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

2023

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 ...

Fourth 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, 2023

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

Background Material

Chapter 2

Existence of Kink-Like Traveling Wave Solutions

2.1 Introduction

The goal of this chapter is to show the existence of the travelling wave solution for the FPUT lattice and describe its profile. From formal calculations and the numerical experiments carried out in (Pace et al., 2019), one expects that the travelling wave solution has a profile given by the kink solution to the mKdV; that is, for $\phi(\xi) = \frac{1}{\sqrt{2}}\tanh(\xi/\sqrt{2})$ we expect to have a travelling wave solution u such that

$$u_n(t) = \epsilon\phi(\epsilon(n + ct)) + \mathcal{O}(\epsilon^3) \tag{2.1}$$

when c is slightly smaller than $V''(0) = 1$.

One would expect that methods used to find the soliton-like solution for the FPUT can also be applied to this case. Notably Friesecke and Pego showed in (Friesecke and Pego, 1999) that there exists a solitary wave solution whose profile is described by the KdV soliton using a fixed-point argument. The argument relies on creating a map from $H^1(\mathbb{R})$ to itself using Fourier multipliers such that the fixed point of the map is the profile of the solitary wave. However, this argument does not seem to extend to our case since the function ϕ is not in a Sobolev space and its Fourier transform is defined only in a distributional sense. Due to this problem, we neglect the functional approach and focus on techniques from bifurcation theory.

One common technique for constructing travelling wave solutions to PDEs is by using the center manifold theorem. For PDEs of one spatial and one temporal variable, the strategy is to assume that the solution is a travelling wave (i.e. of the form $f(x-ct)$) to eliminate the derivative with respect to t and reduce the problem to an ODE with respect to the spatial variable x. Finding bounded solutions of this ODE then results in travelling wave solutions of the PDE. The center manifold is an important tool for finding these solutions since (1) it is finite-dimensional, (2) can typically be approximated by Taylor series up to arbitrary order, and (3) contains all bounded solutions. If a linear operator has an eigenvalue pass through the line $\{\lambda \in \mathbb{C} : \Re\lambda = 0\}$ as a parameter μ varies, then one typically has a center manifold containing small bounded solutions parameterized by μ. Similar techniques can be used to find center manifold for more general semi-dynamical systems defined on Banach spaces (Vanderbauwhede and Iooss, 1992). Such a construction was carried out in (Iooss, 2000) for an abstract ODE representing an advance delay differential equation. The existence of several travelling wave solutions on the FPUT lattice were proved. The bifurcation parameter in this paper was given in part by the wave speed. In fact, [Thm. 5](Iooss, 2000) shows the existence of a heteroclinic orbit on the center manifold when c is slightly smaller than 1. This heteroclinic orbit corresponds to the kink-like solution of the FPUT we are interested in. But no description of its wave profile was given, so obtaining an estimate of the form in eq. (2.1) is still an open problem.

Our argument for getting such an estimate will proceed as follows. We first follow the procedure in (Iooss, 2000) to construct the center manifold parameterized by ϵ, making sure to explicitly compute the dynamics on the center manifold. Making a suitable change of variables, we look for small-amplitude, long-wavelength solutions for the FPUT on the center manifold and show that formally setting $\epsilon = 0$ gives a solution related to the kink solution ϕ. Next we apply results from Fenichel theory to show that this solution persists for $\epsilon > 0$. Lastly we convert our results back to the original formulation of the FPUT lattice and prove an estimate of the form eq. (2.1).

2.2 Construction of Center Manifold

We follow the construction of the center manifold carried out in (Iooss, 2000). Recall that the equations for the FPUT lattice are given by

$$\ddot{x}_n = V'(x_{n+1} - x_n) - V'(x_n - x_{n-1}), \quad n \in \mathbb{Z}. \tag{2.2}$$

We assume that $V(x) = \frac{1}{2}x^2 - \frac{1}{24}x^4 + \mathcal{O}(x^5)$ near $x = 0$. We make the ansatz that

$$x_n(\tilde{t}) = x(n - c\tilde{t}), \tag{2.3}$$

where the $x(t)$ on the right is a function from $\mathbb{R}$ to $\mathbb{R}$. Hence $x(t)$ must satisfy the advance-delay differential equation

$$\ddot{x}(t) = \mu\Big(V'(x(t+1) - x(t)) - V(x(t) - x(t-1))\Big) \tag{2.4}$$

where $\mu = c^{-2}$. For delay differential equations like the one above, we are inherently working with an infinite-dimensional problem; the dynamics of $x(t)$ are determined by its value on an interval of t. Instead of working directly with eq. (2.4), we rewrite the equation as a first-order differential equation in a Banach space. Equation (2.4) cannot be written as a differential equation in a finite-dimensional phase space, and so we use a Banach space to represent a "slice" of the function on the interval $[t-1, t+1]$ for $t \in \mathbb{R}$. We introduce a new variable $v \in [-1, 1]$ and functions $X(t, v) = x(t + v)$. We use the notation $\xi(t) = \dot{x}(t)$, $\delta^1 X(t, v) = X(t, 1)$, and $\delta^{-1} X(t, v) = X(t, -1)$. Then letting $U(t) = (x(t), \xi(t), X(t, v))^T$ represent our solution, eq. (2.4) can be written as follows:

$$\partial_t U = L_\mu U + M_\mu(U) \tag{2.5}$$

where L_μ is the linear operator

$$L_\mu = \begin{pmatrix} 0 & 1 & 0 \\ -2\mu & 0 & \mu(\delta^1 + \delta^{-1}) \\ 0 & 0 & \partial_v \end{pmatrix} \tag{2.6}$$

and

$$M_\mu(U) = \mu(0, g(\delta^1 X - x) - g(x - \delta^{-1} X), 0)^T \tag{2.7}$$

where we define $g(x) = V'(x) - x$. We will also require that $X(t,0) = x(t)$, so that $X(t,v) = x(t+v)$ and solutions of eq. (2.5) correspond with solutions of eq. (2.4). We introduce the following Banach spaces for U:

$$\begin{aligned} \mathbb{H} &= \mathbb{R}^2 \times C[-1,1] \\ \mathbb{D} &= \{(x,\xi,X) \in \mathbb{R}^2 \times C^1[-1,1] \mid X(0) = x\} \end{aligned} \tag{2.8}$$

where the spaces have the usual maximum norms. The operator L_μ is continuous from $\mathbb{D}$ to $\mathbb{H}$. Assuming that $g \in C^4(I)$ where I is an open neighborhood around 0, we have $M_\mu \in C^4(\mathbb{D}, \mathbb{D})$.

The system above has a reversibility symmetry S given by

$$S(x,\xi,X)^T = (-x, \xi, -X \circ s)^T \tag{2.9}$$

where $X \circ s(v) = X(-v)$. That is, eq. (2.5) is reversible and if $U(t)$ is a solution then so is $S \circ U(-t)$. Additionally, if $V'(x)$ is odd, then eq. (2.5) is also odd; this means that if $U(t)$ is a solution then so is $-U(t)$.

Note that eq. (2.5) is not well-posed and solutions may not correspond with the requirement that $X(t,0) = x(t)$. However, we can show that there is a center manifold which contains global solutions and lies in $\mathbb{D}$, and so we will be able to extract the travelling wave solutions that we are interested in.

As shown in (Iooss, 2000, Lem. 1), when $\mu = \mu_0 := 1$ (i.e. when $c = \sqrt{V''(0)} = 1$) the linear operator L_{μ_0} has a quadruple zero eigenvalue with the rest of the spectrum bounded uniformly away from the imaginary axis. This allows for the construction of a four-dimensional center manifold. This construction is not carried out explicitly in (Iooss, 2000), but it follows similarly to the calculations carried out in (Iooss and Kirchgässner, 2000) which relies on results in (Vanderbauwhede and Iooss, 1992).

The four-dimensional eigenspace for $\lambda = 0$ is spanned by the following generalized eigenfunctions:

$$\begin{aligned} \zeta_0 &= (1, 0, 1)^T & \zeta_1 &= (0, 1, v)^T \\ \zeta_2 &= (0, 0, \frac{1}{2}v^2)^T & \zeta_3 &= (0, 0, \frac{1}{6}v^3)^T \end{aligned} \tag{2.10}$$

which satisfy

$$\begin{aligned} L_{\mu_0}\zeta_0 &= 0 \\ L_{\mu_0}\zeta_1 &= \zeta_0 \\ L_{\mu_0}\zeta_2 &= \zeta_1 \\ L_{\mu_0}\zeta_3 &= \zeta_2. \end{aligned} \tag{2.11}$$

The spectral projection onto the eigenspace can be found using the Laurent expansion in $\mathcal{L}(\mathbb{H})$ near $\lambda = 0$

$$(\lambda \mathrm{I} - L_{\mu_0})^{-1} = \frac{D^3}{\lambda^4} + \frac{D^2}{\lambda^2} + \frac{D}{\lambda^2} + \frac{P}{\lambda} - \tilde{L}_{\mu_0}^{-1} + \lambda \tilde{L}_{\mu_0}^{-1} - \cdots \tag{2.12}$$

where P is the spectral projection onto the $\lambda = 0$ eigenspace, $D = L_{\mu_0}P$, and $\tilde{L}_{\mu_0}^{-1}$ is the pseudo-inverse of L_{μ_0} on the subspace $(\mathrm{I} - P)\mathbb{H}$ (see (Kato, 2013)). The spectral projection satisfies

$$\begin{aligned} PW &= ((PW)_x, (PW)_\xi, (PW)_X)^T \\ &= (PW)_x\zeta_0 + (DW)_x\zeta_1 + (D^2W)_x\zeta_2 + (D^3W)_x\zeta_3 \end{aligned} \tag{2.13}$$

The projection operator has an explicit form given by

$$P = \oint_{\gamma} (\lambda I - L_{\mu})^{-1} \, d\lambda \tag{2.14}$$

where γ is a curve going around $\lambda = 0$ counter-clockwise and not intersecting the spectrum of L_{μ}. The projection can be computed by first finding the resolvent $(\lambda \mathrm{I} - L_{\mu})^{-1}$ and then using the residue theorem to compute the integral.

The resolvent operator is straightforward to find. For $F = (f_0, f_1, F_2)^T \in \mathbb{H}$, we want to find $U = (x, \xi, X)^T \in \mathbb{D}$ such that

$$(\lambda \mathrm{I} - L_{\mu})U = F. \tag{2.15}$$

The above is a differential equation with coupled algebraic equations. The operator on the left-hand side is invertible when $N(\lambda; \mu) \neq 0$, where

$$N(\lambda; \mu) = -\lambda^2 + 2\mu(\cosh \lambda - 1). \tag{2.16}$$

Solving for U gives

$$x = -[N(\lambda; \mu)]^{-1}(\lambda f_0 + f_1 + \mu \tilde{f}_{\lambda}) \tag{2.17}$$

$$\xi = -[N(\lambda; \mu)]^{-1}\Big([\lambda^2 + N(\lambda; \mu)]f_0 + \lambda f_1 + \mu \lambda \tilde{f}_{\lambda}\Big) \tag{2.18}$$

$$X(v) = e^{\lambda v} x - \int_0^v e^{\lambda(v-s)} F_2(s) \, ds \tag{2.19}$$

with

$$\tilde{f}_{\lambda} = \int_0^1 [-e^{\lambda(1-s)} F_2(s) + e^{-\lambda(1-s)} F_2(-s)] \, ds. \tag{2.20}$$

The projection can be computed by using the residue theorem. For instance, note that

$$(PF)_x = \mathrm{Res}((\lambda \mathrm{I} - L_{\mu_0}^{-1} F)_x, 0) = \mathrm{Res}(-[N(\lambda; \mu)]^{-1}(\lambda f_0 + f_1 + \mu \tilde{f}_{\lambda}), 0). \tag{2.21}$$

For fixed $F \in \mathbb{H}$, the last term can be found by finding the residue of a meromorphic function in $\mathbb{C}$. Proceeding in this way, we can get

$$(PF)_x = \frac{2}{5}\left(f_0 - \int_0^1 [(1-s) - 5(1-s)^3][F_2(s) + F_2(-s)]\, ds \right) \tag{2.22}$$

$$(DF)_x = (PF)_\xi = \frac{2}{5}\left(f_1 - \int_0^1 [1 - 15(1-s)^2][F_2(s) - F_2(-s)]\, ds \right) \tag{2.23}$$

$$(D^2F)_x = (DF)_\xi = -12\left(f_0 - \int_0^1 (1-s))[F_2(s) + F_2(-s)]\, ds \right) \tag{2.24}$$

$$(D^3F)_x = (D^2F)_\xi = -12\left(f_1 - \int_0^1 [F_2(s) - F_2(-s)]\, ds \right). \tag{2.25}$$

We denote by ζ_j^* the linear continuous forms on $\mathbb{H}$ given for any $F \in \mathbb{H}$ by

$$\begin{aligned}
\zeta_0^*(F) &= (PF)_x \\
\zeta_1^*(F) &= (DF)_x = \zeta_0^*(L_{\mu_0}F) \\
\zeta_2^*(F) &= (D^2F)_x \\
\zeta_3^*(F) &= (D^3F)_x
\end{aligned} \tag{2.26}$$

and we have that

$$\zeta_k^*(\zeta_j) = \delta_{kj} \quad k, j = 0, 1, 2, 3 \tag{2.27}$$

where δ_{kj} is the Kronecker delta.

At this point we could start to compute the four-dimensional center manifold parameterized by μ, but we can do a further simplification. Note that eq. (2.5) is invariant under

$$U \mapsto U + q\zeta_0, \quad \forall q \in \mathbb{R} \tag{2.28}$$

which corresponds to the shift invariance of eq. (2.4). This invariance allows us to reduce the center manifold to a three-dimensional manifold. We first decompose

$U \in \mathbb{H}$ as follows:

$$U = W + q\zeta_0, \quad \zeta_0^*(W) = 0. \tag{2.29}$$

Denote by $\mathbb{H}_1$ the codimension-one subspace of $\mathbb{H}$ where $\zeta_0^*(W) = 0$, and similarly define $\mathbb{D}_1$. Then the system in eq. (2.5) becomes

$$\frac{dq}{dt} = \zeta_0^*(L_\mu W) = \zeta_0^*(L_{\mu_0} W) = \zeta_1^*(W) \tag{2.30}$$

$$\frac{dW}{dt} = \widehat{L}_\mu W + M_\mu(W) \tag{2.31}$$

where $\widehat{L}_\mu W = L_\mu W - \zeta_1^*(W)\zeta_0$. The operator $\widehat{L}_{\mu_0}$ acting on $\mathbb{H}_1$ has the same spectrum as L_{μ_0} except that 0 is now a triple eigenvalue instead of a quadruple eigenvalue. One can check that

$$\widehat{L}_{\mu_0}\zeta_1 = 0, \quad \widehat{L}_{\mu_0}\zeta_2 = \zeta_1, \quad \widehat{L}_{\mu_0}\zeta_3 = \zeta_2, \quad \zeta_3^*(\widehat{L}_{\mu_0} W) = 0. \tag{2.32}$$

Hence we have a three-dimensional center manifold on which solutions are given by

$$W = A\zeta_1 + B\zeta_2 + C\zeta_3 + \Phi_\mu(A, B, C). \tag{2.33}$$

Here Φ_μ takes values in $\mathbb{D}_1$. Note that this implies solutions on the center manifold correspond with solutions of eq. (2.4), as desired. We also have that (1) Φ_μ has the same regularity as V', (2) it satisfies $\zeta_k^*(\Phi_\mu) = 0$ for $k = 1, 2, 3$, and (3) it is at least quadratic in its arguments.

The symmetries noted before in eq. (2.5) are preserved in the center manifold (Vanderbauwhede and Iooss, 1992). The reversibility symmetry S is reduced to following representation on the three-dimensional subspace:

$$S_0 : (A, B, C) \mapsto (A, -B, C). \tag{2.34}$$

The dynamics on the center manifold will be odd when $V'(x)$ is odd, in which case if

$(A(t), B(t), C(t))$ is a solution then so is $(-A(t), -B(t), -C(t))$.

It is at this point that our discussion diverges from the work in (Iooss, 2000). From this point, Iooss uses the reversibility of the vector field and results from normal form theory to study the existence of homoclinic, heteroclinic, and periodic solutions on the center manifold. However, since there is an unspecified change of coordinates, the results in (Iooss, 2000) do not give quantitative estimates but rather qualitative descriptions of the solutions. For our purposes though, we would like to compare the profile of the travelling wave solutions and compare it to the mKdV kink solution, and so we must proceed differently. We shall instead compute the Taylor expansion of Φ_μ up to a certain order and get an explicit representation of the center manifold (up to some specified error).

We assume that Φ_μ can be written as a Taylor series in A, B, C, and μ:

$$\Phi_\mu(A, B, C) = \sum_{i,j,k,\ell} (\mu - 1)^\ell A^i B^j C^k \Phi_{ijk}^{(\ell)} \tag{2.35}$$

Note that the μ terms are centered at $\mu_0 = 1$. We will only need to compute up to some of the cubic terms, so we do not need Φ_μ is analytic as suggested by eq. (2.35). In fact $\Phi_\mu \in C^4$ in a neighborhood of $(\mu, A, B, C) = (1, 0, 0, 0)$ is sufficient and is guaranteed by the regularity we assumed for V' and g.

It is useful to compute $\widehat{L}_\mu$ applied to each eigenvector:

$$\widehat{L}_\mu \zeta_1 = 0 \tag{2.36}$$

$$\widehat{L}_\mu \zeta_2 = \zeta_1 + (\mu - 1) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{2.37}$$

$$\widehat{L}_\mu \zeta_3 = \zeta_2. \tag{2.38}$$

Note that these calculations agree with eq. (2.32) when μ is equal to $\mu_0 = 1$. Now plugging eq. (2.33) into eq. (2.31) gives

$$\begin{aligned}
\dot{A}\zeta_1 + \dot{B}\zeta_2 + \dot{C}\zeta_3 + D\Phi_\mu(A,B,C)\begin{bmatrix}\dot{A}\\ \dot{B}\\ \dot{C}\end{bmatrix} = \\
B\zeta_1 + B(\mu-1)\begin{bmatrix}0\\1\\0\end{bmatrix} + C\zeta_2 + L_{\mu_0}\Phi_\mu(A,B,C) \\
+ \big(2(1-\mu)\Phi_\mu^x + (\mu-1)(\delta^1\Phi_\mu^X + \delta^{-1}\Phi_\mu^X)\big)\begin{bmatrix}0\\1\\0\end{bmatrix} \\
+\mu\left(g\big(A + \frac{1}{2}B + \frac{1}{6}C + (\delta^1\Phi_\mu^X - \Phi_\mu^x)\big) - g\big(A - \frac{1}{2}B + \frac{1}{6}C + (\Phi_\mu^x - \delta^{-1}\Phi_\mu^X)\big)\right)\begin{bmatrix}0\\1\\0\end{bmatrix}
\end{aligned} \tag{2.39}$$

where we represent the components of Φ_μ by $(\Phi_\mu^x, \Phi_\mu^\xi, \Phi_\mu^X)^T$. We now apply the spectral projections ζ_i^* to eq. (2.39) to get the a system of differential equations. Note that we have

$$\begin{bmatrix}0\\1\\0\end{bmatrix} = \frac{2}{5}\zeta_1 - 12\zeta_3 + \begin{bmatrix}0\\ \frac{3}{5}\\ 2v^3 - \frac{2}{5}v\end{bmatrix}, \tag{2.40}$$

where the final term is in the kernel of each ζ_i^*. Thus we get the following system of differential equations:

$$\dot{A} = B + \frac{2}{5}\Big[\cdots\Big] \tag{2.41}$$

$$\dot{B} = C \tag{2.42}$$

$$\dot{C} = -12\Big[\cdots\Big] \tag{2.43}$$

$$D\Phi_\mu(A,B,C)\begin{bmatrix}\dot{A}\\ \dot{B}\\ \dot{C}\end{bmatrix} = L_{\mu_0}\Phi_\mu + \Big[\cdots\Big]\begin{bmatrix}0\\ \frac{3}{5}\\ 2v^3 - \frac{2}{5}v\end{bmatrix}. \tag{2.44}$$

The $\cdots$ within the brackets are given by the following expression

$$\begin{aligned} &B(\mu-1) + 2(1-\mu)\Phi_\mu^x + (\mu-1)(\delta^1\Phi_\mu^X + \delta^{-1}\Phi_\mu^X) \\ &+ \mu\left(g\big(A + \frac{1}{2}B + \frac{1}{6}C + (\delta^1\Phi_\mu^X - \Phi_\mu^x)\big) - g\big(A - \frac{1}{2}B + \frac{1}{6}C + (\Phi_\mu^x - \delta^{-1}\Phi_\mu^X)\big)\right), \end{aligned} \tag{2.45}$$

which we abridged to improve legibility. Equations (2.41) to (2.43) define the dynamics on the center manifold. Equation (2.44) are the components of eq. (2.39) which are in the kernel of the spectral projections. Now using the expression for the derivatives in eqs. (2.41) to (2.43) and plugging into eq. (2.44) gives the following:

$$\frac{\partial\Phi}{\partial A}\left(B + \frac{2}{5}[\cdots]\right) + \frac{\partial\Phi}{\partial B}C + \frac{\partial\Phi}{\partial C}\left(-12[\cdots]\right) = L_{\mu_0}\Phi_\mu + [\cdots]\begin{bmatrix} 0 \\ \frac{3}{5} \\ 2v^3 - \frac{2}{5}v \end{bmatrix}. \tag{2.46}$$

We will now assume Φ_μ has the form given in eq. (2.35). Plugging in the terms of the series into eq. (2.46) above gives use a system of equations we can iteratively solve to get the coefficients. In particular, we will get equations of the form

$$L_{\mu_0}\Phi_{ijk}^{(\ell)} = \cdots \tag{2.47}$$

where the right-hand side will depend on coefficients of order no higher than $\ell+i+j+k$. From the center manifold theorem, we have that the constant and first-order terms are zero. Thus we start by first computing the second-order terms: that is, terms where $\ell + i + j + k = 2$. We get the following set of equations as a result.

$$0 = L_{\mu_0}\Phi^{(2)}_{000} \tag{2.48}$$

$$0 = L_{\mu_0}\Phi^{(1)}_{100} \tag{2.49}$$

$$\Phi^{(1)}_{100} = L_{\mu_0}\Phi^{(1)}_{010} + \begin{bmatrix} 0 \\ \frac{3}{5} \\ 2v^3 - \frac{2}{5}v \end{bmatrix} \tag{2.50}$$

$$\Phi^{(1)}_{010} = L_{\mu_0}\Phi^{(1)}_{001} \tag{2.51}$$

$$0 = L_{\mu_0}\Phi^{(0)}_{200} \tag{2.52}$$

$$2\Phi^{(0)}_{200} = L_{\mu_0}\Phi^{(0)}_{110} \tag{2.53}$$

$$\Phi^{(0)}_{110} = L_{\mu_0}\Phi^{(0)}_{101} \tag{2.54}$$

$$\Phi^{(0)}_{110} = L_{\mu_0}\Phi^{(0)}_{020} \tag{2.55}$$

$$2\Phi^{(0)}_{020} + \Phi^{(0)}_{101} = L_{\mu_0}\Phi^{(0)}_{011} \tag{2.56}$$

$$\Phi^{(0)}_{011} = L_{\mu_0}\Phi^{(0)}_{002} \tag{2.57}$$

Equations (2.48), (2.49) and (2.52) can be solved by noting that ζ_0 is the only zero eigenfunction for L_{μ_0} and $\zeta_0^*(\Phi^{(2)}_{000} = \zeta_0^*(\Phi^{(1)}_{100}) = \zeta_0^*(\Phi^{(0)}_{200}) = 0$ since Φ_μ takes values in $\mathbb{D}_1$, thus $\Phi^{(2)}_{000} = \Phi^{(1)}_{100} = \Phi^{(0)}_{200} = 0$. Then eq. (2.50) is reduced to

$$0 = L_{\mu_0}\Phi^{(1)}_{010} + \begin{bmatrix} 0 \\ \frac{3}{5} \\ 2v^3 - \frac{2}{5}v \end{bmatrix}, \tag{2.58}$$

which can be solved by integrating to get

$$\Phi^{(1)}_{010} = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{2}v^4 + \frac{1}{5}v^2 \end{bmatrix} + k\zeta_0 \tag{2.59}$$

for some $k \in \mathbb{R}$. Imposing the constraint that $\zeta_0^*(\Phi^{(1)}_{010}) = 0$ gives us that

$$k = -13/2100.$$

Similarly integrating eq. (2.51) gives

$$\Phi_{001}^{(1)} = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{10}v^5 + \frac{1}{15}v^3 \end{bmatrix} + k\zeta_1 \tag{2.60}$$

with the same value of k. The remaining terms end up equaling zero, which can be found by substituting in known values into the equations.

One can compute the cubic coefficients in a similar way. In particular, we have that

$$0 = L_{\mu_0}\Phi_{300}^{(0)} \tag{2.61}$$

and so $\Phi_{300}^{(0)} = 0$. We will not need to compute any of the other coefficients. As we will soon see, after a change of variables they end up being in the higher order terms to be neglected. Before proceeding, we will need a new parameterization for the center manifold. We let $\epsilon > 0$ be the new bifurcation parameter such that $c^2 = 1 - \epsilon^2/12$. This choice of parameterization is partially based on the parameterization in (Friesecke and Pego, 1999), and – as we will soon see – the value of ϵ is related to the amplitude of the travelling wave solutions. The bifurcation will now occur at $\epsilon = 0$, which corresponds to the case where $\mu = c^{-2} = 1$. As seen in (Iooss, 2000), the heteroclinic orbits on the center manifold will only exist for c^2 slightly less than 1, and so we will look for these orbits when $\epsilon > 0$. We have that

$$\mu - 1 = c^{-2} - 1 = \frac{1}{1-\epsilon^2/12} - 1 = \frac{\epsilon^2}{12} + \mathcal{O}(\epsilon^4).$$

Since we are looking for ϵ-amplitude waves with wavelength of order ϵ^{-1}, we are motivated to make the following change of variables:

$$A(t) = \epsilon\underline{A}(\epsilon t), \quad B(t) = \epsilon^2\underline{B}(\epsilon t), \quad \epsilon^3\underline{C}(\epsilon t). \tag{2.62}$$

By grouping together orders of ϵ and using the values of $\Phi_{ijk}^{(\ell)}$ computed above, we have the following expansion of the terms in eq. (2.45)

$$\frac{\epsilon^4}{12}(\underline{B} - 6\underline{A}^2\underline{B}) + \frac{\epsilon^5}{6}V^{(5)}(0) \cdot \underline{A}^3\underline{B} + \mathcal{O}(\epsilon^6) \tag{2.63}$$

Then the equations of motion on the center manifold become

$$\begin{aligned}
\underline{A}' &= \underline{B} + \mathcal{O}(\epsilon^2) \\
\underline{B}' &= \underline{C} \\
\underline{C}' &= -\underline{B} + 6\underline{A}^2\underline{B} - 2\epsilon V^{(5)}(0) \cdot \underline{A}^3\underline{B} + \mathcal{O}(\epsilon^2),
\end{aligned} \tag{2.64}$$

where $'$ represents the derivative with respect to the new time variable $s = \epsilon t$. The $\mathcal{O}(\epsilon^2)$ represents functions that are at least C^4 in ϵ, $\underline{A}$, $\underline{B}$, and $\underline{C}$ and can be bounded by a constant times ϵ^2 when we are on bounded domains and $\epsilon > 0$ sufficiently small. Since we will be looking for bounded solutions on the center manifold, these terms can be controlled. Thus the dynamics on the center manifold are controlled up to $\mathcal{O}(\epsilon)$ terms. We may upgrade this to $\mathcal{O}(\epsilon^2)$ if we additionally have $V^{(5)}(0) = 0$.

We shall consider three different assumptions on the potential going forward:

(H1) $V(x) = \frac{1}{2}x^2 - \frac{1}{24}x^4 + \mathcal{O}(x^5)$

(H2) $V(x) = \frac{1}{2}x^2 - \frac{1}{24}x^4 + \mathcal{O}(x^6)$

(H3) $V(x) = \frac{1}{2}x^2 - \frac{1}{24}x^4$

The arguments for each assumption are similar, but stronger assumptions on the potential gives better estimates on the final result.

For (H1), we have that the flow on the center manifold is given by

$$\begin{aligned}
\underline{A}' &= \underline{B} + \epsilon F_1(\underline{A}, \underline{B}, \underline{C}; \epsilon) \\
\underline{B}' &= \underline{C} \\
\underline{C}' &= -\underline{B} + 6\underline{A}^2\underline{B} + \epsilon G_1(\underline{A}, \underline{B}, \underline{C}; \epsilon) \\
\epsilon' &= 0
\end{aligned} \tag{2.65}$$

where F_1 and G_1 will be C^4 for $\epsilon > 0$ and $\mathcal{O}(1)$ as $\epsilon \to 0$. The additional equation $\epsilon' = 0$ is added so that we may use ϵ as an additional coordinate in our results. Note that this will not change the flow on the center manifold since ϵ remains fixed. For (H2) and (H3), we parameterize based on $\eta = \epsilon^2$ and the flow is now given by

$$\begin{aligned}
\underline{A}' &= \underline{B} + \eta F_2(\underline{A}, \underline{B}, \underline{C}; \sqrt{\eta}) \\
\underline{B}' &= \underline{C} \\
\underline{C}' &= -\underline{B} + 6\underline{A}^2\underline{B} + \eta G_2(\underline{A}, \underline{B}, \underline{C}; \sqrt{\eta}) \\
\eta' &= 0
\end{aligned} \tag{2.66}$$

where F_2 and G_2 will be C^4 for $\eta > 0$ and $\mathcal{O}(1)$ as $\eta \to 0$. Reparameterizing to η will ultimately allow us to improve our error from $\mathcal{O}(\epsilon)$ to $\mathcal{O}(\epsilon^2)$. The systems can be extended to C^1 flows for negative values of the parameters: for instance we make the replacement

$$\begin{aligned}
\epsilon F_1(\underline{A}, \underline{B}, \underline{C}; \epsilon) &\to \epsilon F_1(\underline{A}, \underline{B}, \underline{C}; |\epsilon|) \\
\epsilon G_1(\underline{A}, \underline{B}, \underline{C}; \epsilon) &\to \epsilon G_1(\underline{A}, \underline{B}, \underline{C}; |\epsilon|)
\end{aligned} \tag{2.67}$$

to get eq. (2.65) is C^1 for (possibly negative) ϵ near zero. A similar replacement of $\sqrt{\eta} \to \sqrt{|\eta|}$ makes eq. (2.66) C^1 for η near zero. One can get improved regularity of the vector field form (H2) and (H3) by using the original parameter, ϵ, but this sacrifices the ϵ^2 error in the estimate. This trade-off will be necessary to get certain regularity in the error term.

The arguments for the persistence of heteroclinic orbits is similar for eqs. (2.65) and (2.66), so we will focus first on the former system and note where the results differ for the latter system.

2.3 Existence of Heteroclinic Orbit

At this point, our goal is to show the existence of a heteroclinic orbit for eq. (2.64) for $\epsilon > 0$ sufficiently small and to get estimates of the solution. One might expect that the flow on the center manifold for $\epsilon > 0$ small is well approximated by formally setting $\epsilon = 0$. Indeed, if we let $\epsilon = 0$, then the ODEs in eq. (2.64) become equivalent to the third-order differential equation

$$\underline{A}''' + \underline{A}' - 6\underline{A}^2\underline{A}' = 0 \tag{2.68}$$

which has the solution

$$\underline{A}(s) = \frac{1}{\sqrt{2}} \tanh\left(\frac{s}{\sqrt{2}}\right). \tag{2.69}$$

This solution is the profile for the kink solution of the defocusing mKdV, ϕ. This represents a heteroclinic orbit for the system of ODEs since $(\underline{A}(s), \underline{B}(s), \underline{C}(s)) \to (\pm 1/\sqrt{2}, 0, 0)$ as $s \to \pm\infty$. One might expect that for $\epsilon > 0$ that there is also a heteroclinic orbit that is close to the above solution. Thus we want to show that the heteroclinic orbit at $\epsilon = 0$ persists for small perturbations of ϵ, and we want to get estimates of these orbits relative to ϵ. To get these results, we apply Fenichel theory. Review appendix A for the relevant results that will be used.

The idea behind the proof is to show that there is an overflowing invariant set with an unstable manifold and a corresponding inflowing invariant set with a stable manifold. We then show that at $\epsilon = 0$ these manifolds intersect transversally at a point, and that this intersection is given by the above heteroclinic orbit. From there, we show that this intersection is preserved for $\epsilon > 0$ and the heteroclinic orbit remains

$\mathcal{O}(\epsilon)$ or $\mathcal{O}(\epsilon^2)$ close to the original orbit.

2.3.1 The Unstable and Stable Manifolds

We first must find the appropriate overflowing invariant set.[1] From the heteroclinic orbit found for $\epsilon = 0$, we know that $(\underline{A}, \underline{B}, \underline{C}, \epsilon) = (-1/\sqrt{2}, 0, 0, 0)$ should be one point in the set. In fact, for fixed $\epsilon > 0$ we have that multiples of ζ_1 are fixed points for eq. (2.31), which correspond to the linear solutions $x(t) = x_0 + mt$ for eq. (2.4). From the center manifold theorem in (Vanderbauwhede and Iooss, 1992), bounded solutions sufficiently close to the origin will lie exactly on the center manifold. Thus for $\epsilon > 0$ sufficiently close to zero, any closed interval on the $\underline{A}$-axis is composed entirely of fixed points on the center manifold. We will choose $\epsilon_0 > 0$ small enough such that for $\epsilon \in (0, \epsilon_0]$ the $\underline{A}$-axis from $[-1, 1]$ is composed entirely of fixed points.

If we fix a small $\delta > 0$ and set $A_{-\infty} = -1/\sqrt{2}$, then

$$\overline{M} = \{(\underline{A}, 0, 0, \epsilon) \in \mathbb{R}^4 : |(\underline{A} - A_{-\infty}, \epsilon)| \leq \delta\} \tag{2.70}$$

is a smooth manifold with boundary that is invariant under the flow in eq. (2.65). In fact, $\overline{M}$ consists exclusively of fixed points of the flow.

To apply theorem A.1 and get an unstable manifold for $\overline{M}$ we need that

(i) $\overline{M}$ is overflowing invariant, and

(ii) the generalized Lyapunov-type numbers on $\overline{M}$ satisfy the inequalities in theorem A.1.

As written, $\overline{M}$ is *not* an overflowing invariant manifold. However, a common trick in Fenichel is to adjust the flow on the boundary of an invariant manifold so that it

[1] We need also to find the inflowing invariant set, but we can rely on the symmetry of eq. (2.64) to get this. In fact, we will regularly rely on the symmetry of the flow to get many of the results for the inflowing invariant set after working it out for the overflowing invariant set.

becomes overflowing invariant (see (Wiggins, 1994, §6.3)). This will alter the behavior of our dynamical system at the boundary, but elsewhere the dynamics will remain the same. For our case, we may adjust the flow near the boundary $|(\underline{A}-A_{-\infty}, \underline{B}, \underline{C}, \epsilon)| = \delta$ to get $\overline{M}$ is overflowing invariant, but this will not affect the dynamics near the heteroclinic orbit. Thus we can still talk about the existence of the heteroclinic orbit in the unaltered system. This adjustment will need to be done in a way to not greatly affect the generalized Lyapunov-type numbers. For now, we set aside point (i) and address (ii), which is more straightforward.

Since $\overline{M}$ consists only of fixed points, the generalized Lyapunov-type numbers can be computed using the linearization of the flow. Note that since each $(\underline{A}, 0, 0, \epsilon) \in \overline{M}$ is a fixed point, we have that

$$\begin{aligned} F_1(\underline{A}, 0, 0, \epsilon) &= 0 \\ G_1(\underline{A}, 0, 0, \epsilon) &= 0 \end{aligned} \tag{2.71}$$

and the partial derivatives of F_1 and G_1 with respect to $\underline{A}$ or ϵ will be zero. Thus at a point $(\underline{A}, 0, 0, \epsilon) \in \overline{M}$, the linearization of the flow is given by

$$\begin{bmatrix} 0 & 1+\epsilon\frac{\partial F_1}{\partial \underline{B}}(\underline{A}, 0, 0; \epsilon) & \epsilon\frac{\partial F_1}{\partial \underline{C}}(\underline{A}, 0, 0; \epsilon) & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 6\underline{A}^2 - 1 + \epsilon\frac{\partial G_1}{\partial \underline{B}}(\underline{A}, 0, 0; \epsilon) & \epsilon\frac{\partial G_1}{\partial \underline{C}}(\underline{A}, 0, 0; \epsilon) & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{2.72}$$

The tangent space at $p \in M$ is given by $T_pM = \text{span}\{(1, 0, 0, 0), (0, 0, 0, 1)\}$. The vector bundles N^u and N^s will be defined as the unstable and stable subspaces of each fixed point, respectively. That these vector bundles are invariant under the flow and continuous follows immediately from their definition. The two eigenvalues $\lambda_1, \lambda_2 = 0$ correspond with the flow tangent to the manifold. Fixing $\underline{A} = A_{-\infty}$ and $\epsilon = 0$, the other eigenvalues are $\lambda_{3,4} = \pm\sqrt{6A_{-\infty}^2 - 1}$, which correspond with the flow along the vector bundles N^u and N^s, respectively. There at $p_0 = (A_{-\infty}, 0, 0, 0)$ we have the generalized Lyapunov-type numbers given by

$$\lambda^u(p_0) = \nu^s(p_0) = \exp\big(-\sqrt{6A_{-\infty}^2 - 1}\big), \quad \sigma^s(p_0) = 0. \tag{2.73}$$

To have a C^1 unstable manifold, we are required to have $\lambda^u(p), \nu^s(p), \sigma^s(p) < 1$ for each point $p \in M$. By the continuity of eigenvalues, we can guarantee this by choosing δ small enough.

Then condition (ii) is satisfied. Now we want to show that we can alter near ∂M so that (i) is also satisfied without causing (ii) to become invalid. We first introduce a C^∞ bump function, $\chi : [0, \infty) \to \mathbb{R}$, such that

(1) $0 \leq \chi(r) \leq 1$ for $r \in [0, \infty)$

(2) $\chi(r) = 0$ when $r \in [0, \delta - \sigma]$

(3) $\chi(r) = 1$ when $r \in [\delta - \frac{\sigma}{2}, \delta + \frac{\sigma}{2}]$

(4) $\chi(r) = 0$ when $r \in [\delta + \sigma, \infty)$

where $\sigma > 0$ will be a parameter that we can make as small as necessary. We then alter the vector field in eq. (2.65) by setting

$$\begin{aligned}
\underline{A}' &= \underline{B} + \epsilon F_1(\underline{A}, \underline{B}, \underline{C}; \epsilon) + \chi(|(\underline{A} - A_{-\infty}, \underline{B}, \underline{C}, \epsilon)|) \cdot (\underline{A} - A_{-\infty}) \\
\underline{B}' &= \underline{C} \\
\underline{C}' &= -\underline{B} + 6\underline{A}^2\underline{B} + \epsilon G_1(\underline{A}, \underline{B}, \underline{C}; \epsilon) \\
\epsilon' &= \chi(|(\underline{A} - A_{-\infty}, \underline{B}, \underline{C}, \epsilon)|) \cdot \epsilon.
\end{aligned} \tag{2.74}$$

This change keeps the flow C^1 and makes $\overline{M}$ an overflowing invariant vector field. However, a couple things need to be checked before applying theorem A.1: the vector bundles N^u and N^s must be defined on $\chi \neq 0$ and the generalized Lyapunov-type numbers must satisfy the necessary inequalities.

The extension of the normal vector bundles is somewhat technical. We need $TM \oplus N^u$ and $TM \oplus N^s$ invariant under the flow and continuous. In particular, if p is a point in M where $\chi \neq 0$ and $\xi \in N^u_p$, then $D\phi_t(p)\xi \in T_{\phi_t(p)} \oplus N^u_{\phi_t(p)}$ for all t such that $\phi_t(p) \in M$. A similar result should hold for N^s. We also need that the vector bundles are continuous with respect to p. Continuity relies on showing that we can assign the vector bundles in a way so that if $\phi_t(p) \to p'$ as $t \to -\infty$ then $N^u_{\phi_t(p)}$ approaches $N^u_{p'}$. This can be done and the details are carried out in appendix A.4.

For the generalized Lyapunov-type numbers, it can be shown that the values on the altered region of M can be bounded by those on the unaltered region. More generally, we have the following result.

Proposition 2.1. *Let $K \subset M$ be a compact set. If $p \in M$ such that $\phi_{-t}(p) \to K$ as $t \to \infty$, then*

(i) $\lambda^u(p) \leq \lambda^u(K)$,

(ii) $\nu^s(p) \leq \nu^s(K)$, *and*

(iii) if $\nu^s(K) < 1$, *then* $\sigma^s(p) \leq \sigma^s(K)$.

The proof is give in appendix A and follows similarly to the arguments found in (Dieci and Lorenz, 1997).

We can therefore conclude that $W^u_{\text{loc}}(\overline{M})$ exists. If we set $A_\infty = 1/\sqrt{2}$, then an analogous argument holds for showing that

$$\overline{N} = \{(\underline{A}, 0, 0, \epsilon) \in \mathbb{R}^4 : |(\underline{A} - A_\infty, \epsilon)| \leq \delta\} \tag{2.75}$$

has a *stable* manifold, $W^s_{\text{loc}}(\overline{N})$.

2.3.2 Transversal intersection at $\epsilon = 0$

To show a heteroclinic orbit exists for $\epsilon > 0$, we first show that stable and unstable manifolds described above have a transverse intersection at $\epsilon = 0$. This intersection then persists for perturbations in ϵ (since the manifolds are C^1 with respect to ϵ) and thus implies the existence of the heteroclinic orbit.

The heteroclinic orbit at $\epsilon = 0$ can be found explicitly. The dynamics (away from where we modified the vector field) are given by

$$\begin{aligned} \underline{A}' &= \underline{B} \\ \underline{B}' &= \underline{C} \\ \underline{C}' &= -\underline{B} + 6\underline{A}^2\underline{B}. \end{aligned} \tag{2.76}$$

The system of ODEs in eq. (2.76) has two invariants:

$$I_1(\underline{A}, \underline{B}, \underline{C}) = \underline{C} + \underline{A} - 2\underline{A}^3 \tag{2.77}$$

$$I_2(\underline{A}, \underline{B}, \underline{C}) = \frac{1}{2}\underline{B}^2 + \frac{1}{2}\underline{A}^2 - \frac{1}{2}\underline{A}^4 - \underline{A}I_1(\underline{A}, \underline{B}, \underline{C}). \tag{2.78}$$

We then look for solutions on the manifolds given by

$$I_1(A_{-\infty}, 0, 0) = 0 \quad \text{and} \quad I_2(A_{-\infty}, 0, 0) = \frac{1}{8}. \tag{2.79}$$

The above equations and the fact that $\underline{B} = \underline{A}'$ gives us that $\underline{A}$ must satisfy the following first order ODE:

$$(\underline{A}')^2 = \left(\frac{1}{2} - \underline{A}^2\right)^2, \tag{2.80}$$

which can be solved by separation of variables. The solutions are thus

$$\underline{A}(s) = \pm\frac{1}{\sqrt{2}}\tanh\left(\frac{s}{\sqrt{2}}\right) \tag{2.81}$$

up to a shift in the variable s. One can check that these are solutions of eq. (2.76) (taking $\underline{B} = \underline{A}'$ and $\underline{C} = \underline{A}''$) and define two heteroclinic orbits: one traveling from

$A_{-\infty}$ to A_∞ and one traveling from A_∞ to $A_{-\infty}$. Let

$$\gamma_\pm(t) = \begin{bmatrix} \pm\frac{1}{\sqrt{2}}\tanh\left(\frac{s}{\sqrt{2}}\right) \\ \pm\frac{1}{2}\text{sech}^2\left(\frac{s}{\sqrt{2}}\right) \\ \mp\frac{1}{\sqrt{2}}\tanh\left(\frac{s}{\sqrt{2}}\right)\text{sech}^2\left(\frac{s}{\sqrt{2}}\right) \end{bmatrix} = \begin{bmatrix} \pm\phi(s) \\ \pm\phi'(s) \\ \pm\phi''(s) \end{bmatrix} \tag{2.82}$$

denote the two heteroclinic orbits.

The solution corresponding with the choice of $+$ also lies inside the manifolds $W^u_{\text{loc}}(\overline{M})$ and $W^s_{\text{loc}}(\overline{N})$ since it converges to $\overline{M}$ and $\overline{N}$ as $s \to -\infty$ and $s \to +\infty$, respectively. This does not imply the local manifolds intersect since they are only defined in a neighborhood of $\overline{M}$ and $\overline{N}$, but we may extend these manifolds under the flow so that they both contain the point $(\epsilon, \underline{A}, \underline{B}, \underline{C}) = (0, 0, 1/2, 0)$ and thus intersect. We shall refer to the manifolds extended under the flow by $\mathcal{M}_\epsilon$ and $\mathcal{N}_\epsilon$. These extended manifolds are still C^1 with respect to the parameter ϵ.

Now the goal is to demonstrate that this intersection is transverse. That is for $p = (0, 1/2, 0)$ we want to show that $T_p\mathcal{M}_0 + T_p\mathcal{N}_0 = T_p\mathbb{R}^3$. One can explicitly compute each of the tangent spaces at p and show they span $T_p\mathbb{R}^3$. This is done by finding the intersection each of these manifolds make with the $\underline{BC}$-plane. Similar to the construction of the heteroclinic orbit, we find the orbit which approaches some asymptotic value on the $\underline{A}$-axis near $A_{-\infty}$ or A_∞ and find where it intersect the $\underline{BC}$-plane. These orbits lie on the stable and unstable manifolds, and so this shows how the manifolds intersect the plane.

Take ω to be a point near $A_{-\infty} = -1/\sqrt{2}$. The orbit that approaches $(\omega, 0, 0)$ in backwards time lies on the intersection of

$$\begin{aligned} I_1(\underline{A}, \underline{B}, \underline{C}) &= I_1(\omega, 0, 0) = \omega - 2\omega^3 \\ I_2(\underline{A}, \underline{B}, \underline{C}) &= I_2(\omega, 0, 0) = -\frac{1}{2}\omega^2 + \frac{3}{2}\omega^4. \end{aligned} \tag{2.83}$$

Setting $\underline{A} = 0$, we can find that $\mathcal{M}_0$ hits the $\underline{BC}$- plane at

$$m(\omega) = (0, |\omega|\sqrt{3\omega^2 - 1}, \omega - 2\omega^3) \tag{2.84}$$

for ω close to $A_{-\infty}$. In particular, we see $m(A_{-\infty}) = p$. Identical reasoning gives that $\mathcal{N}_0$ intersects the plane at

$$n(\alpha) = (0, |\alpha|\sqrt{3\alpha^2 - 1}, \alpha - 2\alpha^3) \tag{2.85}$$

where α is near $A_\infty = 1/\sqrt{2}$ and $n(A_\infty) = p$.

A tangent vector to the heteroclinic orbit is given by $\gamma_+'(0) = (1/2, 0, -1/2)$, and this vector lies in both $T_p\mathcal{M}_0$ and $T_p\mathcal{N}_0$. Since $m(\omega) \in \mathcal{M}_0$ for ω near $A_{-\infty}$, we have that

$$m'(A_{-\infty}) = \left(0, 1, \frac{1}{\sqrt{2}}\right) \in T_p\mathcal{M}_0. \tag{2.86}$$

Similarly,

$$n'(A_\infty) = \left(0, 1, \frac{-1}{\sqrt{2}}\right) \in T_p\mathcal{N}_0. \tag{2.87}$$

Therefore

$$T_p\mathcal{M}_0 + T_p\mathcal{N}_0 = \text{span}\left\{\left(\frac{1}{2}, 0, \frac{-1}{2}\right), \left(0, 1, \frac{1}{\sqrt{2}}\right), \left(0, 1, \frac{-1}{\sqrt{2}}\right)\right\} = T_p\mathbb{R}^3 \tag{2.88}$$

and the intersection is transverse. This implies that there is a heteroclinic orbit on the intersection of $\mathcal{M}_\epsilon$ and $\mathcal{N}_\epsilon$ for $\epsilon > 0$.

2.3.3 Estimates on the heteroclinic orbit

From the previous section, we have the existence of heteroclinic orbits that are perturbation of $\gamma_\pm$ at $\epsilon = 0$. From the C^1 regularity of the manifolds with respect to the coordinates, we expect that the orbits remain $\mathcal{O}(\epsilon)$ close to the unperturbed orbits in some sense. There are some subtleties to be addressed. The manifolds remain

$\mathcal{O}(\epsilon)$ close in Hausdorff distance, but this does not imply the orbits on the manifolds remain $\mathcal{O}(\epsilon)$ close for all time. The dynamics on the manifolds might change causing orbits on the perturbed manifold to diverge asymptotically despite remaining close initially.

First, let us introduce notation for the perturbed heteroclinic orbits. We shall denote by $\gamma_{\pm,\epsilon} = (A_{\pm,\epsilon}, B_{\pm,\epsilon}, C_{\pm,\epsilon})$ the perturbations of $\gamma_{\pm}$ for $\epsilon > 0$, where we set $\gamma_{\pm,\epsilon}(0)$ to be the point where the orbits cross the $\underline{BC}$-plane. From the continuity of the manifolds with respect to ϵ, we have that $|\gamma_{\pm,\epsilon}(0) - \gamma_{\pm}(0)| = \mathcal{O}(\epsilon)$ for small ϵ. We can extend this estimate onto arbitrarily large finite time scales by applying an argument using the Grönwall inequality. That is, for every $T > 0$ we have for sufficiently small ϵ that $|\gamma_{\pm,\epsilon}(s) - \gamma_{\pm}(s)| = C\epsilon$ for all $s \in [-T, T]$, where $C > 0$ is independent of s.

This argument is insufficient for extending the estimate to all time. To get that the orbits remain close as $s \to \pm\infty$, we can rely on part 7 of theorem A.2. Taking (ϵ, ω) sufficiently close to $(0, A_{-\infty})$ as base points, the theorem states that the unstable manifold is C^1 with respect to (ϵ, ω). Note that the overflowing invariant manifold (away from the bump function) consists only of fixed points, so orbits on the unstable manifold approach a unique fixed point given by (ϵ, ω). Take $T > 0$ large enough so that all the points on $\mathcal{M}_\epsilon$ (for $\epsilon \leq \epsilon_0$) which intersect the $\underline{BC}$-plane are in the local unstable manifold when flowed backward in time by $-T$ units. Then locally, there is a one-to-one correspondence between these points flowed backward in time and the points in $\overline{M}$; furthermore, this correspondence is C^1 due to theorem A.2. This implies that if the points flowed backward are $\mathcal{O}(\epsilon)$ close then their backward limits are $\mathcal{O}(\epsilon)$ close as well. In particular, the backward limits of $\gamma_{+,\epsilon}$ and γ_+ are $\mathcal{O}(\epsilon)$ close. The argument for the stable manifold is analogous. This shows that the heteroclinic orbits remain $\mathcal{O}(\epsilon)$ close for all time. In the case where $V(x) = \frac{1}{2}x^2 - \frac{1}{24}x^4 + \mathcal{O}(\epsilon^6)$, this can be upgraded to $\mathcal{O}(\epsilon^2)$; the proof is similar but we use the regularity of the manifolds with respect to $\eta = \epsilon^2$ instead. Therefore, we have the following.

Proposition 2.2. *There exists $\epsilon_0 > 0$ such that for every $\epsilon \in (0, \epsilon_0]$ there exist two heteroclinic orbits $\gamma_{\pm,\epsilon}$ of eq.* (2.65) *such that*

$$|\gamma_{\pm,\epsilon}(s) - \gamma_{\pm}(s)| \le C\epsilon \quad \textit{for all } s \in \mathbb{R} \tag{2.89}$$

where $\gamma_{\pm}$ are defined in eq. (2.82). *If $V(x) = \frac{1}{2}x^2 - \frac{1}{24}x^4 + \mathcal{O}(\epsilon^6)$, then we instead have the estimate*

$$|\gamma_{\pm,\epsilon}(s) - \gamma_{\pm}(s)| \le C\epsilon^2 \quad \textit{for all } s \in \mathbb{R}. \tag{2.90}$$

By starting to unravel the change of coordinates, we can show the existence of solutions to the advance-delay equation of the form

$$x_{\pm,\epsilon}(t) = q_{\pm,\epsilon}(t) + (\Phi_\mu(\epsilon A_{\pm,\epsilon}(\epsilon t), \epsilon^2 B_{\pm,\epsilon}(\epsilon t), \epsilon^3 C_{\pm,\epsilon}(\epsilon t))_x \tag{2.91}$$

where $q_{\pm,\epsilon}(t)$ satisfies the differential equation

$$\frac{dq_{\pm,\epsilon}}{dt} = \epsilon A_{\pm,\epsilon}(\epsilon t). \tag{2.92}$$

The differential equation for q comes from eq. (2.30) and the fact that $\zeta_1^*(\Phi_\mu) = 0$. Using the fact that for solutions of eq. (2.5) satisfy $\partial_t(U)_x = U_\xi$, we also have that

$$\dot{x}_{\pm,\epsilon}(t) = \epsilon A_{\pm,\epsilon}(\epsilon t) + (\Phi_\mu(\epsilon A_{\pm,\epsilon}(\epsilon t), \epsilon^2 B_{\pm,\epsilon}(\epsilon t), \epsilon^3 C_{\pm,\epsilon}(\epsilon t))_\xi. \tag{2.93}$$

The coordinates $A_{\pm,\epsilon}$, $B_{\pm,\epsilon}$, and $C_{\pm,\epsilon}$ are at least C^5 and Φ_μ is at least C^4.

The main result will be stated by writing eq. (2.2) in terms of the strain variables, $u_n = x_{n+1} - x_n$. That is, we look at the travelling wave solution given by

$$x_{\pm,\epsilon}(n + 1 - c\tilde{t}) - x_{\pm,\epsilon}(n - c\tilde{t}). \tag{2.94}$$

Using a Taylor series expansion centered at $t+\frac{1}{2}$ for $x_{\pm,\epsilon}(t+1)$ and $x_{\pm,\epsilon}(t)$, we get

$$x_{\pm,\epsilon}(t+1/2)+\frac{1}{2}\dot{x}_{\pm,\epsilon}(t+1/2)+\frac{1}{8}\ddot{x}_{\pm,\epsilon}(t+1/2)+\frac{1}{2}\int_0^{1/2}\dddot{x}_{\pm,\epsilon}(t+1/2+s)(s-1/2)^2\,ds \tag{2.95}$$

$$x_{\pm,\epsilon}(t+1/2)-\frac{1}{2}\dot{x}_{\pm,\epsilon}(t+1/2)+\frac{1}{8}\ddot{x}_{\pm,\epsilon}(t+1/2)-\frac{1}{2}\int_0^{1/2}\dddot{x}_{\pm,\epsilon}(t+s)s^2\,ds \tag{2.96}$$

respectively for each term. Thus

$$\begin{aligned}
&x_{\pm,\epsilon}(t+1)-x_{\pm,\epsilon}(t)\\
&\quad=\dot{x}_{\pm,\epsilon}(t+1/2)+\frac{1}{2}\int_0^{1/2}[\,\dddot{x}_{\pm,\epsilon}(t+1/2+s)(s-1/2)^2-\dddot{x}_{\pm,\epsilon}(t+s)s^2]ds\\
&\quad=\pm\epsilon\phi(\epsilon(t+1/2))+\epsilon^2\mathcal{R}_{\epsilon,\pm}(\epsilon(t+1/2))
\end{aligned} \tag{2.97}$$

where $R_{\epsilon,\pm}\in C_b^3$. Thus (after shifting the solution) we have that there are travelling wave like solutions of the FPUT of the form

$$u_n(t)=\pm\epsilon\phi_{\epsilon,\pm}(\epsilon(n-ct))+\epsilon^2\mathcal{R}_{\epsilon,\pm}(\epsilon(n-ct)) \tag{2.98}$$

Additionally, if (H2) holds we improve the error estimate so that there are solutions of the form

$$u_n(t)=\pm\epsilon\phi(\epsilon(n-ct))+\epsilon^3\mathcal{R}_{\epsilon,\pm}(\epsilon(n-ct)). \tag{2.99}$$

To match similar estimates made in (Friesecke and Pego, 1999), one would expect the remainder terms to also be in a Sobolev space like H^1. This is in general not true. The travelling wave solution found above may approach a different limit asymptotically than $\epsilon\phi$, in which case the remainder does not approach zero asymptotically in space. A necessary condition to get $R_{\epsilon,\pm}\in H^1$ would be for u_n to approach the same limits of $\pm\epsilon\phi(\epsilon(n-ct))$ as $|n|\to\infty$.

A useful tool for showing this is the following invariant for eq. (2.4):

$$\dot{x}(t) - \mu \int_t^{t+1} V'(x(s) - x(s-1))\, ds. \tag{2.100}$$

It is easy to check that the above is constant for solutions of the advance-delay differential equation. If $\dot{x}(t) \to r_\infty$ as $t \to \infty$, then eq. (2.100) is equal to

$$r_\infty - \mu V'(r_\infty) \tag{2.101}$$

If we also have that $\dot{x}(t) \to r_{-\infty}$ as $t \to -\infty$, then we have eq. (2.100) is also equal to

$$r_{-\infty} - \mu V'(r_{-\infty}) \tag{2.102}$$

and so the limits $r_{\pm\infty}$ satisfy the equation

$$r_\infty - \mu V'(r_\infty) = r_{-\infty} - \mu V'(r_{-\infty}). \tag{2.103}$$

For arbitrary V, we cannot show that the limits agree with the limits of $\pm\epsilon\phi$. However, if we assume (H3) holds, i.e. that $V(x) = \frac{1}{2}x^2 - \frac{1}{24}x^4$, then we do have the limits agree. This follows in part from the oddness of V' and the reversibility of the system. Recall that the vector field on the center manifold is reversible (given by the reversibility operator S_0) and odd. Therefore, we have that

$$\begin{bmatrix} -A_{\pm,\epsilon}(-s) \\ B_{\pm,\epsilon}(-s) \\ -C_{\pm,\epsilon}(-s) \end{bmatrix} \tag{2.104}$$

is also a solution on the center manifold. One can note that the above solutions lie on the intersection of the stable and unstable manifolds and are in an ϵ-neighborhood of the unperturbed heteroclinic orbits $\gamma_\pm(s)$. This contradicts the uniqueness of the transverse intersection of the manifolds in a neighborhood of the original intersection, and thus the above solutions must in fact be $\gamma_{\pm,\epsilon}(s)$. Hence, comparing the limits at infinity we have that $\lim_{s\to\infty} A_{\pm,\epsilon}(s) = -\lim_{s\to-\infty} A_{\pm,\epsilon}(s)$ and so $r_\infty = -r_{-\infty}$. Therefore, we must have that

$$r_\infty - \mu V'(r_\infty) = 0. \tag{2.105}$$

Given that $V'(x) = x - \frac{1}{6}x^3$, we have that the only solutions to the above equation are $r_\infty = 0, \pm\epsilon/\sqrt{2}$. This implies that the limits of the travelling wave solutions agree with $\pm\epsilon\phi$. Specifically, we must have $A_{\pm,\epsilon}(s) \to \pm 1/\sqrt{2}$ as $s \to \infty$ and $A_{\pm,\epsilon}(s) \to \mp 1/\sqrt{2}$ as $s \to -\infty$.

Now to get the Sobolev estimate, we use the following lemma.

Lemma 2.1. *Suppose that (H3) holds. Then there exist $C > 0$ and $\alpha > 0$ such that*

$$|\gamma_{\pm,\epsilon}(s) - \gamma_\pm(s)| \le Ce^{-\alpha|s|}\epsilon. \tag{2.106}$$

Furthermore, the difference of the heteroclinic orbits are in $H^5(\mathbb{R};\mathbb{R}^3)$ and

$$\|\gamma_{\pm,\epsilon}(s) - \gamma_\pm(s)\|_{H^5(\mathbb{R};\mathbb{R}^3)} \le C\epsilon. \tag{2.107}$$

The proof of lemma 2.1 is given in appendix B. Thus we have

$$\|A_{\pm,\epsilon} - \phi_\pm\|_{H^5(\mathbb{R};\mathbb{R})} \le C\epsilon. \tag{2.108}$$

One can also show from the exponential decay of $\gamma_{\pm,\epsilon}$ to $\gamma_{\pm\infty}$ and the smoothness of Φ_μ that

$$(\Phi_\mu(\epsilon A_{\pm,\epsilon}(s), \epsilon^2 B_{\pm,\epsilon}(s), \epsilon^3 C_{\pm,\epsilon}(s))_\xi \in H^4(\mathbb{R};\mathbb{R}). \tag{2.109}$$

By noticing that the Taylor expansion of $\Phi_\mu(\epsilon A_{\pm,\epsilon}, \epsilon^2 B_{\pm,\epsilon}, \epsilon^3 C_{\pm,\epsilon})$ has no terms of order ϵ^2 or lower, the function is at least of order ϵ^3. Therefore, we have that there is an $R_{\pm,\epsilon} \in H^4(\mathbb{R};\mathbb{R})$ such that

$$\dot{x}_{\pm,\epsilon}(t) = \pm\epsilon\phi(\epsilon t) + \epsilon^2 R_{\pm,\epsilon}(\epsilon t). \tag{2.110}$$

Then converting to strain coordinates as before we have

$$u_n(t) = \pm\epsilon\phi(\epsilon(n - ct)) + \epsilon^2\mathcal{R}_{\pm,\epsilon}(\epsilon(n - ct)) \tag{2.111}$$

where $\mathcal{R}_{\pm,\epsilon} \in H^3(\mathbb{R};\mathbb{R})$.

We state our results as follows.

Theorem 2.1. *There exists $\epsilon_0 > 0$ and $C > 0$ such that for every $\epsilon > (0, \epsilon_0]$ there is a travelling wave solution given by $u_n(t) = u_c(n - ct)$ with positive wave speed $c^2 = 1 - \epsilon^2/12$. Furthermore, we have the additional estimates on the wave profile of u_c.*

(i) If (H1) holds, then

$$\left\| \frac{1}{\epsilon} u_c\left(\frac{\cdot}{\epsilon}\right) - \phi \right\|_{C^3} \leq C\epsilon \tag{2.112}$$

(ii) If (H2) holds, then

$$\left\| \frac{1}{\epsilon} u_c\left(\frac{\cdot}{\epsilon}\right) - \phi \right\|_{C^3} \leq C\epsilon^2 \tag{2.113}$$

(iii) If (H3) holds, then

$$\left\| \frac{1}{\epsilon} u_c\left(\frac{\cdot}{\epsilon}\right) - \phi \right\|_{H^3} \leq C\epsilon \tag{2.114}$$

The same estimates hold for $-\phi$ as the profile wave or for left-moving waves or for both.

Chapter 3

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

3.1 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

$$\ddot{x}_n = V'(x_{n+1} - x_n) - V'(x_n - x_{n-1}), \quad n \in \mathbb{Z}. \tag{3.1}$$

where V is the interaction potential between neighboring particles and $\dot{}$ denotes the derivative with respect to the time $t \in \mathbb{R}$. Equation (3.1) can be rewritten in the strain variables $u_n := x_{n+1} - x_n$ as follows

$$\ddot{u}_n = V'(u_{n+1}) - 2V'(u_n) + V'(u_{n-1}), \quad n \in \mathbb{Z} \tag{3.2}$$

The moving wave solution in eq. (3.1) corresponds to a kink solution in eq. (3.2).

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

$$2\partial_T W + \frac{1}{12}\partial_X^3 W + \partial_X(W^p) = 0, \quad X \in \mathbb{R} \tag{3.3}$$

serves as a modulation equation for solutions of eq. (3.2) (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. (3.3) 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

$$\|u_{\text{in}} - W(\epsilon\cdot, 0)\|_{\ell^2} + \|\dot{u}_{\text{in}} + \epsilon\partial_X W(\epsilon\cdot, 0)\|_{\ell^2} \leq \epsilon^{3/2}, \tag{3.4}$$

the unique solution to eq. (3.2) 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

$$\|u(t) - W(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2(\mathbb{Z})} + \|\dot{u}(t) + \epsilon\partial_X W(\epsilon(\cdot - t), \epsilon^3 t)\|_{\ell^2(\mathbb{Z})} \leq C_0\epsilon^{3/2},$$
$$t \in [-\tau_0\epsilon^{-3}, \tau_0\epsilon^{-3}]. \tag{3.5}$$

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

$$V(u) = \frac{1}{2}u^2 - \frac{1}{24}u^4. \tag{3.6}$$

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. (3.6) (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:

$$\dot{X}(t) = J\mathcal{H}'(X) \tag{3.7}$$

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. (3.7) in the sense that

$$\mathrm{Res}(t) := J[L\tilde{X}_\epsilon(t) + N(\tilde{X}_\epsilon(t))] - \dot{\tilde{X}}_\epsilon(t) \tag{3.8}$$

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

$$\dot{R}(t) = J[L + N'(\tilde{X}_\epsilon(t))]R(t) + \mathrm{Res}(t) + \mathcal{N}(\tilde{X}_\epsilon, R) \tag{3.9}$$

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. (3.9) 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. (3.9). 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

$$|\dot{\mathcal{S}}(t)| \lesssim \epsilon^\alpha + \epsilon^{k+2}\mathcal{S}(t). \tag{3.10}$$

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 $|t| \lesssim r\epsilon^{-(k+2)}|\log(\epsilon)|$.

3.2 Counter-Propagating Waves Ansatz

We make the assumption that solutions of eq. (3.2) can be expressed as a sum of two counter-propagating small-amplitude waves, i.e.,

$$u_n(t) \approx \epsilon f(\epsilon(n+t), \epsilon^3 t) + \epsilon g(\epsilon(n-ct), \epsilon^3 t) + \epsilon^3 \phi(\epsilon n, \epsilon t) \tag{3.11}$$

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

$$c = c(\epsilon, f_\infty) = 1 - \frac{\epsilon^2 f_\infty^2}{4}. \tag{3.12}$$

Plugging in the ansatz in eq. (3.11) back into eq. (3.2) and grouping terms of the same order ϵ together gives

$$
\begin{aligned}
&\epsilon^3\Big(\partial_1^2 f(\cdot,\epsilon^3 t)+\partial_1^2 g(\cdot,\epsilon^3 t)\Big)\\
&+\epsilon^5\Big(2\partial_1\partial_2 f(\cdot,\epsilon^3 t)-2\partial_1\partial_2 g(\cdot,\epsilon^3 t)-\frac{f_\infty^2}{2}\partial_1^2 g+\partial_2^2\phi(\epsilon x,\epsilon t)\Big)\\
&+\mathcal{O}(\epsilon^6)\\
&\qquad= \quad \epsilon^3\Big(\partial_1^2 f(\cdot,\epsilon^3 t)+\partial_1^2 g(\cdot,\epsilon^3 t)\Big)\\
&\qquad\qquad +\epsilon^5\Big(\partial_1^2\phi(\epsilon x,\epsilon t)\\
&\qquad\qquad\quad -\frac{1}{6}\partial_1^2\big[f^3(\cdot,\epsilon^3 t)+3f^2(\cdot,\epsilon^3)g(\cdot,\epsilon^3 t)+3f(\cdot,\epsilon t)g^2(\cdot,\epsilon^3 t)+g^3(\cdot,\epsilon^3 t)\big]\\
&\qquad\qquad\quad +\frac{1}{12}\partial_1^4 f(\cdot,\epsilon^3 t)+\frac{1}{12}\partial_1^4 g(\cdot,\epsilon^3 t)\Big)\\
&\qquad\qquad +\mathcal{O}(\epsilon^6).
\end{aligned}
\tag{3.13}
$$

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

$$2\partial_2 f=-\frac{1}{6}\partial_1(f^3)+\frac{1}{12}\partial_1^3 f \tag{3.14}$$

and

$$-2\partial_2 g=-\frac{1}{6}\partial_1(g^3+3f_\infty g^2)+\frac{1}{12}\partial_1^3 g, \tag{3.15}$$

and

$$
\begin{aligned}
\partial_2^2\phi(\xi,\tau)=\partial_1^2\phi(\xi,\tau)-\frac{1}{6}\partial_1^2\big[3(f^2(\xi+\tau,\epsilon^2\tau)-f_\infty^2)g(\xi-c\tau,\epsilon^2\tau)&\\
+3(f(\xi+\tau,\epsilon^2\tau)-f_\infty)g^2(\xi-c\tau,\epsilon^2\tau)\big]&
\end{aligned}
\tag{3.16}
$$

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

Note that eq. (3.14) is the defocusing mKdV equation and eq. (3.15) 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. (3.11) 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 3.1. *For $k \in \mathbb{N}$, let $\mathcal{X}^k(\mathbb{R})$ be the Banach space*

$$\mathcal{X}^k(\mathbb{R}) := \{f \in L^\infty(\mathbb{R}) \mid f' \in H^{k-1}(\mathbb{R})\} \tag{3.17}$$

with norm

$$\|f\|_{\mathcal{X}^k(\mathbb{R})} := \|f\|_{L^\infty(\mathbb{R})} + \|f'\|_{H^{k-1}(\mathbb{R})}. \tag{3.18}$$

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

$$\mathcal{X}^k(\mathbb{R}) \cong L^\infty(\mathbb{R}) \cap \dot{H}^1(R) \cap \dot{H}^k(\mathbb{R}), \tag{3.19}$$

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. (3.14) has kink solutions of the form

$$f(X,T) = -\sqrt{12v}\tanh\left(\sqrt{12v}(X - vT)\right). \tag{3.20}$$

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

$$-\frac{\epsilon}{\sqrt{2}}\tanh\left(\frac{\epsilon}{\sqrt{2}}\left(n + \left(1 - \frac{\epsilon^2}{24}\right)t\right)\right), \tag{3.21}$$

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

$$\|fg\|_{H^k} \leq C\|f\|_{\mathcal{X}^k}\|g\|_{H^k} \tag{3.22}$$

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

Lemma 3.2. *For non-negative integers k, there is a $C > 0$ such that*

$$\|fg\|_{\mathcal{X}^k} \leq C\|f\|_{\mathcal{X}^k}\|h\|_{\mathcal{X}^k} \tag{3.23}$$

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

See appendix B 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. (3.16) will quickly go to zero, and the equation governing ϕ eq. (3.16) 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 $k, n \in \mathbb{N} \cup \{0\}$

$$H^k_n(\mathbb{R}) := \{g \in H^k(\mathbb{R}) \mid g\langle\cdot\rangle^n \in H^k\} \tag{3.24}$$

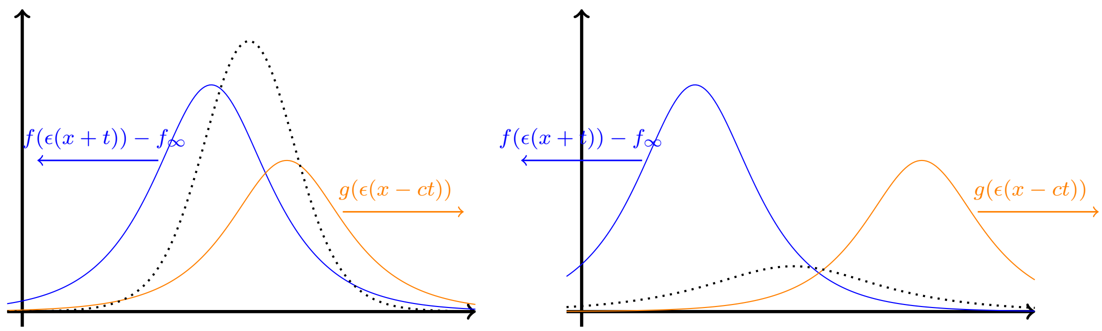

Figure 3·1: The function $f(\epsilon(x+t)) - f_\infty$ (shown in blue) moves to the left while $g(\epsilon(x-ct))$ (shown in orange) moves to the right. Since they are localized, the product (shown by the dotted line) will quickly decay in time.

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

$$\|g\|_{H^k_n(\mathbb{R})} := \|g\langle \cdot \rangle^n\|_{H^k(\mathbb{R})}. \tag{3.25}$$

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

$$c\|\hat{g}\|_{H^n_k} \leq \|g\|_{H^k_n} \leq C\|\hat{g}\|_{H^n_k} \tag{3.26}$$

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

$$\langle x \rangle_+ = \begin{cases} \langle x \rangle, & x > 1 \\ 1, & x < 0 \end{cases} \tag{3.27}$$

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

$$\mathcal{X}^k_{n^+}(\mathbb{R}) := \{f \in \mathcal{X}^k(\mathbb{R}) \mid \lim_{x\to\infty} f(x) = f_\infty \textit{ and } (f - f_\infty)\langle\cdot\rangle_+^n \in \mathcal{X}^k(\mathbb{R})\} \tag{3.28}$$

with norm given by

$$\|f\|_{\mathcal{X}^k_{n^+}(\mathbb{R})} := |f_\infty| + \|(f - f_\infty)\langle\cdot\rangle_+^n\|_{\mathcal{X}^k(\mathbb{R})} \tag{3.29}$$

Similarly,

$$\mathcal{X}^k_{n^-}(\mathbb{R}) := \{f \in \mathcal{X}^k(\mathbb{R}) \mid \lim_{x\to-\infty} f(x) = f_{-\infty} \textit{ and } (f - f_{-\infty})\langle\cdot\rangle_-^n \in \mathcal{X}^k(\mathbb{R})\} \tag{3.30}$$

and

$$\|f\|_{\mathcal{X}^k_{n^-}(\mathbb{R})} := |f_{-\infty}| + \|(f - f_{-\infty})\langle\cdot\rangle_-^n\|_{\mathcal{X}^k(\mathbb{R})} \tag{3.31}$$

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

$$\mathcal{X}^k_n(\mathbb{R}) := \mathcal{X}^k_{n^+}(\mathbb{R}) \cap \mathcal{X}^k_{n^-}(\mathbb{R}), \quad \|f\|_{\mathcal{X}^k_n(\mathbb{R})} := \|f\|_{\mathcal{X}^k_{n^+}(\mathbb{R})} + \|f\|_{\mathcal{X}^k_{n^-}(\mathbb{R})}. \tag{3.32}$$

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

$$(\alpha, f) \mapsto \alpha + f\langle\cdot\rangle_\pm^{-n}. \tag{3.33}$$

One can show that the kink solutions as specified in eq. (3.20) 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 3.3. *For each $k \geq 0$ and $c > 0$, there exists $C > 0$ depending only on k such that*

$$\left\| \frac{1}{\langle \cdot + \tau \rangle_+^2 \langle \cdot - c\tau \rangle^2} \right\|_{C^k} \leq C \sup_{x \in \mathbb{R}} \frac{1}{\langle x + \tau \rangle_+^2 \langle x - c\tau \rangle^2}. \tag{3.34}$$

Furthermore,

$$\int_0^\infty \sup_{x \in \mathbb{R}} \frac{1}{\langle x + \tau \rangle_+^2 \langle x - c\tau \rangle^2} \, d\tau < \infty. \tag{3.35}$$

See appendix B for proof.

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

Proposition 3.1. *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*

$$\sup_{t \in [-\epsilon^{-3}T_0, \epsilon^{-3}T_0]} \|\phi(\cdot, \epsilon t)\|_{H^k} \leq C \left(\sup_{t \in [-\epsilon^{-3}T_0, \epsilon^{-3}T_0]} \left\{ \|f(\cdot, \epsilon^3 t)\|_{\mathcal{X}_2^{k+1}}, \|g(\cdot, \epsilon^3 t)\|_{H_2^{k+1}} \right\} \right)^3 \tag{3.36}$$

and

$$\sup_{t \in [-\epsilon^{-3}T_0, \epsilon^{-3}T_0]} \|\psi(\cdot, \epsilon t)\|_{H^{k-1}} \leq C \left(\sup_{t \in [-\epsilon^{-3}T_0, \epsilon^{-3}T_0]} \left\{ \|f(\cdot, \epsilon^3 t)\|_{\mathcal{X}_2^{k+1}}, \|g(\cdot, \epsilon^3 t)\|_{H_2^{k+1}} \right\} \right)^3, \tag{3.37}$$

where $\psi = \partial_2 \phi$.

Proof. Set $\partial_2 \phi = \psi$. Taking the Fourier transform $\mathcal{F}$ on both sides of eq. (3.16) and writing the ODE as a first order system, we get that

$$\begin{aligned} \partial_2 \begin{bmatrix} \hat{\phi}(k, \tau) \\ \hat{\psi}(k, \tau) \end{bmatrix} &= \begin{bmatrix} \hat{\psi}(k, \tau) \\ -k^2 \hat{\phi}(k, \tau) \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \frac{1}{2} k^2 \mathcal{F}[(f^2(\cdot + \tau), \epsilon^2 \tau) - f_\infty^2) g(\cdot - c\tau, \epsilon^2 \tau) + (f(\cdot + \tau, \epsilon^2 \tau) - f_\infty) g^2(\cdot - c\tau, \epsilon^2 \tau)](k) \end{bmatrix}. \end{aligned} \tag{3.38}$$

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

$$\hat{\phi}(k,T) = \frac{1}{2}\int_0^T k\sin(k(T-\tau))\times \\ \mathcal{F}[(f^2(\cdot+\tau),\epsilon^2\tau) - f_\infty^2)g(\cdot - c\tau,\epsilon^2\tau) + (f(\cdot+\tau,\epsilon^2\tau) - f_\infty)g^2(\cdot - c\tau,\epsilon^2\tau)](k)\, d\tau \tag{3.39}$$

and

$$\hat{\psi}(k,T) = \frac{1}{2}\int_0^T k^2\cos(k(T-\tau))\times \\ \mathcal{F}[(f^2(\cdot+\tau,\epsilon^2\tau) - f_\infty^2)g(\cdot - c\tau,\epsilon^2\tau) + (f(\cdot+\tau,\epsilon^2\tau) - f_\infty)g^2(\cdot - c\tau,\epsilon^2\tau)](k)\, d\tau \tag{3.40}$$

Hence we can get that

$$\begin{aligned}
&\|\phi(\cdot,T)\|_{H^k} \\
&\quad\le C\|\hat{\phi}(\cdot,T)\|_{H_k^0} \\
&\quad\le C\int_0^T \|\partial_1((f^2(\cdot+\tau) - f_\infty^2)g(\cdot - c\tau))\|_{H^k} + \|\partial_1((f(\cdot+\tau) - f_\infty)g^2(\cdot - c\tau))\|_{H^k}\,\mathrm{d}\tau \\
&\quad\le C\int_0^T \|f(\cdot+\tau)\partial_1 f(\cdot+\tau)g(\cdot - c\tau)\|_{H^k} + \|(f^2(\cdot+\tau) - f_\infty^2)\partial_1 g(\cdot - c\tau)\|_{H^k} \\
&\qquad + \|\partial_1 f(\cdot+\tau)g^2(\cdot - c\tau)\|_{H^k} + \|(f(\cdot+\tau) - f_\infty)\partial_1 g(\cdot - c\tau)\|_{H^k}\,\mathrm{d}\tau \\
&\quad\le C\int_0^T \sup_{x\in\mathbb{R}} \frac{1}{\langle x+\tau\rangle_+^2\langle x - c\tau\rangle^2} \times \left(\|f\|^2_{\mathcal{X}_2^{k+1}}\|g\|_{H_2^{k+1}} + \|f\|_{\mathcal{X}_2^{k+1}}\|g\|^2_{H_2^{k+1}}\right)\mathrm{d}\tau,
\end{aligned} \tag{3.41}$$

whence eq. (3.36) follows. The proof for eq. (3.37) is analogous. □

3.3 Setup of Lattice Equations

The scalar second-order differential equation eq. (3.2) with potential V given by eq. (3.6) can be rewritten as the following first-order system:

$$\begin{cases} \dot{u}_n = q_{n+1} - q_n, \\ \dot{q}_n = u_n - u_{n-1} - \frac{1}{6}(u_n^3 - u_{n-1}^3), \end{cases} \qquad n\in\mathbb{Z}. \tag{3.42}$$

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

$$q_n(t) = \sum_{k=-\infty}^{n-1} \dot{u}_k(t) = \sum_{k=-\infty}^{n-1} [\dot{x}_{k+1}(t) - \dot{x}_k(t)] = \dot{x}_n(t) \tag{3.43}$$

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. (3.42) also puts the system into a Hamiltonian framework (when $u, q \in \ell^2(\mathbb{Z})$). Here the equations are of the form

$$\dot{U} = J\mathcal{H}'(U) \tag{3.44}$$

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

$$J = \begin{bmatrix} 0 & e^{\partial} - 1 \\ 1 - e^{-\partial} & 0 \end{bmatrix} \tag{3.45}$$

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. (3.42), but we first must assume certain regularity and decay of f and g.

Assumption 1. *Let f and g be solutions of eqs.* (3.14) *and* (3.15)*, 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

$$u_n(t) = \epsilon f(\epsilon(n+t), \epsilon^3 t) + \epsilon g(\epsilon(n-ct), \epsilon^3 t) + \epsilon^3 \phi(\epsilon n, \epsilon t) + \mathcal{U}_n(t) \tag{3.46}$$

and

$$q_n(t) = \epsilon F(\epsilon(n+t), \epsilon^3 t) + \epsilon G(\epsilon(n-ct), \epsilon^3 t) + \epsilon^3 \Phi(\epsilon n, \epsilon t) - \epsilon F_{-\infty} + \mathcal{Q}_n(t). \quad (3.47)$$

The wave speed c is again given by eq. (3.12).

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

$$\begin{aligned} q_n(t) &= \sum_{k=-\infty}^{n-1} \dot{u}_n(t) \\ &\approx \sum_{k=-\infty}^{n-1} [\epsilon^2 \partial_1 f(\epsilon(k+t)) + \epsilon^4 \partial_2 f(\epsilon(k+t)) \\ &\quad + \epsilon^2 c \partial_1 g(\epsilon(k-ct)) + \epsilon^4 \partial_2 g(\epsilon(k-ct)) \\ &\quad + \epsilon^4 \partial_2 \phi(\epsilon k)]. \end{aligned} \quad (3.48)$$

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

$$\begin{aligned} \epsilon F(\epsilon(n+1+t)) - \epsilon F(\epsilon(n+t)) &= \epsilon^2 \partial_1 f(\epsilon(n+t)) + \epsilon^4 \partial_2 f(\epsilon(n+1)) + \mathcal{O}(\epsilon^6) \\ \epsilon G(\epsilon(n+1-ct)) - \epsilon G(\epsilon(n-ct)) &= \epsilon^2 c \partial_1 g(\epsilon(n-ct)) + \epsilon^4 \partial_2 g(\epsilon(n-ct)) + \mathcal{O}(\epsilon^6) \\ \epsilon^3 \Phi(\epsilon(n+1)) - \epsilon^3 \Phi(\epsilon(n)) &= \epsilon^4 \partial_2 \phi(\epsilon n) + \mathcal{O}(\epsilon^6). \end{aligned} \quad (3.49)$$

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:

$$F := f - \frac{\epsilon}{2}\partial_1 f + \frac{\epsilon^2}{8}\partial_1^2 f - \frac{\epsilon^2}{12} f^3 - \frac{\epsilon^3}{48}\partial_1^3 f + \frac{\epsilon^3}{8} f^2 \partial_1 f \tag{3.50}$$

$$G := -g + \frac{\epsilon}{2}\partial_1 g + \frac{\epsilon^2 f_\infty^2}{4} g + \frac{\epsilon^2}{12}(g^3 + 3 f_\infty g^2) - \frac{\epsilon^2}{8}\partial_1^2 g + \frac{\epsilon^3}{48}\partial_1^3 g \tag{3.51}$$
$$- \frac{\epsilon^3}{24}\partial_1(g^3 + 3 f_\infty g^2) - \frac{\epsilon^3 f_\infty^2}{8}\partial_1 g$$

$$\Phi := \partial_1^{-1}\psi - \frac{\epsilon}{2}\psi. \tag{3.52}$$

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. (3.40). Namely, we have that

$$\begin{aligned}
&\mathcal{F}[\partial_1^{-1}\psi(\cdot, T)](k) = (ik)^{-1}\hat{\psi}(k, T) \\
&\quad = \frac{-i}{2}\int_0^T k\cos(k(T-\tau))\times \\
&\quad \mathcal{F}[(f^2(\cdot+\tau, \epsilon^2\tau) - f_\infty^2) g(\cdot - c\tau, \epsilon^2\tau) + (f(\cdot+\tau, \epsilon^2\tau) - f_\infty) g^2(\cdot - c\tau, \epsilon^2\tau)](k)\, d\tau
\end{aligned} \tag{3.53}$$

and (following the same calculations in eq. (3.41))

$$\|\partial_1^{-1}\psi(\cdot, T)\|_{H^5} \le C\int_0^T \sup_{x\in\mathbb{R}} \frac{1}{\langle x+\tau\rangle_+^2 \langle x - c\tau\rangle^2} \times \left(\|f\|_{\mathcal{X}_2^6}^2 \|g\|_{H_2^6} + \|f\|_{\mathcal{X}_2^6}\|g\|_{H_2^6}^2 \right) \mathrm{d}\tau. \tag{3.54}$$

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. (3.47) does not affect the dynamics of q in eq. (3.42)

An equivalent set of equations to eq. (3.42) are given by

$$\begin{cases} \dot{\mathcal{U}}_n(t) = \mathcal{Q}_{n+1}(t) - \mathcal{Q}_n(t) + \mathrm{Res}_n^{(1)}(t) \\ \dot{\mathcal{Q}}_n(t) = \mathcal{U}_n(t) - \mathcal{U}_{n-1}(t) \\ \qquad - \frac{1}{2}(\epsilon f(\epsilon(n+t)) + \epsilon g(\epsilon(n-ct)) + \epsilon^3\phi(\epsilon n))^2 \mathcal{U}_n(t) \qquad\qquad n \in \mathbb{Z}, \\ \qquad + \frac{1}{2}(\epsilon f(\epsilon(n-1+t)) + \epsilon g(\epsilon(n-1-ct)) + \epsilon^3\phi(\epsilon(n-1)))^2 \mathcal{U}_{n-1}(t) \\ \qquad + \mathrm{Res}_n^{(2)}(t) + \mathcal{B}_n(\epsilon f + \epsilon g + \epsilon^3\phi, \mathcal{U}) \end{cases} \tag{3.55}$$

where

$$\begin{aligned} \mathrm{Res}_n^{(1)}(t) =& \epsilon F(\epsilon(n+1+t)) - \epsilon F(\epsilon(n+t)) \\ &+ \epsilon G(\epsilon(n+1-ct) - \epsilon G(\epsilon(n-ct) + \epsilon^3\Phi(\epsilon(n+1)) - \epsilon^3\Phi(\epsilon n) \\ &- \epsilon^2\partial_1 f(\epsilon(n+t)) - \epsilon^4\partial_2 f(\epsilon(n+t)) \\ &+ \epsilon^2 c\partial_1 g(\epsilon(n-ct)) - \epsilon^4\partial_2 g(\epsilon(n-ct)) - \epsilon^4\partial_2\phi(\epsilon n), \end{aligned} \tag{3.56}$$

$$\begin{aligned} \mathrm{Res}_n^{(2)}(t) =& \epsilon f(\epsilon(n+t)) - \epsilon f(\epsilon(n-1+t)) \\ &+ \epsilon g(\epsilon(n-ct)) - \epsilon g(\epsilon(n-1-ct) + \epsilon^3\phi(\epsilon n) - \epsilon^3\phi(\epsilon(n-1)) \\ &- \epsilon^2\partial_1 F(\epsilon(n+t)) - \epsilon^4\partial_2 F(\epsilon(n+t)) \\ &+ \epsilon^2 c\partial_1 G(\epsilon(n-ct)) - \epsilon^4\partial_2 G(\epsilon(n-ct)) - \epsilon^4\partial_2\Phi(\epsilon n) \\ &- \frac{1}{6}\Big((\epsilon f(\epsilon(n+t)) + \epsilon g(\epsilon(n-ct)) + \epsilon^3\phi(\epsilon n))^3 \\ &\qquad - (\epsilon f(\epsilon(n-1+t)) + \epsilon g(\epsilon(n-1-ct)) + \epsilon^3\phi(\epsilon(n-1)))^3\Big), \end{aligned} \tag{3.57}$$

and

$$\begin{aligned}
&\mathcal{B}_n(\epsilon f + \epsilon g + \epsilon^3 \phi, \mathcal{U}) \\
&\quad = -\frac{1}{6}\Big(3(\epsilon f(\epsilon(n+t)) + \epsilon g(\epsilon(n-ct)) + \epsilon^3 \phi(\epsilon n))\mathcal{U}_n^2(t) \\
&\qquad - 3(\epsilon f(\epsilon(n-1+t)) + \epsilon g(\epsilon(n-1-ct)) + \epsilon^3 \phi(\epsilon(n-1)))\mathcal{U}_{n-1}^2(t) \\
&\qquad + \mathcal{U}_n^3(t) - \mathcal{U}_{n-1}^3(t)\Big).
\end{aligned} \tag{3.58}$$

The terms $\mathcal{U}$ and $\mathcal{Q}$ control the error associated with the ansatz in eqs. (3.46) and (3.47). 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.

3.4 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 3.4. *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 3.5. *Let f and g be solutions of eqs.* (3.14) *and* (3.15)*, 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*

$$\delta := \max\left\{ \sup_{\tau \in [-\tau_0, \tau_0]} \|f(\cdot, \tau)\|_{\mathcal{X}_2^6}, \sup_{\tau \in [-\tau_0, \tau_0]} \|g(\cdot, \tau)\|_{H_2^6} \right\} \tag{3.59}$$

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

$$\|\mathrm{Res}^{(1)}(t)\|_{\ell^2} + \|\mathrm{Res}^{(2)}(t)\|_{\ell^2} \leq C\epsilon^{11/2}(\delta + \delta^5) \tag{3.60}$$

and

$$\|\mathcal{B}_n(\epsilon f + \epsilon g + \epsilon^3\phi, \mathcal{U})\|_{\ell^2} \leq C\epsilon[(\delta + \epsilon^2\delta^3)\|\mathcal{U}\|_{\ell^2}^2 + \|\mathcal{U}\|_{\ell^2}^3] \tag{3.61}$$

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. (3.14), we get the following:

$$\begin{array}{l} \epsilon F(\cdot + \epsilon) - \epsilon F - \epsilon^2\partial_1 f - \epsilon^4\partial_2 f = \\ \qquad \begin{array}{lllll} \epsilon^2\partial_1 f + & \dfrac{\epsilon^3}{2}\partial_1^2 f + & \dfrac{\epsilon^4}{6}\partial_1^3 f & +\dfrac{\epsilon^5}{24}\partial_1^4 f & \\ & -\dfrac{\epsilon^3}{2}\partial_1^2 f & -\dfrac{\epsilon^4}{4}\partial_1^3 f & -\dfrac{\epsilon^5}{12}\partial_1^4 f & \\ & & +\dfrac{\epsilon^4}{8}\partial_1^3 f & +\dfrac{\epsilon^5}{16}\partial_1^4 f & \\ & & -\dfrac{\epsilon^4}{12}\partial_1(f^3) & -\dfrac{\epsilon^5}{24}\partial_1^2(f^3) & \\ & & & -\dfrac{\epsilon^5}{48}\partial_1^4 f & \\ & & & +\dfrac{\epsilon^5}{24}\partial_1^2(f^3) & \\ -\epsilon^2\partial_1 f & & & & \\ & & +\dfrac{\epsilon^4}{12}\partial_1(f^3) & & \\ & & -\dfrac{\epsilon^4}{24}\partial^3 f & & +I_{f,1}(n,t), \end{array} \end{array} \tag{3.62}$$

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

$$\begin{aligned} I_{f,1}(n,t) := {} & \frac{\epsilon^6}{24}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(1-s)^4\,ds - \frac{\epsilon^6}{12}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(1-s)^3\,ds \\ & + \frac{\epsilon^6}{16}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(1-s)^2\,ds - \frac{\epsilon^6}{24}\int_0^1 \partial_1^3(f^3)(\epsilon(n+t+s))(1-s)^2\,ds \\ & - \frac{\epsilon^6}{48}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(1-s)\,ds + \frac{\epsilon^6}{24}\int_0^1 \partial_1^3(f^3)(\epsilon(n+t+s))(1-s)\,ds. \end{aligned} \tag{3.63}$$

Note that all the terms in eq. (3.62) cancel except $I_{f,1}$, and so we are only left with terms of order ϵ^6. Applying lemma 3.4 (and lemmas 3.1 and 3.2 when needed) to the terms in eq. (3.63) gives that the ℓ^2 norm on the left-hand side of eq. (3.62) 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

$$\begin{aligned}
&\epsilon G(\cdot+\epsilon) - \epsilon G + \epsilon^2 c\partial_1 g - \epsilon^4 \partial_2 g = \\
&\qquad -\epsilon^2\partial_1 g \quad -\frac{\epsilon^3}{2}\partial_1^2 g \quad -\frac{\epsilon^4}{6}\partial_1^3 g \quad -\frac{\epsilon^5}{24}\partial_1^4 g \\
&\qquad +\frac{\epsilon^3}{2}\partial_1^2 g \quad +\frac{\epsilon^4}{4}\partial_1^3 g \quad +\frac{\epsilon^5}{12}\partial_1^4 g \\
&\qquad +\frac{\epsilon^4 f_\infty^2}{4}\partial_1 g \quad +\frac{\epsilon^5 f_\infty^2}{8}\partial_1^2 g \\
&\qquad +\frac{\epsilon^4}{12}\partial_1(g^3) \quad +\frac{\epsilon^5}{24}\partial_1^2(g^3) \\
&\qquad +\frac{\epsilon^4}{12}\partial_1(3f_\infty g^2) \quad +\frac{\epsilon^5}{24}\partial_1^2(3f_\infty g^2) \\
&\qquad -\frac{\epsilon^4}{8}\partial_1^3 g \quad -\frac{\epsilon^5}{16}\partial_1^4 g \\
&\qquad +\frac{\epsilon^5}{48}\partial_1^4 g \\
&\qquad -\frac{\epsilon^5}{24}\partial_1^2(g^3) \\
&\qquad -\frac{\epsilon^5}{24}\partial_1^2(3f_\infty g^2) \\
&\qquad -\frac{\epsilon^5 f_\infty^2}{8}\partial_1^2 g \\
&\qquad +\epsilon^2\partial_1 g \\
&\qquad -\frac{\epsilon^4 f_\infty^2}{4}\partial_1 g \\
&\qquad -\frac{\epsilon^4}{12}\partial_1(g^3) \\
&\qquad -\frac{\epsilon^4}{12}\partial_1(3f_\infty g^2) \\
&\qquad +\frac{\epsilon^4}{24}\partial_1^3 g \qquad +I_{g,1}(nt),
\end{aligned} \tag{3.64}$$

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

$$
\begin{aligned}
&I_{g,1}(n,t) := \\
&-\frac{\epsilon^6}{24}\int_0^1 \partial_1^5 g(\epsilon(n-ct+s))(1-s)^4\,ds + \frac{\epsilon^6}{12}\int_0^1 \partial_1^5 g(\epsilon(n-ct+s))(1-s)^3\,ds \\
&+\frac{\epsilon^6 f_\infty^2}{8}\int_0^1 \partial_1^3 g(\epsilon(n-ct+s))(1-s)^2\,ds + \frac{\epsilon^6}{24}\int_0^1 \partial_1^3 (g^3)(\epsilon(n-ct+s))(1-s)^2\,ds \\
&+\frac{\epsilon^6}{24}\int_0^1 \partial_1^3 (3f_\infty g^2)(\epsilon(n-ct+s))(1-s)^2\,ds - \frac{\epsilon^6}{16}\int_0^1 \partial_1^5 g(\epsilon(n-ct+s))(1-s)^2\,ds \\
&+\frac{\epsilon^6}{48}\int_0^1 \partial_1^5 g(\epsilon(n-ct+s))(1-s)\,ds - \frac{\epsilon^6}{24}\int_0^1 \partial_1^3 (g^3)(\epsilon(n-ct+s))(1-s)\,ds \\
&-\frac{\epsilon^6}{24}\int_0^1 \partial_1^3 (3f_\infty g^2)(\epsilon(n-ct+s))(1-s)\,ds - \frac{\epsilon^6 f_\infty^2}{8}\int_0^1 \partial_1^3 g(\epsilon(n-ct+s))(1-s)\,ds
\end{aligned} \tag{3.65}
$$

All terms except those of order ϵ^6 cancel and the terms in eq. (3.65) can be controlled by lemma 3.4.

Similarly we have

$$
\epsilon^3\Phi(\epsilon(n+1),\epsilon t) - \epsilon^3\Phi(\epsilon n,\epsilon t) - \epsilon^4\partial_2\phi_2(\epsilon n,\epsilon t) = \frac{\epsilon^6}{2}\int_0^1 \partial_1^2\psi(\epsilon(n+s),\epsilon t)(1-s)^2\,ds, \tag{3.66}
$$

so the ℓ^2 norm can also be controlled.

Therefore we have

$$
\|\mathrm{Res}^{(1)}(t)\|_{\ell^2} \le C\epsilon^{11/2}(\delta+\delta^3) \tag{3.67}
$$

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

$$
\begin{array}{lllll}
\multicolumn{5}{l}{\epsilon f(\cdot) - \epsilon f(\cdot-\epsilon) - \epsilon^2\partial_1 F - \epsilon^4\partial_2 F - \dfrac{\epsilon^3}{6}(f^3(\cdot) - f^3(\cdot-\epsilon)) =} \\
\quad \epsilon^2\partial_1 f & -\dfrac{\epsilon^3}{2}\partial_1 f & +\dfrac{\epsilon^4}{6}\partial_1^3 f & -\dfrac{\epsilon^5}{24}\partial_1^4 f & \\
\quad -\epsilon^2\partial_1 f & +\dfrac{\epsilon^3}{2}\partial_1^2 f & +\dfrac{\epsilon^4}{12}\partial_1(f^3) - \dfrac{\epsilon^4}{8}\partial_1^3 f & +\dfrac{\epsilon^5}{48}\partial_1^4 f - \dfrac{\epsilon^5}{24}\partial_1^2(f^3) & \\
 & & -\epsilon^4\partial_2 f & +\dfrac{\epsilon^5}{2}\partial_1\partial_2 f & \\
 & & -\dfrac{\epsilon^4}{6}\partial_1(f^3) & +\dfrac{\epsilon^5}{12}\partial_1(f^3) & + I_{f,2}(n,t).
\end{array} \tag{3.68}
$$

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

$$\begin{aligned} I_{f,2}(n,t) := &-\frac{\epsilon^6}{24}\int_0^1 \partial_1^5 f(\epsilon(n+t+s))(s-1)^4\,ds \\ &+\frac{\epsilon^6}{12}\int_0^1 \partial_1^2(f^3)(\epsilon(n+t+s))(s-1)^2\,ds \\ &+\epsilon^6\partial_2\left(\frac{1}{8}\partial_1^2 f - \frac{1}{12}f^3 - \frac{\epsilon}{48}\partial_1^3 f + \frac{\epsilon}{8}f^2\partial_1 f\right) \end{aligned} \tag{3.69}$$

All the above terms in eq. (3.68) cancel except for $I_{f,2}(n,t)$. The integral terms in eq. (3.69) 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. (3.14); then the terms can be controlled by lemma 3.4. Then the left-hand side of eq. (3.68) 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

$$\begin{array}{llll} \epsilon^2\partial_1 g & -\frac{\epsilon^3}{2}\partial_1^2 g & +\frac{\epsilon^4}{6}\partial_1^3 g & -\frac{\epsilon^5}{24}\partial_1^4 g \\ & & +\epsilon^4\partial_1\phi & -\frac{\epsilon^5}{2}\partial_1^2\phi \\ -\epsilon^2\partial_1 g & +\frac{\epsilon^3}{2}\partial_1^2 g & -\frac{\epsilon^4}{8}\partial_1^3 g + \frac{\epsilon^4 f_\infty^2}{4}\partial_1 g & +\frac{\epsilon^5}{48}\partial_1^4 g \\ & & +\frac{\epsilon^4}{12}\partial_1(g^3+3f_\infty g^2) & -\frac{\epsilon^5}{24}\partial_1^2(g^3+3f_\infty g^2) \\ & & & -\frac{\epsilon^5 f_\infty^2}{8}\partial_1^2 g \\ & & +\frac{\epsilon^4 f_\infty^2}{4}\partial_1 g & -\frac{\epsilon^5 f_\infty^2}{8}\partial_1^2 g \\ & & +\epsilon^4\partial_2 g & -\frac{\epsilon^5}{2}\partial_1\partial_2 g \\ & & -\epsilon^4\partial_2\partial_1^{-1}\psi & +\frac{\epsilon^5}{2}\partial_2\psi \\ & & -\frac{\epsilon^4}{6}\partial_1(g^3+3g^2 f+3gf^2) & +\frac{\epsilon^5}{12}\partial_1^2(g^3+3g^2 f+3gf^2), \end{array} \tag{3.70}$$

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

$$
\begin{aligned}
&I_{g,2}(n,t) = \\
&-\frac{\epsilon^6}{24}\int_0^1 \partial_1^5 g(\epsilon(n-s-ct))(s-1)^4\,ds - \frac{\epsilon^6}{2}\int_0^1 \partial_1^3\phi(\epsilon(n-s))(s-1)^2\,ds \\
&-\frac{\epsilon^6 f_\infty^2}{4}\partial_1\left(\frac{f_\infty^2}{4}g + \frac{1}{12}(g^3+3f_\infty g^2) - \frac{1}{8}\partial_1^2 g + \frac{\epsilon}{48}\partial_1^3 g - \frac{\epsilon}{24}\partial_1(g^3+3f_\infty g^2) - \frac{\epsilon f_\infty^2}{8}\partial_1 g\right) \\
&-\epsilon^6\partial_2\left(\frac{f_\infty^2}{4}g + \frac{1}{12}(g^3+3f_\infty g^2) - \frac{1}{8}\partial_1^2 g + \frac{\epsilon}{48}\partial_1^3 g - \frac{\epsilon}{24}\partial_1(g^3+3f_\infty g^2) - \frac{\epsilon f_\infty^2}{8}\partial_1 g\right) \\
&+\frac{\epsilon^6}{12}\int_0^1 \partial_1^3(g^3(\epsilon(n-s-ct)))(s-1)^2 ds \\
&+\frac{\epsilon^6}{12}\int_0^1 \partial_1^3(3g^2(\epsilon(n-s-ct))f(\epsilon(n-s+t)))(s-1)^2\,ds \\
&+\frac{\epsilon^6}{12}\int_0^1 \partial_1^3(3g(\epsilon(n-s-ct))f^2(\epsilon(n-s+t)))(s-1)^2\,ds
\end{aligned}
\tag{3.71}
$$

The terms in eq. (3.70) of order ϵ^3 or lower cancel out. The terms of order ϵ^4 are equal to

$$-\partial_2\partial_1^{-1}\psi + \partial_1\phi - \frac{1}{6}\partial_1(3(f^2-f_\infty^2)g + 3(f-f_\infty)g^2). \tag{3.72}$$

Formally applying ∂_1 implies that the above terms should be constant in space since $\partial_2\psi = \partial_2^2\phi$ satisfies eq. (3.16). 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. (3.72) must be constant. Since all the terms decay to zero at spatial infinity, eq. (3.72) is exactly zero.

The terms of order ϵ^5 can be rewritten as

$$\frac{1}{4}\partial_1(-2\partial_2 g - \frac{1}{12}\partial_1^3 g + \frac{1}{6}(g^3+3f_\infty g^2)) + \frac{1}{2}(\partial_2^2\phi - \partial_1^2\phi + \frac{1}{6}\partial_1^2(3(f-f_\infty)g^2 + 3(f^2-f_\infty^2)g)) \tag{3.73}$$

which is equal to zero since g and ϕ satisfy the PDEs in eqs. (3.15) and (3.16). Thus the right-hand side of eq. (3.70) is equal to $I_{g,2}$. The integral terms in eq. (3.71) are bounded as before. The remaining terms in eq. (3.71) can be bounded by evaluating $\partial_2 g$ using eq. (3.15) and then applying lemma 3.4. We can the get the following bound:

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

Interpolating between powers of δ gives the desired inequality eq. (3.60).

The proof of eq. (3.61) follows immediately. $\square$

To proceed, we construct and energy function for eq. (3.55) to control the ℓ^2 norms of $\mathcal{U}$ and $\mathcal{Q}$. Lemma 3.5 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. (3.55), 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

$$\mathcal{E}(t) = \frac{1}{2}\sum_{n\in\mathbb{Z}} \mathcal{Q}_n^2(t)+\mathcal{U}_n^2(t)-\frac{1}{2}\left(\epsilon f(\epsilon(n+t),\epsilon^3 t)+\epsilon g(\epsilon(n-ct,\epsilon^3 t)+\epsilon^3\phi(\epsilon n,\epsilon t)\right)^2\mathcal{U}_n^2(t) \tag{3.74}$$

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

Lemma 3.6. *Fix $\tau_0 > 0$ and let δ be given by eq.* (3.59) *. 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.* (3.55)*, the energy-type quantity given in eq.* (3.74) *is coercive with the bound*

$$\|\mathcal{Q}(t)\|_{\ell^2}^2 + \|\mathcal{U}(t)\|_{\ell^2}^2 \le 4\mathcal{E}(t), \quad \text{for } t \in (-\tau_0\epsilon^{-3}, \tau_0\epsilon^{-3}). \tag{3.75}$$

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

$$\left|\frac{d\mathcal{E}}{dt}\right| \le C\mathcal{E}^{1/2}\left[\epsilon^{11/2}(\delta+\delta^5)+\epsilon^3\delta^2\mathcal{E}^{1/2}+\epsilon(\delta+\mathcal{E}^{1/2})\mathcal{E}\right] \tag{3.76}$$

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

$$1 - \frac{1}{2}\left(\epsilon\|f\|_{L^\infty} + \epsilon\|g\|_{L^\infty} + \epsilon^3\|\phi\|_{L^\infty}\right)^2 \ge \frac{1}{2}, \tag{3.77}$$

independent on the particular choices of f and g. Hence

$$\mathcal{E}(t) \ge \frac{1}{2}\|\mathcal{Q}\|_{\ell^2}^2 + \frac{1}{4}\|\mathcal{U}\|_{\ell^2}^2 \ge \frac{1}{4}\|\mathcal{Q}\|_{\ell^2}^2 + \frac{1}{4}\|\mathcal{U}\|_{\ell^2}^2 \tag{3.78}$$

and eq. (3.75) follows.

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

$$\begin{aligned}\frac{d\mathcal{E}}{dt} = \sum_{n\in\mathbb{Z}} &\mathcal{Q}_n(t)\mathrm{Res}_n^{(2)}(t) + \mathcal{Q}_n(t)\mathcal{B}_n(\epsilon f + \epsilon g + \epsilon^3\phi, \mathcal{U}(t)) \\ &+ \mathcal{U}_n(t)\mathrm{Res}_n^{(1)}(t)\left(1 - \frac{1}{2}(\epsilon f + \epsilon g + \epsilon^3\phi)^2\right) \\ &+ \mathcal{U}_n^2(t)(\epsilon f + \epsilon g + \epsilon^3\phi) \times (\epsilon^2\partial_1 f + \epsilon^4\partial_2 f - \epsilon^2 c\partial_1 g + \epsilon^4\partial_2 g + \epsilon^4\partial_2\phi).\end{aligned} \tag{3.79}$$

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

$$\begin{aligned}\left|\frac{d\mathcal{E}}{dt}\right| \leq &\|\mathcal{Q}\|_{\ell^2} \times \|\mathrm{Res}^{(2)}(t)\|_{\ell^2} + \|\mathcal{Q}\|_{\ell^2} \times \|\mathcal{B}\|_{\ell^2} + \|\mathcal{U}\|_{\ell^2} \times \|\mathrm{Res}_n^{(1)}(t)\|_{\ell^2} \\ &+ \|\mathcal{U}^2\|_{\ell^1} \times \|(\epsilon f + \epsilon g + \epsilon^3\phi) \times (\epsilon^2\partial_1 f + \epsilon^4\partial_2 f - \epsilon^2 c\partial_1 g + \epsilon^4\partial_2 g + \epsilon^4\partial_2\phi)\|_{\ell^\infty}.\end{aligned} \tag{3.80}$$

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 3.5, we thus have

$$\begin{aligned}\left|\frac{d\mathcal{E}}{dt}\right| \leq &C\Big[\mathcal{E}^{1/2}\epsilon^{11/2}(\delta + \delta^5) + \mathcal{E}^{1/2}\epsilon[(\delta + \epsilon^2\delta^3)\mathcal{E} + \mathcal{E}^{3/2}] \\ &+ \mathcal{E}(\epsilon^3\delta^2 + \epsilon^5\delta^2 + \epsilon^5\delta^4 + \epsilon^7\delta^4 + \epsilon^7\delta^6)\Big],\end{aligned} \tag{3.81}$$

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

$$\left|\frac{d\mathcal{E}}{dt}\right| \leq C\mathcal{E}^{1/2}\left[\epsilon^{11/2}(\delta + \delta^5) + \epsilon^3\delta^2\mathcal{E}^{1/2} + \epsilon(\delta + \mathcal{E}^{1/2})\mathcal{E}\right] \tag{3.82}$$

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

$$u_n(0) \approx \epsilon f(\epsilon n, 0) + \epsilon g(\epsilon n, 0) \tag{3.83}$$

and

$$\dot{u}_n(0) \approx \epsilon \partial_1 f(\epsilon n, 0) - \epsilon^2 g(\epsilon n, 0) \tag{3.84}$$

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*

$$\|u(0) - \epsilon f(\epsilon \cdot, 0) - \epsilon g(\epsilon \cdot, 0)\|_{\ell^2} + \|\dot{u}(0) - \epsilon^2 \partial_1 f(\epsilon \cdot, 0) + \epsilon^2 \partial g(\epsilon \cdot, 0)\|_{\ell^2_2} \leq \epsilon^{5/2} \tag{3.85}$$

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

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

Lemma 3.7. *If $a \in \ell^2_2(\mathbb{Z})$ and*

$$\sum_{k=-\infty}^{n} a_k = 0, \tag{3.86}$$

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

$$\|b\|_{\ell^2} \leq C \|a\|_{\ell^2_2} \tag{3.87}$$

for some $C > 0$ independent of a.

See appendix B for proof.

We can now show the following.

Lemma 3.8. *Let assumptions 2 and 3 hold. Then* $\mathcal{U}(0), \mathcal{Q}(0) \in \ell^2(\mathbb{Z})$ *satisfy*

$$\dot{u}_n(0) = q_{n+1}(0) - q_n(0) \tag{3.88}$$

and

$$\|\mathcal{U}(0)\|_{\ell^2} + \|\mathcal{Q}(0)\|_{\ell^2} \leq C\epsilon^{5/2} \tag{3.89}$$

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. (3.46).

For $q_n(0)$ to satisfy eq. (3.88), 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

$$\begin{aligned}
q_n(0) &= \sum_{k=-\infty}^{n-1} \dot{u}_k(0) \\
&= \sum_{k=-\infty}^{n-1} \left[\dot{u}_k(0) - \epsilon^2\partial_1 f(\epsilon k, 0) - \epsilon^4\partial_1 f(\epsilon k, 0) + \epsilon^2 c\partial_1 g(\epsilon k, 0) - \epsilon^4\partial_2 g(\epsilon k, 0)\right] \\
&\quad + \sum_{k=-\infty}^{n-1} \left[\epsilon^2\partial_1 f(\epsilon k, 0) + \epsilon^4\partial_1 f(\epsilon k, 0) - \epsilon F(\epsilon(k+1), 0) + \epsilon F(\epsilon k, 0)\right] \\
&\quad + \sum_{k=-\infty}^{n-1} \left[-\epsilon^2 c\partial_1 g(\epsilon k, 0) + \epsilon^4\partial_1 g(\epsilon k, 0) - \epsilon G(\epsilon(k+1), 0) + \epsilon G(\epsilon k, 0)\right] \\
&\quad + \epsilon F(\epsilon n, 0) - \epsilon F_{-\infty} + \epsilon G(\epsilon n, 0).
\end{aligned} \tag{3.90}$$

Comparing eq. (3.90) to eq. (3.47), we have that

$$\begin{aligned}
\mathcal{Q}_n(0) &= \sum_{k=-\infty}^{n-1} \left[\dot{u}_k(0) - \epsilon^2\partial_1 f(\epsilon k, 0) - \epsilon^4\partial_1 f(\epsilon k, 0) + \epsilon^2 c\partial_1 g(\epsilon k, 0) - \epsilon^4\partial_2 g(\epsilon k, 0)\right] \\
&\quad + \sum_{k=-\infty}^{n-1} \left[\epsilon^2\partial_1 f(\epsilon k, 0) + \epsilon^4\partial_1 f(\epsilon k, 0) - \epsilon F(\epsilon(k+1), 0) + \epsilon F(\epsilon k, 0)\right] \\
&\quad + \sum_{k=-\infty}^{n-1} \left[-\epsilon^2 c\partial_1 g(\epsilon k, 0) + \epsilon^4\partial_1 g(\epsilon k, 0) - \epsilon G(\epsilon(k+1), 0) + \epsilon G(\epsilon k, 0)\right].
\end{aligned} \tag{3.91}$$

That $\mathcal{Q}_n(0) \to 0$ as $n \to \infty$ is guaranteed by assumption 2. Now lemma 3.7 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. (3.63) and (3.65). This follows from the earlier calculations in lemma 3.5. That $I_{f,1}(k,0)$ and $I_{g,2}(k,0)$ are elements of ℓ_2^2 follows from $f(\cdot,0) \in \mathcal{X}_2^6$ and $g(\cdot,0) \in H_2^6$ and an application of lemma 3.4.

Thus we have eq. (3.89) where the $C > 0$ can be chosen based on the norms of f and g. □

3.5 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 3.1. *Let assumption 1 hold and set*

$$\delta = \max\left\{\sup_{\tau\in\mathbb{R}} \|f(\cdot,\tau)\|_{\mathcal{X}_2^6},\ \sup_{\tau\in\mathbb{R}} \|g(\cdot,\tau)\|_{H_2^6}\right\} \tag{3.92}$$

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. (3.42) *belongs to*

$$C^1([-t_0(\epsilon), t_0(\epsilon)], \ell^\infty(\mathbb{Z})) \tag{3.93}$$

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

$$\begin{aligned}&\|u(t) - \epsilon f(\epsilon(\cdot + t), \epsilon^3 t) - \epsilon g(\epsilon(\cdot - ct), \epsilon^3 t)\|_{\ell^2}\\ &\quad + \|\dot{u}(t) - \epsilon\partial_1 f(\epsilon(\cdot + t), \epsilon^3 t) + \epsilon^2\partial_1 g(\epsilon(\cdot - ct), \epsilon^3 t)\|_{\ell^2} \le C\epsilon^{5/2-r}, \quad t \in [-t_0(\epsilon), t_0(\epsilon)].\end{aligned} \tag{3.94}$$

Proof. Set $\mathcal{S} := \mathcal{E}^{1/2}$ where $\mathcal{E}$ is defined in eq. (3.74). From the results in lemma 3.8, 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 3.6. For fixed constants $r \in (0, 1/2)$, $C > C_0$, and $K > 0$, define the maximal continuation time by

$$T_{C,K,r} := \sup\left\{T_0 \in (0, rK^{-1}\epsilon^{-3}|\log(\epsilon)|] : \mathcal{S}(t) \le C\epsilon^{5/2-r}, t \in [-T_0, T_0]\right\}. \tag{3.95}$$

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

$$\begin{aligned} \left|\frac{d}{dt}\mathcal{S}(t)\right| &= \frac{1}{2\mathcal{E}^{1/2}}\left|\frac{d}{dt}\mathcal{E}(t)\right| \\ &\leq C_1(\delta+\delta^5)\epsilon^{11/2} + C_2\epsilon^3\left[\delta^2+\epsilon^{-2}(\delta+\mathcal{S})\mathcal{S}\right]\mathcal{S} \end{aligned} \tag{3.96}$$

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

$$C_2\left[\delta^2+\epsilon^{-2}(\delta+\mathcal{S})\mathcal{S}\right] \leq C_2\left[\delta^2+\epsilon^{-2}(\delta+C\epsilon^{11/2-r})C\epsilon^{11/2-r}\right], \tag{3.97}$$

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

$$C_2\left[\delta^2+\epsilon^{-2}(\delta+C\epsilon^{11/2-r})C\epsilon^{11/2-r}\right] \leq K. \tag{3.98}$$

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

$$\begin{aligned} \frac{d}{dt}e^{-\epsilon^3Kt}\mathcal{S}(t) &= -\epsilon^3Ke^{-\epsilon^3Kt}\mathcal{S} + e^{-\epsilon^3Kt}\frac{d}{dt}\mathcal{S} \\ &\leq -\epsilon^3Ke^{-\epsilon^3Kt}\mathcal{S} + e^{-\epsilon^3Kt}C_1(\delta+\delta^5)\epsilon^{11/2} \\ &\qquad + e^{-\epsilon^3Kt}C_2\epsilon^3\left[\delta^2+\epsilon^{-2}(\delta+\mathcal{S})\mathcal{S}\right]\mathcal{S} \\ &\leq -\epsilon^3Ke^{-\epsilon^3Kt}\mathcal{S} + e^{-\epsilon^3Kt}C_1(\delta+\delta^5)\epsilon^{11/2} + \epsilon^3Ke^{-\epsilon^3Kt}\mathcal{S} \\ &= e^{-\epsilon^3Kt}C_1(\delta+\delta^5)\epsilon^{11/2}. \end{aligned} \tag{3.99}$$

Integrating gives

$$\begin{aligned} \mathcal{S}(t) &\leq \left(\mathcal{S}(0) + K^{-1}C_1(\delta+\delta^5)\epsilon^{5/2}\right)e^{\epsilon^3Kt} - \epsilon^{5/2}K^{-1}C_1(\delta+\delta^5) \\ &\leq \left(\mathcal{S}(0) + K^{-1}C_1(\delta+\delta^5)\epsilon^{5/2}\right)e^{\epsilon^3Kt} \\ &\leq \left(\mathcal{S}(0) + K^{-1}C_1(\delta+\delta^5)\epsilon^{5/2}\right)e^{K\tau_0(\epsilon)} \\ &\leq \left(C_0 + K^{-1}C_1(\delta+\delta^5)\right)\epsilon^{5/2-r} \end{aligned} \tag{3.100}$$

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

$$C_0 + K^{-1}C_1(\delta + \delta^5) \leq C. \tag{3.101}$$

Note that our earlier choice of K can be enlarged so that eq. (3.98) 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

Fenichel Theory

In this section we give a brief overview of Fenichel theory and give some useful results to be applied in chapter 2. Fenichel theory is concerned with a large class of invariant manifolds which fulfill a certain hyperbolicity condition. Major results include the persistence of the manifolds under perturbations, existence of unstable manifolds, and the foliation of the unstable manifolds. The results are mainly from (Fenichel and Moser, 1971; Fenichel, 1974). Much of the presentation and additional results are taken from (Wiggins, 1994), and for proofs of the theorems and propositions we direct the reader there.

A.1 Overflowing Invariant Manifolds

In this and later sections, we will generally be concerned with an ODE given by

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n \tag{A.1}$$

with a corresponding flow ϕ_t. In the study of dynamics, finding invariant manifolds is often a useful first step to understanding the more complicated behavior. Here we will be primarily be concerned with a less rigid notion of invariant manifolds: overflowing invariant manifolds.

Definition A.1. *Let $\overline{M} = M \cup \partial M$ be a compact, connected C^r manifold with boundary contained in $\mathbb{R}^n$. Then $\overline{M}$ is said to be* overflowing invariant *under eq.* (A.1) *if for every $p \in \overline{M}$, $\phi_t(p) \in \overline{M}$ for all $t \leq 0$ and the vector field eq.* (A.1) *is pointing strictly outward on ∂M.*

There is also a definition for *inflowing invariant manifolds*, but we will exclusively focus on the overflowing variety. Results for inflowing invariant manifolds can be recovered by reversing the direction of time. We also have many results that apply to invariant manifolds as well; to demonstrate this, we typically apply a bump function to the vector field at the boundary of an invariant manifold to make it overflowing while leaving the dynamics the same in the interior. We will discuss this more in appendix A.2.

In the proofs of many of theorems, the technical problem of the boundary of $\overline{M}$ is avoiding by slightly enlarging our set. Let

$$M_1 := \phi_1(M) \quad M_2 := \phi_2(M). \tag{A.2}$$

The manifolds $\overline{M}_1$ and $\overline{M}_2$ are also overflowing invariant.

For many of the proofs found in (Wiggins, 1994), a special set of atlases for the manifold $\overline{M}_2$ is needed.

Proposition A.1. *Let $k = \dim \overline{M}_2$. For every open cover $\mathcal{U}$ of $\overline{M}_2$ there exist atlases*

$$\left\{(U_i^j, \sigma_i) : i = 1, 2, \ldots, s; j = 1, 2, \ldots, 6\right\} \tag{A.3}$$

such that

$$U_i^1 \subset \overline{U}_i^1 \subset U_i^2 \subset \overline{U}_i^2 \subset U_i^3 \subset \overline{U}_i^3 \subset U_i^4 \subset \overline{U}_i^4 \subset U_i^5 \subset \overline{U}_i^5 \subset U_i^6 \subset \overline{U}_i^6 \tag{A.4}$$

with

$$\sigma_i(U_i^j) = \mathcal{D}^j, \quad j = 1, 2, \ldots, 6 \tag{A.5}$$

where $\mathcal{D}^j := \{x \in \mathbb{R}^{n-k} : |x| < j\}$, i.e., the open disc of radius j. Moreover the open covers

$$\mathcal{U}^j = \{U_i^j : i = 1, 2, \ldots, s\} \tag{A.6}$$

are subordinate to $\mathcal{U}$.

A.2 Unstable manifold to overflowing invariant manifolds

Similar to the unstable manifold theorem for hyperbolic fixed points, we will have some invariant manifold approaching our overflowing invariant manifold under the condition that the flow transverse to the manifold is hyperbolic. To make this precise, we split the tangent space on M_2 into a direct sum of vector bundles corresponding with the tangent, stable, and unstable directions. Assume we have the continuous splitting

$$T\mathbb{R}^n|_{M_2} = TM_2 \oplus N^s \oplus N^u \tag{A.7}$$

and associated projections

$$\Pi^s : T\mathbb{R}^n|_{M_2} \to N^s \tag{A.8}$$

$$\Pi^u : T\mathbb{R}^n|_{M_2} \to N^u \tag{A.9}$$

We assume that the subbundles $TM_2 \oplus N^s$ and $TM_2 \oplus N^u$ are invariant under $D\phi_t$ for all $t < 0$.

To characterize the exponential rate of growth/decay in the these bundles under the linearized dynamics, we introduce generalized Lyapunov-type numbers. For a point $p \in M_2$ we consider the following nonzero vectors:

$$\begin{aligned} u_0 &\in N^u_p, \\ w_0 &\in N^s_p, \\ v_0 &\in T_pM_2, \end{aligned} \tag{A.10}$$

and

$$\begin{aligned} u_{-t} &= \Pi^u D\phi_{-t}(p)u_0, \\ w_{-t} &= \Pi^s D\phi_{-t}(p)w_0, \\ v_{-t} &= D\phi_{-t}(p)v_0. \end{aligned} \tag{A.11}$$

Definition A.2. *The* generalized Lyapunov-type numbers *at p are given by*

$$\lambda^u(p) := \inf \left\{ a : \left(\frac{|u_{-t}|}{|u_0|} \right) / a^t \to 0 \ \text{as} \ t \to \infty, \forall u_0 \in N_p^u \right\}, \tag{A.12}$$

$$\nu^s(p) := \inf \left\{ a : \left(\frac{|w_0|}{|w_{-t}|} \right) / a^t \to 0 \ \text{as} \ t \to \infty, \forall w_0 \in N_p^s \right\}. \tag{A.13}$$

If $\nu^s(p) < 1$, then we define

$$\sigma^s(p) = \inf \left\{ b : (|w_0|^b / |v_0|) / (|w_{-t}|^b / |v_{-t}|) \to 0 \ \text{as} \ t \to \infty, \forall v_0 \in T_p M_2, w_0 \in N_p^s \right\}. \tag{A.14}$$

One can also show that these expressions are equal to

$$\lambda^u(p) = \limsup_{t \to \infty} \| \Pi^u D\phi_{-t}(p) \mid_{N_p^u} \|^{1/t} \tag{A.15}$$

$$\nu^s(p) = \limsup_{t \to \infty} \| \Pi^s D\phi_t(\phi_{-t}(p)) \mid_{N^s_{\phi_{-t}(p)}} \|^{1/t} \tag{A.16}$$

$$\sigma^s(p) = \limsup_{t \to \infty} \frac{\log \| D\phi_{-t}(p) \mid_{T_p M} \|}{- \log \| \Pi^s D\phi_t(\phi_{-t}(p)) \mid_{N^s_{\phi_{-t}(p)}} \|}. \tag{A.17}$$

To simplify the notation, we can introduce the linear operators

$$A_t(p) : T_p M \to T_{\phi_{-t}(p)} M, \qquad v \mapsto D\phi_{-t}(p) v \tag{A.18}$$

$$B_t(p) : N^s_{\phi_{-t}(p)} \to N^s_p, \qquad v \mapsto \Pi^s D\phi_t(\phi_{-t}(p)) v \tag{A.19}$$

$$C_t(p) : N^u_p \to N^u_{\phi_{-t}(p)}, \qquad v \mapsto \Pi^u D\phi_{-t}(p) v. \tag{A.20}$$

Then the Lyapunov-type numbers can be rewritten as

$$\lambda^u(p) = \limsup_{t \to \infty} \| C_t(p) \|^{1/t} \tag{A.21}$$

$$\nu^s(p) = \limsup_{t \to \infty} \| B_t(p) \|^{1/t} \tag{A.22}$$

$$\sigma^s(p) = \limsup_{t \to \infty} \frac{\log \| A_t(p) \|}{- \log \| B_t(p) \|}. \tag{A.23}$$

We say that a splitting is hyperbolic if

$$\lambda^u(p) < 1, \quad \nu^s(p) < 1, \quad \forall p \in M. \tag{A.24}$$

We list a couple of useful properties of generalized Lyapunov-type numbers.

Lemma A.1. *(Wiggins, 1994, Lem. 4.1.1) The generalized Lyapunov-type numbers obtain their suprema on M.*

Lemma A.2. *(Wiggins, 1994, Lem. 3.1.2) The generalized Lyapunov-type numbers are constant on orbits, i.e.,*

$$\lambda^u(\phi_{-t}(p)) = \lambda^u(p), \quad \nu^s(\phi_{-t}(p)) = \nu^s(p), \quad \sigma^s(p.) \tag{A.25}$$

Based on the above lemma, one might suspect that the backward limit of an orbit would have the same generalized Lyapunov-type numbers as the points on the orbit. While we do not show exact equality, the backward limit set does provide an upper bound on the numbers. A more narrow result of the kind was proved in (Dieci and Lorenz, 1997, Thm. 2.3), where the backwards limit set was a single point. We will extend this result to the case where the backwards limit set is a compact set.

The proofs for the following lemmas should be similar to the proofs found in (Dieci and Lorenz, 1997).

Lemma A.3. *Let $p \in M$ with $\nu^s(p) < 1$. For $c > \sigma^s(p)$, we have*

$$\|A_t(p)\| \, \|B_t(p)\|^c \to 0 \quad \text{as } t \to \infty. \tag{A.26}$$

Conversely, if eq. (A.26) *holds for some $c \in \mathbb{R}$, then $c \geq \sigma^s(p)$.*

Lemma A.4. *For $p \in M$ and $s, t \geq 0$, we have the following:*

(i) $A_{s+t}(p) = A_t(\phi_{-s}(p))A_s(p)$

(ii) $B_{s+t}(p) = B_s(p)B_t(\phi_{-s}(p))$

(iii) $C_{s+t}(p) = C_t(\phi_{-s}(p))C_s(p)$.

Hence we have the following bounds on the generalized Lyapunov-type numbers.

Proposition 2.1. *Let $K \subset M$ be a compact set. If $p \in M$ such that $\phi_{-t}(p) \to K$ as $t \to \infty$, then*

(i) $\lambda^u(p) \leq \lambda^u(K)$,

(ii) $\nu^s(p) \leq \nu^s(K)$, *and*

(iii) *if* $\nu^s(K) < 1$, *then* $\sigma^s(p) \leq \sigma^s(K)$.

Proof. *(i)* Let $a \in \mathbb{R}$ such that $\lambda^u(K) < a$. For each $q \in K$ there is a $\tau_q > 0$ and an open, precompact neighborhood of q, U_q, such that

$$\|C_{\tau_q}(q')\| < a^{\tau_q} \quad \text{for all } q' \in U_q.$$

Then $\{U_q\}_{q \in K}$ is an open cover of K, and so we can take a finite subcover $\{U_i\}_{i=1}^m$ with associated τ_q values denoted by τ_i for $i = 1, \ldots, m$. Let $U = \bigcup_{i=1}^m U_i$ and assume without loss of generality that $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_m$. Since $\lambda^u(p)$ is constant along trajectories and $\phi_{-t}(p) \to K$ as $t \to \infty$, we can assume that $\phi_{-t}(p) \in U$ for all $t \geq 0$.

We can now break up the orbit of $\phi_{-t}(p)$ into discrete times to keep track of which U_i the orbit lies in. We shall do this inductively. Set $t_0 = 0$. Then $\phi_{-t_0}(p) = p \in U_{i_0}$ for some index $i_0 \in \{1, 2, \ldots, m\}$. Then we can define $t_1 = t_0 + \tau_{i_0}$ and again we have $\phi_{-t_1}(p) \in U_{i_1}$ for some index i_1. We can continue this process. Suppose we have t_k and τ_{i_k}. Then

$$t_{k+1} = t_k + \tau_{i_k}, \qquad \phi_{-t_{k+1}}(p) \in U_{i_{k+1}},$$

and so we have t_{k+1} and $\tau_{i_{k+1}}$ defined. Note that $t_{k+1} - t_k = \tau_{i_k} \leq \tau_m$, so the distance between times does not grow too large. Furthermore, we also have $t_{k+1} - t_k = \tau_{i_k} \geq \tau_1$ and so $t_k \to \infty$ as $k \to \infty$.

Now suppose $t > 0$ is fixed and arbitrary. There is some ℓ such that $t_\ell \leq t < t_{\ell+1}$. Then there is some $s < \tau_m$ such that

$$\begin{aligned} t &= t_\ell + s \\ &= \sum_{k=0}^{\ell-1} \tau_{i_k} + s. \end{aligned} \tag{A.27}$$

Using this decomposition of t along with lemma A.4, we get that

$$\begin{aligned} C_t(p) &= C_{t_\ell+s}(p) \\ &= C_s(\phi_{-t_\ell}(p))C_{t_\ell}(p) \\ &= C_s(\phi_{-t_\ell}(p))C_{\tau_{i_{\ell-1}}}(\phi_{-t_{\ell-1}}(p))C_{\tau_{i_{\ell-2}}}(\phi_{-t_{\ell-2}}(p))\cdots C_{\tau_{i_0}}(p). \end{aligned} \tag{A.28}$$

Thus we have

$$\begin{aligned} \|C_t(p)\| &\le \|C_s(\phi_{-t_\ell}(p))\|a^{\tau_{i_{\ell-1}}} \cdot a^{\tau_{i_{\ell-2}}} \cdots a^{\tau_{i_0}} \\ &= \|C_s(\phi_{-t_\ell}(p))\|a^{t_\ell}. \end{aligned} \tag{A.29}$$

Defining a constant C_1 by

$$C_1 = \max\{a^{-s}\|C_s(q)\| : q \in \overline{U}, 0 \le s \le \tau_m\} \tag{A.30}$$

we can write

$$\|C_t(p)\| \le C_1 a^s a^{t_\ell} = C_1 a^t. \tag{A.31}$$

Since this C_1 is independent of t, raising both sides to $1/t$ and taking the limit as $t \to \infty$ gives us that

$$\limsup_{t\to\infty} \|C_t(p)\|^{1/t} \le a, \tag{A.32}$$

and so $\lambda^u(p) \le a$ for each $a > \lambda^u(K)$. This proves $\lambda^u(p) \le \lambda^u(K)$.

(ii) We follow a similar argument for $\nu^s(p)$. Let $a \in \mathbb{R}$ such that $\nu^s(K) < a$. We can find an open cover of K given by $\{U_i\}_{i=1}^m$ (with each U_i precompact) and positive numbers $\tau_1 \le \tau_2 \le \cdots \le \tau_m$ such that

$$\|B_{\tau_i}(q)\| < a^{\tau_i} \quad \text{for all } q \in U_i. \tag{A.33}$$

The number $\nu^s(p)$ is constant on orbits, so assume that $\phi_{-t}(p) \in U := \cup_{i=1}^m U_i$ for all $t \ge 0$. We can similarly construct the t_k and τ_{i_k} inductively.

Let $t > 0$. Then there is an ℓ such that $t_\ell \le t < t_{\ell+1}$, and we have $0 \le s < \tau_m$ with

$$\begin{aligned} t &= t_\ell + s \\ &= \sum_{k=0}^{\ell-1} \tau_{i_k} + s. \end{aligned} \tag{A.34}$$

Thus

$$B_t(p) = B_{\tau_{i_0}}(p)B_{\tau_{i_1}}(\phi_{-t_1}(p)) \cdots B_{\tau_{i_{\ell-1}}}(\phi_{-t_{\ell-1}}(p))B_s(\phi_{-t_\ell}(p)). \tag{A.35}$$

We can then get

$$\|B_t(p)\| \le \|B_s(\phi_{-t_\ell}(p))\|a^{t-s}. \tag{A.36}$$

Defining a constant C_2 by

$$C_2 = \max\{a^{-s}\|B_s(q)\| : q \in \overline{U}, 0 \le s \le \tau_m\} \tag{A.37}$$

we can write

$$\|B_t(p)\| \le C_2 a^s a^{t_\ell} = C_2 a^t. \tag{A.38}$$

Taking limits gives us $\nu^s(p) \le a$ and thus $\nu^s(p) \le \nu^s(K)$.

(iii) Assume that $\nu^p(K) < 1$. Let $c > \sigma^s(K)$ be arbitrary. For each $q \in K$, there is a τ_q and a precompact, open neighborhood of q, U_q, such that

$$\|A_{\tau_q}(q')\| \, \|B_{\tau_q}(q')\|^c \le \frac{1}{2}, \quad \text{for all } q' \in U_q. \tag{A.39}$$

We again take a finite subcover $\{U_i\}_{i=1}^m$ with corresponding $\tau_1 \le \tau_2 \le \cdots \le \tau_m$. The number $\sigma^s(p)$ is constant on orbits so assume that $\phi_{-t}(p) \in U := \cup_{i=1}^m U_i$ for all $t \ge 0$. The t_k and τ_{i_k} values are constructed the same way as in *(i)*.

For $t > 0$, we have $t_\ell \le t < t_{\ell+1}$ and we can write t as

$$\begin{aligned} t &= t_\ell + s \\ &= \sum_{k=0}^{\ell-1} \tau_{i_k} + s \end{aligned} \tag{A.40}$$

with $0 \le s < \tau_m$. By our product formulas,

$$A_t(p) = A_s(\phi_{-t_\ell}(p))A_{\tau_{i_{\ell-1}}}(\phi_{-t_{\ell-1}}(p))A_{\tau_{i_{\ell-2}}}(\phi_{-t_{\ell-2}}(p)) \cdots A_{\tau_{i_0}}(p) \tag{A.41}$$

and

$$B_t(p) = B_{\tau_{i_0}}(p)B_{\tau_{i_1}}(\phi_{-t_1}(p)) \cdots B_{\tau_{i_{\ell-1}}}(\phi_{-t_{\ell-1}}(p))B_s(\phi_{-t_\ell}(p)). \tag{A.42}$$

Thus

$$\|A_t(p)\| \, \|B_t(p)\||^c \le C_3 \left(\frac{1}{2}\right)^\ell \tag{A.43}$$

where

$$C_3 = \max\{\|A_s(q)\|\,\|B_s(q)\|^c : q \in \overline{U}, 0 \le s \le \tau_m\}. \tag{A.44}$$

As $t \to \infty$, we have $\ell \to \infty$. Therefore $\|A_t(p)\|\,\|B_t(p)|\|^c \to 0$ as $t \to \infty$ and $\sigma^s(p) \le c$. We can then conclude that $\sigma^s(p) \le \sigma^s(K)$. □

At this point, we are nearly ready to state the main theorem concerning the existence of unstable manifolds. One might expect that the unstable manifold will be tangent to the unstable vector bundle, N^u. But we want our unstable manifold to be C^r smooth and N^u is only C^{r-1}. To get around this, the unstable vector bundle is perturbed slightly to increase its regularity.

Proposition A.2. *Suppose N is a C^{r-1} k-dimensional normal vector bundle defined on M_1. Then there is a C^r k-dimensional bundle $N' \subset T\mathbb{R}^n|_{M_1}$, transversal to TM_1. Moreover, given $\epsilon > 0$, for any set U_i^j as constructed in proposition A.1, there exist orthonormal bases*

$$\{e_1^{ij}(p), \ldots, e_k^{ij}(p)\} \quad \textit{for } N|_{U_i^j} \tag{A.45}$$

$$\{f_1^{ij}(p), \ldots, f_k^{ij}(p)\} \quad \textit{for } N'|_{U_i^j} \tag{A.46}$$

such that

$$|e_\ell^{ij}(p) - f_\ell^{ij}(p)| < \epsilon, \quad \ell = 1, \ldots, k. \tag{A.47}$$

The $f_\ell^{ij}(p)$ can be chosen to be C^r functions of $p \in U_i^j$

The main takeaway though is that replacing N^u and N^s with N'^u and N'^s, respectively, allows us to increase the regularity of the vector bundles to C^r. In order to find local coordinates around M_2, we let

$$N_\epsilon'^s := \{(p, v^u) \in N'^s : |v^s| \le \epsilon\}, \tag{A.48}$$

$$N_\epsilon'^u := \{(p, v^u) \in N'^u : |v^u| \le \epsilon\}, \tag{A.49}$$

and set $N'_\epsilon = N'^s_\epsilon \oplus N'^u_\epsilon$. Then for $\epsilon_0 > 0$ suitably small, for any $0 < \epsilon \leq \epsilon_0$ the map

$$\begin{aligned} h : N'_\epsilon &\to \mathbb{R}^n \\ (p, v^s, v^u) &\mapsto p + v^s + v^u \end{aligned} \tag{A.50}$$

is C^r and has an image containing a neighborhood of M_2. We also have local coordinates around the manifold given by

$$\begin{aligned} (\sigma_i \times \tau_i^s \times \tau_i^u) : N'_\epsilon|_{\overline{U}_i^6} \to& \mathbb{R}^{n-(s+u)} \times \mathbb{R}^s \times \mathbb{R}^u, \\ (\sigma_i \times \tau_i^s \times \tau_i^u)(p, v^s, v^u) &= (\sigma_i(p), \tau_i^s(p, v^s), \tau_i^u(p, v^u)) \\ &= (x, y, z) \end{aligned} \tag{A.51}$$

which are C^r diffeomorphisms.

To describe the unstable manifold being tangent to N'^u_ϵ, we define

$$\begin{aligned} h_u : N'^u_\epsilon &\to \mathbb{R}^n \\ (p, v^u) &\mapsto p + v^u \end{aligned} \tag{A.52}$$

so that $h_u(N'^u_\epsilon) \subset \mathbb{R}^n$.

Theorem A.1. *Suppose $\dot{x} = f(x)$ is a C^r vector field on $\mathbb{R}^n$, $r \geq 1$. Let $\overline{M} = M \cup \partial M$ be a C^r, compact connected manifold with boundary overflowing invariant under the vector field $f(x)$. Suppose $\nu^s(p) < 1$, $\lambda^u(p) < 1$, and $\sigma^s(p) < \frac{1}{r}$ for all $p \in M$. Then there exists a C^r overflowing invariant manifold $W^u(\overline{M})$ containing $\overline{M}$ and tangent to $h_u(N'^u_\epsilon)$ along $\overline{M}$ with trajectories in $W^u(\overline{M})$ approaching $\overline{M}$ as $t \to -\infty$.*

A.3 Foliations of unstable manifolds

In addition to the existence of an unstable manifold, we have under certain conditions that the manifold is foliated. We first introduce other generalized Lyapunov-type numbers before stating the theorem.

Definition A.3. *The generalized Lyapunov-type numbers at* p *are given by*

$$\sigma^{cu}(p) := \inf\left\{\rho : ((|u_{-t}|/|v_{-t}|)/(|u_0|/|v_0|))/\rho^t \to 0 \text{ as } t \to \infty, \forall v_0 \in T_pM_2, u_0 \in N_p^u\right\}, \tag{A.53}$$

$$\sigma^{su}(p) := \inf\left\{\rho : ((|u_{-t}|/|w_{-t}|)/(|u_0|/|w_0|))/\rho^t \to 0 \text{ as } t \to \infty, \forall w_0 \in N_p^s, u_0 \in N_p^u\right\}. \tag{A.54}$$

Theorem A.2. *Suppose* $\dot{x} = f(x)$ *is a* C^r *vector field on* $\mathbb{R}^n$, $r \geq 1$. *Let* $\overline{M} = M \cup \partial M$ *be a* C^r *compact connected manifold with boundary, overflowing invariant under the vector field* $f(x)$. *Suppose* $\lambda^u(p) < 1$, $\sigma^{cu}(p) < 1$, *and* $\sigma^{su}(p) < 1$ *for every* $p \in \overline{M}_1$. *Then there exists a* $n - (s+u)$*-parameter family* $\mathcal{F}^u = \cup_{p \in M} f^u(p)$ *of* u*-dimensional surfaces* $f^u(p)$ *(with boundary) such that the following hold:*

1. $\mathcal{F}^u$ *is a negatively invariant family, i.e.,* $\phi_{-t}(f^u(p)) = f^u(\phi_{-t}(p))$ *for any* $t \geq 0$ *and* $p \in M$.

2. *The* u*-dimensional surfaces* $f^u(p)$ *are* C^r.

3. $f^u(p)$ *is tangent to* $h_u(N_p'^u)$ *at* p.

4. *There exists* $C_u, \lambda_u > 0$ *such that if* $q \in f^u(p)$, *then*

$$|\phi_{-t}(q) - \phi_{-t}(p)| < C_u e^{-\lambda_u t}$$

for any $t \geq 0$.

5. *Suppose* $q \in f^u(p)$ *and* $q' \in f^u(p')$. *Then*

$$\frac{|\phi_{-t}(q) - \phi_{-t}(p)|}{|\phi_{-t}(q') - \phi_{-t}(p)|} \to 0 \quad \text{as } t \to \infty$$

unless $p = p'$.

6. $f^u(p) \cap f^u(p') = \emptyset$, *unless* $p = p'$.

7. *If the hypotheses of the unstable manifold theorem hold, i.e., if additionally* $\nu^s(p) < 1$ *and* $\sigma^s(p) < \frac{1}{r}$ *for every* $p \in \overline{M}_1$, *then the* u*-dimensional surfaces* $f^u(p)$ *are* C^r *with respect to the basepoint* p.

8. $\mathcal{F}^u = W^u_{\text{loc}}(M)$.

Lemma A.5. *Suppose $F(\mu, z)$ are a family of C^1 functions from $K \times [0, a]$ into $[0, a]$, where K is a compact set in $\mathbb{R}^n$ and $F(\mu, 0) = 0$ and $F(\mu, z) > 0$. Furthermore, assume*

$$0 < \frac{\partial F}{\partial z}(\mu, 0) < 1, \quad \forall \mu \in K. \tag{A.55}$$

Then there exists $z_0 \in (0, a]$ and constants $C > 0$ and $0 < \rho < 1$ where the family of solutions to the recurrence equations

$$z_{\mu,n+1} = F(\mu, z_{\mu,n}) \tag{A.56}$$

with initial values $z_{\mu,0}$ in the interval $[0, z_0]$ all approach 0 as $n \to \infty$ and

$$|z_{\mu_1,n} - z_{\mu_2,n}| \leq C\rho^n(|\mu_1 - \mu_2| + |z_{\mu_1,n} - z_{\mu_2,n}|) \tag{A.57}$$

for all $\mu_1, \mu_2 \in K$ and $n \geq 0$.

Proof. Define

$$\alpha_0 := \min_{\mu \in K} \frac{\partial F}{\partial z}(\mu, 0) \tag{A.58}$$

and

$$M_0 := \max_{\mu \in K} |D_\mu F(\mu, 0)| \tag{A.59}$$

Set $\alpha = \frac{1+\alpha_0}{2}$ and $M = 2M_0$. Note that $\alpha_0 < \alpha < 1$. Let $z_0 > 0$ such that

$$0 < \frac{\partial F}{\partial z}(\mu, z) \leq \alpha \quad \text{and} \quad |D_\mu F(\mu, z)| \leq M \tag{A.60}$$

for any $z \in [0, z_0]$ and $\mu \in K$.

Now fix $\mu_1, \mu_2 \in K$ and assume $z_{\mu_1,0}, z_{\mu_2,0} \in [0, z_0]$. Then we have that

$$\begin{aligned}
z_{\mu_1,n} - z_{\mu_2,n} &= z_{\mu_1,0} - z_{\mu_2,0} + \sum_{k=0}^{n-1}[(F(\mu_1, z_{\mu_1,k}) - z_{\mu_1,k}) - (F(\mu_2, z_{\mu_2,k}) - z_{\mu_2,k})] \\
&= z_{\mu_1,0} - z_{\mu_2,0} + \sum_{k=0}^{n-1}[(F(\mu_1, z_{\mu_1,k}) - z_{\mu_1,k}) - (F(\mu_1, z_{\mu_2,k}) - z_{\mu_2,k})] \\
&\quad \sum_{k=0}^{n-1}[(F(\mu_1, z_{\mu_2,k}) - z_{\mu_2,k}) - (F(\mu_2, z_{\mu_2,k}) - z_{\mu_2,k})] \\
&\leq z_{\mu_1,0} - z_{\mu_2,0} + \sum_{k=0}^{n-1}(\alpha - 1)(z_{\mu_1,k} - z_{\mu_2,k}) + \sum_{k=0}^{n-1} M|\mu_1 - \mu_2| \\
&= z_{\mu_1,0} - z_{\mu_2,0} + \sum_{k=0}^{n-1}(\alpha - 1)(z_{\mu_1,k} - z_{\mu_2,k}) + M|\mu_1 - \mu_2|n.
\end{aligned}$$

Hence applying the discrete Grönwall inequality gives us

$$z_{\mu_1,n} - z_{\mu_2,n} \leq \alpha^n(z_{\mu_1,0} - z_{\mu_2,0} + M|\mu_1 - \mu_2|n) \tag{A.61}$$

We can then get the final result by choosing $\rho \in (\alpha, 1)$ and $C > 0$ as large as we need. □

Proposition A.3. *Assume the conditions in theorem A.2 hold and that $r \geq 2$ and $u = 1$. Let $p \in M_1$ be a fixed points under the flow and assume that there are fixed points on M_1 approaching p. Then in some neighborhood of p, U, there is $C > 0$ and $\lambda > 0$ such that for any fixed point $p' \in U \cap M_1$ and any $q \in f^u(p)$ and $q' \in f^u(p')$ we have*

$$|(\phi_{-t}(q) - p) - (\phi_{-t}(q') - p')| \leq Ce^{-\lambda t}|q - q'| \tag{A.62}$$

for any $t \geq 0$.

Proof. We use the change of coordinates described before to prove the result. Choose U small enough so that the diffeomorphism is defined on all of U. From theorem A.2, we have C^r functions f_1 and f_2 such that the unstable manifold is given by the graph

$$(x, z) \mapsto (f_1(z; x), f_2(z; x), z) \tag{A.63}$$

Here the point $(x, 0, 0)$ picks out a point on M_1 and the one-dimensional foliation of the unstable manifold is parameterized by z. Let p' be an arbitrary fixed point on $M_1 \cap U$ and denote the local coordinates of p and p' by $(x, 0, 0)$ and $(x', 0, 0)$, respectively. Take $q \in f^u(p)$ and $q' \in f^u(p')$, and let there local coordinates be given by $(f_1(z; x), f_2(z; x), z)$ and $(f_1(z'; x'), f_2(z'; x'), x')$, respectively.

We first want to show that

$$|z(-t) - z'(-t)| \leq Ce^{-\lambda t}(|x - x'| + |z(0) - z'(0)|). \tag{A.64}$$

The argument should be similar to the one given in (Wiggins, 1994) for part 4 of theorem A.2, but we use the Grönwall estimate found before to get dependence on the initial conditions. Fix $T > 0$ and define $z_n = z(-nT)$ and $z'_n = z(-nT)$. These values can be found by iteratively applying the map

$$z \mapsto h(z; x) := \tau^u \circ \phi_{-T} \circ (\sigma \times \tau^s \times \tau^u)^{-1}(f_1(z; x), f_2(x; z), z). \tag{A.65}$$

Thus we have

$$z_{n+1} = h(z_n; x). \tag{A.66}$$

The map h is C^1 for $x \in \overline{U \cap M_1}$ and z sufficiently small. We also have that $0 < \frac{\partial h}{\partial z}(0; x) < 1$. A calculation of this type can be found in (Wiggins, 1994). Therefore we can apply lemma A.5 to get

$$|z_n - z'_n| \leq C\rho^n(|x - x'| + |z_0 - z'_0|) \tag{A.67}$$

for any choice of p'.

The same estimate will hold if we adjust the initial condition of the z_n and z'_n. Thus for any $t \in [0, T)$ we also have

$$|z(-nT - t) - z'(-nT - t)| \leq C\rho^n(|x - x'| + |z(-t) - z'(-t)|). \tag{A.68}$$

By the continuity of the flow, we can replace $|z(-t) - z'(-t)|$ with $C|z(0) - z'(0)|$ in the estimate above. Also, we can choose $\lambda_0 > 0$ small enough so that

$$\max_{0 \leq t < T} e^{\lambda_0(T+t)}\rho < 1 \tag{A.69}$$

so that eq. (A.64) holds.

Now consider the map

$$(x, z) \mapsto g(x, z) := (\sigma \times \tau^s \times \tau^u)^{-1}(f_1(x; z), f_2(x; z), z) - (\sigma \times \tau^s \times \tau^u)^{-1}(x, 0, 0). \quad \text{(A.70)}$$

This gives the difference between a point in $f^u(p)$ and its limit point p when given its local coordinates. The map is C^2 in its arguments. Also, since $g(x, 0) = 0$ for all x we have that

$$\frac{\partial g}{\partial x}(x, 0) = 0. \quad \text{(A.71)}$$

Then we have

$$\begin{aligned} &|(\phi_{-t}(q) - p) - (\phi_{-t}(q') - p')| \\ &\quad = |g(x, z(-t)) - g(x', z'(-t))| \\ &\quad \leq |g(x, z(-t)) - g(x', z(-t))| + |g(x', z(-t)) - g(x', z'(-t))| \\ &\quad \leq C(|z(-t)||x - x'| + |z(-t) - z(-t)|) && \text{(A.72)} \\ &\quad \leq C(e^{-\lambda_u t}|q - q'| + e^{-\lambda_0 t}|q - q'|) && \text{(A.73)} \\ &\quad \leq Ce^{-\lambda t}|q - q'| \end{aligned}$$

for some value of C and $\lambda > 0$. Note that eq. (A.72) follows from the fact that g is C^2 and eq. (A.71). Equation (A.73) follows from theorem A.2, the estimate given in eq. (A.64), and the fact that the map $q \mapsto x$ is a C^1 map. □

A.4 Boundary modifications

In this section we provide the details for the boundary modification we made in section 2.3.1 in order to claim that we had an unstable manifold. In particular, we asserted that the vector bundles could be continuously extended to the region where the perturbation of the vector field was made. If we denote the new vector field in eq. (2.74) by $\tilde{F}$ and its flow by $\tilde{\phi}_t$, then we know that the vectors in $T_p\mathbb{R}^4$ evolve according to the ODE

$$\xi'(s) = D\tilde{F}(\tilde{\phi}_{-s}(p))\xi, \quad \text{(A.74)}$$

a linear non-autonomous ODE. If $\tilde{\phi}_{-s}(p) \to p'$, then ideally we want a way to assign $\xi(0) \in T_p\mathbb{R}^4$ so that span$\{\xi(t)\}$ approaches $N^u_{p'}$ or $N^s_{p'}$ as $s \to \infty$. Thus we are really concerned with solutions of linear ODEs with certain limits at infinity.

This problem can be described more generally as finding solutions of

$$x'(t) = (A + V(t) + R(t))x \tag{A.75}$$

which approach eigenvectors of A as $t \to \infty$. The following theorem gives us a way to find those solutions (Coddington and Levinson, 1955, Chp. 3, Thm. 8.1).

Theorem A.3. *Let A be a constant matrix with characteristic roots μ_j, $j = 1, 2, \ldots, n$, all of which are distinct. Let the matrix V be differentiable and satisfy*

$$\int_0^\infty |V'(t)|\, dt < \infty \tag{A.76}$$

and let $V(t) \to 0$ as $t \to \infty$. Let the matrix R be integrable and let

$$\int_0^\infty |R(t)|\, dt < \infty. \tag{A.77}$$

Let the roots of $\det(A + V(t) - \lambda I) = 0$ *be denoted by $\lambda_j(t)$, $j = 1, 2, \ldots, n$. Clearly, by reordering the μ_j if necessary,* $\lim_{t\to\infty} \lambda_j(t) = \mu_j$. *For a given k, let*

$$D_{kj}(t) = \Re(\lambda_k(t) - \lambda_j(t)). \tag{A.78}$$

Suppose all j, $1 \le j \le n$, fall into one of two classes I_1 and I_2, where

$$j \in I_1 \textit{ if } \int_0^t D_{kj}\, d\tau \to \infty \textit{ as } t \to \infty \textit{ and } \int_{t_1}^{t_2} D_{kj}\, d\tau > -K \textit{ for } t_2 \ge t_1 \ge 0 \tag{A.79}$$

and

$$j \in I_2 \textit{ if } \int_{t_1}^{t_2} D_{kj}(\tau)\, d\tau < K \textit{ for } t_2 \ge t_1 \ge 0 \tag{A.80}$$

where k is fixed and K is a constant. Let p_k be an eigenvector of A associated with μ_k, so that

$$Ap_k = \mu_k p_k \tag{A.81}$$

Then there is a solution φ_k *of eq.* (A.75) *and a* t_0*,* $0 \leq t_0 < \infty$*, such that*

$$\lim_{t\to\infty} \varphi_k(t) \exp\left[-\int_{t_0}^{t} \lambda_k(\tau)\, d\tau\right] = p_k. \tag{A.82}$$

We will first apply the above theorem to

$$\xi'(s) = DF(\tilde{\phi}_{-s}(p))\xi \tag{A.83}$$

before returning to the full problem. Here F denotes the unperturbed vector field as given in eq. (2.65) while $\tilde{\phi}_{-s}$ is the flow under the perturbed vector field. Again, take $\phi_{-s}(p) \to p'$ as $s \to \infty$. Set $A = DF(p')$ so that the linear ODE can be expressed as

$$\xi'(s) = (A + (DF(\tilde{\phi}_{-s}(p)) - A))\xi, \tag{A.84}$$

which is the same form as eq. (A.75) with $V(s) = DF(\tilde{\phi}_{-s}(p)) - A$. One can check that V is differentiable by noting the linearization as computed in eq. (2.72) is C^1 in space. One can also check that $\int_0^\infty |V'(s)|\, ds < \infty$.

There is a slight complication given from the need that the eigenvalues of A are distinct. In our case, we have two eigenvalues that are the same: the two zero eigenvalues. However, we can remove one of our eigenvalues. For the zero eigenvalue with corresponding eigenvector $e_4 = (0, 0, 0, 1)$, we have the following spectral projection

$$Pv = (e_4 \cdot v)e_4. \tag{A.85}$$

The eigenvalue and eigenvector for A are the same as $DF(q)$ for any q in our manifold, and P commutes with $DF(q)$. This allows us to break up the solution ξ of eq. (A.84) into two components as such:

$$\xi = P\xi + (I - P)\xi = \xi_P + \xi_Q \tag{A.86}$$

Then the ODE for ξ_Q is a three-dimensional ODE and $A(I - P)$ has distinct eigenvalues. For simplicity of notation, we will assume that this reduction has been done and drop the subscript Q when talking about the reduced solution.

The conditions on the eigenvalues of $A + V(s) = DF(\tilde{\phi}_{-s}(p))$ follow from the fact that $DF(q)$ has eigenvalues continuous in q and they are distinct.

Thus we can apply theorem A.3 to the reduction of eq. (A.84) to define vector bundles. There are a couple of open question that need to be addressed before we can be satisfied with these vector bundles.

Firstly, one must check that the vector bundles defined this way are well-defined. For instance, one might get different vectors in the vector bundle depending on where the initial condition of the ODE is chosen. Starting at a point p and flowing the vector forward until time t_0 can lead to a different solution than if the vector is assigned by starting at $\tilde{\phi}_{-t_0}(p)$. This would not be unexpected, since there are infinitely many solutions of the linear ODE which approach the unstable bundle in backward time. We sidestep the issue by only assigning a solution to one point in each possible orbit. For our case, we have a radial flow, and so we will assign the vector bundles on a circle in $\overline{M}$ and choose $\mathrm{span}\{\xi(0)\}$ at each point p. The rest of the vector bundles will be defined by simply flowing the vector bundles on the circle backward in time.

Secondly, we need to check that the vector bundles are continuous. Certainly, along an orbit the solutions will be continuous but theorem A.3 does say whether solutions will be continuous in space. To answer this question, we must go the proof of theorem A.3. Essentially, the proof is done by rewriting the ODE using a change of coordinates and variation of parameters and then applying a Picard iteration. We shall look into the details for our specific case. The eigenvalues and eigenvectors of $DF(q)$ will be denoted by $\lambda_k(q)$ and $v_k(q)$, respectively. From the regularity in F, we have that the eigenvectors and values are continuously differentiable in q. Then we define the matrix

$$S(q) = [v_1(q), v_2(q), \ldots, v_n(q)] \tag{A.87}$$

and make the change of coordinates

$$\xi(s) = S(\tilde{\phi}_{-s}(p))\varphi(s). \tag{A.88}$$

The matrix $S(q)$ diagonalizes $DF(q)$ and so

$$\varphi'(s) = \Lambda(s,p)\varphi + \underbrace{[S(\phi_{-s}(p))]' S(\phi_{-s}(q))^{-1}}_{=:R(s,p)} \varphi(s) \tag{A.89}$$

where

$$\Lambda(s,p) = \operatorname{diag}\left\{\lambda_k(\tilde{\phi}_{-s}(p))\right\}. \tag{A.90}$$

Letting $\Psi(s,p)$ be the fundamental matrix solution associated with $\Lambda(s,p)$, the φ we are looking for is the solution to the following integral equation

$$\begin{aligned} \varphi_k(s,p) =& \Psi(s,p)e_k + \int_0^s \Psi_1(s,p)\Psi(\tau,p)^{-1}R(\tau,p)\varphi_k(\tau,p)\,d\tau \\ &- \int_s^\infty \Psi_2(s,p)\Psi(\tau,p)^{-1}R(\tau,p)\varphi_k(\tau,p)\,d\tau. \end{aligned} \tag{A.91}$$

At this point an iteration scheme can be set up, but to show regularity of the solution with respect to p, we set it up as a fixed point problem. We will slightly alter the function on the right-hand side so that the function space for the solution does not depend on p. Let

$$h_k(s,p) := \exp\left[\int_0^s \lambda_k(\sigma)\,d\sigma\right]. \tag{A.92}$$

Then define

$$\begin{aligned} \mathcal{F}_k[\varphi_0,p](s) :=& e_k + h_k(s,p)^{-1}\int_0^s \Psi_1(s,p)\Psi(\tau,p)^{-1}R(\tau,p)h_k(\tau,p)\varphi_0(\tau)\,d\tau \\ &- h_k(s,p)^{-1}\int_s^\infty \Psi_2(s,p)\Psi(\tau,p)^{-1}R(\tau,p)h_k(\tau,p)\varphi_0(\tau)\,d\tau \end{aligned} \tag{A.93}$$

where $\varphi^0 \in C_b^0([0,\infty),\mathbb{R}^n)$. Then by multiplying the above equation through by h_k, we can see that $h_k(\cdot,p)\varphi^0$ is a solution of eq. (A.91) if $\mathcal{F}_k[\varphi_0,p] = \varphi_0$. Choosing our initial condition p sufficiently close to its limit point guarantees that $\mathcal{F}_k[\cdot,p]$ maps into $C_b^0([0,\infty),\mathbb{R}^n)$ and

$$|\mathcal{F}_k[\varphi_1,p](s) - \mathcal{F}_k[\varphi_2,p](s)| \leq \rho|\varphi_1(s) - \varphi_2(s)| \tag{A.94}$$

for some $0 < \rho < 1$. Thus by the Banach fixed point theorem, there is a unique solution in C_b^0.

The regularity of the solutions with respect to p follows from the continuity S and R with respect to p. In particular, we have that $\mathcal{F}_k$ is continuous in p and so the solutions to the fixed point problem $\mathcal{F}_k[\varphi_k^0(\cdot,p),p] = \varphi_k^0(\cdot,p)$ are continuous with respect to p. This means that $p \mapsto \varphi_k(0,p) = \varphi_k^0(0,p)$ is also continuous with respect to p and we can define the vector bundles smoothly. Furthermore, by applying the implicit function theorem, one can get that the regularity of the extended vector bundles is the same as the original vector bundles. That is, if the original bundles were C^k, then the extensions are also C^k.

To summarize, we define the vector bundles on a small circle so that they are invariant under the flow given by eq. (A.83). We now return to the perturbed flow of the tangent vectors given by eq. (A.74). The vector bundles do not remain invariant under the perturbed vector field, but $TM_2 \oplus N^u$ and $TM_2 \oplus N^s$ are both invariant. This is because the perturbation is only in the tangent direction on the M and so the ODE partially decouples. For example, if $\xi(0) \in T_pM_2 \oplus N_p^u$, then we can define the solution to

$$\begin{aligned} u'(s) &= DF(\tilde{\phi}_{-s}(p))u \\ u(0) &= \Pi_p^u \xi(0), \end{aligned} \tag{A.95}$$

which is guaranteed to be in N^u for all time, and

$$\begin{aligned} v'(s) &= (D\tilde{F}(\tilde{\phi}_{-s}(p)) - DF(\tilde{\phi}_{-s}(p)))(v+u) \\ v(0) &= (I - \Pi_p^u)\xi(0) \end{aligned} \tag{A.96}$$

which is in TM_2 for all time, so that $u + v$ is a solution to eq. (A.74).

Appendix B

Proofs of lemmas

Lemma 2.1. *Suppose that (H3) holds. Then there exist $C > 0$ and $\alpha > 0$ such that*

$$|\gamma_{\pm,\epsilon}(s) - \gamma_{\pm}(s)| \leq Ce^{-\alpha|s|}\epsilon. \tag{2.106}$$

Furthermore, the difference of the heteroclinic orbits are in $H^5(\mathbb{R};\mathbb{R}^3)$ and

$$\|\gamma_{\pm,\epsilon}(s) - \gamma_{\pm}(s)\|_{H^5(\mathbb{R};\mathbb{R}^3)} \leq C\epsilon. \tag{2.107}$$

Proof. If (H3) holds, then we have the vector field is C^2 with respect to ϵ. Also, one can check that manifold M is C^∞ and the vector bundles are C^1. Thus the hypotheses for proposition A.3 are satisfied.

Let $\gamma_{\pm}(-\infty)$ be the shared asymptotic limit of $\gamma_{\pm,\epsilon}$ and $\gamma_{\pm}$. Therefore,

$$\begin{aligned}
|\gamma_{\pm,\epsilon}(-s) - \gamma_{\pm}(-s)| &= ((\epsilon, \gamma_{\pm,\epsilon}(-s)) - (\epsilon, \gamma_{\pm}(-\infty))) - ((0, \gamma_{\pm}(-s)) - (0, \gamma_{\pm}(-\infty)))| \\
&\leq Ce^{-\lambda s}|(\epsilon, \gamma_{\pm,\epsilon}(0)) - (0, \gamma_{\pm}(0))| \\
&\leq Ce^{-\alpha s}\epsilon.
\end{aligned}$$

The $s > 0$ direction follows similarly.

The H^5 estimate follows from the fact that $\gamma_{\pm,\epsilon}$ are $\gamma_{\pm}$ are solutions to a C^5 set of ODEs and so derivatives with respect to s are equal to at least C^1 functions with $\gamma_{\pm,\epsilon}$ and $\gamma_{\pm}$ as arguments. □

Lemma 3.1. *For non-negative integers k, there is a $C > 0$ such that*

$$\|fg\|_{H^k} \leq C\|f\|_{\mathcal{X}^k}\|g\|_{H^k} \tag{3.22}$$

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

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

Assuming eq. (3.22) 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 3.2. *For non-negative integers k, there is a $C > 0$ such that*

$$\|fg\|_{\mathcal{X}^k} \leq C\|f\|_{\mathcal{X}^k}\|h\|_{\mathcal{X}^k} \tag{3.23}$$

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

Proof. Using the result from lemma 3.1, we have

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

□

Lemma 3.3. *For each $k \geq 0$ and $c > 0$, there exists $C > 0$ depending only on k such that*

$$\left\|\frac{1}{\langle \cdot + \tau\rangle_+^2\langle \cdot - c\tau\rangle^2}\right\|_{C^k} \leq C\sup_{x\in\mathbb{R}}\frac{1}{\langle x + \tau\rangle_+^2\langle x - c\tau\rangle^2}. \tag{3.34}$$

Furthermore,

$$\int_0^\infty \sup_{x\in\mathbb{R}} \frac{1}{\langle x+\tau\rangle_+^2 \langle x-c\tau\rangle^2}\, d\tau < \infty. \tag{3.35}$$

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

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

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

is a sum of terms of the form

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

where $C \neq 0$ is a constant, m, m_1, m_2 are integers, $0 \leq m_1, m_2 \leq m$, 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 $k \geq 0$. 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. (B.3). That is, the $(k+1)^{\text{st}}$ derivative is a sum of terms of the form

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

Applying the product rule to eq. (B.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. (B.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. (B.3) (and a similar result holds for III). Finally,

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

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

This shows that the $(k+1)^{\text{st}}$ derivative is a sum of terms of the form in eq. (B.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. (B.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. (3.34) 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. (3.35) follows from

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

as $\tau \to \infty$. □

Lemma 3.7. *If $a \in \ell_2^2(\mathbb{Z})$ and*

$$\sum_{k=-\infty}^{n} a_k = 0, \tag{3.86}$$

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

$$\|b\|_{\ell^2} \leq C\|a\|_{\ell_2^2} \tag{3.87}$$

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

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

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

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

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

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

for $n \geq 0$. Therefore,

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

$\square$

References

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

Coddington, E. A. and Levinson, N. (1955). *Theory of ordinary differential equations*. Tata McGraw-Hill Education.

Dieci, L. and Lorenz, J. (1997). Lyapunov-type numbers and torus breakdown: Numerical aspects and a case study. *Numerical Algorithms*, 14(1):79–102.

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.

Fenichel, N. (1974). Asymptotic stability with rate conditions. *Indiana University Mathematics Journal*, 23(12):1109–1137.

Fenichel, N. and Moser, J. (1971). Persistence and smoothness of invariant manifolds for flows. *Indiana University Mathematics Journal*, 21(3):193–226.

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.

Iooss, G. (2000). Travelling waves in the Fermi-Pasta-Ulam lattice. *Nonlinearity*, 13(3):849.

Iooss, G. and Kirchgässner, K. (2000). Travelling waves in a chain of coupled nonlinear oscillators. *Communications in Mathematical Physics*, 211(2):439–464.

Kato, T. (2013). *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media.

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.

Vanderbauwhede, A. and Iooss, G. (1992). Center manifold theory in infinite dimensions. In *Dynamics reported*, pages 125–163. Springer.

Wiggins, S. (1994). *Normally hyperbolic invariant manifolds in dynamical systems*, volume 105. Springer Science & Business Media.