Exact and Approximate Methods for the Computation of the Spectral Stability of Traveling-Wave Solutions

Christopher W. Curtis

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

University of Washington

2009

Program Authorized to Offer Degree: Applied Mathematics

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Christopher W. Curtis

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

_____

Bernard Deconinck

Reading Committee:

_____

Robert E. O'Malley Jr.

_____

J. Nathan Kutz

_____

Anne Greenbaum

_____

Yua Yuan

Date: _____

University of Washington

Abstract

Exact and Approximate Methods for the Computation of the Spectral Stability of
Traveling-Wave Solutions

Christopher W. Curtis

Chair of the Supervisory Committee:
Associate Professor Bernard Deconinck
Applied Mathematics

This thesis addresses the use of various techniques in functional analysis applied to the
problem of determining the spectral stability of a traveling wave solution to a nonlinear
partial differential equation. This work is separated into two parts. In the first part,
a numerical method for the determination of spectral stability, called Hill's Method, is
analyzed. In the second part, I determine, both analytically and numerically, the spectral
stability of a family of solitary wave solutions to a new Boussinesq approximation to the
Euler water wave equations.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I'll get to thanking people when I'm done writing this thing.

DEDICATION

to my beloved wife, Kate

Chapter 1

INTRODUCTION

1.1 Traveling Wave Solutions and Stability

The focus in this thesis is on the stability of traveling-wave solutions (TWS(s)) to evolution equations, *i.e.* nonlinear-partial-differential equations of the form

$$u_t = F(u) \tag{1.1}$$

$$u(x,0) = u_0(x),$$

where $x \in \mathbb{R}$. A TWS is a solution u to (1.1) of the form $u(x,t) = T(ct)u_0(x;c)$, where c is called the wave speed, and T is a semigroup representing the action of a symmetry of the evolution equation, and by $u_0(x;c)$ we mean some appropriate scaling of the initial condition. As an example we look at the focusing nonlinear Schrödinger (NLS) equation:

$$\mathrm{i}u_t = -\frac{1}{2}u_{xx} - u|u|^2. \tag{1.2}$$

It is well known and easy to show that any solution to NLS is invariant under phase transformations *i.e.* if u solves NLS then ue^{is} is also a solution for any real valued number s. Thus if we look for solutions of the form $w(x)e^{ict}$, we get the following problem

$$cw(x) = \frac{1}{2}w_{xx} + w|w|^2 \tag{1.3}$$

Assume w is real and integrate with respect to w to get the differential equation

$$(w_x)^2 = -w^4 + 2cw^2 + d. \tag{1.4}$$

If we assume $d < 0$, $c^2 - d = \frac{1}{4}$, and $c = k^2 - \frac{1}{2}$, then using the elliptic integral [21]

$$\mathrm{cn}^{-1}(x,k) = \int_{kx}^{k}(k^2 - s^2)^{-1/2}((1-k^2)^2 + s^2)^{-1/2}ds, \tag{1.5}$$

we see that $w(x) = k\text{cn}(x, k)$. Thus we can generate the TWS

$$u(x, t) = k\text{cn}(x, k)e^{-i(\frac{1}{2}-k^2)t}, \tag{1.6}$$

where $T(ct) = e^{-i(\frac{1}{2}-k^2)t}$.

Another example illustrating the use of a different symmetry is Korteweg de-Vries (KdV) equation

$$u_t + uu_x + u_{xxx} = 0 \tag{1.7}$$

In this case, it is easy to show that if $u(x, t)$ is a solution to KdV, then $u(x + s, t)$ is also a solution, and thus we look for a TWS of the form $u(x, t) = w(x - ct; c) = T(ct)w(x; c)$. Note, in general one can look at problems with the spatial variable being in any arbitrary dimension. In our case, the spatial variable, x, is restricted to be one-dimensional for ease of presentation.

The interest in TWS's stems both from physical importance and mathematical tractability. In physics, translating profiles have been of interest in a number of fields ranging from fluid dynamics to optics, with the famous sech squared profile of KdV being an excellent example. Of course, a TWS is simply a particular solution to a model of some physical system. Thus the scientist is forced to ask, even if we believe in our model, do we really expect to see a particular solution in nature? The idea of stability addresses this key issue in mathematical modeling.

There are several notions of stability, some stronger or weaker than others. The strongest definition of stability referred to in this paper is known as orbital stability and is defined in the following way. One says that a TWS u is stable if for any solution $\psi(x, t)$ with nearby initial conditions to $u_0(x)$, one can at each time t find some phase shift s such that $\psi(x, t)$ is close to $u_0(x - ct - s)$. In more technical terms, we have if for all $\epsilon > 0$ there exists $\delta > 0$ such that if $\psi(x, t)$ is a solution to the evolution equation with initial condition $\psi(x, 0) = \psi_0(x)$ and

$$||\psi_0(x) - u(x)|| < \delta, \tag{1.8}$$

then

$$\inf_{s \in \mathbb{R}} ||\psi(x, t) - u(x - ct - s)|| < \epsilon \tag{1.9}$$

for $t \gg 1$. Note, $||\cdot||$ is a norm of some Hilbert space, with corresponding inner product $<,>$.

To understand the intuition behind this definition, one can look at the propagation of one-soliton solutions in KdV. Since the amplitude and wave speed are coupled, two one-soliton profiles close in amplitude at $t = 0$ will necessarily diverge in norm as time increases. Thus one might be inclined to say any soliton solution to KdV is unstable. But it is clear that if two soliton solutions begin near each other, they will remain close in norm modulo an appropriate phase shift at every later time t since KdV will preserve the shapes of the initial profiles.

In general, orbital stability is quite difficult to establish since the nonlinearities of the evolution equation prevent any simple understanding of the effects of varying initial conditions. A simpler problem can be defined by way of a TWS $T(ct)w(x;c)$. Plugging this ansatz into (1.1), we get

$$T_c T(ct)w(x;c) = F(T(ct)w(x;c)) \tag{1.10}$$

and so at $t = 0$, noting that $T(0) = 1$, we get $T_c w(x;c) = F(w(x;c))$. Hence, w becomes a stationary solution of the associated PDE

$$u_t = F(u) - T_c u. \tag{1.11}$$

Let $\psi(x,t)$ be a solution to (1.11) such that $\psi(x,0) = w(x) + \epsilon v(x)$. Then formally we can write

$$\psi(x,t) = w(x) + \epsilon \tilde{v}(x,t) + O(\epsilon^2) \tag{1.12}$$

$$\tilde{v}(x,0) = v(x).$$

Substituting (1.13) into (1.11) and collecting in powers of ϵ generates the linear evolution equation

$$\tilde{v}_t = DF(w(x))\tilde{v} - T_c\tilde{v} \tag{1.13}$$

$$\tilde{v}(x,0) = v(x)$$

We will henceforth refer to the linear operator $DF(w(x)) - T_c$ as L_w. If the spectrum of L_w, $\sigma(L_w)$, is contained in the closure of the left-hand side of the complex plane, then $u = T(ct)w(x;c)$ is said to be spectrally stable.

While weaker, spectral stability is easier to establish than orbital stability. Unfortunately, spectral stability does not necessarily imply orbital stability. However, in the important case that the evolution equation is Hamiltonian, *i.e.*

$$u_t = JE'(u), \tag{1.14}$$

where J is skew adjoint, and $E'(u)$ is the derivative of the Hamiltonian $E(u)$, one can directly connect spectral stability to orbital stability. Letting J be invertible, and $u = T(ct)w(x; c)$, in the seminal paper [16], the authors develop necessary and sufficient conditions for establishing orbital stability based on properties of the spectrum of the operator $E''(w) - Q''(w)$. $Q(w) = \frac{1}{2} < J^{-1}T_c w, w >$ is a second conserved quantity of the Hamiltonian system generated by the symmetry used to construct the TWS since $T_c = T'(0)c$, $T'(0)$ being the infinitesimal generator of T. Using our previous notation, we have $F(u) = JE'(u)$, thus $DF(u) = JE''(u)$ and $E''(w) - Q''(w) = J^{-1}(DF(w) - T_c) = J^{-1}L_w$. A great deal of work has been done connecting $\sigma(J^{-1}L_w)$ to $\sigma(L_w)$, and it has been shown for a wide class of Hamiltonian problems that spectrally instability implies orbital stability. Further, being able to compute $\sigma(J^{-1}L_w)$ and $\sigma(L_w)$ provides key hints as to how to establish orbital stability beyond the cases covered in [16].

As an explicit example to make the previous terminology more concrete and to show a prominent case of a Hamiltonian system, we again use focusing NLS. If we separate $u = \psi + i\phi$, then we get the PDE

$$\begin{pmatrix} \psi \\ \phi \end{pmatrix}_t = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2}\psi_{xx} + \psi(\phi^2 + \psi^2) & 0 \\ 0 & \frac{1}{2}\phi_{xx} + \phi(\phi^2 + \psi^2) \end{pmatrix} \tag{1.15}$$

For NLS,

$$E(\phi, \psi) = \int \left(-\frac{1}{2}(\phi_x^2 + \psi_x^2) + \frac{1}{4}(\phi^2 + \psi^2)^2 \right) dx, \tag{1.16}$$

and

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{1.17}$$

It is straightforward to show, where T in this case is the phase shift generating (1.6), that

$$Q(u) = -\frac{c}{2} \int |u|^2 dx. \tag{1.18}$$

Choosing u to be (1.6), $E''(w) - Q''(w)$ is

$$
\begin{pmatrix}
\frac{1}{2}\partial_x^2 + 3w^2(x) + c & 0 \\
0 & \frac{1}{2}\partial_x^2 + w^2(x) + c
\end{pmatrix}
\tag{1.19}
$$

which is just $J^{-1}L_w$ for focussing NLS. For (1.6), one can approximate $\sigma(L_w)$ by way of Hill's method (explained in the following sections), which for $k = .8$, produces the plot
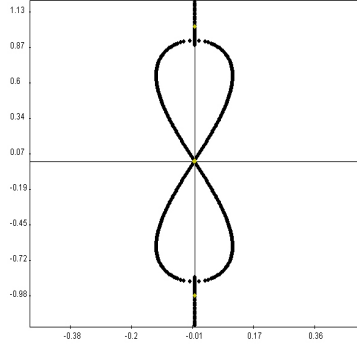
Figure 1.1:

which shows that (1.6) is unstable, both spectrally and thus orbitally.

As for when the given evolution equation is not Hamiltonian, one still attempts to make use of a knowledge of the spectrum of the associated linearization, but no general theorems exist like those in [16]. Instead, one must make recourse to a number of tools from dynamical systems which are usually infinite dimensional versions of the manifold theorems used in finite-dimensional-dynamical systems. Suffice to say though, establishing the spectral stability of solutions to nonlinear problems is the first step any researcher should take when attempting to determine the orbital stability of said particular solution. Thus, this thesis addresses different means at computing the spectrum of linear operators. These are explained in the following sections.

1.2 Hill's Method

Hill's method is used to approximate the spectrum of the operator

$$S_p \psi \equiv \partial_x^p \psi + \sum_{k=0}^{p-1} f_k(x) \partial_x^k \psi, \qquad (1.20)$$

where ψ is in some appropriate space to be defined later. The coefficient functions $f_k(x)$ are smooth, T-periodic functions: $f_k(x + T) = f_k(x)$, $k = 0, \ldots, p - 1$. This is denoted as

$$f_k \in \mathbf{C}^\infty (\mathrm{S}_T). \qquad (1.21)$$

Using Floquet and Fourier theory, our approximation starts by computing a bi-infinite matrix representing a parameter-dependent symbol of S_p. We make the problem finite dimensional by truncating the bi-infinite matrix in both rows and columns; we then compute the eigenvalues of the resulting finite-dimensional matrix. Such an approach is commonly used. This is made more precise in the following section. In modern terminology, this truncation may be called a Galerkin approximation [3], though it is also called a projection method in [5].

In its full generality, Hill's method was first developed in [11]. However, the method has appeared in more specialized contexts as early as 1886, when George Hill published [17]. This paper detailed his investigations into the reduced three-body problem, where an analysis of small perturbations led him to seek solutions to the linear problem

$$\frac{d^2\psi}{dx^2} + \left(\theta_0 + 2 \sum_{n=1}^{\infty} \theta_n \cos(2nx) \right) \phi = 0. \qquad (1.22)$$

Here θ_k, $k = 1, 2, \ldots$, are real parameters. In his analysis, Hill incorporated both Floquet and Fourier theory, which led him to consider infinite-dimensional matrices and their corresponding determinants. Hill used these determinants in a formal way, and he attempted to approximate the spectra of the infinite-dimensional matrices using the spectra of three-by-three truncations. Inspired by Hill's work, a rigorous theory on determinants of infinite matrices was initiated by Poincaré [27] and von Koch [32]. This in turn has led to a modern theory of determinants of operators defined over Banach spaces. The treatise by Gohberg, Goldberg, and Krupnick [14] provides an excellent introduction to both the classical origins

and modern developments of infinite dimensional determinants, and our work relies heavily on the material in [14] (see also [13] and [6]). However, we do not develop this theory any further.

Instead, we focus on proving the validity of Hill's truncation. This problem, in turn, has its own deep and storied history. A wonderful introduction can be found in [5]. Likewise, in the same reference, one can find a number of examples where using finite-dimensional approximations to compute the spectra of infinite-dimensional operators fails spectacularly. For our problem, however, we show that for general S_p, Hill's method never converges to spurious eigenvalues in compact domains. In the case that S_p is self-adjoint, we go further and show, again on any compact domain, that Hill's method converges to the spectrum of S_p restricted to said domain. Further, assuming the convergence of an approximate sequence of eigenvalues to a simple eigenvalue, we show that the corresponding eigenvector approximations converge to a true eigenvector in the L^2-norm.

As shown in [11], Hill's method is exact for constant-coefficient problems. By restricting ourselves to a particular class of self-adjoint operators, which represent the simplest case of non-constant coefficient equations, we show Hill's method approximates the smallest eigenvalue faster than any polynomial power. This restricted class of operators includes classic problems such as Mathieu's equation, and it represents a non-trivial and interesting body of problems for which Hill's method is an excellent approximation scheme.

Another, more abstract but also more general, approach for analyzing Hill's method can be found in the notes of G.M. Vainikko (Chapter 4 of [20]). This approach applies to a more general class of problems than just Hill's method, and once the approach is mastered, its application to Hill's method can be viewed as a corollary. The results in [20] not only allow for establishing the convergence of Hill's method, but the rate of convergence can also be determined. The rate thus found is identical to the one we establish in this paper. In the case of symmetric operators, a convergence proof and rate can also be found in [12]. However, the class of operators considered in [12] is far more restricted than in this paper or [20]. Further, the rate of convergence obtained is far slower than what we or [20] are able to show.

The key to the deeper results in [20] is the notion of the aperture between subspaces

of a Banach space (see also Chapter 4 of [19]). We make no use of this idea, or any other result found in [20]. Instead, a more direct and explicit approach is used, which may be more natural or intuitive if one is interested in Hill's method in its own right, as opposed to regarding it as a special case of a more general problem. Indeed, as mentioned above, Hill's method led to the consideration of determinants of infinite-dimensional operators and the work of [14]. Thus, the methods presented in this paper are new and hopefully insightful.

Remarks.

- The form of the operator (1.20) is restrictive in that we equate the coefficient of the highest-order derivative to one. Were the coefficient a constant, this would not change our results. The affect of a non-constant coefficient on our work is non-trivial. However, in many problems (linear stability, scattering) the spectral problems that arise are of the form used here (see the examples in [11]), although variations occur.

- Numerically computing the eigenvalues of a matrix is a nontrivial problem. It is not a problem we consider in this thesis. The sole interest is in the relation between the finite-dimensional approximations as obtained through Hill's method and the problem they are meant to approximate.

- The work in this paper focuses on spectral problems defined by scalar differential operators (1.20). This restriction is made for ease of presentation. Hill's method, in essence a Galerkin method, works equally well for systems of equations or for problems with multiple independent variables [11]. Our methods of analysis used apply to the system case, but modifications are necessary for the multi-dimensional case.

- Combining the ideas of Floquet decomposition and the truncation of matrix representations of operators is frequently done when considering periodic operator equations. Three contemporary examples of this can be found in [29], [33], and [34]. Special mention should be made with regards to [34].

1.3 Spectral Stability of a Boussinesq Approximation

It has long been a goal of several disciplines to fully understand the Euler surface-water-wave equations, which in scaled form are

$$
\begin{aligned}
\epsilon \Delta \phi + \partial_z^2 \phi &= 0 & 0 \leq z < 1 + \epsilon \eta(x, y) \\
\partial_z \phi &= 0 & z = 0 \\
\partial_t \phi + \eta + \tfrac{1}{2}(\epsilon |\nabla \phi|^2 + |\partial_z \phi|^2) &= 0 & z = 1 + \epsilon \eta(x, y) \\
\partial_t \eta + \epsilon \nabla \phi \cdot \nabla \eta - \tfrac{1}{\epsilon} \partial_z \phi &= 0 & z = 1 + \epsilon \eta(x, y)
\end{aligned}
\tag{1.23}
$$

Here, ϕ is the velocity potential of an incompressible, irrotational fluid (hence Laplace's equation on the interior of the fluid), and η is the height of the wave above the resting state of the fluid, which has been scaled to $z = 1$ in this case. If we denote the unscaled resting height of the fluid as h_0, we have chosen $\epsilon = \frac{a}{h_0}$, where a denotes a characteristic amplitude of the problem. Likewise, we introduce the balance $\frac{\lambda^2}{h_0^2} = \frac{1}{\epsilon}$, where λ is the magnitude of the wave vector in the x and y directions. which will allow us to formally derive long wavelength, shallow water approximations to (1.23).

In this spirit, as in the derivation of KdV (see [1]), we begin by noting that since ϕ must be harmonic on the interior of the fluid, we expand ϕ in z such that

$$
\phi(x, y, z) = \sum_{n=0}^{\infty} \phi_n(x, y) z^n.
\tag{1.24}
$$

Using the boundary condition at the bottom of the fluid, and substituting our expansion for ϕ into the scaled Laplace equation of (1.23) and matching powers of z gives

$$
\phi(x, y, z) = \phi_0(x, y) - \frac{\epsilon}{2} \Delta \phi_0 z^2 + \mathcal{O}(\epsilon^2).
\tag{1.25}
$$

Using this expansion, the equations defined at the surface in (1.23) become

$$
\begin{aligned}
\eta_t + \nabla \cdot \mathbf{w} + \epsilon (\nabla \cdot (\eta \mathbf{w}) - \tfrac{1}{6} \Delta \nabla \cdot \mathbf{w}) &= \mathcal{O}(\epsilon^2) \\
\mathbf{w}_t + \nabla \eta + \epsilon(\tfrac{1}{2} \nabla |\mathbf{w}|^2 - \tfrac{1}{2} \Delta \mathbf{w}_t) &= \mathcal{O}(\epsilon^2),
\end{aligned}
\tag{1.26}
$$

where $\mathbf{w} = \nabla \phi_0$. We now introduce the horizontally scaled velocity $\mathbf{v}$ given by

$$
\mathbf{v} = \mathbf{w} - \frac{\epsilon \theta^2}{2} \Delta \mathbf{w} + \mathcal{O}(\epsilon^2)
\tag{1.27}
$$

where $0 \leq \theta \leq 1$. Formally inverting this equality gives to the order of our approximation

$$\mathbf{w} = \mathbf{v} + \frac{\epsilon\theta^2}{2}\Delta\mathbf{v} + \mathscr{O}(\epsilon^2) \tag{1.28}$$

Substituting this into (1.26), we get

$$\begin{aligned}
\eta_t + \nabla \cdot \mathbf{v} + \epsilon(\nabla \cdot (\eta\mathbf{w})) + \frac{\epsilon}{2}(\theta^2 - \frac{1}{3})\Delta\nabla \cdot \mathbf{w} &= \mathscr{O}(\epsilon^2) \\
\mathbf{v}_t + \frac{\epsilon\theta^2}{2}\Delta\mathbf{v}_t + \nabla\eta + \epsilon(\frac{1}{2}\nabla|\mathbf{v}|^2 - \frac{1}{2}\Delta\mathbf{v}_t) &= \mathscr{O}(\epsilon^2),
\end{aligned} \tag{1.29}$$

To the first order, we have

$$\begin{aligned}
\eta_t + \nabla \cdot \mathbf{v} &= \mathscr{O}(\epsilon) \\
\mathbf{v}_t + \nabla\eta &= \mathscr{O}(\epsilon),
\end{aligned} \tag{1.30}$$

and thus we can introduce the formal identities

$$\begin{aligned}
\Delta\nabla \cdot \mathbf{v} &= \delta\Delta\nabla \cdot \mathbf{v} - (1-\delta)\Delta\eta_t + \mathscr{O}(\epsilon) \\
\Delta\mathbf{v}_t &= -\mu\Delta\nabla\eta + (1-\mu)\Delta\mathbf{v}_t + \mathscr{O}(\epsilon).
\end{aligned} \tag{1.31}$$

If we let

$$\begin{aligned}
a &= \tfrac{1}{2}(\theta^2 - \tfrac{1}{3})\delta & b &= \tfrac{1}{2}(\theta^2 - \tfrac{1}{3})(1-\delta) \\
c &= \tfrac{1}{2}(1-\theta^2)\mu & d &= \tfrac{1}{2}(1-\theta^2)(1-\mu),
\end{aligned} \tag{1.32}$$

and introduce the scalings

$$t = \frac{t}{\sqrt{\epsilon}}, \ (x,y) = (\frac{x}{\sqrt{\epsilon}}, \frac{y}{\sqrt{\epsilon}}), \ \eta = \frac{\eta}{\epsilon}, \ \mathbf{w} = \frac{\mathbf{w}}{\epsilon}, \tag{1.33}$$

and drop all higher order in epsilon terms, we get the following approximation to (1.23)

$$\begin{aligned}
\eta_t + \nabla \cdot \mathbf{v} + \nabla \cdot \eta\mathbf{v} + a\nabla \cdot \Delta\mathbf{v} - b\Delta\eta_t &= 0 \\
\mathbf{v}_t + \nabla\eta + \frac{1}{2}\nabla|\mathbf{v}|^2 + c\nabla(\Delta\eta) - d\Delta\mathbf{v}_t &= 0.
\end{aligned} \tag{1.34}$$

The derivation presented here is a summary of that presented in [8] and [4]. There are several interesting aspects of (1.34) that make it an exciting new approximation to (1.23). As pointed out in [8], by varying the coefficients a, b, c, and d, it appears one can capture a variety of phenomena usually associated with separate approximations to the Euler equation. An excellent example of this in practice can be found in [7], where both the second Kadomstev-Petviashvili (KP-II) equation and an associated analog using Benjamin-Bona-Mahoney (BBM) instead of KdV is derived from (1.34). Thus from [7], we also trivially get BBM and KdV by restricting to one dimension.

Another remarkable aspect of (1.34) is that, at least for the parameter values that interest us in this thesis, not only is (1.34) locally well posed, but solutions of (1.34) will uniformly approximate solutions of (1.23) to $\mathscr{O}(\epsilon)$ on timescales of $\mathscr{O}\left(\dfrac{1}{\epsilon}\right)$ [4]. Puzzling though is that (1.34) is only known to be Hamiltonian in the case that $b = d$. It is not known whether or not (1.34) is Hamiltonian in general. This is an intriguing question given that it is well known that the Euler-water-wave equations are Hamiltonian. Thus if it were the case that (1.34) were not Hamiltonian, it would provide an example where formal multiple scale arguments destroyed the Hamiltonian structure of the parent problem, but which had no bearing on the problem being a valid approximation to said parent problem. It is worth noting that a conserved quantity is found in [4] for cases where $b \neq d$, so perhaps mathematically it is not troublesome to approximate one conservative system with another. However, physically, Hamiltonicity is considered so important that one would not believe non-Hamiltonian versions of (1.34) could ever be a useful model of (1.23).

This issue is not addressed in this thesis. Instead, we attempt to prove the spectral stability of sech^2 solutions to (1.34). In particular we look at solutions of the form

$$\eta = \eta_0 \text{sech}^2 (\lambda \xi) , \tag{1.35}$$

$$v_1 = \eta_0 \sqrt{\frac{3}{\eta_0 + 3}} \text{sech}^2 (\lambda \xi) \tag{1.36}$$

$$v_2 = 0 \tag{1.37}$$

This is a very natural place to begin assessing the phenomenological accuracy of (1.34) given the prominence that the sech^2 profile has in the subject of water waves. It is known that these profiles are stable for both KdV and KP-II, and there is strong experimental evidence (Hammack and Segur) that they describe the profile of surface waves in shallow water. Thus establishing the stability of such solutions to (1.34) is another key step in determining the Boussinesq approximations' validity as a physical model.

However, given the complexity of (1.34), spectral stability is the best that the author can do at present. Even in the case that (1.34) is known to be Hamiltonian, one does not have direct recourse to the results of [16] and its descendants. Analytically, the author can only establish spectral stability with respect to one-dimensional perturbations. Numerically,

we can go much further, and further we also look at particular solutions that look like depression "solitons" and multi "solitons". Note, we use quotes around the word soliton since technically a soliton is a solution to a nonlinear PDE obtained by way of the inverse scattering transform which in turn relies upon the existence of a Lax pair representation of said PDE. Thus no solution of (1.34) is known to be an actual soliton. Soliton is by now though just a quick way to say sech profile.

Chapter 2

HILL'S METHOD

In essence Hill's method combines a Floquet (or Bloch) decomposition with a Fourier expansion so as to reduce the numerical computation of the spectrum of a periodic differential operator to the computation of spectra of a family of (finite-dimensional) matrices. Before continuing, some relevant spaces that will be used throughout the rest of this thesis are defined. Let $L_2\left(S_T\right)$ be defined as the completion of $C\left(S_T\right)$, the space of T-periodic, continuous functions, with respect to the L_2 norm on the interval $[-T/2, T/2]$. Let

$$e_n\left(x\right) = \frac{e^{-i2\pi nx/T}}{\sqrt{T}}, \quad n \in \mathbb{Z}, \tag{2.1}$$

so that for $\phi \in L_2\left(S_T\right)$, we have the associated Fourier series

$$\phi\left(x\right) = \sum_{n=-\infty}^{\infty} \hat{\phi}_n e_n\left(x\right), \tag{2.2}$$

with

$$\hat{\phi}_n = \langle \phi, e_n \rangle = \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} \phi(x) e_n^*(x) dx. \tag{2.3}$$

This allows us to associate with every function $\phi \in L_2\left(S_T\right)$ its Fourier transform

$$\hat{\phi} \equiv \left\{ \hat{\phi}_n \right\}_{n=-\infty}^{\infty}. \tag{2.4}$$

We define the Sobolev spaces $H^p\left(S_T\right)$ in a similar fashion, and in our paper, we define the norm on $H^p\left(S_T\right)$ as ([3], pg. 308)

$$||\phi||_{2,p}^2 \equiv \left| \hat{\phi}_0 \right|^2 + \sum_{|k|>0} \left(\frac{2\pi k}{T} \right)^{2p} \left| \hat{\phi}_k \right|^2. \tag{2.5}$$

2.1   The Floquet-Bloch Decomposition

First, we define S_p over the Sobolev space $\mathrm{H}^p(\mathbb{R})$, with

$$\mathrm{H}^p(\mathbb{R}) = \left\{ f \in \mathrm{L}_2(\mathbb{R}) \mid \sum_{k=0}^{p} \int_{\mathbb{R}} |f^k(x)|^2 dx < \infty \right\}, \tag{2.6}$$

where f^k denotes the kth weak derivative of f. This makes S_p closed and densely defined. We can likewise turn the operator $S_p - \lambda$ into a first-order differential operator defined on $\mathrm{H}^1(\mathbb{R}; \mathbb{C}^p)$, where the notation means that the space $\mathrm{H}^1(\mathbb{R}; \mathbb{C}^p)$ consists of $\mathbb{C}^p$ valued functions with one weak derivative, and for which the function and its derivate have $\mathbb{C}^p$ norms that are both in $\mathrm{L}_2(\mathbb{R})$ (see [28] for more details). Denote the first-order differential operator as $S(x; \lambda) = \frac{d}{dx} - B(x; \lambda)$, where $B(x; \lambda)$ is a $p \times p$ matrix. By definition,

$$\sigma(S_p) = \{\lambda \in \mathbb{C} : S(x; \lambda) \;\; does \;\; not \;\; have \;\; a \;\; bounded \;\; inverse\}. \tag{2.7}$$

Following [30], we use the following decomposition of $\sigma(S_p)$ (see also [9], [24]).

- $\sigma_{pt}(S_p) = \{\lambda \in \mathbb{C} : S(x; \lambda) \;\; is \;\; Fredholm \;\; with \;\; zero \;\; index.\}$

- $\sigma_{ess}(S_p) = \sigma(S_p) \backslash \sigma_{pt}(S_p)$.

Since S_p has only periodic coefficients, we need only compute $\sigma_{ess}(S_p)$ [30]. This reduces to the following problem.

Theorem 1. $\lambda \in \sigma(S_p)$ *if and only if the differential equation*

$$\frac{du}{dx} = B(x; \lambda)u, \; 0 < x < T$$

$$\tag{2.8}$$

$$u(T) = e^{i\mu T} u(0)$$

has a solution for some $\mu \in [0, 2\pi/T)$.

Proof. See [30], page 1001. □

We transform the differential equation in Theorem 1 into

$$\frac{d\psi}{dx} = \tilde{B}(x; \lambda, \mu)\psi, \; 0 < x < T$$

$$\tag{2.9}$$

$$\psi(T) = \psi(0)$$

via the transformation $\psi(x) = e^{-i\mu x} u(x)$. Note, $\tilde{B}(x; \lambda, \mu) = B(x; \lambda) - i\mu$. We can then restate Theorem 1 as

Theorem 2. $\lambda \in \sigma(S_p)$ *if and only if the differential equation*

$$\frac{d\psi}{dx} = \tilde{B}(x; \lambda, \mu)\psi, \quad \psi \in H^1\left(S_T; \mathbb{C}^p\right) \tag{2.10}$$

has a solution for some $\mu \in [0, 2\pi/T)$.

It is easy to show that the pth-order system in Theorem 2 is equivalent to the scalar problem

$$S_p^\mu \phi = \lambda \phi, \quad \phi \in \mathrm{H}^p\left(\mathrm{S}_T\right), \tag{2.11}$$

where

$$S_p^\mu \phi = \mathrm{e}^{-i\mu x} S_p\left(\mathrm{e}^{i\mu x}\phi\right). \tag{2.12}$$

An explicit form for S_p^μ is found in [11]. Theorem 2 implies that we can write $\sigma(S_p)$ as

$$\sigma(S_p) = \bigcup_\mu \sigma(S_p^\mu). \tag{2.13}$$

As implied by (2.11), for each value of μ, $\sigma(S_p^\mu)$ consists only of point spectra. We approximate these point sets numerically for a fixed value of μ.

2.2   The Fourier Decomposition

To reduce the problem to linear algebra, we resort to a Galerkin method [3] using the orthonormal basis e_n given at the beginning of this section. Of course, given any orthonormal basis $\{\varphi_j\}$, we can generate a matrix representation for any linear operator M with entries $\langle M\varphi_j, \varphi_k \rangle$, $(j, k) \in \mathbb{Z}^2$. Our particular choice of basis reflects the boundary conditions of our eigenvalue problem (2.11). We interchangeably refer to the bi-infinite matrix, with entries $\langle S_p^\mu \mathrm{e}_j, \mathrm{e}_k \rangle$, as the Fourier transform or symbol of the linear operator S_p^μ. We denote the symbol (or Fourier transform, or bi-infinite matrix representation) of S_p^μ as $\hat{S}_p^\mu$, where the $(\text{n,m})^{th}$ entry of $\hat{S}_p^\mu$ is denoted by $\hat{S}_{p,nm}^\mu = \langle S_p^\mu \mathrm{e}_m, \mathrm{e}_n \rangle$. We write the Fourier transform of our eigenvalue problem (2.11) as

$$\hat{S}_p^\mu \hat{\phi} = \lambda \hat{\phi}. \tag{2.14}$$

2.3   Finite-Dimensional Projection

The last step of Hill's method requires the introduction of the orthogonal projection operator P_N onto the subspace spanned by the Fourier modes from $-N$ to N. The effect of P_N applied to a periodic function is truncation of the Fourier series *i.e.*

$$P_N \phi(x) = \sum_{n=-N}^{N} \hat{\phi}_n e_n(x). \tag{2.15}$$

Likewise, the action of the symbol of P_N, $\hat{P}_N$, will give

$$\left(\hat{P}_N \hat{\phi}\right)_n = \begin{cases} 0 & |n| > N \\ \\ \hat{\phi}_n & |n| \le N \end{cases}. \tag{2.16}$$

Define the $(2N+1) \times (2N+1)$ matrix $\hat{S}_N^{\mu;\tau}$ via

$$\hat{P}_N \hat{S}_p^\mu \hat{P}_N = \begin{pmatrix} \ddots & & \vdots & & \iddots \\ & 0 & 0 & 0 & \\ \cdots & 0 & \hat{S}_N^{\mu;\tau} & 0 & \cdots \\ & 0 & 0 & 0 & \\ \iddots & & \vdots & & \ddots \end{pmatrix}, \tag{2.17}$$

where the τ emphasizes that $\hat{S}_N^{\mu;\tau}$ is a truncation of a bi-infinite matrix. As a matter of

convention, for any operator A with symbol $\hat{A}$, we define $\hat{A}_N^\tau$ in the same fashion, namely

$$
\hat{P}_N \hat{A} \hat{P}_N = \begin{pmatrix} \ddots & & \vdots & & \iddots \\ & 0 & 0 & 0 & \\ \cdots & 0 & \hat{A}_N^\tau & 0 & \cdots \\ & 0 & 0 & 0 & \\ \iddots & & \vdots & & \ddots \end{pmatrix}.
\tag{2.18}
$$

Likewise we introduce the shorthand $\hat{A}_N = \hat{P}_N \hat{A} \hat{P}_N$.

Finally, we define the approximate eigenvalue problem

$$
\hat{S}_N^{\mu,\tau} \hat{\phi}_N^\tau = \lambda_N \hat{\phi}_N^\tau,
\tag{2.19}
$$

where the subscript N on λ_N reinforces the order of the approximation. A more detailed derivation is presented in [11].

Chapter 3

PROOF OF CONVERGENCE

By the convergence of Hill's method, we mean that the following two properties are satisfied.

1. For a given sequence $\{\lambda_N\}_{N=1}^{\infty}$, $\lambda_N \in \sigma(\hat{S}_N^{\mu,\tau})$, and for any $\epsilon > 0$, there exists an integer M such that any λ_N, $N \geq M$, is in an ϵ-neighborhood of some $\lambda \in \sigma(S_p^{\mu})$.

2. For all $\lambda \in \sigma(S_p^{\mu})$, there exists some sequence $\{\lambda_N\}_{N=1}^{\infty}$, $\lambda_N \in \sigma(\hat{S}_{pN}^{\mu})$, such that $\lambda_N \to \lambda$.

The first condition ensures that Hill's method is accurate, but it leaves open the possibility that the method may not produce all of $\sigma(S_p^{\mu})$. Likewise, the second statement ensures that the method will faithfully reproduce all of $\sigma(S_p^{\mu})$, but it does not rule out that the method will produce spurious information. It is this distinction that leads us to refer to the first condition as the "no-spurious modes" condition.

We are able to prove a slightly restricted version of the no-spurious modes condition for any operator S_p^{μ}. We modify the condition only by requiring the arbitrary sequence $\{\lambda_N\}_{N=1}^{\infty}$ to be confined to a compact subset of the complex plane. The second condition is essentially proved in [28], for self-adjoint operators. We have not been able to improve upon this restriction. However, we present the outline of the proof provided in [28] for the sake of completeness.

3.1   Proof of the No-Spurious-Modes Condition

Our proof of the first condition relies upon one major theorem. Before proving this theorem, we need to develop and explain the basic machinery necessary for our proof. First, for notational ease, we define the operator S^1

- $\mathrm{D}\left(S^1\right) = \mathrm{H}^p\left(\mathrm{S}_T\right), \ \ S^1\phi = S_p^\mu \phi.$

We now provide a brief introduction to the theory of determinants of operators on a separable Hilbert Space, say H. This material was developed in [14], and we reproduce it here only for completeness or to clarify some points made in [14]. Let $\mathrm{B}\,(\mathrm{H})$ denote the space of all bounded operators from H into itself. Let F denote the space of finite-rank operators. For our purposes, it is not sufficient to use the operator norm induced by the norm on H, say $||\cdot||$. Instead, we need to introduce a new norm $||\cdot||_{\mathrm{Z}}$, where Z denotes a sub-algebra of $\mathrm{B}\,(\mathrm{H})$ such that $\mathrm{F} \cap \mathrm{Z}$ is dense in Z and

$$||\cdot|| \leq \mathrm{C}\,||\cdot||_{\mathrm{Z}}, \tag{3.1}$$

where C is a constant. Thus Z is an embedded sub-algebra in $\mathrm{B}\,(\mathrm{H})$. Likewise, if the space of finite-rank operators is dense in Z, this implies every element in Z is compact. Next, define the trace of $K \in \mathrm{F} \cap \mathrm{Z}$ by

$$\mathrm{tr}\,(K) = \sum_{k=1}^{n} \lambda_k, \tag{3.2}$$

and define the determinant of $I + K$ as

$$\det\,(I + K) = \prod_{k=1}^{n}\left(1 + \lambda_k\right), \tag{3.3}$$

where n is the rank of K and λ_k are the eigenvalues of K.

The issue at hand is whether we can find some continuous function that will serve as an extension of the determinant, which has only been defined on $\mathrm{F} \cap \mathrm{Z}$. A necessary and sufficient condition for this (see [14]) is if the trace is a bounded linear functional in the Z norm, *i.e.*

$$|\mathrm{tr}\,(K)| \leq \mathrm{M}\,||K||_{\mathrm{Z}} \tag{3.4}$$

holds for all $K \in \mathrm{F} \cap \mathrm{Z}$, where M is a constant independent of K. If this condition holds, then for $K \in \mathrm{Z}$, we know there exists a sequence of finite-rank operators K_N such that

$$\lim_{N \to \infty} ||K_N - K||_{\mathrm{Z}} = 0, \tag{3.5}$$

and we can define the Z-determinant of $I + K$ as

$$\det{}_{\mathrm{Z}}\,(I + K) = \lim_{N \to \infty} \det\,(I + K_N). \tag{3.6}$$

Using the above definitions, one can prove [14]:

Theorem 3. $(I+K)^{-1}$ *exists if and only if* $det_Z(I+K) \neq 0$.

A space well suited for our purposes was developed by Gohberg *et al* [14]. Define the sub-algebra Ω via:

$$\Omega \equiv \left\{ A \in B\left(L_2\left(S_T\right)\right) : \max\left(\lim_{M \to \infty} \left| \sum_{n=-M}^{M} \hat{A}_{nn} \right|, \left(\sum_{n,m=-\infty}^{\infty} \left| \hat{A}_{nm} \right|^2 \right)^{1/2} \right) < \infty \right\}. \quad (3.7)$$

For $A \in \Omega$, we have the corresponding norm $||A||_\Omega$ defined by

$$||A||_\Omega \equiv \max\left(\left| \lim_{M \to \infty} \sum_{n=-M}^{M} \hat{A}_{nn} \right|, \left(\sum_{n,m=-\infty}^{\infty} \left| \hat{A}_{nm} \right|^2 \right)^{1/2} \right), \quad (3.8)$$

where $\hat{A}_{nm} = \langle A e_m, e_n \rangle$, and e_n is defined as in (2.1). If $A \in \Omega$, we see that

$$||A\phi||_2^2 = \sum_{n=-\infty}^{\infty} \left| \sum_{m=-\infty}^{\infty} \hat{A}_{nm} \hat{\phi}_m \right|^2 \leq ||\phi||_2^2 \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \left| \hat{A}_{nm} \right|^2, \quad (3.9)$$

and therefore $||A||_2 \leq ||A||_\Omega$. The next lemma easily follows.

Lemma 4. *For all finite rank operators* A, *we have*

$$|tr(A)| \leq ||A||_\Omega. \quad (3.10)$$

Proof. Given A has finite rank, let $\psi_1, \cdots, \psi_n$ be an orthonormal basis for the range of A. Then we may write A as

$$A = \sum_{k=1}^{n} \psi_k \langle A \cdot, \psi_k \rangle. \quad (3.11)$$

It is clear that

$$tr\left(\psi_k \langle A \cdot, \psi_k \rangle\right) = \langle A\psi_k, \psi_k \rangle, \quad (3.12)$$

and therefore

$$tr(A) = \sum_{k=1}^{n} \langle A\psi_k, \psi_k \rangle. \quad (3.13)$$

Since $\psi_k \in L_2(S_T)$, we can expand ψ_k as

$$\psi_k = \sum_{j=-\infty}^{\infty} \hat{\psi}_{kj} e_j. \quad (3.14)$$

Note that

$$\langle Ae_j, e_j \rangle = \sum_{k=1}^{n} \hat{\psi}_{kj} \langle Ae_j, \psi_k \rangle. \tag{3.15}$$

Therefore we can rewrite the trace of a finite-rank operator as

$$\begin{aligned}
\text{tr}\,(A) &= \sum_{k=1}^{n} \langle A\psi_k, \psi_k \rangle \\
&= \sum_{k=1}^{n} \sum_{j=-\infty}^{\infty} \hat{\psi}_{kj} \langle Ae_j, \psi_k \rangle \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=1}^{n} \hat{\psi}_{kj} \langle Ae_j, \psi_k \rangle \\
&= \sum_{j=-\infty}^{\infty} \langle Ae_j, e_j \rangle. \tag{3.16}
\end{aligned}$$

Thus from the definition of $||\cdot||_\Omega$ we have the result. $\qquad\square$

Similarly, we have

Lemma 5. *Every finite-rank operator is in Ω.*

Proof. We know that the trace of a finite-rank operator, say A, is bounded, and thus from the previous lemma we know that $\left| \sum_{j=-\infty}^{\infty} \hat{A}_{jj} \right| < \infty$. Likewise we have

$$\begin{aligned}
\sum_{j,k} |\langle Ae_j, e_k \rangle|^2 &= \sum_{j,k} \left| \sum_{l=1}^{n} \hat{\psi}_{lk} \langle Ae_j, \psi_l \rangle \right|^2 \\
&\leq \left(\sum_k \sum_{l=1}^{n} \left| \hat{\psi}_{lk} \right|^2 \right) \left(\sum_j \sum_{l=1}^{n} \left| \langle e_j, A^\dagger \psi_l \rangle \right|^2 \right) \\
&\leq n \sum_{l=1}^{n} \left| \left| A^\dagger \psi_l \right| \right|_2^2 < \infty, \tag{3.17}
\end{aligned}$$

where $A^\dagger$ denotes the adjoint of A. Therefore every finite-rank operator is in Ω. $\qquad\square$

Lastly, we need to show is that F is dense in Ω. We do this, and also establish a useful result for our main theorem, in the following lemma.

Lemma 6. *If $A \in \Omega$, then $\lim_{N \to \infty} ||A - P_N A P_N||_\Omega = 0$.*

Proof. It is clear that

$$\|A - P_N A P_N\|_\Omega \leq \|(I - P_N)A\|_\Omega + \|P_N A(I - P_N)\|_\Omega. \tag{3.18}$$

With

$$\|(I - P_N)A\|_\Omega = \max\left(\left|\sum_{|n|>N} \hat{A}_{nn}\right|, \left(\sum_{|n|>N} \sum_{m=-\infty}^{\infty} \left|\hat{A}_{nm}\right|^2\right)^{1/2}\right), \tag{3.19}$$

and

$$\|P_N A(I - P_N)\|_\Omega^2 = \sum_{|n|\leq N} \sum_{|m|>N} \left|\hat{A}_{nm}\right|^2, \tag{3.20}$$

the result follows, since $\|A\|_\Omega < \infty$. $\qquad\square$

This shows that for $A \in \Omega$, we have

$$\det{}_\Omega (I + A) = \lim_{N\to\infty} \det (I + P_N A P_N), \tag{3.21}$$

where

$$\det (I + P_N A P_N) \equiv \det \left(\hat{I}_N^\tau + \hat{A}_N^\tau\right). \tag{3.22}$$

For omitted proofs and more detail on this material the interested reader is advised to consult [14].

Finally, we need two key facts about operators of the form $I + K$, where I is the identity and K is compact.

- $I + K$ is a Fredholm operator.

- $i(I + K) = 0$.

Note, for any Fredholm operator F,

$$i(F) \equiv \dim(\ker(F)) - \dim(\ker(F^\dagger)), \tag{3.23}$$

where $F^\dagger$ again denotes the adjoint of F. For proof, see [31].

With these tools in hand, we prove the following theorem. This theorem will be the engine to drive the proof of the no-spurious-mode condition.

Theorem 7. *Let $\gamma \in \rho(S^1)$. Then there exists some constant M_γ such that for $N \geq M_\gamma$, $\gamma \in \rho(\hat{S}_N^{1;\tau})$.*

Proof. Define the operator $B : \mathrm{L}_2(\mathrm{S}_T) \to \mathrm{L}_2(\mathrm{S}_T)$ via

$$(\hat{B}\hat{\psi})_n = \begin{cases} \left(\frac{2\pi ni}{T}\right)^{-p} \hat{\psi}_n & n \neq 0 \\ \\ i^{-p}\hat{\psi}_0 & n = 0 \end{cases}, \tag{3.24}$$

where $\hat{\psi}_n$ are the components of the vector $\hat{\psi} \in l_2$, and $\hat{B}$ is the symbol of B. B, when applied to $S^1 - \gamma$, is introduced to nullify the growth along the diagonal of $\hat{S}^1 - \gamma$. Clearly $\mathrm{H}^p(\mathrm{S}_T) \subset \mathrm{R}(B)$. With $\gamma \in \rho(S^1)$, we have that $S^1 - \gamma$ is a bijection from $\mathrm{H}^p(\mathrm{S}_T)$ to $\mathrm{L}_2(\mathrm{S}_T)$ by definition. Therefore, we define the operator A whose symbol is $\hat{B}(\hat{S}^1 - \gamma)$, noting that $\mathrm{H}^p(\mathrm{S}_T) \subset \mathrm{R}(A)$.

Now consider computing the matrix product of $\hat{B}$ and $(\hat{S}^1 - \gamma)$. Clearly this operator is the extension of $\hat{A}$, and we will show that it is a bounded operator on l_2. Therefore, it must be the unique bounded extension of $\hat{A}$ [28]. We refer to the extension of $\hat{A}$ as $\hat{A}$ to economize on notation. Given that δ_{nm} is the Kroenecker delta function, the terms of $\hat{A}$ are then

$$\hat{A}_{nm} = \begin{cases} \left(\frac{2\pi ni}{T}\right)^{-p}\left(\left(i^p\left(\mu + \frac{2\pi n}{T}\right)^p - \gamma\right)\delta_{nm} + \sum_{k=0}^{p-1}\hat{f}_{k,n-m}i^k\left(\mu + \frac{2\pi n}{T}\right)^k\right) & n \neq 0 \\ \\ (\mu^p - \gamma)\delta_{0m} + \sum_{k=0}^{p-1}\hat{f}_{k,-m}\mu^k i^{k-p} & n = 0 \end{cases}. \tag{3.25}$$

See [11], equation 17, for an explicit derivation. Therefore, for $n \neq 0$

$$\hat{A}_{nn} = 1 + \frac{T}{2\pi}\left(p\mu - i\hat{f}_{p-1,0}\right)\frac{1}{n} + \mathrm{O}\left(\frac{1}{n^2}\right), \tag{3.26}$$

which shows

$$\lim_{M\to\infty}\sum_{n=-M}^{M}\left(\hat{A}_{nn} - 1\right) < \infty. \tag{3.27}$$

Likewise, we have also shown

$$\sum_{n=-\infty}^{\infty}\left|\hat{A}_{nn} - 1\right|^2 < \infty. \tag{3.28}$$

For $n \neq m$ and $n \neq 0$, we have

$$\left| \hat{A}_{nm} \right|^2 \leq \left(\frac{T}{2\pi n} \right)^{2p} \left(\sum_{k=0}^{p-1} \left| \hat{f}_{k,n-m} \right| \left| \mu + \frac{2\pi n}{T} \right|^k \right)^2$$

$$\leq \left(\sum_{k=0}^{p-1} \left| \hat{f}_{k,n-m} \right|^2 \right) \left(\sum_{k=0}^{p-1} \left(\frac{T}{2\pi n} \right)^{2p} \left| \mu + \frac{2\pi n}{T} \right|^{2k} \right), \qquad (3.29)$$

while for $n = 0$ we have

$$\left| \hat{A}_{0m} \right|^2 \leq \frac{\mu^{2p} - 1}{\mu^2 - 1} \sum_{k=0}^{p-1} \left| \hat{f}_{k,-m} \right|^2. \qquad (3.30)$$

Therefore

$$\sum_{m \neq n, |n| > 0} \left| \hat{A}_{nm} \right|^2 \leq \left(\sum_{|n| > 0} \sum_{k=0}^{p-1} \left(\frac{T}{2\pi n} \right)^{2p} \left| \mu + \frac{2\pi n}{T} \right|^{2k} \right) \left(\sum_{m=-\infty}^{\infty} \sum_{k=0}^{p-1} \left| \hat{f}_{k,m} \right|^2 \right). \qquad (3.31)$$

The above shows that

$$\sum_{n,m} \left| \hat{A}_{nm} - \delta_{nm} \right|^2 < \infty, \qquad (3.32)$$

and therefore $A - I \in \Omega$. Let $K = A - I$, and so K is compact. It is then clear that $A \in \mathrm{B}\left(\mathrm{L}_2\left(\mathrm{S}_T\right)\right)$, and that A is Fredholm. Therefore the range of A is closed. We know $\mathrm{H}^p\left(\mathrm{S}_T\right) \subset \mathrm{R}\left(A\right)$, $\mathrm{H}^p\left(\mathrm{S}_T\right)$ is dense in $\mathrm{L}_2\left(\mathrm{S}_T\right)$, and so together these facts imply $\mathrm{R}\left(A\right) = \mathrm{L}_2\left(\mathrm{S}_T\right)$. Hence $\dim\left(\ker\left(A^\dagger\right)\right) = 0$, and $i\left(A\right) = 0$, so $\dim\left(\ker\left(A\right)\right) = 0$. Therefore A is a bounded bijection from $\mathrm{L}_2\left(\mathrm{S}_T\right)$ to $\mathrm{L}_2\left(\mathrm{S}_T\right)$, which means A has a bounded inverse by the Open Mapping Theorem.

Knowing that A has a bounded inverse and that $A \in \Omega$, it follows from Theorem 3 that $\det_\Omega\left(A\right) \neq 0$. We have

$$\det_\Omega\left(A\right) = \lim_{N \to \infty} \det\left(I + P_N K P_N\right)$$

$$= \lim_{N \to \infty} \det\left(\hat{I}_N^\tau + \hat{K}_N^\tau\right), \qquad (3.33)$$

and thus there exists constant M_γ such that for $N \geq M_\gamma$, $\det\left(\hat{I}_N^\tau + \hat{K}_N^\tau\right) \neq 0$. Since $\hat{P}_N \hat{B} = \hat{B} \hat{P}_N$,

$$\hat{A}_N^\tau = \hat{B}_N^\tau(\hat{S}_N^{1,\tau} - \gamma \hat{I}_N), \qquad (3.34)$$

which means that $\hat{B}_N^T(\hat{S}_N^{1,\tau} - \gamma \hat{I}_N)$ has trivial kernel. Since $\hat{B}_N$ has trivial kernel, we know that

$$\ker\left(\hat{S}_N^{1,\tau} - \gamma \hat{I}_N\right) = \{0\}, \qquad (3.35)$$

and therefore $\gamma \in \rho(\hat{S}_N^{1;\tau} - \gamma \hat{I}_N)$ for $N \geq M_\gamma$. $\qquad\qquad\square$

Given this theorem, we prove the following corollary.

Corollary 8. *If* $\lambda_{N_j} \in \sigma(\hat{S}_{N_j}^{1,\tau})$ *and* $\lambda_{N_j} \to \gamma$, *then* $\gamma \in \sigma(S^1)$.

Proof. Suppose in contradiction that $\gamma \in \rho(S^1)$. Then, by Theorem 7, we know for some value M that $\gamma \in \rho(\hat{S}_N^{1;\tau})$ for $N \geq M$. Then

$$
\begin{aligned}
\left\|(\hat{S}_N^{1;\tau} - \gamma)^{-1}\right\|_2 &= \left\|(\hat{B}_N^\tau(\hat{S}_N^{1;\tau} - \gamma))^{-1}\hat{B}_N^\tau\right\|_2 \\
&\leq \left\|(\hat{B}_N^\tau(\hat{S}_N^{1;\tau} - \gamma))^{-1}\right\|_2 \left\|\hat{B}_N^\tau\right\|_2 \\
&\leq \left\|(\hat{B}_N^\tau(\hat{S}_N^{1;\tau} - \gamma))^{-1}\right\|_2 \left\|\hat{B}\right\|_2.
\end{aligned}
\tag{3.36}
$$

Following the notation in Theorem 7, $\hat{B}_N^T(\hat{S}_N^{1;\tau}-\gamma) = \hat{I}_N + \hat{K}_N^\tau$. Likewise, per our convention, let $\hat{K}_N$ denote the l_2 operator such that $\hat{K}_N^\tau$ is the $(2N+1)\times(2N+1)$ truncation of $\hat{K}_N$, and $\hat{K}_N = \hat{P}_N \hat{K}_N \hat{P}_N = \hat{P}_N \hat{K} \hat{P}_N$. From Theorem 7, we know that K is compact, and therefore $\hat{K}_N$ converges to K in the uniform operator topology. Clearly

$$
\left\|(\hat{I}_N^\tau + \hat{K}_N^\tau)^{-1}\right\|_2 \leq \left\|(\hat{I} + \hat{K}_N)^{-1}\right\|_2,
\tag{3.37}
$$

and we know that $I + K$ has a bounded inverse. This implies

$$
\left\|(\hat{I} + \hat{K}_N)^{-1}\right\|_2 \leq \left\|(\hat{I} + \hat{K})^{-1}\right\|_2 \left\|(\hat{I} + (\hat{I} + \hat{K})^{-1}(\hat{K}_N - \hat{K}))^{-1}\right\|_2
\tag{3.38}
$$

Since $\hat{K}_N$ converges uniformly to $\hat{K}$, there exists L such that $\left\|(\hat{I} + \hat{K})^{-1}(\hat{K}_N - \hat{K})\right\|_2 < 1/2$ for $N \geq L$, and therefore

$$
\left\|(\hat{I} + (\hat{I} + \hat{K})^{-1}(\hat{K}_N - \hat{K}))^{-1}\right\|_2 \leq 2.
\tag{3.39}
$$

Finally, we know that

$$
\left\|(\hat{S}_N^{1;\tau} - \gamma)^{-1}\right\|_2 \geq \frac{1}{d\left(\gamma, \sigma(\hat{S}_N^{1;\tau})\right)},
\tag{3.40}
$$

where

$$
d(\gamma, \sigma(\hat{S}_N^{1;\tau})) = \inf_{s \in \sigma(\hat{S}_N^{1;\tau})} |\gamma - s|.
\tag{3.41}
$$

This implies that

$$d(\gamma, \sigma(\hat{S}_N^{1;\tau})) \geq \frac{2}{\left|\left|\hat{B}\right|\right|_2 \left|\left|(\hat{I} + \hat{K})^{-1}\right|\right|_2} \tag{3.42}$$

for $N \geq S$. Hence, if $\gamma \in \rho(S^1)$, there can be no subsequence $\lambda_{N_j} \in \sigma(\hat{S}_{N_j}^{1;\tau})$ converging to γ. $\qquad \square$

Now we can prove the restricted no-spurious-mode condition.

Theorem 9. *Let D be some compact set in the complex plane, and let $\{\lambda_N\}_{N=1}^{\infty}$ be a sequence contained in D with $\lambda_N \in \sigma(\hat{S}_N^{1;\tau})$. Then for all $\epsilon > 0$, there exists some integer M such that λ_N is in an ϵ-neighborhood of some value $\lambda \in D \cap \sigma(S^1)$ for $N \geq M$.*

Proof. Suppose instead that there exists a subsequence λ_{N_j} such that $d(\lambda_{N_j}, D \cap \sigma(S^1)) \geq \epsilon > 0$. However, since D is compact, λ_{N_j} must have a convergent subsequence, and this subsequence must converge to some element in $\sigma(S^1)$ by Corollary 8. Hence our original assumption cannot hold, and the theorem is proved. $\qquad \square$

3.2 Proof of the Second Condition

We were able to prove the first condition under quite general assumptions. Specifically, it was not necessary to impose that S^1 was a self-adjoint operator. We are unable to prove the second condition without making this assumption. However, it should be noted that for non-self-adjoint operators, we have been unable to find numerical examples where the second condition appears not to hold.

Our proof relies on a number of results from [28]. To apply these, we need the following lemma.

Lemma 10. *$P_N S^1 P_N$ converges strongly to S^1.*

Proof. Let $\psi \in \mathrm{H}^p(\mathrm{S}_T)$. Then

$$\begin{aligned}
\left|\left|S^1\psi - P_N S^1 P_N \psi\right|\right|_2 &= \left|\left|P_N S^1 (\mathrm{I} - P_N)\psi + (\mathrm{I} - P_N) S^1 \psi\right|\right|_2 \\
&\leq \left|\left|S^1 (\mathrm{I} - P_N)\psi\right|\right|_2 + \left|\left|(\mathrm{I} - P_N) S^1 \psi\right|\right|_2 \tag{3.43} \\
&\leq C \left|\left|(\mathrm{I} - P_N)\psi\right|\right|_{2,p} + \left|\left|(\mathrm{I} - P_N) S^1 \psi\right|\right|_2 .
\end{aligned}$$

This must become arbitrarily small as $N \to \infty$. Therefore the lemma is proved. $\qquad \square$

The results we need from [28] will now be stated for the sake of completeness. Proofs of the lemmas and theorem can be found in [28], pages 290-292.

Definition 11. *For any linear operator T, if $\gamma \in \rho(T)$, the resolvent operator of T is defined as*

$$R_\gamma(T) \equiv (T - \gamma)^{-1}. \tag{3.44}$$

Lemma 12. *If T is a self-adjoint operator, then*

$$||R_\gamma(T)||_2 = \frac{1}{d(\gamma, \sigma(T))}, \tag{3.45}$$

where $d(\gamma, \sigma(T)) = \inf_{s \in \sigma(T)} |\gamma - s|$

Lemma 13. *If T is self adjoint and $Im(\gamma) \neq 0$, then*

$$||R_\gamma(T)||_2 \leq \frac{1}{|Im(\gamma)|}. \tag{3.46}$$

Definition 14. *Given a linear operator T with domain $D(T)$, a core of T is a subset $D \subset D(T)$ such that*

$$\overline{T|_D} = T, \tag{3.47}$$

where $\overline{T|_D}$ is the smallest closed extension of $T|_D$.

Our operator S^1 is closed over $H^p(S_T)$ [23]. Therefore $H^p(S_T)$ is a core for S^1. Likewise, each of the finite-rank operators $P_N S^1 P_N$ is continuous, and consequently, closed on $H^p(S_T)$. This makes $H^p(S_T)$ a common core for S^1 and $P_N S^1 P_N$. We can then use

Lemma 15. *Let $P_N S^1 P_N$ and S^1 be self-adjoint operators on common core D. If $P_N S^1 P_N$ converges strongly to S^1 on D, then $R_\gamma(P_N S^1 P_N)$ converges strongly to $R_\gamma(S^1)$ if $Im(\gamma) \neq 0$.*

Finally, given the above lemma, we use the following theorem.

Theorem 16. *Let $P_N S^1 P_N$ and S^1 be self adjoint on common core D. If $R_\gamma(P_N S^1 P_N)$ converges strongly to $R_\gamma(S^1)$ for $Im(\gamma) \neq 0$, and if $a < b$ and $(a, b) \subset \rho(P_N S^1 P_N)$ for N sufficiently large, then $(a, b) \subset \rho(S^1)$.*

Proof. See [28], page 290. □

Theorem 16 can be modified to accommodate subsequences, since the strong convergence of $P_N S^1 P_N$ to S^1 also holds for subsequences. This lets us prove the second condition. Suppose the second condition were false. This implies that there exists $\lambda \in \sigma(S^1)$ such that

$$d(\lambda, \sigma(\hat{S}_{N_j}^{1;\tau})) \geq \epsilon \tag{3.48}$$

for j sufficiently large. Suppose further that $\lambda \neq 0$. This implies that the disc $B_\lambda(\epsilon) = \{z \in \mathbb{C} : |z - \lambda| < \epsilon\}$ is a subset of $\rho(\hat{S}_{N_j}^{1;\tau})$, which implies $B_\lambda(\tilde{\epsilon}) \subset \rho(P_N S^1 P_N)$, where $\tilde{\epsilon} \leq \epsilon$. Therefore, by Theorem 16, $B_\lambda(\tilde{\epsilon}) \subset \rho(S^1)$. This is a contradiction, which implies the second condition for $\lambda \neq 0$. If $\lambda = 0$, we need only pick some $c \in \rho(S^1)$ and repeat our steps for $S^1 - c$.

Chapter 4

CONVERGENCE OF EIGENFUNCTIONS

We assume in advance that the approximate eigenvalues, $\lambda_N \in \sigma(\hat{S}_N^{1;\tau})$, converge to some $\lambda \in \sigma(S^1)$. Given $\lambda_N \in \sigma(\hat{S}_N^{1;\tau})$, there exists a $(2N+1)$-dimensional vector $\hat{\phi}_N^\tau$ such that

$$\hat{S}_N^{1,\tau}\hat{\phi}_N^\tau = \lambda_N\hat{\phi}_N^\tau, \ \left|\left|\hat{\phi}_N^\tau\right|\right|_2 = 1 \tag{4.1}$$

We prove the following proposition.

Theorem 17. *If $\lambda_N \in \sigma(\hat{S}_N^{1;\tau})$ converges to $\lambda \in \sigma(S^1)$, then there exists a vector $\hat{\phi}$ such that a subsequence of $\hat{\phi}_N$ converges to $\hat{\phi}$ in $||\cdot||_2$ and $S^1\phi = \lambda\phi$.*

Proof. We extend the vectors $\hat{\phi}_N^\tau$ to vectors $\hat{\phi}_N$ so that $\hat{P}_N\hat{\phi}_N = \hat{\phi}_N$. Given that $||\hat{\phi}_N||_2 = 1$, by Alaoglu's theorem [22] there exists a vector $\hat{\phi}$ such that some subsequence of $\hat{\phi}_N$, denoted as $\hat{\phi}_N$, converges weakly to $\hat{\phi}$. Using the operator B from the proof of Theorem 7, and noting that B commutes with the projection operator P_N, we get

$$\lambda_N\hat{B}\hat{\phi}_N = \hat{B}\hat{P}_N\hat{S}^1\hat{P}_N\hat{\phi}_N$$
$$= \hat{P}_N\hat{B}\hat{S}^1\hat{\phi}_N$$
$$= \hat{P}_N(\hat{I} + \hat{K})\hat{\phi}_N$$
$$= \hat{\phi}_N + \hat{P}_N\hat{K}\phi_N. \tag{4.2}$$

B is a compact operator since

$$\left|\left|(\hat{B}\hat{P}_N - \hat{B})\hat{\psi}\right|\right|_2^2 \leq \sum_{|n|>N}\left(\frac{2\pi n}{T}\right)^{-2p}\left|\hat{\psi}_n\right|^2$$
$$\leq \left(\frac{2\pi N}{T}\right)^{-2p}\left|\left|\hat{\psi}\right|\right|_2^2. \tag{4.3}$$

This implies that $\left|\left|\hat{B}\hat{P}_N - \hat{B}\right|\right|_2 \leq \left(\frac{2\pi N}{T}\right)^{-p}$, and so B is a uniform limit of finite-rank operators and is therefore compact. B compact then implies that $\hat{B}\hat{\phi}_N \to \hat{B}\hat{\phi}$. Likewise,

since K is compact, we have $\hat{K}\hat{\phi}_N \to \hat{K}\hat{\phi}$, $\hat{P}_N\hat{K} \to \hat{K}$ uniformly, and

$$\left\|\hat{P}_N\hat{K}\hat{\phi}_N - \hat{K}\hat{\phi}\right\|_2 \leq \left\|\hat{K}\hat{\phi}_N - \hat{K}\hat{\phi}\right\|_2 + \left\|\hat{P}_N\hat{K} - \hat{K}\right\|_2\left\|\hat{\phi}\right\|_2, \tag{4.4}$$

which implies $\hat{P}_N\hat{K}\hat{\phi}_N \to \hat{K}\hat{\phi}$. Therefore, we have

$$\hat{\phi}_N \to \lambda\hat{B}\hat{\phi} - \hat{K}\hat{\phi}. \tag{4.5}$$

Our weakly convergent sequence $\hat{\phi}_N$ has been shown to converge strongly. This implies

$$\hat{\phi}_N \to \hat{\phi}, \tag{4.6}$$

and using (4.5)

$$(\hat{I} + \hat{K})\hat{\phi} = \lambda\hat{B}\hat{\phi}. \tag{4.7}$$

We still need to show that the function ϕ, corresponding to symbol $\hat{\phi}$, is in $\mathrm{D}\left(S^1\right)$, and that it is an eigenfunction. There are two cases to consider. The first is $\lambda \neq 0$. This implies $\hat{S}^1$ is invertible, and we showed in Theorem 7 that the operator $\hat{I} + \hat{K}$ is invertible. If $\hat{I} + \hat{K}$ is invertible, then

$$(\hat{I} + \hat{K})^{-1}(\widehat{BS^1}) = \hat{I}, \tag{4.8}$$

where $\widehat{BS^1}$ denotes the extension of $\hat{B}\hat{S}^1$. Likewise, if S^1 is invertible, then for $\phi \in \mathrm{L}_2\left(\mathrm{S}_T\right)$ there must exist some $\psi \in \mathrm{H}^p\left(\mathrm{S}_T\right)$, with symbol $\hat{\psi}$, such that $\hat{S}^1\hat{\psi} = \hat{\phi}$. This implies

$$\begin{aligned}
\hat{\phi} &= \lambda(\hat{I} + \hat{K})^{-1}\hat{B}\hat{\phi} \\
&= \lambda(\hat{I} + \hat{K})^{-1}\hat{B}\hat{S}^1\hat{\psi} \\
&= \lambda(\hat{I} + \hat{K})^{-1}\widehat{BS^1}\hat{\psi} \\
&= \lambda\hat{\psi}, \tag{4.9}
\end{aligned}$$

and therefore ϕ is an eigenfunction of S^1.

The second case to consider is $\lambda = 0$. In that case let $c \in \rho\left(S^1\right)$ so that the operator $S^1 - c$ is invertible. Repeat the steps for the $\lambda \neq 0$ case. $\qquad\square$

If we assume that the eigenvalue λ is simple, then we see that every subsequence of ϕ_N converges to some unit multiple of ϕ since we claimed every sequence of approximate eigenvectors ϕ_N has a convergent subsequence. We can then say, upon appropriate rescalings,

that the sequence is convergent. The general problem for non-simple eigenvalues appears to be rather difficult, and we do not address it here.

Chapter 5

RATE OF CONVERGENCE

Before proceeding, we need two technical lemmas. The first lemma is from [2], page 69. We include the proof for clarity.

Lemma 18. *If $\phi \in \mathbf{C}^\infty (S_T)$, then $||(I - P_N)\phi||_2 = O(N^{-p})$ for all integer values of $p > 0$.*

Proof. Since $\mathbf{C}^\infty (S_T) \subset \mathrm{H}^p (\mathrm{S}_T)$ for arbitrary p, we can write

$$
\begin{aligned}
||(I - P_N)\phi||_2^2 &= \sum_{|n|>N} \left| \hat{\phi}_n \right|^2 \\
&= \left(\frac{T}{2\pi N} \right)^{2p} \sum_{|n|>N} \left(\frac{2\pi N}{T} \right)^{2p} \left| \hat{\phi}_n \right|^2 \\
&\leq \left(\frac{T}{2\pi N} \right)^{2p} \sum_{|n|>N} \left(\frac{2\pi n}{T} \right)^{2p} \left| \hat{\phi}_n \right|^2 \\
&\leq \left(\frac{T}{2\pi N} \right)^{2p} ||\phi||_{2,p}^2 .
\end{aligned}
\tag{5.1}
$$

Therefore

$$
||(I - P_N)\phi||_2 = \mathrm{O}\left(N^{-p} \right), \ p > 0.
\tag{5.2}
$$

$\square$

The second lemma relies upon a restriction of the self-adjoint operator S^1 to the form

$$
S^1 = \partial_x^p + \sum_{k=1}^{p-1} c_k \partial_x^k + f(x),
\tag{5.3}
$$

where the c_i are constants. We denote the constant-coefficient differential operator as D_p^μ. This restriction greatly simplifies our work since the operator P_N commutes with D_p^μ. We now prove our second technical lemma:

Lemma 19. *If ϕ_N and ϕ are the approximate and true eigenfunctions, respectively, of the operator $D_p^\mu + f(x)$, then convergence in $||\cdot||_2$ implies convergence in $||\cdot||_{2,k}$ for all positive integers k.*

Proof. We have $P_N S^1 \phi_N = D_p^\mu \phi_N + P_N f(x)\phi_N$. As shown in the previous section, if approximate eigenfunctions ϕ_N converge to ϕ in the $||\cdot||_2$ norm, then

$$\lim_{N \to \infty} ||P_N f(x)\phi_N - f(x)\phi||_2 = 0. \tag{5.4}$$

This implies

$$\lim_{N \to \infty} \left||D_p^\mu \phi_N - D_p^\mu \phi\right||_2 = 0, \tag{5.5}$$

which means that ϕ_N converges to ϕ in the graph norm associated with S^1, *i.e.*

$$\lim_{N \to \infty} \left(||\phi_N - \phi||_2 + \left||S^1 \phi_N - S^1 \phi\right||_2\right) = 0. \tag{5.6}$$

The graph norm associated with the operator S^1 is equivalent to the pth Sobolev norm [23], which implies that

$$\lim_{N \to \infty} ||\phi_N - \phi||_{2,p} = 0. \tag{5.7}$$

Convergence in $||\cdot||_{2,p}$ implies $\partial_x \phi_N \to \partial_x \phi$ in $||\cdot||_2$, and thus $\partial_x(D_p^\mu \phi_N + f(x)\phi_N)$ converges to $\partial_x S^1 \phi$ in $||\cdot||_2$. This implies that ϕ_N converges to ϕ in $||\cdot||_{2,p+1}$. Proceeding this way, we see that ϕ_N converges to ϕ in $||\cdot||_{2,k}$ for all integers $k > 0$. $\square$

Finally, we need the following min-max theorem [18]:

Theorem 20. *Suppose the self-adjoint operator S^1 has least eigenvalue $\lambda_0 > -\infty$. Then*

$$\lambda_0 = \inf_{||\psi||_2=1} \left\langle S^1 \psi, \psi \right\rangle, \tag{5.8}$$

where ψ is understood to be in the domain of S^1.

Using Theorem 20 and our technical lemmas, we prove the next theorem.

Theorem 21. *Let $\lambda = \min \sigma(S^1) > -\infty$, λ simple. Then there exists a sequence $\lambda_N \to \lambda$, $\lambda_N \in \sigma(\hat{S}_N^{1;\tau})$, and*

$$|\lambda_N - \lambda| = O\left(N^{-q}\right), \ q \geq 1. \tag{5.9}$$

Proof. By Theorem 20, we have

$$\lambda = \inf_{||\psi||_2=1} \left\langle S^1\psi, \psi \right\rangle. \tag{5.10}$$

Define the sequence $\{\lambda_N\}_{N=1}^{\infty}$ via

$$\lambda_N = \inf_{||\hat\psi_N^\tau||_2=1} \left\langle \hat S_N^{1,\tau}\hat\psi_N^\tau, \hat\psi_N^\tau \right\rangle, \tag{5.11}$$

where $\hat\psi_N^\tau \in \mathbb{C}^{2N+1}$. Let $\hat\psi_N$ denote the extension of $\hat\psi_N^\tau$, *i.e.* $\hat P_N\hat\psi_N = \hat\psi_N$. We can equivalently define λ_N as

$$\lambda_N = \inf_{||\psi_N||_2=1} \left\langle S^1\psi_N, \psi_N \right\rangle. \tag{5.12}$$

This shows that $\lambda_N \geq \lambda_{N+1} \geq \cdots \geq \lambda$. Since the λ_N's are a monotone sequence, they must have a limit, say $\tilde\lambda$. Suppose $\tilde\lambda > \lambda$. Since we know $\lambda \in \sigma(S^1)$ and S^1 is self adjoint, by our second condition for convergence, we know there exists a sequence $\gamma_N \in \sigma(\hat S_N^{1,\tau})$ such that $\gamma_N \to \lambda$. This implies for N large enough that $\gamma_N < \tilde\lambda$, but this would imply that $\gamma_N < \lambda_N$, which is impossible. Therefore $\tilde\lambda = \lambda$ and $\lambda_N \to \lambda$.

Let $c > \lambda_1$, $c \neq 0$. We can alter our definition of λ_N to

$$\lambda_N = \inf_{||\psi||_2=1} \left(\left\langle S^1 P_N\psi, P_N\psi \right\rangle + c\left\langle (I - P_N)\psi, \psi \right\rangle \right). \tag{5.13}$$

We introduce this alteration in order to take infimums over the same domain. Let each eigenvalue λ_N have corresponding eigenvector ϕ_N, and let λ have eigenvector ϕ. We showed in Section 4 that $\phi_N \to \phi$ in $||\cdot||_2$, so we can restrict ourselves to the set

$$E_N = \{\phi_j\}_{j\geq N}. \tag{5.14}$$

Assume that $\lambda > 0$, which implies $\lambda_N > 0$. Consider the difference

$$\frac{1}{\lambda} - \frac{1}{\lambda_N} = \sup_{E_N} \frac{1}{\left\langle S^1\psi, \psi \right\rangle} - \sup_{E_N} \frac{1}{\left(\left\langle S^1 P_N\psi, P_N\psi \right\rangle + c\left\langle (I - P_N)\psi, \psi \right\rangle \right)}. \tag{5.15}$$

We have

$$\left\langle S^1\psi, \psi \right\rangle = \left\langle S^1 P_N\psi, P_N\psi \right\rangle + R_N\left(\psi\right), \tag{5.16}$$

$$R_N\left(\psi\right) \equiv \left\langle (I - P_N)\psi, S^1 P_N\psi \right\rangle + \left\langle S^1\psi, (I - P_N)\psi \right\rangle, \tag{5.17}$$

and so

$$\frac{1}{\lambda} - \frac{1}{\lambda_N} \leq \frac{1}{\lambda\lambda_N}\sup_{E_N}\left|-R_N\left(\psi\right) + c\left\langle(I - P_N)\psi, \psi\right\rangle\right|, \tag{5.18}$$

or

$$\lambda_N - \lambda \leq \sup_{E_N}\left|-R_N\left(\psi\right) + c\left\langle(I - P_N)\psi, \psi\right\rangle\right|. \tag{5.19}$$

Using Cauchy-Schwartz,

$$\left|R_N\left(\psi\right)\right| \leq \left(\left|\left|S^1 P_N\psi\right|\right|_2 + \left|\left|S^1\psi\right|\right|_2\right)\left|\left|(I - P_N)\psi\right|\right|_2, \tag{5.20}$$

$$\left|c\left\langle(I - P_N)\psi, \psi\right\rangle\right| \leq c\left|\left|\psi\right|\right|_2\left|\left|(I - P_N)\psi\right|\right|_2. \tag{5.21}$$

Given the result of Lemma 19, we bound $\left(\left|\left|S^1 P_N\psi\right|\right|_2 + \left|\left|S^1\psi\right|\right|_2\right)$ by some constant M. Since E_N is closed, there must be some vector $\phi_K \in E_N$ such that

$$\sup_{E_N}\left|-R_N\left(\psi\right) + c\left\langle(I - P_N)\psi, \psi\right\rangle\right| \leq (M + c)\left|\left|(I - P_N)\phi_K\right|\right|_2. \tag{5.22}$$

Knowing that each ϕ_K is smooth, Lemma 18 implies

$$\lambda_N - \lambda \leq (M + c)\left(\frac{T}{2\pi N}\right)^q \left|\left|\phi_K\right|\right|_{2,q} \tag{5.23}$$

for all $q > 0$. Lemma 19 shows that $\left|\left|\phi_K\right|\right|_{2,q} \to \left|\left|\psi\right|\right|_{2,q}$. So, for a given ϵ, there must be some value L such that $\left|\left|\phi_K\right|\right|_{2,q} \leq (1 + \epsilon)\left|\left|\psi\right|\right|_{2,q}$ for all $K \geq L$. Hence, for $N \geq L$, we have

$$\lambda_N - \lambda \leq (M + c)(1 + \epsilon)\left(\frac{T}{2\pi N}\right)^q \left|\left|\psi\right|\right|_{2,q}. \tag{5.24}$$

In the case that $\lambda \leq 0$, pick α such that $\alpha + \lambda > 0$. Likewise we see that

$$\alpha + \lambda = \inf_{||\psi||_2=1}\left\langle(S^1 + \alpha)\psi, \psi\right\rangle, \tag{5.25}$$

$$\alpha + \lambda_N = \inf_{||\psi_N||_2=1}\left\langle(S^1 + \alpha)\psi_N, \psi_N\right\rangle. \tag{5.26}$$

Then we repeat our argument from above. $\qquad\qquad\square$

Note, in the case that $\inf \sigma(S^1) = -\infty$, but $\sup \sigma(S^1) < \infty$, we can apply the theorem just proved to the operator $-S^1$.

Chapter 6

THE BOUSSINESQ APPROXIMATION

6.1 Linearization and Spectral Problem

Again, the Boussinesq system is given by

$$\eta_t + \nabla \cdot \mathbf{v} + \nabla \cdot \eta \mathbf{v} + a\nabla \cdot \Delta \mathbf{v} - b\Delta \eta_t = 0,$$

$$\mathbf{v}_t + \nabla \eta + \frac{1}{2}\nabla |\mathbf{v}|^2 + c\nabla(\Delta \eta) - d\Delta \mathbf{v}_t = 0. \tag{6.1}$$

Further, In the case that $b \geq 0$ and $d \geq 0$, we can rewrite (1.34) to fit our definition of evolution equations as in (1.1). This restriction will be imposed throughout the rest of this thesis, and further a and c will be kept negative and equal. As shown though in [8] and [4], this restriction is not so severe and in fact represents the majority of the physically interesting cases where (1.34) might be used. The spectral stability of traveling line solitary wave solutions under two-dimensional small perturbations (transverse stability) are now studied.

In specific, we linearize the system around the solution $(\eta^*, u^*, 0)$ which is the line (independent of y) traveling solution to (6.1), i,e.

$$\eta^*(x, y, t) = \eta(\xi), \quad u^*(x, y, t) = u(\xi)$$

with $\xi = x - kt$. Using the fact $\partial_t = -k\partial_\xi + \partial_t$ and let the solution be of the form

$$\eta(\xi, y, t) = \eta^* + \epsilon \bar{\eta}(\xi, y, t)$$

$$u(\xi, y, t) = u^* + \epsilon \bar{u}(\xi, y, t) \tag{6.2}$$

$$v(\xi, y, t) = 0 + \epsilon \bar{v}(\xi, y, t)$$

By substituting (6.2) into (6.1) and dropping the ϵ^2 terms, one obtains the equations for $(\bar{\eta}, \bar{u}, \bar{v})$ which read (the bars on the dependent variables are dropped for simplicity of

notation)

$$-\eta_t + b\Delta\eta_t = -k\eta_\xi + bk\Delta\eta_\xi + \nabla\cdot\mathbf{v} + a\nabla\cdot\Delta\mathbf{v} + \nabla(\eta^*\mathbf{v} + \mathbf{v}^*\eta)$$

$$-\mathbf{v}_t + d\Delta\mathbf{v}_t = -k\mathbf{v}_\xi + dk\Delta\mathbf{v}_\xi + \nabla\eta + c\nabla(\Delta\eta) + \nabla(\mathbf{v}\cdot\mathbf{v}^*)$$

(6.3)

Since the equations are linear, it is natural to consider the solutions with respect to each mode. Assume the solutions are in the form of

$$\eta(\xi,y,t) = N(\xi,\rho,\Omega)e^{i\rho y + \Omega t} + \text{ complex conjugate}$$

$$\mathbf{v}(\xi,y,t) = \mathbf{V}(\xi,\rho,\Omega)e^{i\rho y + \Omega t} + \text{ complex conjugate}$$

(6.4)

Substituting (6.4) into (6.3) and denote $\mathbf{V} = (U,V)^t$, one obtains the equation

$$\Omega\begin{pmatrix} L_1 & 0 & 0 \\ 0 & K_1 & 0 \\ 0 & 0 & K_1 \end{pmatrix}\begin{pmatrix} N \\ U \\ V \end{pmatrix} = \begin{pmatrix} kL_1\partial_\xi + u^*_\xi + u^*\partial_\xi & M_2\partial_\xi + \eta^*_\xi & i\rho M_2 \\ M_1\partial_\xi & kK_1\partial_\xi + u^*_\xi + u^*\partial_\xi & 0 \\ i\rho M_1 & i\rho u^* & kK_1\partial_\xi \end{pmatrix}\begin{pmatrix} N \\ U \\ V \end{pmatrix}$$

(6.5)

where

$$K_1 = d\partial_{\xi\xi} - d\rho^2 - 1, \quad L_1 = b\partial_{\xi\xi} - b\rho^2 - 1,$$

$$M_1 = c\partial_{\xi\xi} - c\rho^2 + 1, \quad M_2 = a\partial_{\xi\xi} - a\rho^2 + 1 + \eta^*,$$

(6.6)

The corresponding eigenvalue problem is finding eigenvalues Ω and corresponding eigenfunctions for each ρ. The study of the case $\rho = 0$ is associated to the one-dimensional stability of the line solutions and the study of $\rho \neq 0$ is associated to the transverse stability.

6.2 Spectral Stability of Elevation Soliton

We wish to establish the spectral stability of the elevated soliton like solution

$$\eta^* = \eta_0\text{sech}^2(\lambda\xi),$$

$$u^* = \eta_0\sqrt{\frac{3}{\eta_0 + 3}}\text{sech}^2(\lambda\xi),$$

$$v^* = 0$$

(6.7)

for parameter values

$$a = c < 0, b > -a, d > 0,$$

(6.8)

with $2a + b + d = \frac{1}{3}$, and

$$\lambda = \frac{1}{2}\sqrt{\frac{2\eta_0}{3(a-b) + 2b(\eta_0 + 3)}}, \quad k = \frac{3 + 2\eta_0}{\sqrt{3(3 + \eta_0)}}.$$

Note, in all subsequent calculations in this section and in the appendix, we substitute $-a$ for a. First, we transform (6.5) into a first-order system

$$\frac{d}{d\xi}\begin{pmatrix} \phi \\ \psi \end{pmatrix} = (A(\Omega, \eta_0, \rho) + i\rho W(\eta_0, \rho) + C(\xi, \eta_0, \rho))\begin{pmatrix} \phi \\ \psi \end{pmatrix}, \tag{6.9}$$

where

$$\phi = (N, U, N', U', N'', U'')^T, \tag{6.10}$$

$$\psi = \qquad (V, V', V'')^T, \tag{6.11}$$

and

$$A(\Omega, \eta_0, \rho) = \begin{pmatrix} A_0(\eta_0, \rho) + \Omega A_1(\eta_0, \rho) & 0 \\ & \\ 0 & B_0(\eta_0, \rho) + \Omega B_1(\eta_0, \rho) \end{pmatrix}. \tag{6.12}$$

The explicit entries in the above matrices are provided in the appendix.

While the first-order system is not an eigenvalue problem *per se*, it is equivalent to our original eigenvalue problem in the following way. Define the operator $T(\Omega)$ such that

$$\begin{aligned} \mathrm{D}\left(T(\Omega)\right) &= \mathrm{H}^1\left(\mathbb{R}; \mathbb{C}^9\right), \\ T(\Omega) &= \frac{d}{d\xi} - (A(\Omega, \eta_0, \rho) + C(\xi, \eta_0, \rho) + i\rho W(\eta_0, \rho)). \end{aligned} \tag{6.13}$$

Given some value Ω, if $T(\Omega)$ has a bounded inverse then Ω is not in the spectrum of our linearization. Likewise, if $T(\Omega)$ does not have a bounded inverse, then Ω is in the spectrum of our linearization. For more details and exposition, see [30]. A means of determining the invertibility of $T(\Omega)$ is provided by the theory of exponential dichotomies. Following [10] and [30], we say that $T(\Omega)$ generates an exponential dichotomy on $\mathbb{R}$ if given the fundamental solution matrix of $T(\Omega)$, say $X(\xi)$, there exists a projection operator P, and positive constants K, L, α, and β such that the inequalities

$$\begin{aligned} \left\|X(t)PX^{-1}(s)\right\| &\leq Ke^{-\alpha(t-s)}, \quad \text{for } t \geq s \\ \left\|X(t)(I-P)X^{-1}(s)\right\| &\leq Le^{-\beta(s-t)}, \quad \text{for } s \geq t \end{aligned} \tag{6.14}$$

are satisfied for all values of t and s in $\mathbb{R}$.

A result of Palmer ([25], [26]) shows that $T(\Omega)$ has a bounded inverse if and only if there exists an exponential dichotomy on $\mathbb{R}$. While such a condition is difficult to show for $T(\Omega)$, for the constant-coefficient operator

$$\tilde{T}(\Omega) \equiv \frac{d}{d\xi} - A(\Omega, \eta_0, \rho), \tag{6.15}$$

$\tilde{T}(\Omega)$ will have an exponential dichotomy on $\mathbb{R}$ if and only if $A(\Omega, \eta_0, \rho)$ is hyperbolic [30]. Thus we might hope to make use of the "roughness" theorem from [10].

Theorem 22. *If the ODE $\frac{d}{d\xi} - F(\xi)$ generates an exponential dichotomy on $\mathbb{R}$, then if the matrix $B(\xi)$ satisfies the inequality*

$$\sup_{\xi \in \mathbb{R}} ||B|| \leq \frac{\alpha}{4K^2}, \tag{6.16}$$

where α and K are the constants in (6.14), then $\frac{d}{d\xi} - F(\xi) - B(\xi)$ generates an exponential dichotomy on $\mathbb{R}$ as well.

It can be shown (see Appendix, section A.1), that $A(\Omega, \eta_0, \rho)$ is hyperbolic if $\mathrm{Re}\,(\Omega) \neq 0$, which is the region in which we are interested. Then $A(\Omega, \eta_0, \rho)$ has three eigenvalues, say $\mu_1(\Omega, \eta_0, \rho)$, $\mu_2(\Omega, \eta_0, \rho)$, and $\mu_3(\Omega, \eta_0, \rho)$ in the left-hand plane, and each are simple and analytic for $\mathrm{Re}\,(\Omega) \geq 0$ (see Appendix, Lemma 28). Let Γ denote a smooth, closed curve containing only these eigenvalues, and define the projection associated with the exponential dichotomy for $A(\Omega)$ by

$$P(\Omega, \eta_0, \rho) = -\frac{1}{2\pi i} \oint_\Gamma \left(A(\Omega, \eta_0, \rho) - z \right)^{-1} dz. \tag{6.17}$$

This projection can be decomposed into the projections P_i associated with the individual eigenvalues μ_i. Further, the P_i are mutually orthogonal [19], and so we have (letting the variables Ω, η_0, and ρ be understood)

$$\left|\left| e^{At} P e^{-As} \right|\right| = \left|\left| P_1 e^{\mu_1(t-s)} + P_2 e^{\mu_2(t-s)} + P_3 e^{\mu_3(t-s)} \right|\right|$$
$$\leq e^{-\alpha(\Omega, \eta_0, \rho)(t-s)}, \tag{6.18}$$

where $\alpha(\Omega, \eta_0, \rho) = \min_i \{|\mathrm{Re}\,(\mu_i(\Omega, \eta_0, \rho))|\}$.

This provides the following modification of the "roughness" theorem from [10].

Theorem 23. *If*

$$\sup_{\xi \in \mathbb{R}} ||C(\xi, \eta_0, \rho) + i\rho W(\eta_0, \rho)|| \leq \frac{1}{4}\alpha(\Omega, \eta_0, \rho) \tag{6.19}$$

then $T(\Omega)$ will have an exponential dichotomy on $\mathbb{R}$ and therefore have a bounded inverse.

Hence, if we can control $\alpha(\Omega, \eta_0, \rho)$ for $\mathrm{Re}\,(\Omega) > 0$, we can establish a range of η_0 and ρ values satisfying the inequality of Theorem 23 and thus prove spectral stability for the given range of parameter values. Let $\mu_1(\Omega, \eta_0, \rho)$ be the eigenvalue in the left-hand plane with smallest real part. We know that $\mu_1(\Omega, \eta_0, \rho)$ can never be strictly imaginary (see appendix, Lemma 26), but for Theorem 23 to be useful, we need to find some lower bound on the magnitude of $\mathrm{Re}\,(\mu_1)$. From basic perturbation theory, we know for $|\Omega| \ll 1$ that $\mu_1(\Omega, \eta_0, \rho)$ must be within $O(\Omega)$ of $\mu_1(0, \eta_0, \rho)$.

We can now prove the following theorem.

Theorem 24. *If $Re\,(\Omega) \geq 0$, a, b, d are in some neighborhood of $-1/9, 1/3$ and $2/9$, $\rho = 0$, and η_0 is sufficiently small, then*

$$\alpha(\Omega, \eta_0, \rho) \leq \mu_1(0, \eta_0, 0). \tag{6.20}$$

Proof. From the appendix, we know that $\mu_2 = -\sqrt{\rho^2 + 1/d}$, and so it satisfies (6.23). Therefore, we only need to study the remaining two eigenvalues. From the asymptotic results presented in the appendix, we know there exists constants M and N such that (6.23) holds for $|\Omega| \leq M$ and $|\Omega| \geq N$. Therefore, we need only establish the inequality in the intermediate region D2 from Figure 1. Given that $\mu_1(\Omega, \eta_0, \rho)$ is analytic in the right-half plane, this makes $\mathrm{Re}\,(\mu_1(\Omega, \eta_0, \rho))$ a harmonic function. So, to prove the inequality on D2, we only need to prove it along $\partial(\mathrm{D2})$, and then invoke the maximum principle for harmonic functions. Clearly, by construction, this means we only need to prove the inequality on $\partial(\mathrm{D2}) \cap i\mathbb{R}$. A final simplification can be made by noting that if we let $\Omega = i\omega$, $\omega \in \mathbb{R}$, then $\overline{\sigma(A(i\omega))} = \sigma(A(-i\omega))$, so then one only need prove the bound on $\partial(\mathrm{D2}) \cap i\mathbb{R}^+$.

Given the above, denote the characteristic polynomial of $A_0 + \Omega A_1$ as $p(\mu)$. If we show that the associated polynomial
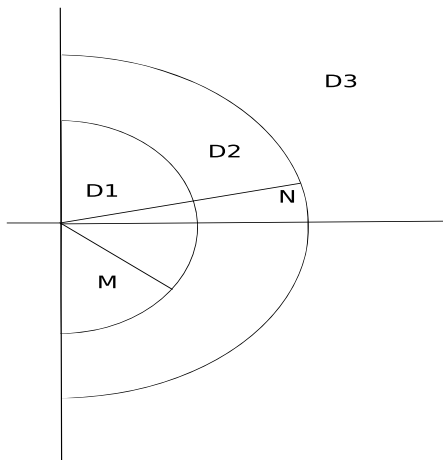
$$b(z) \equiv p(\mu_1(0, \eta_0, 0) + iz) \tag{6.21}$$

Figure 6.1:

has no real roots in z, then the asymptotic behavior of both μ_1 and μ_3 as $\Omega \to \infty$ guarantees (6.23) must be satisfied. The four eigenvalues of $A(\Omega, \eta_0, \rho)$ are either imaginary or have positive real part (see Lemma 27). This ensures that there can be at most two real roots for $b(z)$.

So, we must show that $b(z)$ cannot have a real root. To do this, first let $p_1(x) = \operatorname{Re}(b(x))$, and $p_2(x) = \operatorname{Im}(b(x))$, where x is real. Hence, for b to have real root, p_1 and p_2 must have a simultaneous root. Thus, we compute the resolvent (see [15] for definition and relevant theorems), $\mathscr{R}(p_1, p_2)$, which is zero if and only if p_1 and p_2 have a common root. We can compute $\mathscr{R}(p_1, p_2)$ using MAPLE, which gives

$$\mathscr{R}(p_1, p_2) = 64\mu_1^6(0, \eta_0, 0)(k^2 bd - a^2)\omega^2 g(\omega, k, \mu_1(0, \eta_0, 0)). \tag{6.22}$$

Unfortunately, g is a rather complicated polynomial expression in all of its variables. Thankfully though, $\mu_1(0, \eta_0, 0) \to 0$ and $k \to 1$ as $\eta_0 \to 0$, thus we can expand g as

$$g(\omega, k, \mu_1(0, \eta_0, 0)) = 3\omega^2(aw^2 + 1)^2 \tilde{g}(w) + \mathscr{O}((k-1), \mu_1(0, \eta_0, 0)). \tag{6.23}$$

Hence, if we show that $\tilde{g}$ is positive for $\omega > 0$, we are done, since we only need to look at values of ω on a finite interval. Of course, even $\tilde{g}(\omega)$ is difficult to analyze, but if we insert the particular values for a, b, and d as given in the statement of the theorem, we get

$$\tilde{g}(\omega) = \frac{16}{1594323}\omega^{10} - \frac{728}{1594323}\omega^8 + \frac{14200}{1594323}\omega^6 + \frac{5849}{177147}\omega^4 + \frac{19718}{19683}\omega^2 + 15. \tag{6.24}$$

One can factor $\frac{16}{1594323}\omega^{10} - \frac{728}{1594323}\omega^8 + \frac{14200}{1594323}\omega^6$ into $\frac{8\omega^6}{1594323}(2\omega^4 - 91\omega^2 + 1775)$, which only has a real root at $\omega = 0$. Since $2\omega^4 - 91\omega^2 + 1775$ is positive at the origin, and $\frac{5849}{177147}\omega^4 + \frac{19718}{19683}\omega^2 + 15$ is always positive, we see that $\tilde{g}(\omega) > 0$, and thus for η_0 sufficiently small, $g(\omega, k, \mu_1(0, \eta_0, 0)) > 0$. Thus p_1 and p_2 have no common roots and the theorem is proved. $\qquad\square$

To now make use of the roughness theorem for exponential dichotomies, we must estimate the size of $C(\xi, \eta_0, 0)$. For C, we use the matrix 2-norm, which for an $n \times n$ matrix A is

$$||A||_2^2 = \sum_{i,j=1}^{n} |a_{ij}|^2 \tag{6.25}$$

It is straightforward to show

$$||C(\xi, \eta_0, 0)||_2 \leq \eta_0 \sqrt{\frac{\alpha(1 + 4\lambda^2)}{k^2 bd - a^2}} \operatorname{sech}(\lambda\xi) \tag{6.26}$$
$$\alpha = k^2 d^2 + a^2 + (kd + a)^2 + (kb + a)^2.$$

and hence $\sup_{\xi \in \mathscr{R}} ||C(\xi, \eta_0, 0)||_2 = \mathscr{O}(\eta_0)$. As shown in Lemma 24, $|\mu_1(0, \eta_0, 0)| = \mathscr{O}(\sqrt{\eta_0})$. Hence, we must necessarily have $\sup_{\xi \in \mathscr{R}} ||C(\xi, \eta_0, 0)||_2 \leq |\mu_1(0, \eta_0, 0)|/4$ for small η_0.

Appendix A

A.1   Matrix Entries

$$A_0(\eta_0, \rho) = \begin{pmatrix} 0 & I & 0 \\ 0 & 0 & I \\ 0 & kZ & 0 \end{pmatrix}, \tag{A.1}$$

$$A_1(\eta_0, \rho) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -Z & 0 & E \end{pmatrix}, \tag{A.2}$$

with

$$Z = \frac{1}{k^2 bd - a^2} \begin{pmatrix} kd(b\rho^2 + 1) & a(d\rho^2 + 1) \\ a(b\rho^2 + 1) & kb(d\rho^2 + 1) \end{pmatrix}, \quad E = \frac{1}{k^2 bd - a^2} \begin{pmatrix} kbd & ad \\ ab & kbd \end{pmatrix}. \tag{A.3}$$

In V, we have

$$B_0(\eta_0, \rho) + \Omega B_1(\eta_0, \rho) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\frac{\Omega}{k}(\rho^2 + \frac{1}{d}) & (\rho^2 + \frac{1}{d}) & \frac{\Omega}{k} \end{pmatrix}. \tag{A.4}$$

For the constant coefficient coupling dominated by ρ, we have

$$W(\eta_0, \rho) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & W_1 \\ W_2 & W_3 & 0 \end{pmatrix}, \tag{A.5}$$

with

$$W_1 = \frac{-1}{k^2bd - a^2} \begin{pmatrix} 0 & 0 & 0 \\ kd(a\rho^2 + 1) & 0 & -kad \\ a(a\rho^2 + 1) & 0 & -a^2 \end{pmatrix}, \quad W_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \dfrac{-(a\rho^2 + 1)}{kd} & 0 & 0 \end{pmatrix}$$

$$(A.6)$$

$$W_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \dfrac{a}{kd} & 0 \end{pmatrix}.$$

Finally we have

$$C(\xi, \eta_0, \rho) = \begin{pmatrix} 0 & 0 & 0 \\ C_1 & C_2 & C_3 \\ C_4 & 0 & 0 \end{pmatrix}, \tag{A.7}$$

with, dropping the $*$ superscript,

$$C_1 = \frac{1}{k^2bd - a^2} \begin{pmatrix} 0 & 0 & 0 \\ kdu_\xi & kd\eta_\xi + au_\xi & kdu \\ au_\xi & a\eta_\xi + kbu_\xi & au \end{pmatrix}, \quad C_2 = \frac{1}{k^2bd - a^2} \begin{pmatrix} 0 & 0 & 0 \\ kd\eta + au & 0 & 0 \\ a\eta + kbu & 0 & 0 \end{pmatrix},$$

$$C_3 = \frac{i\rho}{k^2bd - a^2} \begin{pmatrix} 0 & 0 & 0 \\ kd\eta & 0 & 0 \\ a\eta & 0 & 0 \end{pmatrix}, \quad C_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \dfrac{i\rho}{kd}u & 0 \end{pmatrix}$$

$$(A.8)$$

A.2 The Hyperbolicity of $A(\Omega, \eta_0, \rho)$

We are interested in finding where the matrix $A(\Omega, \eta_0, \rho)$ is hyperbolic, which amounts to computing the spectrum of the constant-coefficient operator

$$\begin{pmatrix} L_1 & 0 & 0 \\ 0 & K_1 & 0 \\ 0 & 0 & K_1 \end{pmatrix}^{-1} \begin{pmatrix} kL_1\partial_\xi & M_1\partial_\xi & 0 \\ M_1\partial_\xi & kK_1\partial_\xi & 0 \\ 0 & 0 & kK_1\partial_\xi \end{pmatrix}. \tag{A.9}$$

We know that L_1 and K_1 have bounded inverse on $L_2(\mathbb{R})$, so we can instead study

$$F \equiv \begin{pmatrix} k\partial_\xi & L_1^{-1}M_1\partial_\xi & 0 \\ K_1^{-1}M_1\partial_\xi & k\partial_\xi & 0 \\ 0 & 0 & k\partial_\xi \end{pmatrix}. \tag{A.10}$$

We can compute the spectrum of F by taking its Fourier transform, $\hat{F}(w)$, and then finding the zeros of $\det(\hat{F} - \Omega)$. We have

$$\det(\hat{F} - \Omega) = (-ikw - \Omega)\left((-ikw - \Omega)^2 + w^2 \frac{(aw^2 + a\rho^2 + 1)^2}{(bw^2 + b\rho^2 + 1)(dw^2 + d\rho^2 + 1)}\right) \tag{A.11}$$

So, we clearly see that $i\mathbb{R}$ is in the spectrum of F, but we still need to find the roots of

$$(-ikw - \Omega)^2 + w^2 \frac{(aw^2 + a\rho^2 + 1)^2}{(bw^2 + b\rho^2 + 1)(dw^2 + d\rho^2 + 1)}. \tag{A.12}$$

If we separate real and imaginary parts we must have

$$\mathrm{Re}\,(\Omega)^2 - (kw + \mathrm{Im}\,(\Omega))^2 + w^2 \frac{(aw^2 + a\rho^2 + 1)^2}{(bw^2 + b\rho^2 + 1)(dw^2 + d\rho^2 + 1)} = 0$$

$$\mathrm{Re}\,(\Omega)\,(kw + \mathrm{Im}\,(\Omega)) = 0. \tag{A.13}$$

If we let $\mathrm{Re}\,(\Omega) \neq 0$, then we must have $\mathrm{Im}\,(\Omega) = -kw$, but then we must also have

$$\mathrm{Re}\,(\Omega)^2 + w^2 \frac{(aw^2 + a\rho^2 + 1)^2}{(bw^2 + b\rho^2 + 1)(dw^2 + d\rho^2 + 1)} = 0, \tag{A.14}$$

which is impossible. Therefore $\sigma(F) = i\mathbb{R}$, and $A(\Omega, \eta_0, \rho)$ is hyperbolic if and only if $\mathrm{Re}\,((\Omega, \eta_0, \rho)) \neq 0$.

A.3 Results about the Characteristic Polynomial of $A(\Omega, \eta_0, \rho)$

We will now list a number of lemmas concerning the characteristic polynomial, say $p_A(\mu)$, of the matrix (6.12). Obviously we can factor this polynomial into the characteristic polynomials for $A_0 + \Omega A_1$ and $B_0 + \Omega B_1$. The eigenvalues for $B_0 + \Omega B_1$ are

$$\left\{ \pm\sqrt{\rho^2 + \frac{1}{d}}, \frac{\Omega}{k} \right\}. \tag{A.15}$$

Denote the characteristic polynomial for $A_0 + \Omega A_1$ by $p(\mu; \Omega)$. There is of course no general formula to compute the roots of this polynomial since it is sixth order. However, it is not difficult to show though that

$$p(\mu; \Omega) = \mu^2(k^2 p_4(\mu) + \tilde{p}_4(\mu)) - 2k\Omega\mu p_4(\mu) + \Omega^2 p_4(\mu), \qquad (A.16)$$

where

$$p_4(\mu) = (b\mu^2 - (b\rho^2 + 1))(d\mu^2 - (d\rho^2 + 1))$$
$$\tilde{p}_4(\mu) = -a^2(\mu^2 - (\rho^2 + \frac{1}{a}))^2. \qquad (A.17)$$

This allows us to prove:

Lemma 25. *If $Re\,(\Omega) \geq 0$ and $Im\,(\Omega) \neq 0$, then $p(\mu; \Omega)$ does not have any real, negative, roots.*

Proof. Assuming $\mu \in \mathbb{R}$, then separating $p(\mu; \Omega) = 0$ into real and imaginary parts gives

$$\mu^2(k^2 p_4(\mu) + \tilde{p}_4(\mu)) - 2k\mathrm{Re}\,(\Omega)\,\mu p_4(\mu) + (\mathrm{Re}\,(\Omega)^2 - \mathrm{Im}\,(\Omega)^2)p_4(\mu) = 0$$

$$(\mathrm{Re}\,(\Omega) - k\mu)\mathrm{Im}\,(\Omega)\,p_4(\mu) = 0. \qquad (A.18)$$

We only have one possibility and that is $p_4(\mu) = 0$. This would then imply that $\mu^2 \tilde{p}_4(\mu) = 0$, but this is obviously impossible. $\qquad \square$

We also need:

Lemma 26. *If Ω is purely imaginary, then the eigenvalues are symmetric about the imaginary axis.*

Proof. Let $p(\mu; i\omega) = 0$. Then

$$p(-\bar{\mu}; i\omega) = \overline{p(\mu; i\omega)} = 0. \qquad (A.19)$$

$\qquad \square$

A.4 Asymptotics for $\sigma(A(\Omega, \eta_0, \rho))$

A.4.1 Small Ω Results

To establish results for $|\Omega| \ll 1$, we start by computing the roots of $p(\mu; 0)$, which are

$$\left\{ 0, \pm\sqrt{\rho^2 + \frac{\frac{k^2}{3} + 2a(k^2 - 1) \pm k\sqrt{k^2(b-d)^2 + 4(a-d)(a-b)}}{2(k^2bd - a^2)}} \right\}. \tag{A.20}$$

We can then immediately provide the estimates on $\mu_1(0, \eta_0, 0)$ used in the main text.

Lemma 27.

$$|\mu_1(0, \eta_0, 0)| = \sqrt{\frac{\eta_0}{2(k^2bd - a^2)}} (\beta + \mathcal{O}(\eta_0))^{1/2} \tag{A.21}$$

$$\beta = \tfrac{1}{3}(\tfrac{1}{2} + 6a - \tfrac{1}{18}(b - d)^2)$$

Proof. By performing a Taylor expansion, one has that

$$k = 1 + \frac{\eta_0}{2} + \mathcal{O}(\eta_0^2) \tag{A.22}$$

which also gives $k^2 = 1 + \eta_0 + \mathcal{O}(\eta_0^2)$. As can readily be shown

$$\mu_1^2(0, \eta_0, 0) = \frac{\frac{k^2}{3} + 2a(k^2 - 1) - k\sqrt{k^2(b-d)^2 + 4(a-d)(a-b)}}{2(k^2bd - a^2)}, \tag{A.23}$$

and so by inserting the expansion in k, k^2, and using the identity $-2a + b + d = \frac{1}{3}$, the estimate is obtained. $\square$

The root at zero is a double root, and we will denote each root as $\mu_5(0, \eta_0, \rho)$ and $\mu_6(0, \eta_0, \rho)$. Let the two negative roots be denoted by $\mu_3(0, \eta_0, \rho)$ and $\mu_1(0, \eta_0, \rho)$, with $\mu_3(0, \eta_0, \rho) < \mu_1(0, \eta_0, \rho)$. There is some question as to what will happen to $\mu_5(\Omega, \eta_0, \rho)$ and $\mu_6(\Omega, \eta_0, \rho)$ for $\mathrm{Re}\,(\Omega) > 0$. Thankfully, zero is a semi-simple eigenvalue of A_0, and we can use the reduction method of Kato [19]. This gives the expansions

$$\mu_5(\Omega, \eta_0, \rho) = \Omega\frac{k(b\rho^2 + 1)(d\rho^2 + 1) + (1 + a\rho^2)\sqrt{(b\rho^2 + 1)(d\rho^2 + 1)}}{k^2(b\rho^2 + 1)(d\rho^2 + 1) - (1 + a\rho^2)^2} + O(\Omega^2),$$

$$\mu_6(\Omega, \eta_0, \rho) = \Omega\frac{k(b\rho^2 + 1)(d\rho^2 + 1) - (1 + a\rho^2)\sqrt{(b\rho^2 + 1)(d\rho^2 + 1)}}{k^2(b\rho^2 + 1)(d\rho^2 + 1) - (1 + a\rho^2)^2} + O(\Omega^2). \tag{A.24}$$

This shows that for $\mathrm{Re}\,(\Omega) > 0$, the two eigenvalues that begin at the origin move into the right-half plane. Using the same asymptotic methods as above, we also prove the following key lemma.

Lemma 28. *There exists some $M > 0$ such that $\mu_1(\Omega, \eta_0, \rho)$ and $\mu_3(\Omega, \eta_0, \rho)$ satisfy (6.23) for $|\Omega| \leq N$. Further, $\mu_1(\Omega, \eta_0, \rho)$ and $\mu_3(\Omega, \eta_0, \rho)$ are analytic for $\mathrm{Re}\,(\Omega) \geq 0$.*

Proof. We know that $A(\Omega, \eta_0, \rho)$ has no strictly imaginary eigenvalues for $\mathrm{Re}\,(\Omega) \neq 0$. Thus $\mu_3(\Omega, \eta_0, \rho)$ and $\mu_1(\Omega, \eta_0, \rho)$ are the only eigenvalues of $A_0 + \Omega A_1$ in the left-hand plane. Further, since $\mu_3(\Omega, \eta_0, \rho)$ and $\mu_1(\Omega, \eta_0, \rho)$ are simple at $\Omega = 0$, we know both functions are analytic in Ω about the origin [19]. Let

$$\mu_1(\Omega, \eta_0, \rho) = \mu_1(0, \eta_0, \rho) + \alpha_1 \Omega + O(\Omega^2)$$
$$\mu_3(\Omega, \eta_0, \rho) = \mu_3(0, \eta_0, \rho) + \alpha_3 \Omega + O(\Omega^2). \tag{A.25}$$

We then have

$$\alpha_1 = -\frac{(d(\mu_1^2(0) - \rho^2) + 1)(b(\mu_1^2(0) - \rho^2) + 1))}{\mu_1^2(0)\sqrt{k^2(b-d)^2 + 4(a-d)(a-b)}}$$
$$\alpha_3 = \frac{(d(\mu_3^2(0) - \rho^2) + 1)(b(\mu_3^2(0) - \rho^2) + 1))}{\mu_3^2(0)\sqrt{k^2(b-d)^2 + 4(a-d)(a-b)}}, \tag{A.26}$$

and $\alpha_1 < 0$ while $\alpha_3 > 0$. Thus for Ω sufficiently small, (6.23) holds. As for the remainder of the lemma, if $\mathrm{Im}\,(\Omega) > 0$, then for small enough Ω we know $\mathrm{Im}\,(\mu_1(\Omega, \eta_0, \rho)) < 0$ and $\mathrm{Im}\,(\mu_3(\Omega, \eta_0, \rho)) > 0$. By Lemma 26, we see then that neither $\mu_1(\Omega, \eta_0, \rho)$ nor $\mu_3(\Omega, \eta_0, \rho)$ can ever cross the real axis, and therefore they can never intersect. We can repeat a similar argument for $\mathrm{Im}\,(\Omega) < 0$. Since we need analyticity for $\mathrm{Re}\,(\Omega) \geq 0$, we still have to examine the case where $\Omega > 0$, which is to say that we need to show that μ_1 and μ_3 never intersect for $\Omega > 0$. To do this, first note that for real Ω, the roots of $p(\mu; \Omega)$ must be symmetric with respect to the real axis. Since μ_1 and μ_3 are simple for small Ω, this means they must be real. Further, because of symmetry, they can only leave the negative real axis if they intersect. We can then look for the positive roots of $p(-\mu; \Omega)$ using Descartes rule of signs. It is straightforward to show that there can only be an even number of roots, and so therefore the μ_1 and μ_3 can never intersect for $\Omega > 0$. We are then finished. $\qquad\square$

A.4.2 Large Ω Results

In order to study the behavior of the eigenvalues in the left-hand plane for $|\Omega| \gg 1$, we first introduce the substitution $t = \frac{1}{\Omega}$. This transforms the eigenvalue problem

$$(A_0 + \Omega A_1)\phi = \mu(\Omega)\phi \tag{A.27}$$

into

$$(tA_0 + A_1)\phi = t\,\mu\left(\frac{1}{t}\right)\phi. \tag{A.28}$$

Repeating the methods used in the previous section, again see [19], for the eigenvalues with negative real part, we obtain the following expansions.

$$\mu_1 = -\sqrt{\rho^2 + \frac{1}{b}} + \mathrm{O}\left(\frac{1}{\Omega}\right)$$
$$\mu_3 = -\sqrt{\rho^2 + \frac{1}{d}} + \mathrm{O}\left(\frac{1}{\Omega}\right) \tag{A.29}$$

Thus we see there must exist some value $N > 0$, such that for $|\Omega| \geq N$, both μ_1 and μ_3 will satisfy the inequality (6.23).

50

BIBLIOGRAPHY

[1] M. Ablowitz and H. Segur. *Solitons and the Inverse Scattering Transform*. SIAM, Philadelphia, PN, 1981.

[2] K. Atkinson. *The Numerical Solution of Integral Equations of the Second Kind*. Cambridge University Press, Cambridge, 1997.

[3] K. Atkinson and W. Han. *Theoretical Numerical Analysis: A Functional Analysis Framework*. Springer, New York, NY, 2005.

[4] J.L. Bona, T. Colin, and D. Lannes. Long wave approximations for water waves. *Arch. Rational Mech. Anal.*, 178:373–410, 2005.

[5] A. Böttcher. Infinite matrices and projection methods. In *Lectures on Operator Theory and Its Applications*, pages 3–72. American Mathematical Society, Providence, RI, 1996.

[6] A. Böttcher and B. Silbermann. *Analysis of Toeplitz Operators*. Springer-Verlag, New York, NY, 1990.

[7] M. Chen. From Boussinesq systems to KP-type equations. *Canadian Applied Mathematics Quarterly*, 15:367–373, 2004.

[8] M. Chen, J.L. Bona, and J-C. Saut. Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media. part I: Derivation and the linear theory. *Journal of Nonlinear Science*, 12:283–318, 2002.

[9] J.B. Conway. *A Course in Functional Analysis*. Springer, New York, NY, 2nd edition, 1990.

[10] W.A. Coppel. *Dichotomies in Stability Theory*. Springer, Berlin, 1978.

[11] B. Deconinck and J. N. Kutz. Computing spectra of linear operators using the Floquet-Fourier-Hill method. *J. Comput. Phys.*, 219:296–321, 2006.

[12] L.M. Delves and Freeman T.L. *Analysis of Global Expansion Methods: Weakly Asymptotically Diagonal Systems*. Academic Press, New York, NY, 1981.

[13] I. Gohberg, S. Goldberg, and M.A. Kaashoek. *Classes of Linear Operators: Vol. I*. Birkhauser-Verlag, Basel, 1990.

[14] I. Gohberg, S. Goldberg, and N. Krupnick. *Traces and Determinants of Linear Operators*. Birkhauser-Verlag, Basel, 2000.

[15] P.A. Griffiths. *Introduction to Algebraic Curves*. AMS, 1991.

[16] M. Grillakis, J. Shatah, and W.A. Strauss. Stability theory of solitary waves in the presence of symmetry. *Journal of Functional Analysis*, 74:160–197, 1987.

[17] G. W. Hill. On the part of the motion of the lunar perigee which is a function of the mean motions of the sun and moon. *Acta Math.*, 8(1):1–36, 1886.

[18] P.D. Hislop and I.M. Sigal. *Introduction to Spectral Theory: With Applications to Schröedinger Operators*. Springer, New York, NY, 1996.

[19] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, N.Y., 2nd edition, 1976.

[20] M.A. Krasnosel'skii, G.M. Vainikko, P.P. Zabreiko, Ya. B. Rutitskii, and V. Ya. Stetsenko. *Approximate Solution of Operator Equations*. Wolters-Noordhoff, Groningen, 1972.

[21] D.F. Lawden. *Elliptic Functions and Applications*. Springer-Verlag, New York, NY, 1989.

[22] P.D. Lax. *Functional Analysis*. Wiley, New York, NY, 2002.

[23] J. Locker. *Functional Analysis and Two-Point Differential Operators*. Longman Scientific and Technical, New York, NY, 1986.

[24] A.W. Naylor and G.R. Sell. *Linear Operator Theory in Engineering and Science*. Springer-Verlag, New York, NY, 1982.

[25] K.J. Palmer. Exponential dichotomies and transversal homoclinic points. *Journal of Differential Equations*, pages 225–256, 1984.

[26] K.J. Palmer. Exponential dichotomies and fredholm operators. *Proceedings of the American Mathematical Society*, pages 149–156, 1988.

[27] H. Poincaré. Sur les déterminants d'ordre infini. *Bull. Soc. Math. France*, 14:77–90, 1886.

[28] M. Reed and B. Simon. *Methods of Modern Mathematical Physics I, Functional Analysis*. Academic Press, New York, NY, 1972.

[29] H. Sandberg, E. Mollerstedt, and B. Bernhardsson. Frequency-domain analysis of linear time-periodic systems. *IEEE Trans. of Auto. Control*, 50:1971–1983, 2005.

[30] B. Sandstede. Stability of traveling waves. In B. Fiedler, editor, *Handbook of Dynamical Systems, VOL. 2*, pages 983–1055. Elsevier Science B.V., 2002.

[31] M. Schechter. *Principles of Functional Analysis*. American Mathematical Society, Providence, RI, 2nd edition, 2001.

[32] H. Von Koch. Sur les systèmes d'une infinité d'équations linéaires à une infinité d'inconnues. *C. R. Cong. d. Math. Stockholm*, pages 43–61, 1910.

[33] J. Zhou. Zeros and poles of linear continuous-time periodic systems: Definitions and properties. *IEEE Trans. of Auto. Control*, 53:1998–2011, 2008.

[34] J. Zhou, T. Hagiwara, and M. Araki. Spectral characteristics and eigenvalues computation of the harmonic state operators in continuous-time periodic systems. *Systems and Control Letters*, 53:141–155, 2004.