_1^3 g & -\frac{\epsilon^5}{24}\partial_1^4 g & \\
 & +\frac{\epsilon^3}{2}\partial_1^2 g & +\frac{\epsilon^4}{4}\partial_1^3 g & +\frac{\epsilon^5}{12}\partial_1^4 g & \\
 & & +\frac{\epsilon^4 f_\infty^2}{4}\partial_1 g & +\frac{\epsilon^5 f_\infty^2}{8}\partial_1^2 g & \\
 & & +\frac{\epsilon^4}{12}\partial_1(g^3) & +\frac{\epsilon^5}{24}\partial_1^2(g^3) & \\
 & & +\frac{\epsilon^4}{12}\partial_1(3f_\infty g^2) & +\frac{\epsilon^5}{24}\partial_1^2(3f_\infty g^2) & \\
 & & -\frac{\epsilon^4}{8}\partial_1^3 g & -\frac{\epsilon^5}{16}\partial_1^4 g & \\
 & & & +\frac{\epsilon^5}{48}\partial_1^4 g & \\
 & & & -\frac{\epsilon^5}{24}\partial_1^2(g^3) & \\
 & & & -\frac{\epsilon^5}{24}\partial_1^2(3f_\infty g^2) & \\
 & & & -\frac{\epsilon^5 f_\infty^2}{8}\partial_1^2 g & \\
+\epsilon^2\partial_1 g & & & & \\
 & & -\frac{\epsilon^4 f_\infty^2}{4}\partial_1 g & & \\
 & & -\frac{\epsilon^4}{12}\partial_1(g^3) & & \\
 & & -\frac{\epsilon^4}{12}\partial_1(3f_\infty g^2) & & \\
 & & +\frac{\epsilon^4}{24}\partial_1^3 g & & +I_{g,1}(nt),
\end{array}
\tag{1.64}
$$

where $I_{g,1}$ contains the integral remainder terms.

$$
\begin{aligned}
&I_{g,1}(n,t) := \\
&-\frac{\epsilon^6}{24}\int_0^1 \partial_1^5 g(\epsilon(n-ct+s))(1-s)^4\,ds + \frac{\epsilon^6}{12}\int_0^1 \partial_1^5 g(\epsilon(n-ct+s))(1-s)^3\,ds \\
&+\frac{\epsilon^6 f_\infty^2}{8}\int_0^1 \partial_1^3 g(\epsilon(n-ct+s))(1-s)^2\,ds + \frac{\epsilon^6}{24}\int_0^1 \partial_1^3 (g^3)(\epsilon(n-ct+s))(1-s)^2\,ds \\
&+\frac{\epsilon^6}{24}\int_0^1 \partial_1^3 (3f_\infty g^2)(\epsilon(n-ct+s))(1-s)^2\,ds - \frac{\epsilon^6}{16}\int_0^1 \partial_1^5 g(\epsilon(n-ct+s))(1-s)^2\,ds \\
&+\frac{\epsilon^6}{48}\int_0^1 \partial_1^5 g(\epsilon(n-ct+s))(1-s)\,ds - \frac{\epsilon^6}{24}\int_0^1 \partial_1^3 (g^3)(\epsilon(n-ct+s))(1-s)\,ds \\
&-\frac{\epsilon^6}{24}\int_0^1 \partial_1^3 (3f_\infty g^2)(\epsilon(n-ct+s))(1-s)\,ds - \frac{\epsilon^6 f_\infty^2}{8}\int_0^1 \partial_1^3 g(\epsilon(n-ct+s))(1-s)\,ds
\end{aligned}
\tag{1.65}
$$

All terms except those of order ϵ^6 cancel and the terms in eq. (1.65) can be controlled by lemma 4.

Similarly we have

$$
\epsilon^3\Phi(\epsilon(n+1),\epsilon t) - \epsilon^3\Phi(\epsilon n,\epsilon t) - \epsilon^4\partial_2\phi_2(\epsilon n,\epsilon t) = \frac{\epsilon^6}{2}\int_0^1 \partial_1^2\psi(\epsilon(n+s),\epsilon t)(1-s)^2\,ds, \tag{1.66}
$$

so the ℓ^2 norm can also be controlled.

Therefore we have

$$
\|\mathrm{Res}^{(1)}(t)\|_{\ell^2} \le C\epsilon^{11/2}(\delta+\delta^3) \tag{1.67}
$$

The bound on $\mathrm{Res}^{(2)}(t)$ can be approached similarly. Focusing on the terms with

f and F in $\mathrm{Res}^{(2)}(t)$, we have

$$
\begin{aligned}
&\epsilon f(\cdot) - \epsilon f(\cdot - \epsilon) - \epsilon^2 \partial_1 F - \epsilon^4 \partial_2 F - \frac{\epsilon^3}{6}(f^3(\cdot) - f^3(\cdot - \epsilon)) = \\
&\quad \epsilon^2 \partial_1 f \quad -\frac{\epsilon^3}{2}\partial_1 f \quad +\frac{\epsilon^4}{6}\partial_1^3 f \qquad\qquad\qquad -\frac{\epsilon^5}{24}\partial_1^4 f \\
&\quad -\epsilon^2 \partial_1 f \quad +\frac{\epsilon^3}{2}\partial_1^2 f \quad +\frac{\epsilon^4}{12}\partial_1(f^3) - \frac{\epsilon^4}{8}\partial_1^3 f \quad +\frac{\epsilon^5}{48}\partial_1^4 f - \frac{\epsilon^5}{24}\partial_1^2(f^3) \\
&\qquad\qquad\qquad\qquad\quad -\epsilon^4 \partial_2 f \qquad\qquad\qquad\quad +\frac{\epsilon^5}{2}\partial_1 \partial_2 f \\
&\qquad\qquad\qquad\qquad\quad -\frac{\epsilon^4}{6}\partial_1(f^3) \qquad\qquad\quad +\frac{\epsilon^5}{12}\partial_1(f^3) \qquad\qquad + I_{f,2}(n,t).
\end{aligned}
\tag{1.68}
$$

where the integral remainder terms and the other terms of order ϵ^6 are contained in $I_{f,2}$:

$$
\begin{aligned}
I_{f,2}(n,t) := &-\frac{\epsilon^6}{24}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(s-1)^4\,ds \\
&+\frac{\epsilon^6}{12}\int_0^1 \partial_1^2(f^3)(\epsilon(n+t+s))(s-1)^2\,ds \\
&+\epsilon^6 \partial_2\left(\frac{1}{8}\partial_1^2 f - \frac{1}{12}f^3 - \frac{\epsilon}{48}\partial_1^3 f + \frac{\epsilon}{8}f^2\partial_1 f\right)
\end{aligned}
\tag{1.69}
$$

All the above terms in eq. (1.68) cancel except for $I_{f,2}(n,t)$. The integral terms in eq. (1.69) can be controlled like before. The non-integral term can be controlled by first evaluating the derivative in time, ∂_2, and replacing the terms $\partial_2 f$ using eq. (1.14); then the terms can be controlled by lemma 4. Then the left-hand side of eq. (1.68) can be bounded by a term of the form

$$C\epsilon^{11/2}(\delta + \delta^3).$$

Taylor expanding the remaining terms in $\mathrm{Res}^{(2)}(t)$ leads to

$$
\begin{array}{llll}
\epsilon^2\partial_1 g & -\dfrac{\epsilon^3}{2}\partial_1^2 g & +\dfrac{\epsilon^4}{6}\partial_1^3 g & -\dfrac{\epsilon^5}{24}\partial_1^4 g \\
 & & +\epsilon^4\partial_1\phi & -\dfrac{\epsilon^5}{2}\partial_1^2\phi \\
-\epsilon^2\partial_1 g & +\dfrac{\epsilon^3}{2}\partial_1^2 g & -\dfrac{\epsilon^4}{8}\partial_1^3 g + \dfrac{\epsilon^4 f_\infty^2}{4}\partial_1 g & +\dfrac{\epsilon^5}{48}\partial_1^4 g \\
 & & +\dfrac{\epsilon^4}{12}\partial_1(g^3+3f_\infty g^2) & -\dfrac{\epsilon^5}{24}\partial_1^2(g^3+3f_\infty g^2) \\
 & & & -\dfrac{\epsilon^5 f_\infty^2}{8}\partial_1^2 g \\
 & & +\dfrac{\epsilon^4 f_\infty^2}{4}\partial_1 g & -\dfrac{\epsilon^5 f_\infty^2}{8}\partial_1^2 g \\
 & & +\epsilon^4\partial_2 g & -\dfrac{\epsilon^5}{2}\partial_1\partial_2 g \\
 & & -\epsilon^4\partial_2\partial_1^{-1}\psi & +\dfrac{\epsilon^5}{2}\partial_2\psi \\
 & & -\dfrac{\epsilon^4}{6}\partial_1(g^3+3g^2f+3gf^2) & +\dfrac{\epsilon^5}{12}\partial_1^2(g^3+3g^2f+3gf^2),
\end{array}
\tag{1.70}
$$

where the integral remainder terms and other terms of order ϵ^6 are contained in $I_{g,2}$:

$$
\begin{aligned}
&I_{g,2}(n,t) = \\
&-\frac{\epsilon^6}{24}\int_0^1 \partial_1^5 g(\epsilon(n-s-ct))(s-1)^4\,ds - \frac{\epsilon^6}{2}\int_0^1 \partial_1^3\phi(\epsilon(n-s))(s-1)^2\,ds \\
&-\frac{\epsilon^6 f_\infty^2}{4}\partial_1\left(\frac{f_\infty^2}{4}g + \frac{1}{12}(g^3+3f_\infty g^2) - \frac{1}{8}\partial_1^2 g + \frac{\epsilon}{48}\partial_1^3 g - \frac{\epsilon}{24}\partial_1(g^3+3f_\infty g^2) - \frac{\epsilon f_\infty^2}{8}\partial_1 g\right) \\
&-\epsilon^6\partial_2\left(\frac{f_\infty^2}{4}g + \frac{1}{12}(g^3+3f_\infty g^2) - \frac{1}{8}\partial_1^2 g + \frac{\epsilon}{48}\partial_1^3 g - \frac{\epsilon}{24}\partial_1(g^3+3f_\infty g^2) - \frac{\epsilon f_\infty^2}{8}\partial_1 g\right) \\
&+\frac{\epsilon^6}{12}\int_0^1 \partial_1^3(g^3(\epsilon(n-s-ct)))(s-1)^2 ds \\
&+\frac{\epsilon^6}{12}\int_0^1 \partial_1^3(3g^2(\epsilon(n-s-ct))f(\epsilon(n-s+t)))(s-1)^2\,ds \\
&+\frac{\epsilon^6}{12}\int_0^1 \partial_1^3(3g(\epsilon(n-s-ct))f^2(\epsilon(n-s+t)))(s-1)^2\,ds
\end{aligned}
\tag{1.71}
$$

The terms in eq. (1.70) of order ϵ^3 or lower cancel out. The terms of order ϵ^4 are equal to

$$-\partial_2\partial_1^{-1}\psi + \partial_1\phi - \frac{1}{6}\partial_1(3(f^2-f_\infty^2)g + 3(f-f_\infty)g^2). \tag{1.72}$$

Formally applying ∂_1 implies that the above terms should be constant in space since $\partial_2\psi = \partial_2^2\phi$ satisfies eq. (1.16). However, one should be careful with this calculation due to the differences in scaling of the spatial variables: for example, ϕ and ψ's spatial variable is rescaled to ϵn while f's is rescaled to $\epsilon(n+t)$. Taking a derivative with respect to $\xi = \epsilon x$ gives that eq. (1.72) must be constant. Since all the terms decay to zero at spatial infinity, eq. (1.72) is exactly zero.

The terms of order ϵ^5 can be rewritten as

$$\frac{1}{4}\partial_1(-2\partial_2 g - \frac{1}{12}\partial_1^3 g + \frac{1}{6}(g^3+3f_\infty g^2)) + \frac{1}{2}(\partial_2^2\phi - \partial_1^2\phi + \frac{1}{6}\partial_1^2(3(f-f_\infty)g^2 + 3(f^2-f_\infty^2)g)) \tag{1.73}$$

which is equal to zero since g and ϕ satisfy the PDEs in eqs. (1.15) and (1.16). Thus the right-hand side of eq. (1.70) is equal to $I_{g,2}$. The integral terms in eq. (1.71) are bounded as before. The remaining terms in eq. (1.71) can be bounded by evaluating

$\partial_2 g$ using eq. (1.15) and then applying lemma 4. We can the get the following bound:

$$\|\mathrm{Res}^{(2)}(t)\|_{\ell^2} \leq C\epsilon^{11/2}(\delta + \delta^3 + \delta^5).$$

Interpolating between powers of δ gives the desired inequality eq. (1.60).

The proof of eq. (1.61) follows immediately. □

To proceed, we construct and energy function for eq. (1.55) to control the ℓ^2 norms of $\mathcal{U}$ and $\mathcal{Q}$. Lemma 5 essentially states that $\mathrm{Res}^{(1)}(t)$, $\mathrm{Res}^{(2)}(t)$, and $\mathcal{B}$ remain appropriately small. If one drops the residual and nonlinear terms from eq. (1.55), then we are left with a linear (non-autonomous) Hamiltonian system. Hence, an appropriate choice of an energy function would simply be the Hamiltonian for this reduced system (as suggested in our earlier proof sketch). Define

$$\mathcal{E}(t) = \frac{1}{2}\sum_{n\in\mathbb{Z}} \mathcal{Q}_n^2(t) + \mathcal{U}_n^2(t) - \frac{1}{2}\left(\epsilon f(\epsilon(n+t), \epsilon^3 t) + \epsilon g(\epsilon(n-ct, \epsilon^3 t) + \epsilon^3\phi(\epsilon n, \epsilon t)\right)^2 \mathcal{U}_n^2(t) \tag{1.74}$$

The following lemma gives us that $\mathcal{E}$ can be used to control $\mathcal{U}$ and $\mathcal{Q}$.

Lemma 6. *Fix $\tau_0 > 0$ and let δ be given by eq. (1.59) . There exists $\epsilon_0 = \epsilon_0(\delta) > 0$ sufficiently small such that for every $\epsilon \in (0, \epsilon_0)$ and for every local solution $(\mathcal{U}, \mathcal{Q}) \in C^1([-\tau_0\epsilon^{-3}, \tau_0\epsilon^{-3}], \ell^2(\mathbb{Z}))$ of eq. (1.55), the energy-type quantity given in eq. (1.74) is coercive with the bound*

$$\|\mathcal{Q}(t)\|_{\ell^2}^2 + \|\mathcal{U}(t)\|_{\ell^2}^2 \leq 4\mathcal{E}(t), \quad \textit{for } t \in (-\tau_0\epsilon^{-3}, \tau_0\epsilon^{-3}). \tag{1.75}$$

Moreover, there exists $C > 0$ independent of ϵ and δ such that

$$\left|\frac{d\mathcal{E}}{dt}\right| \leq C\mathcal{E}^{1/2}\left[\epsilon^{11/2}(\delta + \delta^5) + \epsilon^3\delta^2\mathcal{E}^{1/2} + \epsilon(\delta + \mathcal{E}^{1/2})\mathcal{E}\right] \tag{1.76}$$

for every $t \in [-\tau_0\epsilon^{-3}, \tau_0\epsilon^{-3}]$ and $\epsilon \in (0, \epsilon_0)$.

Proof. Note that $\delta > 0$ can be used to control the $L^\infty(\mathbb{R})$ norms of f, g, and ψ. Thus we can choose ϵ_0 small enough so that for $\epsilon \in (0, \epsilon_0)$ we have

$$1 - \frac{1}{2}\left(\epsilon\|f\|_{L^\infty} + \epsilon\|g\|_{L^\infty} + \epsilon^3\|\phi\|_{L^\infty}\right)^2 \geq \frac{1}{2}, \tag{1.77}$$

independent on the particular choices of f and g. Hence

$$\mathcal{E}(t) \geq \frac{1}{2}\|\mathcal{Q}\|_{\ell^2}^2 + \frac{1}{4}\|\mathcal{U}\|_{\ell^2}^2 \geq \frac{1}{4}\|\mathcal{Q}\|_{\ell^2}^2 + \frac{1}{4}\|\mathcal{U}\|_{\ell^2}^2 \tag{1.78}$$

and eq. (1.75) follows.

Now we take the time derivative of $\mathcal{E}$ to get that

$$\begin{aligned}\frac{d\mathcal{E}}{dt} = \sum_{n\in\mathbb{Z}} & \mathcal{Q}_n(t)\mathrm{Res}_n^{(2)}(t) + \mathcal{Q}_n(t)\mathcal{B}_n(\epsilon f + \epsilon g + \epsilon^3\phi, \mathcal{U}(t)) \\ & + \mathcal{U}_n(t)\mathrm{Res}_n^{(1)}(t)\left(1 - \frac{1}{2}(\epsilon f + \epsilon g + \epsilon^3\phi)^2\right) \\ & + \mathcal{U}_n^2(t)(\epsilon f + \epsilon g + \epsilon^3\phi) \times (\epsilon^2\partial_1 f + \epsilon^4\partial_2 f - \epsilon^2 c\partial_1 g + \epsilon^4\partial_2 g + \epsilon^4\partial_2\phi).\end{aligned} \tag{1.79}$$

Then using the Cauchy inequality and the Hölder inequality for $p = 1$ and $q = \infty$ we get that

$$\begin{aligned}\left|\frac{d\mathcal{E}}{dt}\right| \leq & \|\mathcal{Q}\|_{\ell^2} \times \|\mathrm{Res}^{(2)}(t)\|_{\ell^2} + \|\mathcal{Q}\|_{\ell^2} \times \|\mathcal{B}\|_{\ell^2} + \|\mathcal{U}\|_{\ell^2} \times \|\mathrm{Res}_n^{(1)}(t)\|_{\ell^2} \\ & + \|\mathcal{U}^2\|_{\ell^1} \times \|(\epsilon f + \epsilon g + \epsilon^3\phi) \times (\epsilon^2\partial_1 f + \epsilon^4\partial_2 f - \epsilon^2 c\partial_1 g + \epsilon^4\partial_2 g + \epsilon^4\partial_2\phi)\|_{\ell^\infty}.\end{aligned} \tag{1.80}$$

Note that if $a \in \ell^2$, then $a \in \ell^\infty$ and $\|a\|_{\ell^\infty} \leq \|a\|_{\ell^2}$. Thus we can replace the ℓ^∞ norms above with ℓ^2 norms. Using the results in lemma 5, we thus have

$$\begin{aligned}\left|\frac{d\mathcal{E}}{dt}\right| \leq & C\Big[\mathcal{E}^{1/2}\epsilon^{11/2}(\delta + \delta^5) + \mathcal{E}^{1/2}\epsilon[(\delta + \epsilon^2\delta^3)\mathcal{E} + \mathcal{E}^{3/2}] \\ & + \mathcal{E}(\epsilon^3\delta^2 + \epsilon^5\delta^2 + \epsilon^5\delta^4 + \epsilon^7\delta^4 + \epsilon^7\delta^6)\Big],\end{aligned} \tag{1.81}$$

where the $C > 0$ is independent of ϵ and δ. The right-hand side of the above inequality can be simplified by taking ϵ_0 smaller. That is, taking ϵ_0 sufficiently small (dependent on δ), we can absorb higher orders of ϵ into lower orders. For example, $\epsilon^3\delta^2 + \epsilon^5\delta^2 \leq 2\epsilon^3\delta^2$ for ϵ small enough. Thus we arrive at

$$\left|\frac{d\mathcal{E}}{dt}\right| \leq C\mathcal{E}^{1/2}\left[\epsilon^{11/2}(\delta + \delta^5) + \epsilon^3\delta^2\mathcal{E}^{1/2} + \epsilon(\delta + \mathcal{E}^{1/2})\mathcal{E}\right] \tag{1.82}$$

as desired. □

Lastly, before we can prove our main result, we must show that for appropriate initial conditions that $\mathcal{U}(0)$ and $\mathcal{Q}(0)$ are suitably small. In particular, we want our initial conditions to be "close to" the traveling wave ansatz in the sense that

$$u_n(0) \approx \epsilon f(\epsilon n, 0) + \epsilon g(\epsilon n, 0) \tag{1.83}$$

and

$$\dot{u}_n(0) \approx \epsilon \partial_1 f(\epsilon n, 0) - \epsilon^2 g(\epsilon n, 0) \tag{1.84}$$

where the higher-order ϵ terms are neglected. Recall that we assume ϕ and $\partial_1 \phi$ to have initial conditions exactly equal to zero, so those terms drop. A seemingly appropriate notion of "closeness" would be in the ℓ^2 norm, as used in (Khan and Pelinovsky, 2017; Schneider and Wayne, 2000). However, since $q_n(0) = \sum_{k=-\infty}^{n-1} \dot{u}_k(0)$, we may lose some decay due to the summation and $\mathcal{Q}(0)$ will not be in ℓ^2. To counter this, we need some extra localization assumptions on $\dot{u}_n(0)$.

Assumption 3. *Suppose that the initial conditions for u satisfy*

$$\|u(0) - \epsilon f(\epsilon \cdot, 0) - \epsilon g(\epsilon \cdot, 0)\|_{\ell^2} + \|\dot{u}(0) - \epsilon^2 \partial_1 f(\epsilon \cdot, 0) + \epsilon^2 \partial g(\epsilon \cdot, 0)\|_{\ell^2_2} \leq \epsilon^{5/2} \tag{1.85}$$

and that $f(\cdot, 0) \in \mathcal{X}_2^6$ and $g(\cdot, 0) \in H_2^6$

The ℓ^2_2 norm will be sufficient to get that the summation is in ℓ^2 based on the following lemma.

Lemma 7. *If $a \in \ell^2_2(\mathbb{Z})$ and*

$$\sum_{k=-\infty}^{n} a_k = 0, \tag{1.86}$$

then $b_n = \sum_{k=-\infty}^{n} a_k$ is in $\ell^2(\mathbb{Z})$ and

$$\|b\|_{\ell^2} \leq C \|a\|_{\ell^2_2} \tag{1.87}$$

for some $C > 0$ independent of a.

See appendix A for proof.

We can now show the following.

Lemma 8. *Let assumptions 2 and 3 hold. Then* $\mathcal{U}(0), \mathcal{Q}(0) \in \ell^2(\mathbb{Z})$ *satisfy*

$$\dot{u}_n(0) = q_{n+1}(0) - q_n(0) \tag{1.88}$$

and

$$\|\mathcal{U}(0)\|_{\ell^2} + \|\mathcal{Q}(0)\|_{\ell^2} \leq C\epsilon^{5/2} \tag{1.89}$$

with $C > 0$ *independent of* ϵ.

Proof. That $\|\mathcal{U}(0)\|_{\ell^2} \leq C\epsilon^{5/2}$ follows immediately from applying assumption 3 to eq. (1.46).

For $q_n(0)$ to satisfy eq. (1.88), it must equal $\sum_{k=-\infty}^{n-1} \dot{u}_k(0)$ (modulo a constant which we assume without loss of generality to be zero). Thus we have

$$\begin{aligned} q_n(0) &= \sum_{k=-\infty}^{n-1} \dot{u}_k(0) \\ &= \sum_{k=-\infty}^{n-1} \left[\dot{u}_k(0) - \epsilon^2 \partial_1 f(\epsilon k, 0) - \epsilon^4 \partial_1 f(\epsilon k, 0) + \epsilon^2 c \partial_1 g(\epsilon k, 0) - \epsilon^4 \partial_2 g(\epsilon k, 0)\right] \\ &\quad + \sum_{k=-\infty}^{n-1} \left[\epsilon^2 \partial_1 f(\epsilon k, 0) + \epsilon^4 \partial_1 f(\epsilon k, 0) - \epsilon F(\epsilon(k+1), 0) + \epsilon F(\epsilon k, 0)\right] \\ &\quad + \sum_{k=-\infty}^{n-1} \left[-\epsilon^2 c \partial_1 g(\epsilon k, 0) + \epsilon^4 \partial_1 g(\epsilon k, 0) - \epsilon G(\epsilon(k+1), 0) + \epsilon G(\epsilon k, 0)\right] \\ &\quad + \epsilon F(\epsilon n, 0) - \epsilon F_{-\infty} + \epsilon G(\epsilon n, 0). \end{aligned} \tag{1.90}$$

Comparing eq. (1.90) to eq. (1.47), we have that

$$\begin{aligned} \mathcal{Q}_n(0) &= \sum_{k=-\infty}^{n-1} \left[\dot{u}_k(0) - \epsilon^2 \partial_1 f(\epsilon k, 0) - \epsilon^4 \partial_1 f(\epsilon k, 0) + \epsilon^2 c \partial_1 g(\epsilon k, 0) - \epsilon^4 \partial_2 g(\epsilon k, 0)\right] \\ &\quad + \sum_{k=-\infty}^{n-1} \left[\epsilon^2 \partial_1 f(\epsilon k, 0) + \epsilon^4 \partial_1 f(\epsilon k, 0) - \epsilon F(\epsilon(k+1), 0) + \epsilon F(\epsilon k, 0)\right] \\ &\quad + \sum_{k=-\infty}^{n-1} \left[-\epsilon^2 c \partial_1 g(\epsilon k, 0) + \epsilon^4 \partial_1 g(\epsilon k, 0) - \epsilon G(\epsilon(k+1), 0) + \epsilon G(\epsilon k, 0)\right]. \end{aligned} \tag{1.91}$$

That $\mathcal{Q}_n(0) \to 0$ as $n \to \infty$ is guaranteed by assumption 2. Now lemma 7 can be applied to get the result if the summands are in ℓ_2^2 and of order $\epsilon^{5/2}$. The first summand satisfies this condition because of assumption 3. Note that the latter summands are equal to $-I_{f,1}(k,0)$ and $-I_{g,1}(k,0)$, as defined in eqs. (1.63) and (1.65). This follows from the earlier calculations in lemma 5. That $I_{f,1}(k,0)$ and $I_{g,2}(k,0)$ are elements of ℓ^2_2 follows from $f(\cdot,0) \in \mathcal{X}^6_2$ and $g(\cdot,0) \in H^6_2$ and an application of lemma 4.

Thus we have eq. (1.89) where the $C > 0$ can be chosen based on the norms of f and g. $\square$

1.5 Proof of Long-Time Stability

Now with the setup complete, the main result of this chapter can be shown. The result and proof are analogous to those of (Khan and Pelinovsky, 2017, Thm. 1).

Theorem 1. *Let assumption 1 hold and set*

$$\delta = \max\left\{ \sup_{\tau\in\mathbb{R}} \|f(\cdot,\tau)\|_{\mathcal{X}^6_2},\ \sup_{\tau\in\mathbb{R}} \|g(\cdot,\tau)\|_{H^6_2} \right\} \tag{1.92}$$

For fixed $r \in (0, 1/2)$, there exists positive constants ϵ_0, C, and K such that for all $\epsilon \in (0, \epsilon_0)$, when initial data $(u(0), \dot{u}(0))$ satisfy assumptions 2 and 3, the unique solution (u, q) to the FPU equation eq. (1.42) belongs to

$$C^1([-t_0(\epsilon), t_0(\epsilon)], \ell^\infty(\mathbb{Z})) \tag{1.93}$$

with $t_0(\epsilon) := rK^{-1}\epsilon^{-3}|\log(\epsilon)|$ and satisfies

$$\begin{aligned} &\|u(t) - \epsilon f(\epsilon(\cdot + t), \epsilon^3 t) - \epsilon g(\epsilon(\cdot - ct), \epsilon^3 t)\|_{\ell^2} \\ &\quad + \|\dot{u}(t) - \epsilon\partial_1 f(\epsilon(\cdot + t), \epsilon^3 t) + \epsilon^2 \partial_1 g(\epsilon(\cdot - ct), \epsilon^3 t)\|_{\ell^2} \le C\epsilon^{5/2-r}, \quad t \in [-t_0(\epsilon), t_0(\epsilon)]. \end{aligned} \tag{1.94}$$

Proof. Set $\mathcal{S} := \mathcal{E}^{1/2}$ where $\mathcal{E}$ is defined in eq. (1.74). From the results in lemma 8, we get that $\mathcal{S}(0) \le C_0\epsilon^{5/2}$ for some constant $C_0 > 0$ and ϵ_0 as chosen in lemma 6. For fixed constants $r \in (0, 1/2)$, $C > C_0$, and $K > 0$, define the maximal continuation time by

$$T_{C,K,r} := \sup\left\{ T_0 \in (0, rK^{-1}\epsilon^{-3}|\log(\epsilon)|] : \mathcal{S}(t) \le C\epsilon^{5/2-r}, t \in [-T_0, T_0] \right\}. \tag{1.95}$$

We also define the maximal evolution time of the mKdV equation as $\tau_0(\epsilon) = rK^{-1}|\log(\epsilon)|$. The goal is then to pick C and K so that $T_{C,K,r} = \epsilon^{-3}\tau_0(\epsilon)$.

We have that

$$\begin{aligned}\left|\frac{d}{dt}\mathcal{S}(t)\right| &= \frac{1}{2\mathcal{E}^{1/2}}\left|\frac{d}{dt}\mathcal{E}(t)\right| \\ &\leq C_1(\delta+\delta^5)\epsilon^{11/2} + C_2\epsilon^3\left[\delta^2+\epsilon^{-2}(\delta+\mathcal{S})\mathcal{S}\right]\mathcal{S}\end{aligned} \tag{1.96}$$

where $C_1, C_2 > 0$ are independent of δ and ϵ. While $|t| \leq T_{C,K,r}$,

$$C_2\left[\delta^2+\epsilon^{-2}(\delta+\mathcal{S})\mathcal{S}\right] \leq C_2\left[\delta^2+\epsilon^{-2}(\delta+C\epsilon^{11/2-r})C\epsilon^{11/2-r}\right], \tag{1.97}$$

where the right-hand side is continuous in ϵ for $\epsilon \in [0, \epsilon_0]$. Thus the right-hand side can be uniformly bounded by a constant independent of ϵ. Choose $K > 0$ (dependent on C) sufficiently large so that

$$C_2\left[\delta^2+\epsilon^{-2}(\delta+C\epsilon^{11/2-r})C\epsilon^{11/2-r}\right] \leq K. \tag{1.98}$$

Hence, we can get that for $t \in [-T_{C,K,r}, T_{C,K,r}]$

$$\begin{aligned}\frac{d}{dt}e^{-\epsilon^3Kt}\mathcal{S}(t) &= -\epsilon^3Ke^{-\epsilon^3Kt}\mathcal{S} + e^{-\epsilon^3Kt}\frac{d}{dt}\mathcal{S} \\ &\leq -\epsilon^3Ke^{-\epsilon^3Kt}\mathcal{S} + e^{-\epsilon^3Kt}C_1(\delta+\delta^5)\epsilon^{11/2} \\ &\qquad + e^{-\epsilon^3Kt}C_2\epsilon^3\left[\delta^2+\epsilon^{-2}(\delta+\mathcal{S})\mathcal{S}\right]\mathcal{S} \\ &\leq -\epsilon^3Ke^{-\epsilon^3Kt}\mathcal{S} + e^{-\epsilon^3Kt}C_1(\delta+\delta^5)\epsilon^{11/2} + \epsilon^3Ke^{-\epsilon^3Kt}\mathcal{S} \\ &= e^{-\epsilon^3Kt}C_1(\delta+\delta^5)\epsilon^{11/2}.\end{aligned} \tag{1.99}$$

Integrating gives

$$\begin{aligned}\mathcal{S}(t) &\leq \left(\mathcal{S}(0)+K^{-1}C_1(\delta+\delta^5)\epsilon^{5/2}\right)e^{\epsilon^3Kt} - \epsilon^{5/2}K^{-1}C_1(\delta+\delta^5) \\ &\leq \left(\mathcal{S}(0)+K^{-1}C_1(\delta+\delta^5)\epsilon^{5/2}\right)e^{\epsilon^3Kt} \\ &\leq \left(\mathcal{S}(0)+K^{-1}C_1(\delta+\delta^5)\epsilon^{5/2}\right)e^{K\tau_0(\epsilon)} \\ &\leq \left(C_0+K^{-1}C_1(\delta+\delta^5)\right)\epsilon^{5/2-r}\end{aligned} \tag{1.100}$$

for $t \in [-T_{C,K,r}, T_{C,K,r}]$, where the last line follows in part from the definition of $\tau_0(\epsilon)$. Now choose $C > C_0$ sufficiently large so that

$$C_0 + K^{-1}C_1(\delta+\delta^5) \leq C. \tag{1.101}$$

Note that our earlier choice of K can be enlarged so that eq. (1.98) still holds as well as the above inequality. Therefore, with these choices of C and K, the maximal interval can be extended to $T_{C,K,r} = \epsilon^{-3}\tau_0(\epsilon)$. $\square$

Appendix A

Proofs of lemmas

Lemma 1. *For non-negative integers k, there is a $C > 0$ such that*

$$\|fg\|_{H^k} \leq C\|f\|_{\mathcal{X}^k}\|g\|_{H^k} \tag{1.22}$$

for any $f \in \mathcal{X}^k(\mathbb{R})$ and $g \in H^k(\mathbb{R})$.

Proof. The result follows from induction on k.

For $k = 0$, we have

$$\|fg\|_{H^0} \leq \|f\|_{L^\infty}\|g\|_{H^0}. \tag{A.1}$$

Assuming eq. (1.22) holds for $k \geq 0$, we have that

$$\begin{aligned} \|fg\|_{H^{k+1}} &\leq C\left(\|fg\|_{H^k} + \|\partial^{k+1}(fg)\|_{L^2}\right) \\ &\leq C\left(\|f\|_{\mathcal{X}^k}\|g\|_{H^k} + \|\partial^{k+1}(fg)\|_{L^2}\right), \end{aligned}$$

where the second term can be bounded by

$$\begin{aligned} \|\partial^{k+1}(fg)\|_{L^2} &\leq \|\partial^k(\partial^1 fg)\|_{L^2} + \|\partial^k(f\partial^1 g)\|_{L^2} \\ &\leq \|\partial^1 fg\|_{H^k} + \|f\partial^1 g\|_{H^k} \\ &\leq \|\partial^1 f\|_{H^k}\|g\|_{H^k} + \|f\|_{\mathcal{X}^k}\|\partial^1 g\|_{H^k} \\ &\leq \|f\|_{\mathcal{X}^{k+1}}\|g\|_{H^{k+1}} + \|f\|_{\mathcal{X}^{k+1}}\|g\|_{H^{k+1}} \\ &= 2\|f\|_{\mathcal{X}^{k+1}}\|g\|_{H^{k+1}}. \end{aligned}$$

This completes the induction. □

Lemma 2. *For non-negative integers k, there is a $C > 0$ such that*

$$\|fg\|_{\mathcal{X}^k} \leq C\|f\|_{\mathcal{X}^k}\|h\|_{\mathcal{X}^k} \tag{1.23}$$

for any $f, g \in \mathcal{X}^k(\mathbb{R})$.

Proof. Using the result from lemma 1, we have

$$\begin{aligned} \|fg\|_{\mathcal{X}^k} &\le \|fg\|_{L^\infty} + \|(fg)'\|_{H^{k-1}} \\ &\le \|f\|_{L^\infty}\|g\|_{L^\infty} + \|f'g\|_{H^{k-1}} + \|fg'\|_{H^{k-1}} \\ &\le \|f\|_{L^\infty}\|g\|_{L^\infty} + C\|f'\|_{H^{k-1}}\|g\|_{\mathcal{X}^{k-1}} + \|f\|_{\mathcal{X}^{k-1}}\|g'\|_H^{k-1} \\ &\le C\|f\|_{\mathcal{X}^k}\|g\|_{\mathcal{X}^k}. \end{aligned} \tag{A.2}$$

□

Lemma 3. *For each $k \ge 0$ and $c > 0$, there exists $C > 0$ depending only on k such that*

$$\left\| \frac{1}{\langle \cdot + \tau\rangle_+^2 \langle \cdot - c\tau\rangle^2} \right\|_{C^k} \le C \sup_{x\in\mathbb{R}} \frac{1}{\langle x + \tau\rangle_+^2 \langle x - c\tau\rangle^2}. \tag{1.34}$$

Furthermore,

$$\int_0^\infty \sup_{x\in\mathbb{R}} \frac{1}{\langle x + \tau\rangle_+^2 \langle x - c\tau\rangle^2}\, d\tau < \infty. \tag{1.35}$$

Proof. The main argument of the proof is given by showing the following claim holds:

Claim: For each integer $k \ge 0$,

$$\frac{\partial^k}{\partial x^k}\left[\frac{1}{\langle x + \tau\rangle_+^2 \langle x - c\tau\rangle^2}\right]$$

is a sum of terms of the form

$$\frac{C}{\langle x + \tau\rangle_+^{2+m} \langle x - c\tau\rangle^{2+m}} \langle x + \tau\rangle_+^{m_1} \langle x - c\tau\rangle^{m_2} F(x,\tau), \tag{A.3}$$

where $C \ne 0$ is a constant, m, m_1, m_2 are integers, $0 \le m_1, m_2 \le m$, and $F \in C_b^n(\mathbb{R}\times\mathbb{R})$ for every $n \in \mathbb{N}$.

This can be proved inductively. We have the $k = 0$ case immediately by setting $C = 1$, $m = m_1 = m_2 = 0$, and $F(x) = 1$. Now we assume that the claim holds for $k \ge 0$. To get the form of the $(k+1)^{\text{st}}$ derivative, we can use linearity and look at the derivative of each term of the form eq. (A.3). That is, the $(k+1)^{\text{st}}$ derivative is a sum of terms of the form

$$\frac{\partial}{\partial x}\left[\frac{C}{\langle x + \tau\rangle_+^{2+m} \langle x - c\tau\rangle^{2+m}} \langle x + \tau\rangle_+^{m_1} \langle x - c\tau\rangle^{m_2} F(x,\tau)\right]. \tag{A.4}$$

Applying the product rule to eq. (A.4) gives us

$$\begin{aligned}
\frac{\partial}{\partial x}\left[\frac{C}{\langle x+\tau\rangle_+^{2+m}\langle x-c\tau\rangle^{2+m}}\langle x+\tau\rangle_+^{m_1}\langle x-c\tau\rangle^{m_2}F(x,\tau)\right] = \\
\underbrace{\frac{\partial}{\partial x}\left[\frac{C}{\langle x+\tau\rangle_+^{2+m}\langle x-c\tau\rangle^{2+m}}\right]\langle x+\tau\rangle_+^{m_1}\langle x-c\tau\rangle^{m_2}F(x,\tau)}_{I} \\
+\underbrace{\frac{C}{\langle x+\tau\rangle_+^{2+m}\langle x-c\tau\rangle^{2+m}}\frac{\partial}{\partial x}\left[\langle x+\tau\rangle_+^{m_1}\right]\langle x-c\tau\rangle^{m_2}F(x,\tau)}_{II} \\
+\underbrace{\frac{C}{\langle x+\tau\rangle_+^{2+m}\langle x-c\tau\rangle^{2+m}}\langle x+\tau\rangle_+^{m_1}\frac{\partial}{\partial x}\left[\langle x-c\tau\rangle^{m_2}\right]F(x,\tau)}_{III} \\
+\underbrace{\frac{C}{\langle x+\tau\rangle_+^{2+m}\langle x-c\tau\rangle^{2+m}}\langle x+\tau\rangle_+^{m_1}\langle x-c\tau\rangle^{m_2}\frac{\partial}{\partial x}\left[F(x,\tau)\right].}_{IV}
\end{aligned}$$

We now go term-by-term. For the first term, we have

$$\begin{aligned}
I =& \frac{-(2+m)C}{\langle x+\tau\rangle_+^{2+(m+1)}\langle x-c\tau\rangle^{2+(m+1)}}\langle x+\tau\rangle_+^{m_1+1}\langle x+\tau\rangle^{m_2}\Big(\langle x-c\tau\rangle_+'F(x,\tau)\Big) \\
&-\frac{(2+m)C}{\langle x+\tau\rangle_+^{2+(m+1)}\langle x-c\tau\rangle^{2+(m+1)}}\langle x+\tau\rangle_+^{m_1}\langle x-c\tau\rangle^{m_2+1}\Big(\langle x-c\tau\rangle'F(x,\tau)\Big),
\end{aligned}$$

where $\langle\cdot\rangle'$ denotes the derivative of $\langle\cdot\rangle$. It's clear that both of these are of the form in eq. (A.3).

Also, we have

$$II = \frac{Cm_1}{\langle x+\tau\rangle_+^{2+m}\langle x-c\tau\rangle^{2+m}}\langle x+\tau\rangle_+^{m_1-1}\langle x-c\tau\rangle^{m_2}\Big(\langle x+\tau\rangle_+'F(x,\tau)\Big).$$

The above is again of the form in eq. (A.3) (and a similar result holds for III). Finally,

$$IV = \frac{C}{\langle x+\tau\rangle_+^{2+m}\langle x-c\tau\rangle^{2+m}}\langle x+\tau\rangle_+^{m_1}\langle x-c\tau\rangle^{m_2}\frac{\partial F}{\partial x}(x,\tau), \tag{A.5}$$

which of the form in eq. (A.3).

This shows that the $(k+1)^{\text{st}}$ derivative is a sum of terms of the form in eq. (A.3) and proves the claim.

Now the proposition can be proved fairly straight-forwardly from the claim. The k^{th} derivative is a sum of terms of the form in eq. (A.3), each of which can be bounded as

$$\left|\frac{C}{\langle x+\tau\rangle_+^{2+m}\langle x-c\tau\rangle^{2+m}}\langle x+\tau\rangle_+^{m_1}\langle x-c\tau\rangle^{m_2}F(x,\tau)\right| \\ \leq C\|F\|_{C^0(\mathbb{R}\times\mathbb{R})}\sup_{x\in\mathbb{R}}\frac{1}{\langle x+\tau\rangle_+^2\langle x-c\tau\rangle^2}.$$

The constant in eq. (1.34) can be chosen to be the sum of the constants in the above inequality. Note that there is no τ dependence since we are taking the supremum of F over all x and τ.

The result in eq. (1.35) follows from

$$\sup_{x\in\mathbb{R}}\frac{1}{\langle x+\tau\rangle_+^2\langle x-c\tau\rangle^2} = \mathcal{O}(1/\tau^2) \tag{A.6}$$

as $\tau\to\infty$. □

Lemma 7. *If $a\in\ell_2^2(\mathbb{Z})$ and*

$$\sum_{k=-\infty}^{n} a_k = 0, \tag{1.86}$$

then $b_n = \sum_{k=-\infty}^{n} a_k$ is in $\ell^2(\mathbb{Z})$ and

$$\|b\|_{\ell^2} \leq C\|a\|_{\ell_2^2} \tag{1.87}$$

for some $C>0$ independent of a.

Proof. Let $E_n := \{k\in\mathbb{Z} \mid k\leq n\}$ so that the characteristic function χ_{E_n} satisfies

$$\chi_{E_n}(k) = \begin{cases} 1, & k\leq n \\ 0, & k>0 \end{cases}. \tag{A.7}$$

Then applying the Cauchy-Schwarz inequality, we get that

$$\begin{aligned}\left|\sum_{k=-\infty}^{n} a_k\right| &= \left|\sum_{k=-\infty}^{\infty} \langle k\rangle^2 a_k \frac{\chi_{E_n}(k)}{\langle k\rangle^2}\right| \\ &\leq \|a\|_{\ell^2_2}\left(\sum_{k=-\infty}^{\infty} \frac{\chi_{E_n}(k)}{\langle k\rangle^4}\right)^{1/2} \\ &= \|a\|_{\ell^2_2}\left(\sum_{k=-\infty}^{n} \frac{1}{\langle k\rangle^4}\right)^{1/2}.\end{aligned}$$

By comparing the final sum to the integral $\int_{-\infty}^{n} 1/\langle x\rangle^4\, dx$, we have that there is a constant $C > 0$ independent of a such that

$$\left|\sum_{k=-\infty}^{n} a_k\right| \leq C\|a\|_{\ell^2_2} \times \frac{1}{\langle n\rangle^{3/2}} \tag{A.8}$$

for $n \leq 0$. By noting that $\sum_{k=-\infty}^{n} a_k = -\sum_{k=n+1}^{\infty} a_k$, an identical argument can be applied to get that

$$\left|\sum_{k=n}^{\infty} a_k\right| \leq C\|a\|_{\ell^2_2} \times \frac{1}{\langle n\rangle^{3/2}} \tag{A.9}$$

for $n \geq 0$. Therefore,

$$\|b\|_{\ell^2} \leq C\left(\sum_{n=-\infty}^{\infty} \frac{1}{\langle n\rangle^3}\right)^{1/2} \|a\|_{\ell^2_2}. \tag{A.10}$$

□

References

Bambusi, D. and Ponno, A. (2006). On metastability in FPU. *Communications in mathematical physics*, 264(2):539–561.

Dumas, E. and Pelinovsky, D. (2014). Justification of the log–KdV equation in granular chains: The case of precompression. *SIAM Journal on Mathematical Analysis*, 46(6):4075–4103.

Friesecke, G. and Pego, R. L. (1999). Solitary waves on FPU lattices: I. qualitative properties, renormalization and continuum limit. *Nonlinearity*, 12(6):1601.

Hong, Y., Kwak, C., and Yang, C. (2021). On the korteweg–de vries limit for the fermi–pasta–ulam system. *Archive for Rational Mechanics and Analysis*, 240(2):1091–1145.

Khan, A. and Pelinovsky, D. E. (2017). Long-time stability of small FPU solitary waves. *Discrete & Continuous Dynamical Systems*, 37(4):2065.

Pace, S. D., Reiss, K. A., and Campbell, D. K. (2019). The β fermi-pasta-ulam-tsingou recurrence problem. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(11):113107.

Schneider, G. and Wayne, C. E. (2000). Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model. In *Equadiff 99: (In 2 Volumes)*, pages 390–404. World Scientific.