Bachelor's project

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

Jonas Pedersen Vean

Bifurcation of Weakly Dispersive Partial Differential Equations

Bachelor's project in Mathematics
Supervisor: Mats Ehrnström
June 2020

NTNU
Norwegian University of
Science and Technology

Jonas Pedersen Vean

Bifurcation of Weakly Dispersive Partial Differential Equations

Bachelor's project in Mathematics
Supervisor: Mats Ehrnström
June 2020

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of Science and Technology

BIFURCATION OF WEAKLY DISPERSIVE PARTIAL DIFFERENTIAL EQUATIONS

BACHELOR PROJECT

JONAS P. VEAN

Abstract. In this thesis we explore the use of local bifurcation theory to show existence of small-amplitude traveling wave solutions to nonlinear dispersive partial differential equations that in a sense are generalizations of the Korteweg–de Vries and Whitham equations of hydrodynamics. Of special note is the equation given by $\partial_t u + L\partial_x u + \partial_x(u)^{p+1} = 0$, whose traveling wave solutions are found to be small perturbations in the direction of $\cos(\xi_0 x)$ in the Hölder space $C^{0,\alpha}(\mathbb{R})$ viewed as a bifurcation space for the problem. One of the main goals of the thesis was to provide a coherent exposition to the material needed to understand everything discussed.

Contents

1 Introduction 2
2 A Primer on Calculus on Banach Spaces 2
3 Local Bifurcation Theory 7
3.1 Fredholm Operators and the Lyapunov–Schmidt Reduction Method 8
3.2 Bifurcation of A Single Eigenvalue 9
4 Background Material 11
4.1 A Brief Primer on the Korteweg–de Vries and Whitham Equations 11
4.2 Some Functional Analysis, Fourier Theory 11
4.3 Distribution Theory 13
4.4 Hölder Spaces 15
4.5 Classical Symbols 17
5 Nonlinear Dispersive Equations Inspired by KdV and Whitham 18
5.1 On Generalizing to Arbitrary Classical Symbols 22

Date: January 2020 - June 2020, NTNU, Norwegian University of Science and Technology, Department of Mathematical Sciences, 7491 Trondheim, Norway.

1. Introduction

How do we solve partial differential equations? There happens to be a vast multitude of methods to employ, and indeed no two methods need be equivalent when considering the same equation. In this thesis, we shall explore the method of reducing nonlinearities down to what might essentially be linear systems on function spaces - should we be so lucky! To this end, we will be needing something called local bifurcation theory, which we explore in Section 3 of this article. A fair bit of the buildup required to understand this is included, and therefore a review of calculus on Banach spaces is included at the very beginning.

From there we move on to the essentials of the functional analysis we need to understand the machinery involved with the partial differential equation. This is the material of Section 4.

Finally, we put local bifurcation theory, functional analysis and theory of some of the theory behind pseudodifferential operators together when exploring nonlinear dispersive equations in Section 5.

2. A Primer on Calculus on Banach Spaces

In analysis one would like to work over spaces whose structure is well-behaved and practically simple, such as metric spaces, vector spaces and the like. At a fundamental level, we would like to work with Banach spaces, which we recall are complete normed vector spaces. In this section we review some concepts about calculus on Banach spaces.

This section will for the most part be inspired by Buffoni and Toland's exposition [1], in particular their discussion on Banach space theory and the like, with adapted notation and some reorganizing.

Definition 2.1. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces and $U \subseteq X$ open. A map $F\colon U \to Y$ is called *continuous at $x \in U$* if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for every $y \in Y$ with $\|x - y\|_X < \delta$ we have $\|F(x) - F(y)\|_Y < \varepsilon$. If F is continuous at each and every point $x \in U$ we simply call F continuous. In this case we may write $F \in C(U, Y)$ or $F \in C^0(U, Y)$.

For normed spaces and metric spaces this generalization of continuity from elementary calculus is readily available. However, how does one define a derivative of such a map? Recall from the case $F\colon \mathbb{R} \to \mathbb{R}\,; x \mapsto F(x)$ that F has a derivative $\partial_x F(a) = A$ at $a \in \mathbb{R}$ if

$$A = \lim_{x \to a} \frac{F(x) - F(a)}{x - a} \tag{2.1}$$

exists. Notice, however, that Eqn. (2.1) could be rephrased as: for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $|x - a| < \delta$ then

$$\left| \frac{F(x) - F(a)}{x - a} - A \right| < \varepsilon \quad \text{or} \quad \frac{|F(x) - F(a) - A(x - a)|}{|x - a|} < \varepsilon$$

by the definition of a limit. Should the derivative of F exist at the point $a \in \mathbb{R}$, then for $|x - a|$ sufficiently small the approximation $F(x) \approx F(a) + \partial_x F(a) \cdot (x - a)$ is valid in the above sense. In other words, being differentiable at a point means we can locally approximate the function as a linearization around the point. This motivates the definition of the Fréchet derivative.

Definition 2.2. (Fréchet differentiation)
Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces and $U \subseteq X$ open. We say that a map $F\colon U \to Y$ is *Fréchet differentiable at* $x_0 \in U$ if there exists a linear map $A \in \mathcal{L}(X, Y)$ such that

$$\lim_{x \to x_0} \frac{\|F(x) - F(x_0) - A(x - x_0)\|_Y}{\|x - x_0\|_X} = 0. \tag{2.2}$$

In this case we call A the Fréchet derivative of F at x_0 and write $A = \mathrm{d}F[x_0]$. If F is Fréchet differentiable at every point in X, then the map

$$\mathrm{d}F\colon X \to \mathcal{L}(X, Y)\,; x \mapsto \mathrm{d}F[x]$$

is well-defined and the evaluation $\mathrm{d}F[x_0](x)$ acts as a directional derivative of F at x_0 "along" the vector $x \in X$.

Remark. One may as well use a local definition of the Fréchet derivative, where instead of having X as domain we consider open sets $U \subseteq X$ and maps $F\colon U \to Y$.

Proposition 2.1. (Chain rule of the Fréchet derivative)
Let X, Y and Z be Banach spaces and let $U \subseteq X$, $V \subseteq Y$ be open sets. If $F\colon U \to Y$ and $G\colon V \to Z$ are Fréchet differentiable maps and $F(U) \subseteq V$, then

(i) the composition $G \circ F\colon X \to Z$ is Fréchet differentiable
(ii) the Fréchet derivative obeys the chain rule. If $x_0 \in U$ then

$$\mathrm{d}(G \circ F)[x_0] = \mathrm{d}G[F(x_0)] \circ \mathrm{d}F[x_0].$$

Definition 2.3. (Partial Derivatives)
Let X, Y and Z be Banach spaces, $U \subseteq X \times Y$ be open in the product topology, and $F\colon U \to Z$ a function. Consider the projection maps $\pi_X(x, y) = x$, $\pi_Y(x, y) = y$, then set $U_{x_0} = \pi_X^{-1}(x_0) \cap U$ and $U_{y_0} = \pi_Y^{-1}(y_0) \cap U$ for $(x_0, y_0) \in U$. If $F(\cdot, y_0)$ has a Fréchet derivative at x_0 on U_{y_0} we denote it by $\partial_x F[(x_0, y_0)] \in \mathcal{L}(X, Z)$ and call it the *partial derivative* of F with respect to x at $(x_0, y_0) \in U$. Similarly for $y_0 \in U_{x_0}$, F Fréchet differentiable on U_{x_0} with $\partial_y F[(x_0, y_0)] \in \mathcal{L}(Y, Z)$.

Definition 2.4. (Higher Derivatives)
Let X and Y be Banach spaces, suppose that $F\colon U \to Y$, $U \subseteq X$ open, is continuously Fréchet differentiable on U. If $\mathrm{d}F\colon U \to \mathcal{L}(X, Y)$ is itself differentiable at $x_0 \in U$, then we say that the *second (order) Fréchet derivative* exists and is denoted by $\mathrm{d}(\mathrm{d}F)[x_0] \in \mathcal{L}(X, \mathcal{L}(X, Y))$. Higher k-order Fréchet derivatives are defined similarly when the previous order is defined and continuously differentiable, namely through a k-fold multilinear scheme: $\mathrm{d}(\mathrm{d} \cdots (\mathrm{d}F))[x_0] \in \mathcal{L}(X, \mathcal{L}(\cdots \mathcal{L}(X, Y)))$. A function that is k times continuously Fréchet differentiable on $U \subseteq X$ is said to be of class $C^k(U, Y)$.

Definition 2.5. (Homeomorphisms, Diffeomorphisms)
Let X and Y be Banach spaces, $U \subseteq X$ open, $F\colon U \to Y$ a continuous function. The function F is called a *homeomorphism* if it is bijective and if F^{-1} is continuous on Y. Furthermore, if $F \in C^k(U, Y)$ is k times continuously Fréchet differentiable and bijective with $F^{-1} \in C^k(Y, U)$, then we say that F is a C^k-diffeomorphism.

Definition 2.6. A subset $U \subseteq X$ of a vector space X is called *convex* if every pair of points $x_1, x_2 \in U$ can be connected via. a line segment between them, i.e. we have a parametrized curve

$$\gamma\colon [0, 1] \longrightarrow X\,; t \longmapsto \gamma(t) = (1 - t)x_1 + tx_2$$

which lies entirely in U.

Lemma 2.1. Let X and Y be Banach spaces, and let $U \subseteq X$ be a convex open set. If $F\colon U \to Y$ is Fréchet differentiable at each point of U with the property that

$$\sup_{x \in U} \|\mathrm{d}F[x]\|_{\mathcal{L}(X,Y)} = m < \infty.$$

Then we have that F is Lipschitz on U:

$$\|F(x_2) - F(x_1)\|_Y \leq m\|x_2 - x_1\|_X, \quad x_1,\, x_2 \in U. \tag{2.3}$$

Proof. We fix two points $x_1, x_2 \in U$ and use the convexity property of U: the line γ satisfies $\gamma(t) = (1-t)\,x_1 + t\,x_2 \in U$ for $t \in [0,1]$. We may find a real-valued linear functional $\varphi \in Y^*$ such that

$$\varphi(F(x_2) - F(x_1)) = \|F(x_2) - F(x_1)\|_Y, \quad \|\varphi\|_{Y^*} = 1.$$

Note that $\varphi\colon Y \to \mathbb{R}$ is continuous, so that the function $g(t)\colon [0,1] \to \mathbb{R}$ defined by $g(t) = \varphi \circ F(\gamma(t))$ is continuous on $[0,1]$. Differentiating we obtain

$$g'(t) = \varphi(\mathrm{d}F[\gamma(t)](x_2 - x_1))$$

which when combined with $\|\varphi\|_{Y^*} = 1$ then leads to

$$|g'(t)| \leq \|\mathrm{d}F[\gamma(t)]\|_{\mathcal{L}(X,Y)}\|x_2 - x_1\|_X \leq m\|x_2 - x_1\|_X. \tag{2.4}$$

Using the mean value theorem for one-variable functions with $t^* \in (0,1)$ we obtain

$$g'(t^*) = \frac{\varphi(F(\gamma(1)) - F(\gamma(0))))}{1-0} = \varphi(F(x_2) - F(x_1)) = \|F(x_2) - F(x_1)\|_Y. \tag{2.5}$$

Combining the estimate from (2.4) with the latter equation (2.5) gives the desired result.

□

Theorem 2.1. (Inverse Function Theorem)
Let X and Y be Banach spaces, $x_0 \in U$ be an open neighborhood of $U \subseteq X$ and let $F \in C^1(U,Y)$ such that the Fréchet derivative $\mathrm{d}F[x_0] \in \mathcal{L}(X,Y)$ is a homeomorphism. Then there exists a connected open set $\tilde{U} \subset U$ with $x_0 \in \tilde{U}$ such that $F|_{\tilde{U}}\colon \tilde{U} \to V$ for some $V \subseteq Y$ open with $F(x_0) \in V$ is a local C^1-diffeomorphism.

Remark. If one instead assumes $F \in C^k(U,Y)$, then F with the above assumptions becomes a local C^k-diffeomorphism.

Proof. Consider first the map $\Phi\colon X \to X$ given by $\Phi(x) = \mathrm{d}F[x_0]^{-1}(F(x) - F(x_0))$ and note that $\mathrm{d}\Phi[x_0] = I \in \mathcal{L}(X,X)$, the linear identity operator on X, and that $\Phi(x_0) = 0$.

We may then choose an $r \in (0,1)$ such that if $\|x - x_0\|_X \leq r$ then $x \in U$ and $\|\mathrm{d}\Phi[x] - I\|_{\mathcal{L}(X,X)} \leq 1/4$. Assume $y \in X$ satisfies $\Phi(x) = y$ for some $x \in X$. We show that one can find a sequence $(x_n)_{n\in\mathbb{N}_0}$ converging to a solution x of $\Phi(x) = y$. Consider the sequence given by

$$x_{n+1} = y + x_n - \Phi(x_n), \quad n \in \mathbb{Z}_{\geq 0}. \tag{2.6}$$

Note that $x_1 - x_0 = y$, which implies that $\|x_1 - x_0\|_X = \|y - 0\|_X$. If we consider the latter norm to be sufficiently small, we can see that $x_n \in B(x_0; r)$ by the estimates

$$\begin{aligned}
\|x_{n+1} - x_n\|_X &= \|\Phi(x_{n-1}) - x_{n-1} - (\Phi(x_n) - x_n)\|_X \\
&\leq \sup_{0\leq t\leq 1} \|\mathrm{d}F[x_{n-1} + t(x_n - x_{n-1})] - I\|_{\mathcal{L}(X,X)} \cdot \|x_n - x_{n-1}\|_X \\
&\leq \frac{1}{4}\|x_n - x_{n-1}\|_X
\end{aligned}$$

which when applied successively results in the inequality $\|x_{n+1} - x_n\|_X \leq 4^{-n}\|y\|_X$. Using this inequality with the triangle inequality we obtain

$$\begin{aligned}\|x_n - x_0\|_X &\leq \|x_n - x_{n-1}\|_X + \|x_{n-1} - x_{n-2}\|_X + \cdots + \|x_1 - x_0\|_X \\ &\leq \left(\frac{1}{4^{n-1}} + \frac{1}{4^{n-2}} + \cdots + 1\right)\|y\|_X \\ &< \left(\sum_{k=0}^{\infty} \frac{1}{4^k}\right)\|y\|_X = \frac{4}{3}\|y\|_X = \frac{4}{3}\|y - 0\|_X.\end{aligned}$$

Therefore if we choose $\|y - 0\|_X < 3r/4$ we see that $\|x_n - x_0\|_X < r$ for all $n \geq 0$. The sequence $(x_n)_{n\in\mathbb{N}_0}$ is Cauchy since for all $m > n$ we have by the triangle inequality

$$\|x_m - x_n\|_X \leq \sum_{k=n}^{m-1} \|x_{k+1} - x_k\|_X \leq \sum_{k=n}^{m-1} \frac{1}{4^k}\|y\|_X \leq 4^{-n+1}\|y\|_X$$

which can be made small for sufficiently large n. Thus the sequence $(x_n)_{n\in\mathbb{N}_0}$ converges to some $x \in X$ with $\|x - x_0\|_X \leq 4\|y\|_X/3 < r$. By continuity we established that $x = y + x - \Phi(x)$, i.e. $\Phi(x) = y$.

We can now define open subsets

$$V = \{y \in X \mid \|y\|_X < 3r/4\}, \quad \tilde{U} = \{x \in X \mid \|x - x_0\|_X < r,\, \Phi(x) \in V\}$$

which makes $\Phi|_{\tilde{U}} \colon \tilde{U} \to V$ a bijection. Surjectiveness comes a priori from the definition of $\tilde{U}$, and injectiveness stems from the uniqueness of the limit of the sequence $(x_n)_{n\in\mathbb{N}_0}$. Let y_1, $y_2 \in V$ and x_1, $x_2 \in \tilde{U}$ such that $\Phi(x_1) = y_1$, $\Phi(x_2) = y_2$, then

$$\begin{aligned}&\|y_2 - y_1\|_X = \|\Phi(x_2) - \Phi(x_1)\|_X \\ &= \|(x_2 - x_1) + (\mathrm{d}\Phi[x_2] - I)(x_2 - x_1) + (\Phi(x_2) - \Phi(x_1) - \mathrm{d}\Phi[x_2](x_2 - x_1))\|_X \\ &\geq \|x_2 - x_1\|_X - \|(\mathrm{d}\Phi[x_2] - I)(x_2 - x_1)\|_X - \|\Phi(x_2) - \Phi(x_1) - \mathrm{d}\Phi[x_2](x_2 - x_1)\|_X \\ &\geq \frac{3}{4}\|x_2 - x_1\|_X - \|\Phi(x_2) - \Phi(x_1) - \mathrm{d}\Phi[x_2](x_2 - x_1)\|_X.\end{aligned}$$

In order to show continuity of Φ^{-1} we first prove the lemma

Lemma. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. Assume $U \subset X$ is open convex and $F \colon U \to Y$ is Fréchet differentiable on U. If there exists $A \in \mathcal{L}(X, Y)$ such that $\|\mathrm{d}F[x] - A\|_{\mathcal{L}(X,Y)} \leq m$ for all $x \in U$, then for any pair x_1, $x_2 \in U$ we have that

$$\|F(x_2) - F(x_1) - \mathrm{d}F[x_2](x_2 - x_1)\|_Y \leq 2m\|x_2 - x_1\|_X \tag{2.7}$$

Proof of Lemma. Fix the point $x_2 \in U$. Define the function $G(x) = F(x) - \mathrm{d}F[x_2](x)$ for $x \in U$ and observe that

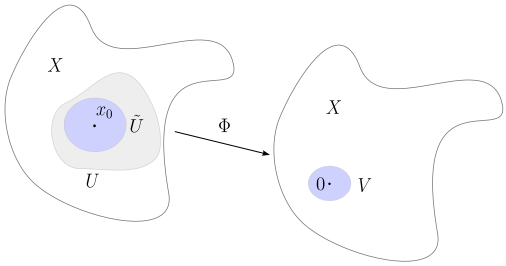
$$\begin{aligned}\|\mathrm{d}G[x]\|_{\mathcal{L}(X,Y)} &= \|\mathrm{d}F[x] - \mathrm{d}F[x_2]\|_{\mathcal{L}(X,Y)} \\ &\leq \|\mathrm{d}F[x] - A\|_{\mathcal{L}(X,Y)} + \|\mathrm{d}F[x_2] - A\|_{\mathcal{L}(X,Y)} \leq 2m\end{aligned}$$

which when using Lemma 2.1 on $G \colon X \to Y$ we obtain the desired result.

Continuing with the proof of the theorem, we now see that the previous chain of inequalities results in

$$\frac{1}{4}\|x_2 - x_1\|_X \leq \|y_2 - y_1\|_X \iff \|\Phi^{-1}(y_2) - \Phi^{-1}(y_1)\|_X \leq 4\|y_2 - y_1\|_X \tag{2.8}$$

which shows that Φ^{-1} is Lipschitz continuous on V. Taking the pre-image of V then shows that $\tilde{U}$ is also connected, but not necessarily path-connected or even convex.



Finally, to prove that Φ^{-1} is continuously Fréchet differentiable on V we consider the setting where $\Phi(x) = y$, $\Phi(x') = y'$ such that $x, x' \in \tilde{U}$. By the definition of $\tilde{U}$, we have $\|x - x_0\|_X < r$ and thus $\mathrm{d}\Phi[x] = \mathrm{d}\Phi[\Phi^{-1}(y)]$ is a homeomorphism due to the following lemma:

Lemma. Let X, Y be Banach spaces and let $S, T \in \mathcal{L}(X, Y)$ with T a homeomorphism and $\|S - T\|_{\mathcal{L}(X,Y)} < \|T^{-1}\|^{-1}_{\mathcal{L}(X,Y)}$. Then S is a homeomorphism.

Proof of Lemma. The condition on the norm is equivalent to $\|T^{-1}(S - T)\|_{\mathcal{L}(X,Y)} < 1$. Because of the norm, one can consider a power series $\sum_{k=0}^{\infty}(T^{-1}(S - T))^k$ with the property that

$$(I - T^{-1}(S - T)) \left(\sum_{k=0}^{\infty} (T^{-1}(S - T))^k \right) = I$$

so then we observe that, for $S \neq T$, the inverse

$$(I - T^{-1}(S - T))^{-1} = \sum_{k=0}^{\infty} (T^{-1}(S - T))^k \tag{2.9}$$

is in $\mathcal{L}(Y, X)$ since each term in the sum on the right hand side is bounded such that the sum itself is well-defined and bounded. From this we see that $\|-T^{-1}(S-T)\|_{\mathcal{L}(X,Y)} < 1$ as well and that

$$(I + T^{-1}(S - T))^{-1} = \sum_{k=0}^{\infty} (-1)^k (T^{-1}(S - T))^k$$

is bounded absolutely by Eqn. (2.9) and thus $I + T^{-1}(S - T)$ has a power series inverse and is therefore a homeomorphism. A final observation is that S can be expressed as $S = T(I + T^{-1}(S - T))$, which is a composition of homeomorphisms and thus a homeomorphism in it self. ☐

With $x \in \tilde{U}$ and $\|I - \mathrm{d}\Phi[x]\|_{\mathcal{L}(X,X)} < 1/4$ we have by the previous lemma that $\mathrm{d}\Phi[x]$ is a homeomorphism. Consider then the natural candidate for the Fréchet derivative for Φ^{-1} at $y \in V$ which is $(\mathrm{d}\Phi[\Phi^{-1}(y)])^{-1}$, for which we have

$$\begin{aligned}
&\frac{\|\Phi^{-1}(y) - \Phi^{-1}(y') - (\mathrm{d}\Phi[\Phi^{-1}(y)])^{-1}(y - y')\|_X}{\|y - y'\|_X} \\
&= \frac{\|x - x' - (\mathrm{d}\Phi[x])^{-1}(\Phi(x) - \Phi(x'))\|_X}{\|y - y'\|_X} \\
&= \frac{\|(\mathrm{d}\Phi[x])^{-1}\{\Phi(x) - \Phi(x') - \mathrm{d}\Phi[x](x - x')\}\|_X}{\|x - x'\|_X} \cdot \frac{\|x - x'\|_X}{\|y - y'\|_X} \\
&\leq \|\mathrm{d}\Phi[x]^{-1}\|_{\mathcal{L}(X,X)} \cdot \frac{\|\Phi(x) - \Phi(x') - \mathrm{d}\Phi[x](x - x')\|_X}{\|x - x'\|_X} \cdot \frac{\|x - x'\|_X}{\|y - y'\|_X}
\end{aligned}$$

letting $\|y - y'\|_X \to 0$ with the Lipschitz continuity of Φ^{-1} from Eqn. (2.8), the entire latter expression vanishes in the limit, which shows that Φ^{-1} is Fréchet differentiable on $V \subset X$. The continuity of this Fréchet derivative is guaranteed by the previous lemma. □

With the inverse function theorem, we may establish the implicit function theorem as a corollary.

Theorem 2.2. (Implicit Function Theorem)
Let X, Y and Z be Banach spaces and let $U \subseteq X \times Y$ be open in the product topology. Let $(X_0, y_0) \in U$. Assume $F\colon U \to Z$ is of class $F \in C^k(U, Z)$ such that $F(x_0, y_0) = z_0$ and $\partial_x F[(x_0, y_0)] \in \mathcal{L}(X, Z)$ is a homeomorphism. Then there exists an open ball $B(y_0; r)$, $r > 0$, and a connected open set $V \subseteq U$ and a mapping $\phi \in C^k(B(y_0; r), X)$ such that

$$(x_0, y_0) \in V \text{ and } F^{-1}(z_0) \cap V = \{(\phi(y), y) \mid y \in B(y_0; r)\}.$$

Proof. First define a new function $G \in C^k(U, Z \times X)$ by $G(x, y) = (F(x, y), y)$. We then have that $G(x_0, y_0) = F(z_0, y_0)$ and that for $(x, y) \in X \times Y$

$$\mathrm{d}G[(x_0, y_0)](x, y) = (\partial_x F[(x_0, y_0)]x + \partial_y F[(x_0, y_0)]y, y)$$

since the Fréchet derivative acts like the total derivative from the calculus of differential forms, namely $\mathrm{d}F = \partial_x F \,\mathrm{d}\pi_X + \partial_y F \,\mathrm{d}\pi_Y$ where π_X, π_Y are the standard projections into X and Y respectively. We observe that another representation is

$$\mathrm{d}G[(x_0, y_0)] = \begin{bmatrix} \partial_x F[(x_0, y_0)] & \partial_y F[(x_0, y_0)] \\ 0 & I \end{bmatrix}.$$

The determinant of this matrix is non-zero since $\partial_x F[(x_0, y_0)]$ is a homeomorphism, and thus $\mathrm{d}G[(x_0, y_0)]$ is invertible with

$$\mathrm{d}G[(x_0, y_0)]^{-1} = \begin{bmatrix} \partial_x F[(x_0, y_0)]^{-1} & -\partial_x F[(x_0, y_0)]^{-1}\partial_y F[(x_0, y_0)] \\ 0 & I \end{bmatrix}.$$

and is thus clearly bounded in $\mathcal{L}(X \times Y, Z \times Y)$. Using the inverse function theorem (Theorem 2.1) on G we may find a connected open set $V \subseteq U$ with $(x_0, y_0) \in V$ and an open ball $B((z_0, y_0); R) \subseteq Z \times Y$ with $R > 0$. By the theorem, $G\colon V \to B((z_0, y_0); R)$ is a C^k-diffeomorphism.

Declare $W = \{y \in Y \mid (z_0, y) \in B((z_0, y_0); R)\}$ and define $\phi(y) = x$ for $y \in W$ if and only if $G^{-1}(z_0, y) = (x, y) \in V$. In this case, $(x, y) = (\phi(y), y)$. Note that G^{-1} is of class C^k on $B((z_0, y_0); R)$, which in particular means that G^{-1} is of class C^k on W also. Note that the projection π_X is smooth on $X \times Y$, so since $\phi(y) = \pi_X \circ G^{-1}(z_0, y)$ it follows that ϕ is of class C^k on W. □

3. Local Bifurcation Theory

The setup for bifurcation theory is a nonlinear function $F\colon \mathbb{F} \times X \to Y$ with $F(\lambda, 0)) = 0$ for all $\lambda \in \mathbb{F}$, given a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and Banach spaces X and Y. Our goal is to find solutions $x \in X$ to $F(\lambda, x) = 0$.

As with general mathematical problems, we are inclined to reduce our nonlinear problem to a problem that is in some way solvable by conventional methods or well-studied theory. One way to go about doing this in our present case is through Lyapunov-Schmidt reduction of the equation.

This entire section is heavily based on the material covered by Kielhöfer [11].

3.1. Fredholm Operators and the Lyapunov–Schmidt Reduction Method

Definition 3.1. (Nonlinear Fredholm Operators)
Let X and Z be Banach spaces, $U \subset X$ open, $F\colon U \to Z$ Fréchet differentiable. Assume furthermore that $\mathrm{d}F[x]$, $x \in U$ satisfies

(i) $\dim\ker(\mathrm{d}F[x]) < \infty$, the kernel is finite dimensional
(ii) $\operatorname{codim}\operatorname{im}(\mathrm{d}F[x]) = \dim\mathcal{L}(X,Z) - \dim\operatorname{im}(\mathrm{d}F[x]) < \infty$
(iii) the image $\operatorname{im}(\mathrm{d}F[x])$ is closed in Z

then we call F a *nonlinear Fredholm operator* with *Fredholm index* given by the integer $\dim\ker(\mathrm{d}F[x]) - \operatorname{codim}\operatorname{im}(\mathrm{d}F[x])$.

Considering the function $F\colon U \to Z$ for $U \subset X \times Y$ open, we may consider the conditions $F(x_0, y_0) = 0$, $F \in C(U, Z)$ and $\partial_x F \in C(U, \mathcal{L}(X, Z))$. Furthermore, we assume that $F(\cdot, y_0)$ is a nonlinear Fredholm operator with respect to x for some $y_0 \in V$.

We may decompose the Banach spaces X and Z into

$$X = \ker(\partial_x F[(x_0, y_0)]) \oplus X_0 \quad \text{and} \quad Z = \operatorname{im}(\partial_x F[(x_0, y_0)]) \oplus Z_0.$$

Defining projections $P\colon X \to \ker(\partial_x F[(x_0, y_0)])$ and $Q\colon Z \to Z_0$ in the natural way, by the open mapping theorem both of these maps are in particular continuous.

Theorem 3.1. (Lyapunov–Schmidt Method of Reduction)
Let X, Y and Z be Banach spaces, $F\colon U \to Z$ as above with $U \subset X \times Y$ open, and P, Q projections onto $\ker(\partial_x F[(x_0, y_0)])$ and Z_0 respectively. Then there is an open neighborhood $\tilde{U}$ of (x_0, y_0) in $U \subset X \times Y$ such that our problem $F(x, y) = 0$ with $(x, y) \in \tilde{U}$ is equivalent to a finite-dimensional problem

$$\Phi(\xi, y) = 0 \qquad (\xi, y) \in U_0 \times V \subset \ker(\partial_x F[(x_0, y_0)]) \times Y \tag{3.1}$$

where $\Phi\colon U_0 \times V \to Z_0$ is continuous with $\Phi(\xi_0, y_0) = 0$.

Proof. Observe first that with the projection maps, the equation $F(x, y) = 0$ is equivalent to the system

$$QF(Px + (I - P)x, y) = 0$$
$$(I - Q)F(Px + (I - P)x, y) = 0$$

where due to the properties of projections we may write $x = \xi + \eta$ for $\xi = Px \in \ker(\partial_x F[(x_0, y_0)])$ and $(I - P)x = \eta \in X_0$. Our aim is to obtain a function whose properties satisfy the conditions of the implicit function theorem (Theorem 2.2). To this end, define the function $G\colon U_0 \times W \times V \to \operatorname{im}(\partial_x F[(x_0, y_0)])$ by

$$G(\xi, \eta, y) = (I - Q)F(\xi + \eta, y).$$

Furthermore, if $\xi_0 = Px_0 \in U_0$ and $\eta_0 = (I - P)x_0 \in W$, then $G(\xi_0, \eta_0, y_0) = 0$. Our assumptions provide us with the existence of the partial derivative $\partial_\eta G[(\xi_0, \eta_0, y_0)]\colon X_0 \to \operatorname{im}(\partial_x F[(x_0, y_0)])$, so we obtain

$$\partial_\eta G[(\xi_0, \eta_0, y_0)] = (I - Q)\partial_x F[(x_0, y_0)] \in \mathcal{L}(X_0, \operatorname{im}(\partial_x F[(x_0, y_0)]))$$

where we note that due to the setup this is a bijection. Thus $\partial_\eta G[(\xi_0, \eta_0, y_0]$ is a homeomorphism, and we may then use the implicit function theorem on G. Our equation $G(\xi, \eta, y) = 0$ for $(\xi, \eta, y) \in U_0 \times W \times V$ is equivalent to $\xi = \phi(\eta, y)$ with $\phi \in C(U_0 \times V, W)$ such that $\phi(\eta_0, y_0) = \xi_0$. For $(\xi, y) \in U_0 \times V$ we obtain

$$\Phi(\xi, y) = QF(\xi + \phi(\xi, y), y) = 0$$

with $\Phi \in C(U_0 \times V, Z_0)$ and $\Phi(\xi_0, y_0) = 0$ as desired. □

Remark. We call the function $\Phi\colon U_0 \times V \to Z$ as in the latter theorem and proof the *bifurcation function* for the problem considered.

A property that will be useful for us is the preservation of regularity of F onto its locally-defined bifurcation map.

Corollary 3.1. (Regularity of Bifurcation Functions)
If $F\colon U \to Z$ as in Theorem 3.1 has regularity $F \in C^k(U, Z)$, then for the functions ϕ, Φ as in the proof of said theorem, we have $\phi \in C^k(U_0 \times V, X_0)$ and $\Phi \in C^k(U_0 \times V, Z_0)$. Furthermore,

$$\partial_\xi \phi[(\xi_0, y_0)] = 0 \quad \text{and} \quad \partial_\xi \Phi[(\xi_0, y_0)] = 0.$$

Proof. The regularity of ϕ and Φ follow from the implicit function theorem when the regularity of F is assumed to be C^k for $k \geq 1$. Recall that the bifurcation function is given by $\Phi(\xi, y) = QF(\xi + \phi(\xi, y), y)$. We differentiate $I - \Phi$ with respect to ξ to obtain, for $(\xi, y) \in U_0 \times V$

$$(I - Q)\partial_x F[(\xi + \phi(\xi, y), y)](I_{\ker(\partial_x F[(x_0,y_0)])} + \partial_\xi \phi[(\xi, y)]) = 0 \tag{3.2}$$

where if we evaluate at $(\xi, y) = (\xi_0, y_0)$ then

$$\partial_x F[(\xi_0 + \phi(\xi_0, y_0), y_0)] \circ I_{\ker(\partial_x F[(x_0,y_0)])} = \partial_x F[(x_0, y_0)] \circ I_{\ker(\partial_x F[(x_0,y_0)])} = 0$$

which when combined with Eqn. (3.2) leads us to conclude that

$$(I - Q)\partial_x F[(x_0, y_0)]\, \partial_\xi \phi[(\xi_0, y_0)] = 0.$$

Note that $\partial_\xi \phi[(\xi_0, y_0)]\colon U_0 \times V \to X_0$, so if $\partial_x F[(x_0, y_0)]\, \partial_\xi \phi[(\xi_0, y_0)] = 0$ we are forced to conclude that $\partial_\xi \phi[(\xi_0, y_0)] = 0$ since X_0 is the complement subspace of the kernel.

Differentiating $\Phi(\xi, y) = QF(\xi + \phi(\xi, y), y)$ with respect to ξ, using what is shown from previous calculations, we obtain

$$\partial_\xi \Phi[(\xi_0, y_0)] = Q\partial_x F[(x_0, y_0)]\, I_{\ker(\partial_x F[(x_0,y_0)])} = 0.$$

□

3.2. Bifurcation of A Single Eigenvalue

We return to our bifurcation problem with $F(\lambda, x) = 0$. Assume that a given solution (λ_0, x_0) has two distinct solution curves passing through this point. What conditions are needed for two such solution curves to exist? Indeed, a necessary condition for the existence of two such curves has to be a non-bijective partial derivative $\partial_x F[(\lambda_0, x_0)]$, since if it were bijective we would be able to use the implicit function theorem locally around (λ_0, x_0) in such a way where the bifurcation cannot possibly occur.

An essential trick for our analysis is to normalize one of the solution curves to that of a trivial solution line $\lambda \times \{0\} \subset \mathbb{R} \times X$. To realize that this is always possible, consider the solution curve $\gamma(s) = (\lambda(s), x(s))$, satisfying $F(\gamma(s)) = 0$ and set $G(s, x) = F(\lambda(s), x(s) + x)$ which in turn means that $G(s, 0) = 0$ for all applicable $s \in \mathbb{R}$. This makes the trivial solution line one of the solution curves when considering G instead of F. In the proceeding matter, we assume that F has a trivial solution line.

For a single eigenvalue, our kernel of the partial derivative has to have dimension equal one: $\dim \ker(\partial_x F[(\lambda_0, 0)])$. In addition to this, we make the assumption that our function $F\colon \mathbb{R} \times U \to Z$ is a nonlinear Fredholm operator of index zero, meaning that we assume $\operatorname{codim} \operatorname{im}(\partial_x F[(\lambda_0, 0)]) = 1$. Also, assume F satisfies the criteria in the Lyapunov-Schmidt reduction method in Theorem 3.1.

Theorem 3.2. (Crandall–Rabinowitz)
Assume $F \in C^2(V \times U, Z)$ is a nonlinear Fredholm operator for $0 \in U \subset X$ and $\lambda_0 \in V \subset \mathbb{R}$ open, along with the normalized assumptions as outlined above. Furthermore, assume that

$$\ker(\partial_x F[(\lambda_0, 0)]) = \operatorname{span}\{v_0\}, \quad v \in X, \quad \|v_0\|_X = 1$$

and that the second mixed partial derivatives commute and satisfy

$$\partial^2_{x\lambda} F[(\lambda_0, 0)] v_0 \notin \operatorname{im}(\partial_x F[(\lambda_0, 0)]).$$

Then there is a second, distinct solution curve $\gamma \colon (-\delta, \delta) \to V \times U$ through $\gamma(0) = (\lambda_0, 0)$ which is continuously differentiable and solves $F(\gamma(s)) = 0$ for all $s \in (-\delta, \delta)$.

Finally, there are only two solutions intersecting at the bifurcation point $(\lambda_0, 0)$, namely the trivial solution line curve and γ as above.

Proof. We have assumed that F satisfies the assumptions required for Lyapunov-Schmidt reduction, so there exists $\Phi \in C^2(V_0 \times U_0, Z_0)$ given by

$$\Phi(\lambda, \xi) = QF(\lambda, \xi + \phi(\lambda, \xi)) \tag{3.3}$$

where $\phi \in C^2(V_0 \times U_0, X_0)$, $\Phi(\lambda, \xi) = 0$ locally around $(\lambda_0, 0) \in V_0 \times U$. Note that due to the zero Fredholm index, one necessarily has $\dim Z_0 = 1$. Due to the trivial solution line, $F(\lambda, 0) = 0$ for all $\lambda \in \mathbb{R}$, we have for all $\lambda \in V_0$

$$\phi(\lambda, 0) = 0 \quad \text{and} \quad \partial_\lambda \phi[(\lambda, 0)] = 0.$$

Evaluating $\Phi(\lambda, \xi)$ at $(\lambda, 0)$, by Eqn. (3.3) we have $\Phi(\lambda, 0) = 0$ for all $\lambda \in V_0$. Because of this we may represent $\Phi(\lambda, \xi)$ by

$$\Phi(\lambda, \xi) = \Phi(\lambda, \xi) - \Phi(\lambda, 0) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} \Phi(\lambda, t\xi)\, \mathrm{d}t = \int_0^1 \partial_\xi \Phi(\lambda, t\xi)\, \xi \, \mathrm{d}t.$$

Now let $\xi = s v_0 \in U_0 \subset \ker(\partial_x F[(\lambda_0, 0)])$ for $s \in (-\varepsilon, \varepsilon)$ and consider the modified equation

$$\tilde{\Phi}(\lambda, s) \equiv \int_0^1 \partial_\xi \Phi(\lambda, s t v_0)\, v_0 \, \mathrm{d}t = 0$$

for $s \neq 0$. Then regularity assumptions imply that $\tilde{\Phi} \in C^1((-\varepsilon, \varepsilon) \times U_0, Z_0)$, additionally we have $\tilde{\Phi}(\lambda_0, 0) = 0$. Taking another derivative with respect to λ, we compute

$$\begin{aligned}
\partial_\lambda(\partial_\xi \Phi[(\lambda, \xi)] v_0) &= \partial_\lambda(Q \partial_x F[(\lambda, \xi + \phi(\lambda, \xi))](v_0 + \partial_\xi \phi(\lambda, \xi))\, v_0) \\
&= Q \partial^2_x F[(\lambda, \xi + \phi(\lambda, \xi))](v_0 + \partial_\xi \phi(\lambda, \xi)\, v_0, \partial_\lambda \phi(\lambda, \xi)) \\
&\quad + Q \partial_x F[(\lambda, \xi + \phi(\lambda, \xi))]\, \partial^2_{\lambda\xi} \phi(\lambda, \xi)\, v_0 \\
&\quad + Q \partial^2_{x\lambda} F[(\lambda, \xi + \phi(\lambda, \xi))](v_0 + \partial_\xi \phi(\lambda, \xi)\, v_0)
\end{aligned} \tag{3.4}$$

where derivatives involving λ are identified with scalars and linear maps in the following sense:

$$\partial_\lambda F[(\lambda, x)]\, 1 = \partial_\lambda F[(\lambda, x)] \in Z, \quad \partial^2_{x\lambda} F[(\lambda, x)](1, x) = \partial^2_{x\lambda} F[(\lambda, x)] x \in \mathcal{L}(X, Z).$$

Note that if $F \in C^2(V \times U, Z)$ in the Fréchet sense, we immediately know that both $\partial^2_{x\lambda} F$ and $\partial^2_{\lambda x} F$ exist and are equal as operators. Evaluating Eqn. (3.4) at $(\lambda, \xi) = (\lambda_0, 0)$ we obtain, due to the porjection Q and that $\partial_\xi \phi[(\lambda_0, 0)] = 0$ from Corollary 3.1, that

$$\partial_\lambda \tilde{\Phi}[(\lambda_0, 0)] = Q \partial^2_{x\lambda} F[(\lambda_0, 0)]\, v_0 \in Z.$$

This derivative is identified with a non-zero element of Z since Q projects onto the complement of the image of $\partial_x F[(\lambda_0, 0)]$ and $\partial^2_{x\lambda} F[(\lambda_0, 0)]\, v_0 \notin \operatorname{im}(\partial_x F[(\lambda_0, 0)])$ by the theorem's assumptions. Thus $\partial_\xi \tilde{\Phi}[(\lambda_0, 0)]$ is a homeomorphism and by the implicit function theorem, Theorem 2.2, there exists $(-\delta, \delta) \subset (-\varepsilon, \varepsilon)$ and a function $\varphi\colon (-\delta, \delta) \to V_0$ satisfying $\varphi(0) = \lambda_0$ and $\tilde{\Phi}(\varphi(s), s) = 0$ for all $s \in (-\delta, \delta)$.

Our desired non-trivial solution to the bifurcation function then is

$$\Phi(\varphi(s), sv_0) = s\tilde{\Phi}(\varphi(s), s) = 0.$$

Define the curve $\gamma\colon (-\delta, \delta) \to V \times U$ by $\gamma(s) = (\varphi(s), sv_0 + \phi(\varphi(s), sv_0))$. Then $\gamma(0) = (\lambda_0, 0)$ and $F(\gamma(s)) = 0$ for all $s \in (-\delta, \delta)$, as desired. □

4. Background Material

4.1. A Brief Primer on the Korteweg–de Vries and Whitham Equations

In this section we give a preliminary to the Korteweg-de Vries (KdV) equation and the function spaces involved in the analysis of the behaviour of this equation. Briefly put, the KdV equation describes solitary waves of a fluid. Solitary waves are dispersive waves which do not change their shape over time, and in particular when they do not obey the linear superposition principle when two or more such solitary waves collide.

A dimensional version of the KdV-equation, or one variant thereof, is given by

$$\partial_t \eta + c_0\, \partial_x \eta + \frac{3}{2}\frac{c_0}{h_0}\eta\, \partial_x \eta + \frac{1}{6} c_0 h_0^2\, \partial_x^3 \eta = 0 \tag{4.1}$$

where $c_0 = \sqrt{g h_0}$ and g is the gravitational acceleration, h_0 is the height from the surface of the fluid to the fluid floor, which is assumed to be entirely flat. We may rescale the (t, x)-coordinates and shift the function η such that we can write down an equivalent, dimensionless version of the KdV-equation

$$\partial_t u + u\, \partial_x u + \partial_x^3 u = 0. \tag{4.2}$$

A slight modification to the dimensional KdV equation, Eqn. (4.1), proposed by Gerald B. Whitham exhibits the possibility of solutions with wave breaking and peaking - informally speaking waves whose profile may look like sharp peaks. This modification is introduced through the convolution kernel given by

$$K_{\text{Whitham}} = \mathcal{F}^{-1}\left(\sqrt{\frac{g\, \tanh h_0 \xi}{\xi}}\right)$$

such that we instead obtain the (dimensional) Whitham equation

$$\eta_t + \frac{3}{2}\frac{c_0}{h_0}\eta\, \eta_x + K_{\text{Whitham}} * \eta_x = 0. \tag{4.3}$$

Existence of small-amplitude periodic traveling waves through bifurcation theory are investigated in [5].

4.2. Some Functional Analysis, Fourier Theory

To further explore the main problem of this thesis, we will for convenience and rigour lay down some conventions and fundamental theorems. Readers familiar with function space theory and Fourier theory may choose to skip this subsection.

Spaces of p-integrable functions $f\colon \Omega \subseteq X \to \mathbb{C}$ are denoted by $L^p(\Omega, \mathbb{C})$ (or respectively to $\mathbb{R}$ for real-valued functions) and are normed vector spaces with norm given by the Lebesgue integral

$$\|f\|_{L^p(\Omega,\mathbb{C})} = \left(\int_\Omega |f|^p \,\mathrm{d}x\right)^{\frac{1}{p}}.$$

Furthermore, we adopt the convention that for $p = \infty$ we obtain the normed vector space $L^\infty(\Omega, \mathbb{C})$ of all measurable functions that are bounded essentially: if $\|f\|_{L^\infty(\Omega,\mathbb{C})} := \operatorname{ess\,sup}_{x\in\Omega} |f(x)| < +\infty$ then there exists $M \geq 0$ such that $f(x) \leq M$ for almost every $x \in \Omega$. Implicitly, we have identified functions that agree almost everywhere, otherwise we would not have a normed vector space since a non-trivial family of functions satisfy $\|f\|_{L^p(\Omega,\mathbb{C})} = 0$ besides the zero function. For instance, consider the family of characteristic functions of single points where $f \in \{\chi_{\{q\}}\}_{q\in\mathbb{Q}}$ for $\Omega = X = \mathbb{R}$ all have integral zero, but are all almost everywhere equal to the zero function. For a detailed exposure to measure theory, consider reading Tao's book [13].

Our conventions for Fourier transformations are as follows, provided $f, \hat{f} \in L^1(\mathbb{R}, \mathbb{C})$:

$$\mathcal{F}f(\xi) := \int_{\mathbb{R}} f(x) \exp(-ix\xi)\, \mathrm{d}x$$
$$\mathcal{F}^{-1}\hat{f}(x) := \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) \exp(ix\xi)\, \mathrm{d}\xi$$

Fubini's theorem and the dominated convergence theorem (cf. Tao [13]) guarantees that these two transformations are mutually compatible in a natural way, namely that for $f \in L^1(\mathbb{R}, \mathbb{C})$ and $\mathcal{F}f \in L^1(\mathbb{R}, \mathbb{C})$ we have

$$\mathcal{F}^{-1}\{\mathcal{F}f\} = \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}f(\xi) \exp(ix\xi)\, \mathrm{d}\xi = \lim_{\varepsilon\to 0} \frac{1}{2\pi} \int_{\mathbb{R}} \mathcal{F}f(\xi) \exp\left(-\frac{1}{4\pi}\varepsilon^2\xi^2 + ix\xi \right) \mathrm{d}\xi$$

Note that $\exp(-\varepsilon^2\xi^2/4\pi + ix\xi) \in L^1(\mathbb{R}, \mathbb{C})$, so we may use Fubini's theorem to flip the Fourier transform in the integral:

$$\int_{\mathbb{R}} \mathcal{F}f(\xi) \exp\left(-\frac{1}{4\pi}\varepsilon^2\xi^2 + ix\xi \right) \mathrm{d}\xi = \int_{\mathbb{R}}\int_{\mathbb{R}} f(z) \exp\left(-\frac{1}{4\pi}\varepsilon^2\xi^2 - i\xi(z-x) \right) \mathrm{d}\xi\, \mathrm{d}z$$
$$= \int_{\mathbb{R}} f(z)\, \mathcal{F}\left(\exp\left(-\frac{1}{4\pi}\varepsilon^2\xi^2 \right) \right)(z-x)\, \mathrm{d}z = \int_{\mathbb{R}} f(z) \frac{2\pi}{\varepsilon} \exp(-\pi(z-x)^2/\varepsilon^2)\, \mathrm{d}z$$

which then by the dominated convergence theorem amounts to

$$\mathcal{F}^{-1}\{\mathcal{F}f\} = \lim_{\varepsilon\to 0} \int_{\mathbb{R}} f(z) \frac{1}{\varepsilon} \exp(-\pi(z-x)^2/\varepsilon^2)\, \mathrm{d}z = \lim_{\varepsilon\to 0} (\varphi_\varepsilon * f)(x) = f(x)$$

where $\varphi_\varepsilon(t) = 1/\varepsilon \exp(-\pi t^2/\varepsilon^2)$ has the property that $\lim_{\varepsilon\to 0}(\varphi_\varepsilon * f)(x) = f(x)$ for any $f \in L^1(\mathbb{R}, \mathbb{C})$. Thus we have the desired property that $\mathcal{F}^{-1}\{\mathcal{F}f\} = f$. This justifies the name of inverse Fourier transformation. Note that Fourier inversion only works provided both functions f and $\mathcal{F}f$ are L^1-integrable.

Recall that Fourier coefficients for $2L$-periodic functions $f \in L^p((-L, L), \mathbb{C})$ are defined by

$$\hat{f}_k := \frac{1}{2L} \int_{-L}^{L} f(x) \exp(-ix\tfrac{k\pi}{L})\, \mathrm{d}x$$

Theorem 4.1. (Carleson–Hunt)
Let $f \in L^p((-L, L), \mathbb{C})$ be a $2L$-periodic function for $L > 0$ and $p \in (1, +\infty)$. Then for $\hat{f}_k$ Fourier coefficients of f we have

$$\sum_{k\in\mathbb{Z}} \hat{f}_k \exp(ixk) = f(x) \quad a.e. \tag{4.4}$$

A proof of the Carleson–Hunt theorem can be found in [2] and [9] as Lennart Carleson proved the L^2 case and Richard A. Hunt generalized this for the case of L^p for $p > 1$. An overview of the Carleson and Carleson–Hunt theorems can be found in [12].

Definition 4.1. (Schwartz Space)
A smooth function $f \in C^\infty(\mathbb{R}^n, \mathbb{C})$ is called a Schwartz function if for all pairs of multiindices $\alpha, \beta \in \mathbb{Z}^n_{\geq 0}$ one has bounded Schwartz seminorm

$$\rho_{\alpha,\beta} := \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty.$$

The space of all Schwartz functions is called the Schwartz space and is denoted $\mathscr{S}(\mathbb{R}^n, \mathbb{C})$. A sequence f_k in $\mathscr{S}$ is said to converge to $f \in \mathscr{S}$ if and only if $\lim_{k\to\infty} \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta (f_k - f)| = 0$ for every pair of multiindices $\alpha, \beta \in \mathbb{Z}^n_{\geq 0}$

$$f_k \longrightarrow f \text{ in } \mathscr{S} \iff \lim_{k\to\infty} \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta (f_k - f)| = 0 \quad \forall \alpha, \beta \in \mathbb{Z}^n_{\geq 0}.$$

Informally, one can interpret the Schwartz space as the functions whose derivatives decay rapidly at infinity. Note that the space of compactly supported smooth functions $C_0^\infty(\mathbb{R}^n, \mathbb{C})$ is included in the Schwartz space by virtue of the compact support.

We are now going to list a few key properties of the Schwartz space that will prove useful, but we shall withhold their proofs.

Proposition 4.1. The Schwartz space is a dense subspace of $L^p(\mathbb{R}^n, \mathbb{C})$ for $1 \leq p < \infty$.

Proof. This follows from the fact that the smooth, compactly supported functions can be shown to be dense in L^p, and that the Schwartz space is contained in $L^p(\mathbb{R}^n, \mathbb{C})$ due to the rapid decay of the functions which may bound their p-norm integral, which then converges. □

Proposition 4.2. The Fourier transform is a one-to-one and onto map on the Schwartz space. Furthermore, Fourier inversion always holds:

$$\mathcal{F}\mathcal{F}^{-1} = I_{\mathscr{S}} = \mathcal{F}^{-1}\mathcal{F}.$$

These properties and their proofs, along with a slew of other useful facts and properties, may be found in [6].

4.3. Distribution Theory

Our equations need a treatment of functions which are not readily analyzable using standard Fourier analysis.To this end, we shall barely scratch the surface of fruitful theory called distribution theory. To better understand the premise, it may be best to consider a preliminary example.

Example 4.1. Consider a function $u \colon \mathbb{R} \to \mathbb{R}$ such that $\mathcal{F}(u) = \hat{u}$ and $\partial_x u$ exist in a manner such that we may write

$$\mathcal{F}(\partial_x u) = \int_{\mathbb{R}} \partial_x u \exp(-ix\xi)\,\mathrm{d}x = \int_{\mathbb{R}} (-i\xi) u(x) \exp(-ix\xi)\,\mathrm{d}x = -i\xi\, \hat{u}(\xi).$$

Likewise, for $u \colon \mathbb{R}^n \to \mathbb{R}$ we may, given suitable assumptions on the regularity and decay of the derivatives, we may write

$$\mathcal{F}(\partial_x^\alpha u)(\xi) = (-i\xi)^{|\alpha|}\, \hat{u}(\xi).$$

The interesting thing here, of course, is that we may write this formally as the action of a Fourier multiplier $m(\xi)$ on the Fourier side, namely by

$$\mathcal{F}(Lu)(\xi) = \mathcal{F}(\partial_x^\alpha u)(\xi) = (-i\xi)^{|\alpha|} \hat{u}(\xi) = m(\xi)\, \hat{u}(\xi).$$

However, the inverse Fourier transform of $m(\xi)$ is not defined, as its action on the Fourier side fails to even be L^∞-bounded on $\mathbb{R}^n$. Even when this is the case, we observe that its action in the grander picture is far from unreasonable given suitable regularity on u - we are simply taking the derivative! Although it is unreasonable for $m(\xi)$ to have an inverse Fourier transform *as a function*, we hastily introduce the core material of distributions to show that in the appropriate setting, it does make sense to speak of Fourier transformations of such objects.

Denote the space of test functions over the (open) set $\Omega \subseteq \mathbb{R}^n$ by

$$\mathcal{D}(\Omega) = \{\varphi \in C^\infty(\Omega) \mid \operatorname{supp}\varphi \subset \Omega \text{ compactly}\}.$$

The space $\mathcal{D}(\Omega)$ is equipped with a topology which is characterized thusly: consider a sequence of functions $(\varphi_j) \subset \mathcal{D}(\Omega)$. If there exists a compact set $K \subset \Omega$ and a $\varphi \in \mathcal{D}(\Omega)$ such that

$$\operatorname{supp}\varphi_j \subset K \quad \forall j \in \mathbb{Z}_{\geq 0} \text{ and } \sup_{x\in K} |\partial^\alpha \varphi_j(x) - \partial^\alpha \varphi(x)| \to 0 \text{ as } j \to \infty$$

then we say that $\varphi_j \to \varphi$ in $\mathcal{D}(\Omega)$.

A *distribution* T over $\mathcal{D}(\Omega)$ is a continuous linear functional $T\colon \mathcal{D}(\Omega) \to \mathbb{C}$, whereby linear we mean for all $a, b \in \mathbb{C}$ and $\varphi, \psi \in \mathcal{D}(\Omega)$ it is true that

$$T(a\,\varphi + b\,\psi) = a\,T(\varphi) + b\,T(\psi).$$

By *continuous* linear functional we mean that

$$T(\varphi_j) \to T(\varphi) \text{ as } j \to \infty \text{ provided } \varphi_j \to \varphi \text{ as } j \to \infty \text{ in } \mathcal{D}(\Omega).$$

The space of all continuous linear functionals over $\mathcal{D}(\Omega)$ shall be denoted by $\mathcal{D}'(\Omega)$. The topology on the space of continuous linear functionals $\mathcal{D}'(\Omega)$ is given as the following: consider a sequence $(T_j) \subset \mathcal{D}'(\Omega)$ for which we have

$$T_j(\varphi) \to T(\varphi) \text{ for all } \varphi \in \mathcal{D}(\Omega).$$

In this case we say T_j converges to T in $\mathcal{D}'(\Omega)$. The convergence criteria are included since presenting these spaces without them would be pointless. Regardless, we are not going to be needing these criteria.

Example 4.2. What do these distributions look like formally? Consider a linear functional $T_f \in \mathcal{D}(\mathbb{R})$ for which we have

$$T_f(\varphi) = \int_{\mathbb{R}} f(x)\,\varphi(x)\,\mathrm{d}x$$

where $f \in L^1_{\mathrm{loc}}(\mathbb{R}) = \{f\colon \mathbb{R} \to \mathbb{R} \mid f|_K \in L^1(K) \text{ for any } K \subset \mathbb{R} \text{ compact}\}$. This is indeed a continuous linear functional, but notice that f need not be continuous. As we shall see, this kind of distribution will prove useful.

Remark. We sometimes use the notation $\langle T, \varphi\rangle = T(\phi)$ to better signify the action of the distribution $T \in \mathcal{D}(\Omega)$, especially when T has a regular form as in the preceding example.

We may have continuous linear functionals on the Schwartz space $\mathscr{S}(\Omega)$ as well, by the exact same definition and topology as for $\mathcal{D}(\Omega)$, except now over $\mathscr{S}(\Omega)$. We call the space of continuous linear functionals $\mathscr{S}'(\Omega)$ the space of *tempered distributions*. We see then that $\mathcal{D}(\Omega)$ extends to the Schwartz space $\mathscr{S}(\Omega)$ in the sense that $\mathcal{D}(\Omega) \subset \mathscr{S}(\Omega)$. However, this means that on the distributional side we have $\mathscr{S}'(\Omega) \subset \mathcal{D}'(\Omega)$. In other words, every tempered distribution is also a distribution.

Definition 4.2. (Fourier transform of tempered distributions)
Let $T \in \mathscr{S}'(\mathbb{R}^n)$. Then the Fourier transform of T denoted $\mathcal{F}T$ is defined formally by

$$\mathcal{F}T(\varphi) = T(\mathcal{F}\varphi), \quad \varphi \in \mathscr{S}(\mathbb{R}^n).$$

Example 4.3. What is the Fourier transform of the unit function $f(x) = 1$? This seems a nonsensical question, which it is in the sense of taking Fourier transformations of functions, but for distributions the Fourier transform makes sense:

$$\begin{aligned}\mathcal{F}1(\varphi) = \langle 1, \mathcal{F}\varphi\rangle &= \int_{\mathbb{R}^n} \hat{\varphi}(\xi)\,\mathrm{d}\xi = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \varphi(x)\,\exp(-ix\xi)\,\mathrm{d}x\,\mathrm{d}\xi \\ &= \int_{\mathbb{R}^n} \exp(i0\xi)\int_{\mathbb{R}^n} \varphi(x)\,\exp(-ix\xi)\,\mathrm{d}x\,\mathrm{d}\xi = (2\pi)^n\mathcal{F}^{-1}\mathcal{F}(\varphi)(0) = (2\pi)^n\varphi(0)\end{aligned}$$

In fact, one may feasibly extend this notion of Fourier transform to all essentially bounded functions $f \in L^\infty(\mathbb{R}^n)$.

Convolutions of distributions are readily definable from our building blocks covered thus far, and will be important later for checking the consistency of equations that arise from using Fourier multipliers.

Definition 4.3. (Convolutions on Tempered Distributions)
Given $\psi \in \mathscr{S}(\mathbb{R}^n)$ and $f \in \mathscr{S}'(\mathbb{R}^n)$ we define the distribution $\psi * f$ by

$$\langle \psi * f, \varphi\rangle = \langle f, \tilde{\psi} * \varphi\rangle \quad \text{for } \varphi \in \mathscr{S}(\mathbb{R}^n)$$

where $\tilde{\psi}(x) = \psi(-x)$.

4.4. Hölder Spaces

Our choice of bifurcation space on our problem will be what is known as a Hölder space. In case of unfamiliarity, we write its definition and key properties in this subsection. The primary references for this section is [3] and [8].

Definition 4.4. (Hölder Spaces)
Let $\Omega \subseteq \mathbb{R}^n$ be open, and denote the space of bounded, continuous functions over Ω as $BC(\Omega)$, and likewise with $BC^k(\Omega)$ for k-times differentiable, bounded continuous functions. We say a function $f \in BC^k(\Omega)$ is *Hölder k-times continuously differentiable with exponent* $0 < \alpha \leq 1$ if each derivative of f up to order k has finite $C^{0,\alpha}$-norm given by

$$\|f\|_{C^{0,\alpha}(\Omega)} := \sup_{x\in\Omega}|f(x)| + \sup_{\substack{x,\,y\in\Omega\\ x\neq y}} \frac{|f(x)-f(y)|}{\|x-y\|^\alpha}, \quad [f]_\alpha := \sup_{\substack{x,\,y\in\Omega\\ x\neq y}} \frac{|f(x)-f(y)|}{\|x-y\|^\alpha}.$$

Furthermore, the norm of $C^{k,\alpha}(\Omega)$ is given by

$$\|f\|_{C^{k,\alpha}(\Omega)} = \sum_{|\beta|\leq k} \|\partial^\beta f\|_{BC(\Omega)} + \sum_{|\beta|=k} [\partial^\beta f]_\alpha.$$

The space of all Hölder continuous functions over Ω with exponent α is then the Hölder space $C^{0,\alpha}(\Omega) = \{f \in BC(\Omega) \mid \|f\|_{C^{0,\alpha}(\Omega)} < \infty\}$.

Remark. Our notation avoids the possible confusion between C^k-spaces and Hölder spaces $C^{k,\alpha}$. We note that for $f \in C^{0,\alpha}(\Omega)$ with exponent $\alpha = 1$ there exists an $M > 0$ such that for $x, y \in \Omega$ and $x \neq y$

$$\frac{|f(x)-f(y)|}{\|x-y\|} \leq M$$

which then clearly implies that f is Lipschitz continuous on Ω. Hence unit exponent functions are all Lipschitz. Furthermore, if we allow the exponent α to be strictly greater than 1, we see by the same token that

$$|f(x) - f(y)| \leq M\|x - y\|^\alpha \leq M\|x - y\|$$

whenever $\|x - y\| \leq 1$, so therefore $f \in C^{0,\alpha}(\Omega)$ is locally Lipschitz about every point $x \in \Omega$. By Rademacher's theorem (cf. [8]) we have that every Lipschitz function $f\colon U \subset \Omega \to \mathbb{R}$ is almost everywhere differentiable, and thus we can speak of a total derivative almost everywhere locally on our domain Ω. Therefore we notice that by the definition of the total derivative/Fréchet derivative that

$$\lim_{\|x-y\|\to 0} \frac{|f(x) - f(y) - \mathrm{d}f[y](x-y)|}{\|x - y\|} = 0$$

for almost every $y \in U$, which, after applying the inverse triangle inequality and Cauchy-Schwarz turns to

$$\left| \frac{|f(x) - f(y)|}{\|x - y\|} - \|\mathrm{d}f[y]\| \right| \leq \frac{|f(x) - f(y) - \mathrm{d}f[y](x-y)|}{\|x - y\|}.$$

Applying the limit $\|x - y\| \to 0$ on both sides necessarily forces the Fréchet derivative to be zero at the point, and thus f is constant on each connected component since we may find a parametrized curve $\gamma\colon [0,1] \to \Omega$ such that $\gamma(0) = x$, $\gamma(1) = y$ and

$$f(x) - f(y) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) \cdot \mathrm{d}f(\gamma(t))\, \mathrm{d}t = 0.$$

Therefore, every function is constant on each (path) connected component of Ω when $\alpha > 1$, which is not of much interest when considering connected domains.

Proposition 4.3. The Hölder spaces $C^{0,\alpha}(\Omega)$ for $0 < \alpha \leq 1$ are Banach spaces.

Proof. Write the following for $f \in C^{0,\alpha}(\Omega)$

$$\|f\|_{C^{0,\alpha}(\Omega)} = \|f\|_{BC(\Omega)} + [f]_\alpha.$$

We first need to show that $\|\cdot\|_{C^{0,\alpha}(\Omega)}$ is indeed a norm on the vector space $C^{0,\alpha}(\Omega)$. The normed space $(BC(\Omega), \|\cdot\|_{BC(\Omega)})$ with norm given by $\|f\|_{BC(\Omega)} = \sup_{x\in\Omega} |f(x)|$ has the required non-degeneracy condition in the definition of our proposed norm for $C^{0,\alpha}(\Omega)$, therefore if the remaining term involving $[f]_\alpha$ constitutes a semi-norm (that is, regardless of degeneracy), then $\|\cdot\|_{C^{0,\alpha}(\Omega)}$ is a norm. Clearly, for any scalar $\lambda \in \mathbb{R}$ we have $\|\lambda f\|_{C^{0,\alpha}(\Omega)} = |\lambda|\|f\|_{C^{0,\alpha}(\Omega)}$, thus it remains to calculate a triangle inequality for our semi-norm term:

$$\begin{aligned}
[f + g]_\alpha &= \sup_{\substack{x,\, y\in\Omega \\ x\neq y}} \frac{|(f(x) + g(x)) - (f(y) - g(y))|}{\|x - y\|^\alpha} = \sup_{\substack{x,\, y\in\Omega \\ x\neq y}} \frac{|f(x) - f(y) + g(x) - g(y)|}{\|x - y\|^\alpha} \\
&\leq \sup_{\substack{x,\, y\in\Omega \\ x\neq y}} \frac{|f(x) - f(y)| + |g(x) - g(y)|}{\|x - y\|^\alpha} = [f]_\alpha + [g]_\alpha
\end{aligned}$$

which therefore shows that $(C^{0,\alpha}(\Omega), \|\cdot\|_{C^{0,\alpha}(\Omega)})$ is a normed space.

Regarding completeness, first recall that $(BC(\Omega), \|\cdot\|_{BC(\Omega)})$ as before is a Banach space and that functions $f \in C^{0,\alpha}(\Omega)$ are also by definition bounded and continuous. Let $(f_n) \subset C^{0,\alpha}(\Omega)$ be a Cauchy sequence, which then means that for every $\varepsilon > 0$ there exists $N(\varepsilon) > 0$ such that $\|f_n - f_m\|_{C^{0,\alpha}(\Omega)} < \varepsilon$ whenever $m, n \geq N(\varepsilon)$. In particular, this means that for $x \neq y$ in Ω we have

$$\varepsilon > \|f_n - f_m\|_{C^{0,\alpha}(\Omega)} \geq [f_n - f_m]_\alpha = \frac{|f_n(x) - f_n(y) - (f_m(x) - f_m(y))|}{\|x - y\|^\alpha}.$$

By this we immediately note that if the Cauchy sequence (f_n) converges to $f \in BC(\Omega)$, we may then write

$$\begin{aligned}\frac{|f_n(x) - f(x) - (f_n(y) - f(y))|}{\|x - y\|^\alpha} &= \lim_{m\to\infty} \frac{|f_n(x) - f_m(x) - (f_n(y) - f_m(y))|}{\|x - y\|^\alpha} \\ &\leq \limsup_{m\to\infty} \|f_n - f_m\|_{C^{0,\alpha}(\Omega)}\end{aligned}$$

since (f_n) was assumed to be Cauchy, and therefore we have $[f_n - f]_\alpha \to 0$ as $n \to \infty$ which implies $\lim_{n\to\infty}\|f_n - f\|_{C^{0,\alpha}(\Omega)} = 0$, concluding the proof. ☐

Definition 4.5. (Banach Algebra)
Let $(X, \|\cdot\|_X)$ be a Banach space. If furthermore X is an associative algebra over the real (or complex) numbers such that for $f, g \in X$

$$\|f \cdot g\|_X \leq \|f\|_X \cdot \|g\|_X$$

then X is called a *Banach algebra.*

Proposition 4.4. The Hölder space $C^{0,\alpha}(\Omega)$, $\Omega \subseteq \mathbb{R}^n$ is a Banach algebra.

Proof. We have already established that $(C^{0,\alpha}(\Omega), \|\cdot\|_{C^{0,\alpha}(\Omega)})$ is a Banach space. Furthermore, for $f, g \in C^{0,\alpha}(\Omega)$ there exists $M_f, M_g \geq 0$ such that

$$\begin{aligned}|(fg)(x) - (fg)(y)| &\leq |f(x)(g(x) - g(y))| + |g(y)(f(x) - f(y))| \\ &\leq \|f\|_{BC(\Omega)}\, M_g\|x - y\|^\alpha + \|g\|_{BC(\Omega)}\, M_f\|x - y\|^\alpha \\ &= (\|f\|_{BC(\Omega)}\, M_g + \|g\|_{BC(\Omega)}\, M_f)\|x - y\|^\alpha\end{aligned}$$

which shows that we may take the quotient and supremum to get finite $C^{0,\alpha}$-norm, thus concluding the proof. ☐

4.5. Classical Symbols

The theory of pseudodifferential operators gives us operators that on the Fourier side act as if they were in some sense differential operators in the physical space. The rough idea is to manipulate the functions on the Fourier side in ways that mimic the behaviour of differential operators.

Pseudodifferential operators can be viewed as convolution kernels of the form

$$K(x) = \int_{\mathbb{R}^n} a(x, \xi)\, \exp\left(i\varphi(x, \xi)\right) \mathrm{d}\xi$$

where $\varphi(x, \xi)$ is a phase function and $a(x, \xi)$ is a *symbol.* These functions may be regarded quite generally.

Kazuaki [10] gives a good exposition to the general classification of symbol classes.

Definition 4.6. (Symbol Classes)
Let Ω be an open subset of $\mathbb{R}^n$. If $s \in \mathbb{R}$ and $0 \leq \delta < \rho \leq 1$ we let $S^s_{\rho,\delta}(\Omega \times \mathbb{R}^n)$ be the set of all functions $a(x, \xi)$ such that for any compact $K \subset \Omega$ and multi-indices α, β there exists constants $C_{K,\alpha,\beta} > 0$ such that for all $x \in K$ and $\xi \in \mathbb{R}^n$ one has

$$|\partial^\alpha_\xi \partial^\beta_x a(x, \xi)| \leq C_{K,\alpha,\beta}(1 + |\xi|)^{s-\rho|\alpha|+\delta|\beta|}.$$

We call $S^s_{\rho,\delta}(\Omega \times \mathbb{R}^n)$ the *symbol class of order s.*

Of particular interest to us are the *classical symbols* given by $S^s_{1,0}(\mathbb{R},\mathbb{R})$. An important family of such symbols are the *Bessel symbols* given by

$$m(\xi) = (1+\xi^2)^{\frac{s}{2}}, \quad s \in \mathbb{R} \setminus \{0\}.$$

These symbols when acted upon a function have the ability to increase or decrease the regularity/smoothness of said function, depending as the order s is strictly negative or positive respectively.

5. Nonlinear Dispersive Equations Inspired by KdV and Whitham

Given our build-up of theory, we are now ready to tackle our main problem. First we present our equation of interest, and then we begin the analysis of its local bifurcation solutions after having examined some rudimentary properties. The following analysis is heavily inspired by the work of Ehrnström and Kalisch [5].

Our main focus will be the family of equations given by

$$\partial_t u + L\partial_x u + \partial_x(u^{p+1}) = 0, \quad p \in \mathbb{Z}_{\geq 2}. \tag{5.1}$$

Here, the Fourier multiplier L will be assumed to be a Bessel symbol on the Fourier side

$$m(\xi) = (1+\xi^2)^{\frac{s}{2}}, \quad s \in (-\infty, 0).$$

This is a classical symbol. Note that it is also real and symmetric as a function.

Remark. Note that for $m(\xi) = -\xi^2$ and $p = 1$ one obtains the dimensionless KdV equation, Eqn. (4.2). In fact, using the convolution theorem

$$\mathcal{F}(K_{\text{Whitham}} * \partial_x u) = \mathcal{F}(K_{\text{Whitham}}) \cdot \mathcal{F}(\partial_x u) = \sqrt{\tfrac{g \tanh h_0 \xi}{\xi}}\, \mathcal{F}(\partial_x u)$$

we see that even the Whitham equation (4.3) may be rephrased as an equation with a Fourier multiplier, meaning that Eqn. (5.1) specializes to the Whitham equation as well. The proof that $K_{\text{Whitham}} \in L^1(\mathbb{R})$ can be found in [5]. Indeed, this in part justifies calling the family of equations Eqn. (5.1) a kind of generalization of KdV and Whitham. Because of this, we delimit solutions and their analysis to the heuristics of physical limits - even though our equation need not be physically inspired per se.

We know that by our previously established theory, if we first impose the ansatz of traveling solutions of the form $u(t,x) = \eta(x-ct)$ with propagation speed $c > 0$ we may rewrite Eqn. (5.1) as

$$-c\,\eta' + L\eta' + \eta^p\,\eta' = 0$$

where we recall by our previously established theory that L may have its action written by a convolution since $\mathcal{F}(Lf)(\xi) = m(\xi)\,\hat{f}(\xi) = \mathcal{F}(\mathcal{F}^{-1}m * f)(\xi)$, at least in the distributional sense. Since we may pass an integral through the convolution, our equation may be integrated to the following equation

$$-c\,\eta + L\eta + \eta^{p+1} = B$$

where we may normalize $B = 0$ due to the expected convergent properties of the solution as $|x| \to \infty$. With all of this, we have arrived at the normalized (weak) version of our equation

$$-c\,\eta + L\eta + \eta^{p+1} = 0. \tag{5.2}$$

Remark. For Bessel symbols of negative order, the convolution kernel K in $Lf = K * f$ happens to have unit L^p-norm for $1 \leq p < \infty$. A proof and further examination of this can be found in Grafakos [7].

The following theorem, lemma and consequent proof of theorem is heavily inspired by the article of Ehrnström and Kalisch [5].

Theorem 5.1. (Main theorem - Existence of small amplitude solutions to Eqn. (5.1))

For a given $L > 0$ there exists a local bifurcation curve consisting of $2L$-periodic, even and continuous solutions to the weak normalized equation (5.2). These solutions are perturbations in the direction of $\cos(\pi x/L)$ in the appropriate bifurcation space, and their maximal wave speed $c_{\max}$ is determined by

$$\int_{\mathbb{R}} K(x)\,\mathrm{d}x = c_{\max}$$

which for Fourier multipliers of Bessel symbols of negative order happens to evaluate to $c_{\max} = 1$. Furthermore, owing to the dispersion relation $m(\xi)$ of the equation, the wave speed at the bifurcation point is given by

$$c^* = \left(1 + \frac{\pi^2}{L^2}\right)^{\frac{s}{2}}$$

where in particular as $L \to \infty$ one has $c \to 1$.

Recall from Section 3.2 that Fréchet derivatives on parameters λ are identified with scalars. Therefore it also naturally commutes with the derivative on the bifurcation space.

Lemma 5.1. (Crandall-Rabinowitz Revisited)
Let W be a Banach algebra, and let $c \in (0,1)$ be a parameter. Let $\mathcal{L}\colon W \to W$ be the Fréchet derivative at $0 \in W$ with respect to the function u of the map

$$\mathcal{J}\colon u \longmapsto -cu + Lu + u^{p+1}. \tag{5.3}$$

Suppose also that both $\mathcal{L}$ and $\partial_c\mathcal{L}$ exist and are continuous on and onto W, and that for some specific parameter $c^* \in (0,1)$ the following conditions hold:

(i) $\dim\ker(\mathcal{L}) = 1$;
(ii) $W = \ker(\mathcal{L}) \oplus \operatorname{im}(\mathcal{L})$;
(iii) $(\partial_c\mathcal{L})\ker(\mathcal{L}) \cap \operatorname{im}(\mathcal{L}) = 0$.

Then there exists $\varepsilon > 0$ and a continuous bifurcation curve $\{(\phi_s, c_s) \mid |s| < \varepsilon\}$ with $c_s|_{s=0} = c^*$. Furthermore ϕ_0 is the vanishing solution of the normalized equation (5.2) and $\{\phi_s\}_s$ are nontrivial solutions to the normalized equation with corresponding wave speeds $\{c_s\}_s$. In addition to all of this, we have for all solutions $\phi_s \in W$ that

$$\operatorname{dist}(\phi_s, \ker(\mathcal{L})) = o(s).$$

The reader is encouraged to work through the connections of the original Crandall-Rabinowitz formulation and the formulation presented above to show equivalency in our context.

Before we turn to proving the theorem, we have to analyze how the convolution acts on functions $f \in L^\infty(\mathbb{R})$ that are both even and $2L$-periodic. Assume that the Fourier multiplier L can be written as a convolution when acting on the function f, so

$Lf = K * f$. Then the following integral makes sense

$$\int_{\mathbb{R}} K(x-y)f(y)\,\mathrm{d}y = \sum_{k\in\mathbb{Z}} \int_{-L}^{L} K(x-y+2kL)f(y)\,\mathrm{d}y$$
$$= \int_{-L}^{L} \left(\sum_{k\in\mathbb{Z}} K(x-y+2kL) \right) f(y)\,\mathrm{d}y =: \int_{-L}^{L} A(x-y)f(y)\,\mathrm{d}y$$

where we immediately note through inspection that $A(x)$ is even, $2L$-periodic and continuous on $[-L, L] \setminus \{0\}$. Furthermore, we exploit Minkowski's inequality

Lemma 5.2. (Minkowski's Inequality)
Let $1 \leq p < \infty$ and assume $f, g \in L^p(\mathbb{R})$. Then Minkowski's inequality is given by

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

which can be used to show that $A(x) \in L^p(-L, L)$ for $1 \leq p < 2$. Then, according to Carleson-Hunt (Theorem 4.1), we may approximate $A(x)$ by its Fourier series pointwise almost everywhere:

$$A(x) = \frac{1}{L} \sum_{k=0}^{\infty}{}' \hat{A}_k \cos\left(\tfrac{k\pi x}{L}\right) \quad a.e.$$

where the prime on the sum indicates that the first term in the sum be multiplied by $1/2$. The Fourier coefficients of A are given by

$$\hat{A}_k = \int_{-L}^{L} \sum_{j\in\mathbb{Z}} K(x+2jL)\, \exp\left(-\tfrac{ixk\pi}{L}\right) \mathrm{d}x$$
$$= \sum_{j\in\mathbb{Z}} \int_{-L}^{L} K(x+2jL)\, \exp\left(-\tfrac{i(x+2jL)k\pi}{L}\right) \mathrm{d}x$$
$$= \int_{\mathbb{R}} K(x)\, \exp\left(-\tfrac{ixk\pi}{L}\right) \mathrm{d}x = \hat{K}\left(\tfrac{k\pi}{L}\right).$$

Finally, the convolution we initially began with can now be written as

$$K * f(x) = \frac{1}{L} \sum_{k=0}^{\infty}{}' \hat{f}_k \hat{A}_k \cos\left(\tfrac{xk\pi}{L}\right) = \frac{1}{L} \sum_{k=0}^{\infty}{}' \hat{f}_k \hat{K}\left(\tfrac{k\pi}{L}\right) \cos\left(\tfrac{xk\pi}{L}\right).$$

Having established these properties, we turn to prove Theorem 5.1.

Proof. Firstly, note that by our previous examinations, we have that $Lu = K * u$ for some kernel K (distributional if need be) and that since $m(\xi) = (1+\xi^2)^{\frac{s}{2}}$ for $s < 0$ we have that

Linearization of the main equation gives

$$\mathcal{L}\psi := \psi - \frac{1}{c} K * \psi = 0 \tag{5.4}$$

where if $\psi \in L^\infty(\mathbb{R})$ we see that in the distributional sense we have

$$\hat{\psi} \left(1 - \frac{1}{c} m(\xi) \right) = 0.$$

Note that $\hat{\psi}$, $\frac{1}{c}\widehat{K * \psi}$ and $\frac{1}{c}\hat{K}$ all exist as tempered distributions in the space $\mathscr{S}'(\mathbb{R})$. Furthermore, given that our Fourier multiplier is a Bessel symbol of strictly negative order we have that $1 - m(\xi)/c$ is both essentially bounded and smooth, so taking the product with the distribution $\hat{\psi}$ we may let this product act on Schwartz functions $\varphi \in$

$\mathscr{S}(\mathbb{R})$. Again, since the product is a distribution, we may then use the (distributional) convolution theorem to obtain the relation

$$\frac{1}{c}\widehat{K * \psi}(\varphi) = \frac{1}{c}(\hat{\psi}\hat{K})(\varphi), \qquad \text{for any } \varphi \in \mathscr{S}(\mathbb{R}).$$

This establishes

Given our equation as above, we start to examine whenever $\hat{\psi}$ vanishes. Given $c < 1$ we see that the equation

$$1 - \frac{1}{c}(1+\xi^2)^{\frac{s}{2}} = 0$$

has two solutions $\pm\xi_0$ since the Bessel function is in particular always decreasing and symmetric about $\xi = 0$. For $c = 1$ we have only one solution, namely $\xi = 0$. Lastly, for $c > 1$ we have no solutions to the above equation - which immediately implies that the distribution $\hat{\psi}(\varphi)$ has to vanish for all φ when $c > 1$. Using these results, we may formally find the inverse Fourier transforms of the distribution by use of the Dirac delta distribution given by

$$\langle \delta, \varphi \rangle = \varphi(0), \quad \varphi \in \mathscr{S}(\mathbb{R})$$

and shifting this distribution by $\xi \in \mathbb{R}$ to the distribution δ_ξ such that $\langle \delta_\xi, \varphi \rangle = \varphi(\xi)$. Then it turns out that the nontrivial solutions to the linearized equation (5.4) are given by the functions

$$\begin{cases} \psi(x) = C, & c = 1, \\ \psi(x) = C\,\cos(\xi_0 x), & c < 1, \end{cases} \tag{5.5}$$

for constants $C \in \mathbb{R} \setminus \{0\}$. Because of the physical nature of the context from which we derive our PDE, we shall not be including the non-zero constant functions as part of our analysis. Allowing the parameter $c > 0$ to be the bifurcation parameter for fixed wavelength $L > 0$ (which also fixes the period), we set out to perform local bifurcation theory on our nontrivial solutions. Clearly, from Eqn. (5.5) we see that in the case of $2L$-periodic and even solutions to our linearized equation we have

$$\dim\ker(\mathcal{L}) = 1 \text{ if and only if } \xi_0 = k\pi/L \text{ for } k \in \mathbb{Z}_{\geq 1}.$$

Now, choose the lowest mode of frequency $k = 1$ as above. This ensures uniqueness of c in the dispersion relation of our equation, and also allows us to establish the proposed c^* as in the theorem.

Our choice of bifurcation space will be the Hölder space $C^{0,\alpha}(\mathbb{R})$ for $\alpha > 1/2$. Note that we have already shown every $C^{0,\alpha}$-space to be a Banach algebra, so in particular we have that multiplication of functions is a continuous operation on the Hölder space. The nonlinear term is then not a problem in terms of continuity.

As we have seen, we may write as the action of $\mathcal{L}$ the following relation

$$\mathcal{L}u = \frac{1}{L}\sum_{k=0}^{\infty}{}' \hat{u}(k)\left(1 - \tfrac{1}{c}\hat{A}(k)\right)\cos\left(\tfrac{xk\pi}{L}\right) \quad a.e. \tag{5.6}$$

on the interval $[-L, L]$. Due to the Riemann-Lebesgue lemma one has the convergent property that $\hat{A}(k) \to 0$ as $k \to \infty$, so therefore we have that $\mathcal{L}u$ is contained in our Banach algebra $C^{0,\alpha}(\mathbb{R})$ and is hence a continuous map. Furthermore, notice that

$$\|\mathcal{L}u\|_{C^{0,\alpha}(\mathbb{R})} \leq C\|u\|_{C^{0,\alpha}(\mathbb{R})}$$

for some constant $C > 0$ since the Bessel symbols of negative order happen to increase the regularity of the functions they act upon. Therefore we have that Eqn. (5.6) is an equality due to continuity.

We have already established that $\ker(\mathcal{L}) = \mathrm{span}_{\mathbb{R}}(\cos(\pi x/L))$, which corresponds to

$$\hat{A}(1) = c \qquad \text{and} \qquad \hat{A}(k) \neq c \text{ for } k \neq 1.$$

Therefore, based on Theorem 6.2 in [4] by Ehrnström et al. we may take some given $u \in C^{0,\alpha}(\mathbb{R})$ and $u^{\perp} \in C^{0,\alpha}(\mathbb{R})$ with $\widehat{u^{\perp}}(1) = 0$, also $\widehat{u^{\perp}}(k) = \hat{u}(k)$ for $k \neq 1$. The aforementioned theorem makes sense of the following function in our Hölder space

$$v(x) := \frac{1}{L}\sum_{k=0}^{\infty}{}' \frac{\widehat{u^{\perp}}(k)}{1 - \frac{1}{c}\hat{A}(k)} \cos\left(\tfrac{k\pi x}{L}\right)$$

which then readily implies that $v(x) = \mathcal{L}^{-1}u^{\perp}$. Moreover, this implies that

$$u(x) = \mathcal{L}v + \frac{\hat{u}(1)}{L} \cos\left(\tfrac{k\pi x}{L}\right)$$

which indeed shows that $C^{0,\alpha}(\mathbb{R}) = \ker(\mathcal{L}) \oplus \mathrm{im}(\mathcal{L})$. The derivative of the Fréchet derivative with respect to the bifurcation parameter is then

$$(\partial_c \mathcal{L})u = -(\partial_c \frac{1}{c} K) * u = \frac{1}{c^2} K * u.$$

By the same arguments as above we therefore have

$$(\partial_c \mathcal{L})u = \frac{1}{Lc^2}\sum_{k=0}^{\infty}{}' \hat{u}(k)\, \hat{A}(k) \cos\left(\tfrac{k\pi x}{L}\right)$$

as a bounded map on $C^{0,\alpha}(\mathbb{R})$ and hence continuous. Finally, we also have that

$$(\partial_c \mathcal{L}) \ker(\mathcal{L}) \cap \mathrm{im}(\mathcal{L}) = \ker(\mathcal{L}) \cap \mathrm{im}(\mathcal{L}) = 0.$$

$\square$

5.1. On Generalizing to Arbitrary Classical Symbols

By now it should be quite clear that generalizing to other classical symbols with negative order, explicitly or formally given as functions, should be possible if they satisfy the various L^1-integrability conditions, the correct dispersion relation properties, and the various other ingredients that went into the proof of our main theorem.

However, in generality it might prove difficult to verify all of these properties *a priori* without more information about the functions we are looking at.

References

[1] B. Buffoni and J. Toland. *Analytic Theory of Global Bifurcation.* Princeton Series in Applied Mathematics. Princeton University Press, 2003.

[2] L. Carleson. Convergence and growth of partial sums of fourier series. *Acta Math.*, (116):135–157, 1966.

[3] B. K. Driver. Analysis tools with applications. *Springer/University of California San Diego Online*, June 2003. Retrieved online.

[4] M. Ehrnström, J. Escher, and B.-V. Matioc. Well-posedness, instabilities, and bifurcation results for the flow in a rotating hele–shaw cell. *Journal of Mathematical Fluid Mechanics*:13(2):271–293, 2011.

[5] M. Ehrnström and H. Kalisch. Traveling waves for the whitham equation. *Differential and Integral Equations*, 22(11-12):1193–1210, 2009.

[6] F. J. Gjestland. *Distributions, Schwartz Space and Fractional Sobolev Spaces.* Master's thesis, Norwegian University of Technology and Science, July 2013.

[7] L. Grafakos. *Modern Fourier Analysis.* Graduate Texts in Mathematics. Springer Verlag, third edition, 2014.

[8] J. Heinonen. Lectures on lipschitz analysis. *Lectures at the 14th Jyväskylä Summer School*, Aug. 2004. Retrieved online.

[9] R. A. Hunt. On the convergence of fourier series. *Proc. Conf. Orthogonal Expansions and their Continuous Analogues, Southern Illinois Univ. Press*:234–255, 1968.

[10] Kazuaki. *Semigroups, Boundary Value Problems and Markov Processes.* Springer Monographs in Mathematics. Springer Verlag, second edition, 2014.

[11] H. Kielhöfer. *Bifurcation Theory - An Introduction with Applications to PDEs*, volume 156 of *Applied Mathematical Sciences.* Springer Verlag, 2011.

[12] M. Lacey. Carleson's theorem: proof, complements, variations, 2003. arXiv: `math/0307008 [math.CA]`.

[13] T. Tao. *An Introduction to Measure Theory*, volume 126 of *Graduate Studies in Mathematics.* Amer. Math. Soc., second edition, 2011.

NTNU

Norwegian University of Science and Technology