Three Stratified Fluid Models: Benjamin-Ono, Tidal Resonance, and Quasi-Geostrophy

by

Alfredo N. Wetzel

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied and Interdisciplinary Mathematics)
in The University of Michigan
2015

Doctoral Committee:

Professor Peter D. Miller, Co-Chair
Assistant Professor Brian K. Arbic, Co-Chair
Professor Jinho Baik
Professor John Boyd
Professor Sijue Wu

To my wife, my mother, and the memory of my father.

ACKNOWLEDGEMENTS

I am personally and professionally indebted to my advisors Professor Peter D. Miller and Professor Brian K. Arbic, for their patience, guidance, and support. This manuscript would not only not have been possible without their assistance, but I am certain that the research leading to it would not have been as enjoyable, challenging, or worthwhile.

I am very grateful to Professor John P. Boyd, Professor Jinho Baik, and Professor Sijue Wu for agreeing to be on my thesis committee. I truly appreciate their help in the development of the content of this manuscript. I would particularly like to thank Professor Smadar Karni for her support and unjustified kindness throughout my time at the University of Michigan. I would also like to thank the many professors of the Mathematics department at the University of Michigan whose doors were always open (even if figuratively) and created the very welcoming atmosphere evident at the department.

Lastly, I would like to thank the staff of the Mathematics department who did all of what they could to help me with my administrative duties, in spite of my best efforts to not answer e-mails or co-operate.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

LIST OF ABBREVIATIONS

BO Benjamin-Ono

DSW Dispersive Shock Wave

HIM Hallberg Isopycnal Model

HYCOM HYbrid Coordinate Ocean Model

ILW Intermediate Long Wave

IST Inverse-Scattering Transform

KdV Korteweg-de Vries

QG Quasi-Geostrophic

rKS rational Klaus-Shaw

RMS Root Mean Square

SQG surface Quasi-Geostrophy

CHAPTER I

Introduction

Stratified fluid models are pivotal in understanding numerous physical phenomena such as the dynamics of the ocean and atmosphere. Ocean dynamics can be understood approximately as consisting of a buoyant upper layer overlying a more dense and stable lower layer which encompasses most of the water column. For these reasons, much of the ocean may be crudely modeled as a two-layer system of warmer, lighter water resting over colder, denser water. The story is much the same for the lower atmosphere; except that the stable layer lies atop a bottom boundary layer. Most dramatically, in the troposphere, it is the dynamics between these layers which leads to the formation of clouds and ultimately weather. Of course, even at this level of simplification the complexity of the flow patterns is daunting. It is thus customary to restrict the dynamics even further to focus on different regimes of interest, e.g., western boundary currents, large-scale circulation, internal waves, tides, etc.

In this dissertation, we introduce three seemingly disparate topics with wide ranging applications: the Benjamin-Ono (BO) equation in Chapter II, a two-layer tidal model in Chapter III, and the Quasi-Geostrophic (QG) equations in Chapter IV. Each of these topics may be viewed, however, as a special regime of inviscid, incompressible, stably stratified fluid dynamics. The BO equation describes the evolution of the internal interface of a one-dimensional, dispersive, two-layer fluid bounded above by a rigid lid whose infinitely deep bottom layer is irrotational. In particular, the BO equation has been used to model internal

waves in deep water [15], the atmospheric roll cloud wave-train known as the morning glory [65], and nonlinear Rossby waves in shear flow [64], to name a few examples. Beyond fluid dynamics, the BO equation has been observed to model the spectral dynamics of incoherent shocks in nonlinear optics [27]. Our tidal model describes the effects of astronomical forcing on the tidal elevations in a two-layer, one-dimensional, linearized, non-rotating, damped, shallow water system with arbitrary topography and basin width. The two-layer tidal model is intended to provide a simple model to study the effect that stratification has on the surface tidal elevation. A similar model was used by Arbic, Karsten, and Garrett [6] to model resonant interactions of an ocean basin coupled with a shelf. Finally, the QG equations describe the circulation of turbulent nearly two-dimensional eddies in the atmosphere and ocean, under the effects of the earth's rotation, in hydrostatic balance and near geostrophic balance. The equations essentially describe perturbations from exact geostrophic flow and are widely used to model synoptic scale (scale of large-scale eddying motion: $\sim 1,000$ km in the atmosphere and $\sim$100 km in the ocean [81, Table 2.1]) phenomena in both the atmosphere and ocean. In this dissertation, we will be considering mostly a multi-level QG model. Figure 3.3 provides cartoon sketches of the three models described above (BO, QG, and tidal).

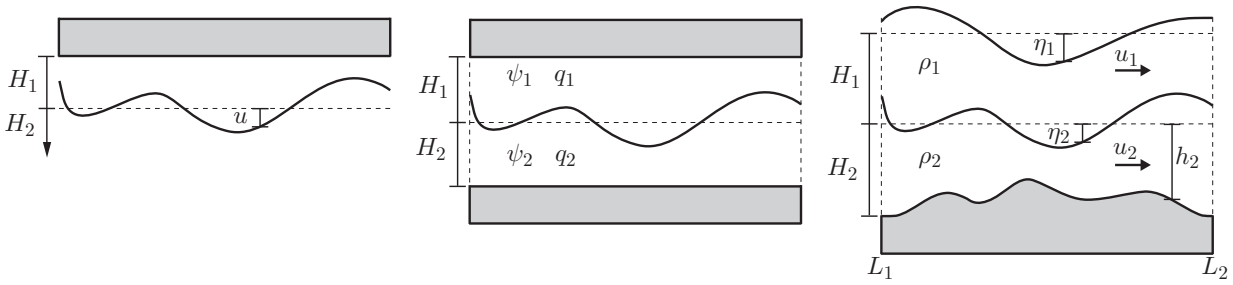

Figure 1.1: Sketch of BO model (left), QG model (middle), and two-layer tidal model (right). The labels H_1 and H_2 correspond to the resting depths of the upper and lower layers. Further explanations of labels in the figures can be found in the individual chapters.

In the remainder of this chapter we introduce each topic in more detail to discuss our motivation and goals.

1.1 The Benjamin-Ono Equation

The initial value problem of the BO equation

$$\frac{\partial u}{\partial t} + 2u\frac{\partial u}{\partial x} + \epsilon\mathcal{H}\left[\frac{\partial^2 u}{\partial x^2}\right] = 0, \quad -\infty < x < \infty, \quad t > 0 \tag{1.1}$$

describes the weakly nonlinear evolution of one-dimensional internal gravity waves in a stratified fluid [9, 17, 63], where u corresponds to the wave profile, $\epsilon > 0$ is a measure of the effects of dispersion, and the operator $\mathcal{H}$ denotes the Hilbert transform defined by the Cauchy principal value integral

$$\mathcal{H}[u](x,t) := \frac{1}{\pi}\fint_{-\infty}^{\infty} \frac{u(\xi,t)}{\xi - x}\, d\xi. \tag{1.2}$$

Even though the BO equation has not yet been satisfactorily understood from the point of view of inverse scattering techniques, it serves as a significant example of a nonlocal *integrable* evolution equation; for a treatment of the initial value problem using dispersive partial differential equation techniques the reader is referred to [79]. As such the BO equation has many useful properties: it arises from a Lax pair [12] (cf. (2.1) and (2.2)), it has nontrivial Bäcklund transformations [62], it possesses infinitely many soliton solutions [14], and it belongs to an infinite hierarchy of Hamiltonian flows (infinitely many constants of motion in involution) [14, 25], to name a few. Perhaps most importantly, the BO equation serves as a conceptual bridge between one- and multi-dimensional integrable systems, where nonlocality is a common feature. This nonlocality appears both in the physical domain through the Hilbert transform and also in the inverse scattering transform domain through a nonlocal "jump" condition in the associated Riemann-Hilbert problem (cf. Riemann-Hilbert Problem 1). Fortunately, for the BO equation this Riemann-Hilbert problem is a scalar one, presumably simplifying the subsequent analysis.

Remark I.1. For the values $\nu/\epsilon \in \mathbb{N}$ equation (1.3) below is a special case of a multisoliton solution of the BO equation; see equation (5) in [14], which solves a differently scaled version

of the BO equation (1.1) with $\epsilon = 1$, or Section 2.3 in [90] for a discussion including ϵ.

The Cauchy problem of the BO equation (1.1) is solvable, for suitably decaying initial conditions $u_0(x)$ as $|x| \to \infty$, by an Inverse-Scattering Transform (IST) procedure associated to a spectral problem (cf. (2.1)) in which u_0 appears as a potential. In Section 2.2, we construct explicitly the Jost solutions and "bound state" eigenfunctions for this spectral problem with generic (simple poles) rational initial data u_0, and are thus able to deduce the corresponding scattering data. This work was inspired by a brief note in a paper of Kodama, Ablowitz, and Satsuma [43, Section 7] on the Intermediate Long Wave (ILW) equation from which BO arises as a limiting case. In [43], the spectral problem (2.1) with $\epsilon = 1$ is analyzed for a "Lorentzian" potential

$$u_0(x) = \frac{2\nu}{x^2+1} = -\frac{i\nu}{x+i} + \text{c.c.} \quad \text{for} \quad \nu \in \mathbb{R} \tag{1.3}$$

and it is shown that the eigenvalues are given, for integer ν, by the roots of the Laguerre polynomial of degree ν. Some complications and essential difficulties in extending this remarkable calculation to more general initial conditions were highlighted by Xu [90, Section 2.3].

In the small-dispersion limit ($\epsilon \to 0$), numerical experiments indicate that, for t sufficiently small (independent of ϵ), the solution of the BO equation with smooth ϵ-independent initial data u_0 is well approximated by the solution of the inviscid Burgers equation (Hopf equation; equation (1.1) with $\epsilon = 0$). At the advent of a shock in this small-dispersion regime, the solution of (1.1) is then regularized by the formation of a Dispersive Shock Wave (DSW). This DSW has the well-known structure of an $O(1)$-amplitude modulated periodic traveling wave with $O(\epsilon)$ wavelength. In the DSW region the solution of the BO equation may be formally approximated using Whitham modulation theory; see Section 2.1.5. Surprisingly, and unlike the case of the seemingly simpler Korteweg-de Vries (KdV) equation, the modulation equations for the BO equation are fully uncoupled [19], consisting of several independent copies of the inviscid Burgers equation with corresponding matching conditions. A proce-

dure to match the Whitham modulation approximation for the DSW onto inviscid Burgers solutions outside the DSW was developed and applied by Matsuno [51, 52] and Jorge, Minzoni, and Smyth [37] to analyze the Cauchy problem for (1.1). Partial confirmation of these results was given by Miller and Xu [55], who rigorously computed the weak limit of the solution of the Cauchy problem for (1.1) for a class of positive initial data using IST techniques and approximations of the scattering data motivated by the IST analysis, by developing an analogue for the BO equation of a method first invented for KdV by Lax and Levermore [46]. Avoiding the (a priori unjustified) approximation of the scattering data requires its careful analysis in the small-dispersion limit by direct means. In Section 2.3 we use the exact formulas obtained in Section 2.2 to study the asymptotic behavior of the scattering data for the BO equation (1.1) in the limit $\epsilon \to 0$.

1.2 Two-Layer Tidal Model

In Chapter III, we utilize a simplified two-layer tidal model (cf. (3.1)–(3.4)) to examine the effects that oceanic stratification and changes in oceanic stratification have on long term surface tidal trends. This work is primarily motivated by a desire to understand global realistic-domain numerical model results, briefly noted in [5, 75] and commented on in Section 3.1 of this dissertation, that demonstrate changes in the large-scale tides induced by the presence of stratification. Thus, stratification not only introduces the expected small-scale (often referred to as "baroclinic") tides [89], but also alters the large-scale tide. The large-scale tide is often referred to as the "barotropic" tide, and is associated with the depth-averaged flow. The global numerical simulations by themselves, however, do not explain why the stratification affects the large-scale tides. Instead, we use the global numerical simulations in part to motivate our simple analytical model, whose parameter space can be explored much more fully. We use the analytical model to examine the effect of deep ocean stratification and climatic perturbations to this stratification on both the large-horizontal-scale and small-horizontal-scale components of the surface tidal elevation. In addition, we use the

analytical model to determine key parameters affecting surface tidal elevations, and we compare sensitivities in our analytical model to sensitivities seen in the global realistic-domain numerical simulations. This work thus complements the extensive literature on barotropic-to-baroclinic tidal conversion, which usually takes the barotropic tide as a given and aims to determine perturbations from this base state. In contrast, here we solve directly for the barotropic (large-scale) tide as a function of several input parameters including stratification.

The effects of seasonal changes in stratification to the tides have been documented in models of coastal regions by, for instance, Kang, Foreman, Lie, Lee, Cherniawsky, and Yum [39], Müller [56], and Müller, Cherniawsky, Foreman, and von Storch [58]. Numerous studies have determined that surface tidal elevations have undergone secular (long term) changes at many coastal locations over the last century [13, 16, 21, 35, 57, 68, 69, 86, 87, 91]. As noted in several of these papers, in some locations, the secular changes in tides are of comparable magnitude to the secular changes in mean sea level. The fact that changes in stratification affect the large-scale tides suggests that climatic changes in oceanic stratification, over time scales varying from seasonal to centennial, may contribute to observed changes in tides.

For simplicity, the analytical model used in Chapter III assumes two-layer, linearized, non-rotating, shallow-water dynamics in one horizontal dimension. We assume astronomical forcing and allow for arbitrary bottom topography in a finite basin. We also allow for linear damping, which intends to mimic (parameterize) the energy loss due to breaking internal gravity waves in the deep ocean. For convenience, all numerical results of the analytical model utilize a topography consisting of a Gaussian bump in the center of the basin. Small-scale tides are generated by large-scale tidal flow over the topographic feature. We obtain solutions to the analytical model using both a Fourier series and a Neumann series expansion. Importantly, the Neumann series contains built-in large-horizontal and small-horizontal scale modes allowing both the tidal velocities and elevations to be naturally decomposed into these modes. This analytical model is thus an extension of the classic model of half-wavelength resonance in a basin with two closed boundaries [18, 66] that includes two-layer stratification,

astronomical forcing, damping, and bottom topography. Our analytical model is meant to be as simple as possible while still capturing impacts of stratification, damping, and topography on large- and small-scale tides in a closed basin meant to roughly represent conditions in the global stratified ocean.

The key sensitivity found in the analytical model arises from the fact that stratification introduces a perturbation to the large-scale gravity wave speeds or, equivalently, wavelength [28]. While this phase speed perturbation is small in and of itself, nevertheless, relatively large sensitivities are seen in the analytical model as a function of this perturbation, when appropriate conditions of bottom topography and damping are taken into account. Notably, the introduction of stratification requires the model user to choose whether damping acts on the bottom-most layer, the top-most layer, or both, and this choice significantly affects model behavior. Changes in the large-scale surface tidal elevation are of the same order as changes in the small-scale (baroclinic) response. The sensitivity to phase speed perturbations may be used to interpret our preliminary global numerical results, and also to suggest a suite of new numerical simulations to be performed in follow-up work.

Our study is intended to be general in scope and is not focused on particular regions or time scales. Related regional studies include [39], which examined seasonal changes in stratification and tides in the Yellow and East China Seas, and [16], which examined the impact of internal tides on the secular variations in surface tidal elevations recorded in Hawai‘ian tide gauges. Another related study is that of Müller [56], who used an idealized model to study the effects of seasonally varying stratifications on coastal tides. Müller [56], however, considered changes in eddy viscosities rather than gravity wave speeds and focused on the barotropic transport rather than on large-scale tidal elevations. Lastly, the study in [58] was global in scale, but focused on seasonal changes which are small in the open-ocean.

1.3 Forced-Dissipated Quasi-Geostrophy

Many interesting oceanic phenomena take place in the near balance of gravity with the vertical pressure gradient (*hydrostatic balance*) and the Coriolis force with the horizontal pressure gradient (*geostrophic balance*). The quasi-geostrophic (QG) equations arise as a simplification of the three-dimensional equations of oceanic motion in the case when the hydrostatic balance is exact and the horizontal scales are roughly the size of the Rossby *deformation radius* L_d; the scale at which the effects of the earth's rotation become dynamically important. At mid-latitudes, in a stratified ocean, $L_d \approx 100$ km. With these assumptions, we neglect effects that are considered relatively unimportant in mid-scale planetary fluid motion. Namely, we remove gravity waves and sound waves from our analysis and tease out the dynamics of large-scale (synoptic scale) fluid phenomena [81].

Though strictly three-dimensional, the QG model behaves much like a two-dimensional system in one important way. The QG system may cascade energy to large scales [4, 72]; analogously to the well known result of two-dimensional turbulence obtained from Kolmogorov theory [81]. That is, energy input at an imposed forcing scale is moved by nonlinear interactions to smaller wavenumbers. The QG system, thus, may slowly develop eddies of length scales larger than those of the original forcing. This process is known as an inverse cascade and at a basic level it is driven by the conserved quantities of the QG system. The effects of boundary conditions on the system may, however, strongly influence the dynamics of the system. Namely, strong bottom friction in a two-layer QG system arrests the inverse cascade of energy by significantly damping the dynamics of the bottom layer [4]. This suggests that the parametric dependence of the QG system on boundary data may induce dynamics fundamentally different from those of either two- or three-dimensional turbulent flow.

In Chapter IV, we investigate the effects of bottom friction on the oceanic eddies and the energy budget in forced-dissipated QG dynamics with multiple layers. In part, we accomplish this study by using a modal decomposition recently proposed by Smith and Vanneste [77] that has been well known in Classical Mechanics. The decomposition is naturally suited

to handle the addition of more realistic boundary conditions that include the advection of buoyancy, and hence is well suited to aid in the description of modal QG dynamics. In this dissertation, we describe some of the mathematical richness of this method and elaborate on the discussion given by Smith and Vanneste [77]. Lastly, we generalize some results on energy cascades found by Arbic, Flierl, and Scott [4].

1.4 Notation

In this dissertation, we intend for the definitions of variables and functions to be consistent only within each major topic. For instance, the definition of a variable or function in the BO will only apply in the sections concerning BO. This arrangement should hopefully not cause much confusion as the contextual usages of the functions will be quite different between topics and, moreover, some repetition of letters seems inevitable.

Throughout the dissertation we will denote complex conjugation with the use of an asterisk, e.g., for the function $\varphi(x)$ the complex conjugate is denoted $\varphi(x)^*$. The usage of an over line, $\overline{\varphi}$, does not denote conjugation and is merely a symbol; its usage should be clear from context. For instance, in Chapter II we use the over line merely as a distinctive mark, while in Chapter IV we use the over line to denote the vertical mean. We use capital boldface roman letters, $\mathbf{A}$, to denote matrices and lower case boldface roman, $\mathbf{b}$, to denote vectors. We use the symbol $+(-)$ as a superscript to denote a function that has an analytic continuation into the upper(lower)-half x-plane (the physical variable), while a subscript denotes analytic continuation into the upper(lower)-half λ-plane (the spectral variable). Thus, $\varphi^+(x;\lambda)$ denotes the boundary value of a function analytic in the upper-half x-plane for a given value of λ, while $\varphi_-(x;\lambda)$ denotes the boundary value of a function analytic in the lower-half λ-plane for a given value of x. Lastly, we use Log to denote the principal value of the complex logarithm, i.e., the branch such that $\mathrm{Im}\{\mathrm{Log}(z)\} \in (-\pi, \pi)$. This is the only form of the logarithm we will use. Thus, z^p always denotes $e^{p\mathrm{Log}(z)}$ when p is not an integer.

CHAPTER II

Direct Scattering for the Benjamin-Ono Equation with Rational Initial Data

The contents of this chapter have been submitted for publication [54, 53] and are presented here with modifications and additions to improve readability.

2.1 Background

In this section we briefly introduce relevant previous work on the Benjamin-Ono (BO) equation (1.1). The reader familiar with the IST and modulational solutions of the BO equation may skip directly to Section 2.2.

2.1.1 The Lax Pair

The BO equation (1.1) arises as the compatibility condition of the Lax Pair

$$i\epsilon\frac{\partial w^+}{\partial x} + \lambda(w^+ - w^-) = -uw^+, \tag{2.1}$$

$$i\frac{\partial w^\pm}{\partial t} - 2i\lambda\frac{\partial w^\pm}{\partial x} + \epsilon\frac{\partial^2 w^\pm}{\partial x^2} - 2i\mathcal{C}^\pm\left[\frac{\partial u}{\partial x}\right]w^\pm = -\rho w^\pm \tag{2.2}$$

for $x \in \mathbb{R}$ and a spectral parameter $\lambda \in \mathbb{C}$ [12, 62]. In other words, the system (2.1)–(2.2) above for $w^\pm$ is consistent if and only if the potential $u(x,t)$ is a solution to the BO equation

(1.1). The superscripts $\pm$ denote the boundary values taken by the function $w(x)$, analytic on $\mathbb{C}\backslash\mathbb{R}$, from the upper- and lower-half x-plane on the real line. The parameter ρ is an arbitrary constant that in practice is chosen differently for different simultaneous solutions in order to ensure a simple time dependence of the scattering data.

The Cauchy operators $\pm\mathcal{C}^{\pm}$ are bounded operators on $L^2(\mathbb{R})$ and are complementary orthogonal projections from $L^2(\mathbb{R})$ onto its Hardy subspaces $\mathbb{H}^{\pm}(\mathbb{R})$ of functions analytic in the upper $(+)$ and lower $(-)$ half-planes, respectively. That is, they satisfy the identities

$$\mathcal{C}^+ \circ \mathcal{C}^- = \mathcal{C}^- \circ \mathcal{C}^+ = 0, \quad \mathcal{C}^+ \circ \mathcal{C}^+ = \mathcal{C}^+, \quad (-\mathcal{C}^-) \circ (-\mathcal{C}^-) = -\mathcal{C}^-. \tag{2.3}$$

They may be defined by the singular integrals

$$\mathcal{C}^{\pm}[u](x) = \lim_{\delta \downarrow 0} \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{u(y)}{y - (x \pm \delta i)}\, dy. \tag{2.4}$$

In addition, the Cauchy operators may be similarly written using the *Sokhotski formulas* as

$$\pm\, \mathcal{C}^{\pm} = \frac{1}{2}\mathbb{I} \pm \frac{1}{2i}\mathcal{H}. \tag{2.5}$$

In the Fourier domain $\pm\mathcal{C}^{\pm}$ may be equivalently defined as multiplication by the characteristic functions of complementary half-lines. Immediately from these results we have the identities

$$\mathcal{C}^+ - \mathcal{C}^- = \mathbb{I} \quad \text{and} \quad \mathcal{C}^+ + \mathcal{C}^- = -i\mathcal{H}, \tag{2.6}$$

where the first identity is called the *Plemelj formula*. We will also use the fact that the operators $\mathcal{C}^{\pm}$ commute with differentiation with respect to x. For details on singular integrals in general, and the Cauchy operators in particular, we refer the interested reader to the books by Gakhov [26] and Muskhelishvili [61].

Remark II.1. Throughout this dissertation the Cauchy operators $\mathcal{C}^{\pm}$ will only be applied

to functions of the spatial variable $x \in \mathbb{R}$ at fixed time t; see for example (2.2) and (2.4). The appearance of the Cauchy operators in the Lax pair arises directly from the nonlocality of the BO equation (1.1) in the x variable.

The fact that the BO equation (1.1) is the compatibility condition for the Lax pair (2.1)–(2.2) can be seen easily by rewriting the Lax pair in the form

$$\lambda w^- = A w^+ \quad \text{and} \quad \frac{\partial w^\pm}{\partial t} = B^\pm w^\pm, \tag{2.7}$$

for the operators

$$A := i\epsilon \frac{\partial}{\partial x} + \lambda + u \quad \text{and} \quad B^\pm := 2\lambda \frac{\partial}{\partial x} + i\epsilon \frac{\partial^2}{\partial x^2} + 2\mathcal{C}_\pm \left[\frac{\partial u}{\partial x} \right] + i\rho. \tag{2.8}$$

From the Plemelj formula (2.6) we may note that

$$B^+ = B^- + 2\frac{\partial u}{\partial x} \quad \text{and} \quad AB^+ = B^+ A - 2\lambda \frac{\partial u}{\partial x} - 2i\epsilon \frac{\partial u}{\partial x} \frac{\partial}{\partial x} + \epsilon \mathcal{H} \left[\frac{\partial^2 u}{\partial x^2} \right]. \tag{2.9}$$

Then, we may use equations (2.7) and identities (2.9) to write

$$0 = \left(\frac{\partial}{\partial t} - B^- \right) \lambda w^- = \left(\frac{\partial}{\partial t} + 2\frac{\partial u}{\partial x} - B^+ \right) A w^+ = \left(\frac{\partial u}{\partial t} + 2u \frac{\partial u}{\partial x} + \epsilon \mathcal{H} \left[\frac{\partial^2 u}{\partial x^2} \right] \right) w^+, \tag{2.10}$$

after some manipulation.

2.1.2 The Direct Scattering Problem

Solving the direct scattering problem of the BO equation requires, in essence, the solution of equation (2.1) of the Lax pair for fixed time t, and the determination of how the solutions depend on the spectral parameter λ. Indeed, it is enough to consider the spectral problem (2.1) at time $t = 0$ with

$$u(x, 0) = u_0(x); \tag{2.11}$$

see Section 2.1.3. This procedure was first proposed in a general setting by Fokas and Ablowitz [24]. Subsequently, Kaup and Matsuno [40] complemented and simplified this construction in the case that the initial condition is known to be a real function. We sketch out the most salient points of this process here for reference.

To solve equation (2.1) it is convenient to first recast it with the use of the projection operators $\pm\mathcal{C}^{\pm}$. Namely, using the Plemelj formula (2.6) and Liouville's theorem of complex analysis we can write (2.1) as a set of two equations:

$$i\epsilon\frac{\partial w^+}{\partial x} + \lambda\left(w^+ - w_0\right) = -\mathcal{C}^+[u_0 w^+] \tag{2.12}$$

and

$$\lambda\left(w^- - w_0\right) = -\mathcal{C}^-\left[u_0 w^+\right], \tag{2.13}$$

where w_0 is a suitable constant [90]. According to equation (2.13), the constant w_0 can be understood as the the limit of the function w^- as $x \to \infty$ in the closed lower-half x-plane. This simplification allows us to forgo equation (2.1) involving both w^+ and w^-, as it is now enough to study the spectral theory of (2.12) which only involves the function w^+.

Traditionally, the spectral analysis of (2.12) is accomplished using the so-called *Jost solutions* — sometimes referred as Jost functions or "left"/"right" eigenfunctions (depending on the boundary information) — to construct the scattering data of the problem. The Jost solutions are certain solutions of (2.12) (analytic and bounded for $\mathrm{Im}\{x\} > 0$) that are well-defined for $\lambda > 0$; in the literature [1, 24, 40, 73] they are denoted by $w^+ = M(x;\lambda)$ and $\overline{N}(x;\lambda)$ (for $w_0 = 1$) and $w^+ = \overline{M}(x;\lambda)$ and $N(x;\lambda)$ (for $w_0 = 0$). They may be characterized for fixed $\lambda > 0$ via their asymptotic behavior for large real x as follows:

$$\begin{aligned} M \to 1,\ \overline{M}e^{-i\lambda x/\epsilon} \to 1 \quad \text{as} \quad x \to -\infty, \\ \overline{N} \to 1,\ Ne^{-i\lambda x/\epsilon} \to 1 \quad \text{as} \quad x \to +\infty. \end{aligned} \tag{2.14}$$

The Jost solutions solve Fredholm integral equations of the second kind [24].

Importantly, while the Jost solutions are initially defined only for $\lambda > 0$, M and $\overline{N}$ have analytic extensions into the complex λ-plane while in general $\overline{M}$ and N do not [24]. Indeed, for each fixed x with $\mathrm{Im}\{x\} \geq 0$, M and $\overline{N}$ can be shown to be the boundary values on $\mathbb{R}^+$ from the upper- and lower-half λ-planes respectively of a single function W meromorphic on $\mathbb{C} \setminus \mathbb{R}^+$. The function W has as its only singularities in λ on $\mathbb{C} \setminus \mathbb{R}^+$ a discrete set of simple poles located precisely at the eigenvalues $\lambda = \lambda_j < 0$; see [2, Theorem 2.1] and [88]. We will henceforth use the notation W_+ (resp., W_-) to denote the Jost solution M (resp., $\overline{N}$). The auxiliary Jost solutions $\overline{M}$ and N will not appear beyond this section of the dissertation.

It is then possible to show that W_+, W_-, and N satisfy the *scattering relation*

$$W_+(x;\lambda) - W_-(x;\lambda) = \beta(\lambda) N(x;\lambda), \quad \lambda \in \mathbb{R}^+, \tag{2.15}$$

determining a function $\beta : \mathbb{R}^+ \longrightarrow \mathbb{C}$ (independent of x) called the *reflection coefficient*. From (2.15) one may solve for the reflection coefficient in terms of $W_\pm$ by using the Jost solutions' asymptotic behavior (2.14) for large real x; see (2.84). For the inverse theory, it is then important that $N(x;\lambda)$ may be fully eliminated from (2.15) to yield a *nonlocal jump condition* relating the boundary values $W_\pm$ of W across its branch cut $\mathbb{R}^+$. In this way a scattering relation linking $W_+(x;\lambda)$ and $W_-(x;\lambda)$ directly is obtained:

$$W_+(x;\lambda) = W_-(x;\lambda) + \beta(\lambda) e^{i\lambda x/\epsilon} \int_0^\lambda \frac{\beta(k)^*}{2\pi i k} e^{-ikx/\epsilon} W_-(x;k)\, dk \quad \text{for} \quad \lambda \in \mathbb{R}^+. \tag{2.16}$$

Since the functions $W_\pm$ represent boundary values of a meromorphic function W on $\mathbb{C} \setminus \mathbb{R}^+$, the condition above can be interpreted as a *jump condition* across the branch cut $\mathbb{R}^+$ of W.

Remark II.2. Fokas and Ablowitz [24] originally derived a form of equation (2.16) that required knowledge of an additional function of $\lambda > 0$, not merely the reflection coefficient. With the help of an "adjoint" spectral problem, Kaup and Matsuno [40] subsequently showed that for real potentials u_0 this additional function is related — essentially through complex

conjugation — to the reflection coefficient β.

For the remainder of the scattering data, a set of phase constants γ_j associated with the negative poles $\lambda = \lambda_j$ of $W(x;\lambda)$, we introduce the "bound state" *eigenfunction* $w^+ = \Phi_j(x) \in \mathbb{H}^+(\mathbb{R})$ satisfying (2.12) with $w_0 = 0$ for a given *eigenvalue* $\lambda = \lambda_j < 0$ and normalized by the condition

$$x\Phi_j(x) \to 1 \quad \text{as} \quad |x| \to \infty \quad \text{(uniformly for } \mathrm{Im}\{x\} \geq 0), \tag{2.17}$$

or equivalently, as can be shown asymptotically from (2.12),

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} u_0(x)\Phi_j(x)\, dx = \lambda_j. \tag{2.18}$$

The meromorphic function W then has a Laurent expansion about each eigenvalue λ_j of the form:

$$W(x;\lambda) = -i\epsilon \frac{\Phi_j(x)}{\lambda - \lambda_j} + (x + \gamma_j)\, \Phi_j(x) + O(\lambda - \lambda_j), \tag{2.19}$$

where $\gamma_j \in \mathbb{C}$ is the *phase constant* associated with the eigenvalue λ_j. In addition, if (2.17) holds in the sense that

$$\Phi_j(x) = \frac{1}{x} + O\left(\frac{1}{x^2}\right) \quad \text{as} \quad |x| \to \infty \quad \text{(uniformly in the closed upper-half } x\text{-plane)}, \tag{2.20}$$

it is possible to explicitly define the phase constants γ_j in terms of the eigenfunctions [40] of the problem as

$$\gamma_j := \frac{\epsilon}{2\pi\lambda_j} \int_{-\infty}^{\infty} \Phi_j(x)^* \left(x\Phi_j(x) - 1\right)\, dx, \tag{2.21}$$

or, appropriately interpreting the integral using the Cauchy principal value at $x = \infty$ and using (2.20),

$$\gamma_j := \frac{\epsilon}{2\pi\lambda_j} \int_{-\infty}^{\infty} x\, |\Phi_j(x)|^2\, dx - \frac{i\epsilon}{2\lambda_j}. \tag{2.22}$$

Formula (2.22) highlights the fact that only the real part of the phase constant γ_j is truly an independent piece of the scattering data; the imaginary part is fully determined by the respective eigenvalue. In passing we note that some care is required when surveying the literature since some sources use the notation γ_j to refer *only* to the real part of the phase constant presented here.

Remark II.3. The fact that the eigenvalues λ_j are negative is discussed in [2] and is essentially due to the self-adjointness of the operator $\mathcal{L} = -i\epsilon\partial/\partial x - \mathcal{C}^+ u_0 \mathcal{C}^+$ of the spectral problem (2.12).

Remark II.4. We bring up one general fact which is, to our knowledge, not mentioned in the literature. When the potential is an even function ($u_0(x) = u_0(-x)$ for $x \in \mathbb{R}$), the real part of every phase constant γ_j is identically zero. This follows from the fact that for these potentials the identity $\Phi_j(x) = \Phi_j(-x)^*$ for $x \in \mathbb{R}$ holds for every eigenfunction $\Phi_j(x)$ normalized by (2.17), as can be seen easily by taking the complex conjugate of equation (2.12) and making the substitution $x = -x$.

In this manner the scattering data for real potentials u_0 is constructed from the Jost solutions and eigenfunctions of (2.12). We list all the required scattering data here for clarity.

Table 1. Scattering Data of the Benjamin-Ono Lax Pair.

- Reflection coefficient $\beta(\lambda)$ for $\lambda \in \mathbb{R}^+$: defined by (2.15); see also (2.84).
- Eigenvalues $\{\lambda_j < 0\}_{j=1}^N$: determined from the spectral problem (2.12).
- Phase constants $\{\gamma_j\}_{j=1}^N$: defined in terms of the eigenvalues and the corresponding normalized eigenfunctions by (2.19) or, equivalently, (2.21).

Remark II.5. The phase constant γ_j associated to an eigenvalue λ_j can be equally obtained from (2.21)–(2.22) with the knowledge of the relevant eigenfunction or from the Laurent

series expansion of the function W as shown in (2.19). We will use both formulations in this dissertation.

Remark II.6. The number N, denoting the number of eigenvalues associated with u_0, here should not be confused with the Jost solution $N(x;\lambda)$. We do not require the function $N(x;\lambda)$ in the subsequent discussion and hence do not use it outside of this section.

2.1.3 Time Evolution of the Scattering Data

As is standard with inverse scattering theory — assuming that the potential u in equation (2.12) is no longer fixed but rather depends on t as $u = u(x,t)$ evolving according to the compatibility condition (the BO equation (1.1)) for the Lax pair — the time evolution of the scattering data can be deduced from the time equation in the Lax pair, namely, equation (2.2). As with other integrable systems, the time evolution of the scattering data for the BO equation is not merely explicit but also nearly trivial. The following results are well-known and can be found, for example, in [24]; we have only modified the formulas to include the dispersion parameter $\epsilon > 0$.

First, it is a fundamental property arising from the Lax pair that the eigenvalues λ_j are invariants of motion; see equation (2.10). That is,

$$\frac{d\lambda_j}{dt} = 0 \quad \text{for} \quad j = 1, \ldots, N. \tag{2.23}$$

Second, both functions W_+ and W_- must satisfy equation (2.2) independently for an appropriate choice of the constant ρ. Thus, using the asymptotics of these functions for large real x and appropriate manipulation yields

$$\beta(\lambda,t) = \beta(\lambda,0)e^{-i\lambda^2 t/\epsilon} \quad \text{for} \quad \lambda \in \mathbb{R}^+. \tag{2.24}$$

Third, using the fact that $\Phi_j(x)$ satisfies equation (2.2) we obtain

$$\gamma_j(t) = 2\lambda_j t + \gamma_j(0) \quad \text{for} \quad j = 1, \ldots, N. \tag{2.25}$$

2.1.4 The Inverse Scattering Problem

The last remaining step in the inverse scattering procedure is to reconstruct the potential $u(x,t)$ from the time-evolved scattering data. One manner to accomplish this involves the construction of the function $W(x;\lambda)$ through a Riemann-Hilbert problem. In this setting, we think of W as a function of the complex variable λ depending parametrically on x (and t via the time-evolution formulas (2.23)–(2.25) for the scattering data), and we therefore suppress the parameters and write $W = W(\lambda)$.

Riemann-Hilbert Problem 1. Find a complex valued function $W(\lambda)$ with the following properties.

1. **Analyticity:** $W(\lambda)$ is analytic for $\lambda \notin \mathbb{R}^+ \cup \{\lambda_1, \ldots, \lambda_N\}$. Each point $\{\lambda_j\}_{j=1}^N$ is a simple pole of W, and $W(\lambda)$ takes continuous boundary values $W_\pm$ on $\mathbb{R}^+$ from $\mathbb{C}^\pm$.

2. **Residue condition:** The residue at each simple pole λ_j of $W(\lambda)$ satisfies

$$\operatorname*{Res}_{\lambda=\lambda_j} W(\lambda) = -\frac{i\epsilon}{x+\gamma_j(t)} \frac{\partial}{\partial\lambda}\left[W(\lambda)(\lambda-\lambda_j)\right]\Big|_{\lambda=\lambda_j}. \tag{2.26}$$

3. **Jump condition:** The boundary values $W_\pm(\lambda)$ are related for $\lambda > 0$ by the nonlocal jump condition

$$W_+(\lambda) = W_-(\lambda) + \beta(\lambda) e^{i\lambda x/\epsilon} \int_0^\lambda \frac{\beta(k)^*}{2\pi i k} e^{-ikx/\epsilon} W_-(k)\, dk, \quad \text{where} \quad \lambda \in \mathbb{R}^+. \tag{2.27}$$

4. **Normalization:** $W(\lambda)$ is normalized at infinity;

$$W(\lambda) \to 1 \quad \text{as} \quad \lambda \to \infty. \tag{2.28}$$

Remark II.7. Riemann-Hilbert Problem 1 for W, though scalar, has a significant complication arising from the fact that the jump condition (2.27) is nonlocal. Nonetheless, some insight into this problem can be gained by considering the "reflectionless" and "eigenvalueless" limiting cases. Namely, when the scattering data is reflectionless ($\beta(\lambda) = 0$ for all $\lambda > 0$) the function W is meromorphic and a solution to Riemann-Hilbert Problem 1 is explicitly attainable through partial fraction expansions; this procedure reproduces the general N-soliton solution formula first obtained by Matsuno using the Hirota bilinear method [48]. Similarly, when the scattering data does not include eigenvalues (W analytic on $\mathbb{C}\backslash\mathbb{R}^+$), a solution can be sought using standard methods of singular integral equation theory [26, 61] (and the resulting integral equations can be solved by Neumann series if β is sufficiently small in a suitable sense).

Assuming that the Riemann-Hilbert problem can be solved uniquely for $W(\lambda) = W(x,t;\lambda)$, we can recover the function $u = u(x,t)$ that solves the initial-value problem for (1.1) as follows. From the asymptotic behavior of (2.12) as $\lambda \to \infty$ we can deduce that

$$\lambda\left(W(x,t;\lambda) - 1\right) = -\mathcal{C}^+[u](x,t) + o(1) \quad \text{as} \quad \lambda \to \infty. \tag{2.29}$$

In addition, when the solution $u(x,t)$ is real, Fokas and Ablowitz [24] pointed out that using the Plemelj formula (2.6) we can write

$$u(x,t) = \mathcal{C}^+[u](x,t) + \mathcal{C}^+[u](x,t)^* = 2\mathrm{Re}\left\{\mathcal{C}^+[u](x,t)\right\}. \tag{2.30}$$

These two equations combined give us a direct path to the reconstruction of $u(x,t)$. Namely, combining (2.29) and (2.30), the potential $u(x,t)$ can be found by computing

$$u(x,t) = \lim_{\lambda\to\infty} 2\mathrm{Re}\left\{\lambda(1 - W(x,t;\lambda))\right\}. \tag{2.31}$$

2.1.5 Modulation Theory for the Benjamin-Ono Equation

Modulation theory presents a formal approach to obtain oscillatory solutions of the BO initial-value problem in the zero-dispersion limit ($\epsilon \to 0$). We provide a brief discussion of this topic here for reference, to present a few mild generalizations, and for contrast with our rigorous results in Section 2.3. In Section 2.3, as a necessary first step to obtain rigorous solutions of BO using IST theory, we present rigorous analysis of the asymptotics of the BO scattering data.

The essence of Whitham's modulation theory [85] is to asymptotically describe the evolution of an oscillatory wave-train; it provides a nonlinear analogue of the WKB method. To accomplish this, the method exploits the separation of scales between the slowly varying envelope of oscillations and the "fast" oscillations themselves. The key observation for the applicability of this method is that in cases where there is such a clear separation of scales, quantities such as the wave-number, phase speed, and amplitude are locally well defined. The method then entails obtaining suitable evolution equations for the parameters describing the dynamics of the wave-train. In this section, we describe how Whitham's modulation theory applies to the BO equation using a variational principle and discuss some of the properties of this envelope of oscillations. Much of the information in this section can be found, with suitable modifications, in the work of Matsuno [51, 52] and Xu [90].

The BO equation (1.1) may be obtained from the variation $\delta F/\delta\varphi = 0$, under the replacement $u = \epsilon\partial\varphi/\partial x$, of the functional

$$F[\varphi] = \iint L\left(\frac{\partial\varphi}{\partial t}, \frac{\partial\varphi}{\partial x}; \epsilon\right)\, dx\, dt \tag{2.32}$$

with Lagrangian density L given by

$$L\left(\frac{\partial\varphi}{\partial t}, \frac{\partial\varphi}{\partial x}; \epsilon\right) = \frac{1}{2}\frac{\partial\varphi}{\partial t}\frac{\partial\varphi}{\partial x} + \frac{\epsilon}{3}\left(\frac{\partial\varphi}{\partial x}\right)^3 + \frac{\epsilon}{2}\frac{\partial\varphi}{\partial x}\mathcal{H}\left[\frac{\partial^2\varphi}{\partial x^2}\right]. \tag{2.33}$$

To derive modulation and envelope equations we initially assume that the potential function φ describes a uniform wave-train given by

$$\varphi(x,t;\epsilon) = \frac{\psi(x,t)}{\epsilon} + \Phi\left(\frac{\theta(x,t)}{\epsilon}\right), \tag{2.34}$$

where

$$\psi(x,t) = \beta x - \gamma t \quad \text{and} \quad \theta(x,t) = kx - \omega t. \tag{2.35}$$

Without loss of generality we assume that $k \geq 0$.

Remark II.8. The functions β and γ in this section should not be confused with the reflection coefficient and phase constants used in the rest of this dissertation.

It follows from direct computation that

$$\frac{\partial\psi}{\partial t} = -\gamma, \quad \frac{\partial\psi}{\partial x} = \beta, \quad \frac{\partial\theta}{\partial t} = -\omega, \quad \frac{\partial\theta}{\partial x} = k \tag{2.36}$$

and

$$u = \epsilon\frac{\partial\varphi}{\partial x} = \beta + k\Phi'\left(\frac{\theta}{\epsilon}\right), \tag{2.37}$$

where the prime on Φ denotes differentiation with respect to the argument. The function β should not be confused with the reflection coefficient from the previous section. A closed form solution of the BO equation may then be obtained by substituting equation (2.37) into (1.1) and solving for the 2π-periodic Φ'. Using Fourier expansions and some knowledge of the Hilbert transform we find that this solution takes the form

$$u = \beta + \frac{k^2}{\sqrt{H^2+k^2} - H\cos(\theta/\epsilon)}, \tag{2.38}$$

where

$$H = H(\gamma,\beta,\omega,k) := \frac{1}{2}\left(u_{\max} - u_{\min}\right) \tag{2.39}$$

and $u_{\max}$ ($u_{\min}$) represents the greatest (least) value that the solution $u(x,t)$ attains in the period. The phase speed of the wave can also be found to satisfy the condition

$$c = \frac{\omega}{k} = \sqrt{H^2 + k^2} + 2\beta. \tag{2.40}$$

Following Whitham's modulation theory, we now assume that the parameters γ, β, ω, and k (not all independent following equation (2.40)) are slowly varying and derive corresponding evolution equations. Immediately we may derive the equations

$$\frac{\partial \beta}{\partial t} + \frac{\partial \gamma}{\partial x} = 0 \quad \text{and} \quad \frac{\partial k}{\partial t} + \frac{\partial \omega}{\partial x} = 0 \tag{2.41}$$

for β and k from equality of mixed partials in (2.36). It was shown by Whitham [85] that evolution equations for γ and ω can be obtained from the Lagrangian density L averaged over the phase's period. Without loss of generality this period can be taken to be 2π and the expression for the averaged Lagrangian takes the form

$$\langle L \rangle := \frac{1}{2\pi} \int_0^{2\pi} L \, d\tilde{\theta}, \quad \tilde{\theta} = \theta/\epsilon. \tag{2.42}$$

For the Lagrangian (2.33), this expression can be shown [51, 52] with some work to be

$$\langle \epsilon^2 L \rangle = \frac{k^3}{12} - \frac{k}{4}\left(\frac{\omega^2}{k^2} - 2\beta\frac{\omega}{k} + 2\gamma\right) + \frac{\beta^3}{3} - \frac{\beta\gamma}{2}. \tag{2.43}$$

The evolution equations then take the form

$$\frac{\partial}{\partial t}\frac{\partial \langle L \rangle}{\partial \omega} - \frac{\partial}{\partial x}\frac{\partial \langle L \rangle}{\partial k} = 0 \quad \text{and} \quad \frac{\partial}{\partial t}\frac{\partial \langle L \rangle}{\partial \gamma} - \frac{\partial}{\partial x}\frac{\partial \langle L \rangle}{\partial \beta} = 0 \tag{2.44}$$

which explicitly are

$$\frac{\partial}{\partial t}\left(\beta-\frac{\omega}{k}\right)+\frac{\partial}{\partial x}\left(\frac{k^2}{2}+\frac{\omega^2}{2k^2}-\gamma\right)=0 \quad \text{and} \quad \frac{\partial}{\partial t}\left(k+\beta\right)+\frac{\partial}{\partial x}\left(\omega+2\beta^2-\gamma\right)=0. \tag{2.45}$$

However, the set of evolution equations (2.41) and (2.45) are not independent since the expression

$$\frac{\partial}{\partial t}\left(k+\beta\right)=\frac{\partial k}{\partial t}+\frac{\partial \beta}{\partial t}=-\frac{\partial}{\partial x}\left(\omega+\gamma\right), \tag{2.46}$$

which follows using (2.41), implies that the right hand side of (2.45) reduces to

$$\frac{\partial}{\partial x}\left(\beta^2-\gamma\right)=0. \tag{2.47}$$

Moreover, since only x-derivatives of γ appear in the evolution equations we may take without loss of generality that $\gamma=\beta^2$. So, after some manipulation we may write the evolution equations (2.41) and (2.45) simply as

$$\frac{\partial \beta}{\partial t}+2\beta\frac{\partial \beta}{\partial x}=0, \quad \frac{\partial a}{\partial t}+2a\frac{\partial a}{\partial x}=0, \quad \frac{\partial b}{\partial t}+2b\frac{\partial b}{\partial x}=0 \tag{2.48}$$

where $a=(c-k)/2$ and $b=(c+k)/2$ using the fact that $c=\omega/k$. Now, from the fact that $k\geq 0$ it follows that $c-k\leq c+k$ and from equation (2.40) we can deduce that $k+2\beta\leq c$. Thus, we must always have the relation

$$\beta\leq a\leq b \tag{2.49}$$

for the modulation parameters. This is essentially the same system of equations obtained with a different method by Dobrokhotov and Krichever [19] for a single phase asymptotic approximation.

Remark II.9. The modulation equations (2.48) may be contrasted with those that arise from the KdV equation in which case the modulation equations form a set of coupled tran-

scendental equations; see [84] or [85, Section 16.15].

For certain initial conditions in the small dispersion case, the modulation parameters β, a, b can be easily related to the solution of the BO equation with zero dispersion ($\epsilon = 0$; Hopf equation or inviscid Burgers' equation). Consider the initial condition

$$u(x,0) = u_0(x) \tag{2.50}$$

such that the solution to inviscid Burgers' equation has at most three branches and let us label the three branches $u_0^B \leq u_1^B \leq u_2^B$ in the multivalued region $x_l \leq x \leq x_r$; see Figure 2.1.

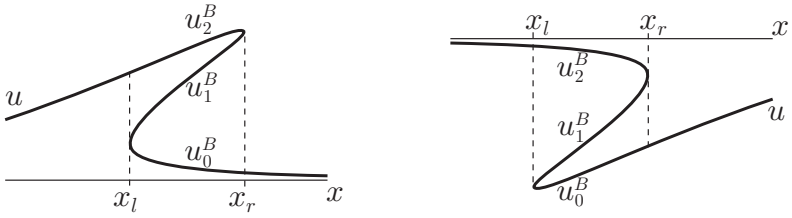

Figure 2.1: Labelling of the solution branches of the inviscid Burgers' equation in the multivalued region $x_l \leq x \leq x_r$.

In the multivalued region $x_l \leq x \leq x_r$, the small dispersion BO problem develops a DSW manifesting oscillations which are (formally) well approximated by the modulation equations of Whitham's method. Precisely, for $x_l \leq x \leq x_r$, the BO initial-value problem is approximated by (2.38), where the functions H, β, k, and θ are determined from the modulation equations (2.48) with the imposition of appropriate initial conditions. Outside of the DSW region the solution should asymptotically match the single valued inviscid Burgers' solution. In particular, at x_l we require for the envelope of oscillations to vanish and the mean value of u to match continuously with the inviscid Burgers' solution. Concisely this means that

$$H = 0 \quad \text{and} \quad \langle u \rangle = u_2^B \quad \text{at} \quad x = x_l. \tag{2.51}$$

This implies from equation (2.40) that at $x = x_l$ we have $a = \beta$, using the fact that

$\langle u \rangle = k + \beta = \beta - a + b$, $\langle u \rangle = u_2^B = b$ at $x = x_l$. In similar fashion, at x_r we require for the solution to approximate a solitary wave and for the mean value of u to match continuously with the inviscid Burgers' solution; or

$$k = 0 \quad \text{and} \quad \langle u \rangle = u_0^B \quad \text{at} \quad x = x_r. \tag{2.52}$$

This implies that $a = b$ and $\langle u \rangle = u_0^B = \beta$ at $x = x_r$.

To determine the evolution of the functions β, a, b in terms of the initial condition u_0, we proceed as follows. The solutions of the modulation equations may be written implicitly in the form

$$\beta = g_0(x - 2\beta t), \quad a = g_1(x - 2at), \quad b = g_2(x - 2bt), \tag{2.53}$$

for some arbitrary functions g_0, g_1, and g_2. From the asymptotic matching conditions discussed in the previous paragraph, at x_l we have

$$u_0(x_l - 2bt) = g_2(x_l - 2bt), \quad g_0(x_l - 2\beta t) = g_1(x_l - 2\beta t). \tag{2.54}$$

Similarly, at x_r we have the matching

$$u_0(x_r - 2\beta t) = g_0(x_r - 2\beta t), \quad g_1(x_r - 2at) = g_2(x_r - 2at). \tag{2.55}$$

Since t is arbitrary, these conditions put together imply that the functions $g_0(x), g_1(x), g_2(x)$ correspond to the sections of the initial condition $u_0(x)$ which give rise to the branches u_0^B, u_1^B, u_2^B. This matching guarantees the fact that

$$\beta = u_0^B, \quad a = u_1^B, \quad b = u_2^B. \tag{2.56}$$

We may equally write the modulation solution (2.38) in terms of the slowly varying parameters β, a, and b or branches u_0^B, u_1^B, and u_2^B. Using equation (2.40) and the fact that

the average of the modulation solution is given by $\langle u \rangle = \beta + k$, we may write the solution (2.38) as

$$u = \beta + k + \frac{k\left(k - (c-2\beta) + \sqrt{(c-2\beta)^2 - k^2}\cos(\theta/\epsilon)\right)}{c - 2\beta - \sqrt{(c-2\beta)^2 - k^2}\cos(\theta/\epsilon)}. \tag{2.57}$$

After some manipulating we may finally write

$$u = u_0^B - u_1^B + u_2^B + \frac{(u_2^B - u_1^B)\left(2u_0^B - 2u_1^B + 2\sqrt{(u_2^B - u_0^B)(u_1^B - u_0^B)}\cos(\theta/\epsilon)\right)}{u_1^B + u_2^B - 2u_0^B - 2\sqrt{(u_2^B - u_0^B)(u_1^B - u_0^B)}\cos(\theta/\epsilon)}, \tag{2.58}$$

where according to modulation theory the phase θ is found by solving the system of equations

$$\frac{\partial\theta}{\partial x} = u_2^B - u_1^B \quad \text{and} \quad \frac{\partial\theta}{\partial t} = (u_1^B)^2 - (u_2^B)^2. \tag{2.59}$$

Remark II.10. The equations (2.59) are compatible due to the equality of mixed partials and the fact that each u_i^B for $i = 0, 1, 2$ satisfies the Hopf (inviscid Burgers') equation.

Formula (2.58) is easily used to determine upper and lower bounds to the envelope of oscillations in the DSW region $x_l \leq x \leq x_r$. Namely, by noting that $\cos(\theta/\epsilon) = 1$ and $\cos(\theta/\epsilon) = -1$ give the upper and lower bounds of the expression for $u(x,t)$ we obtain

$$\begin{aligned} u_{\max,\min} &= u_0^B + \frac{(u_2^B - u_1^B)^2}{u_1^B + u_2^B - 2u_0^B \mp 2\sqrt{(u_2^B - u_0^B)(u_1^B - u_0^B)}} \\ &= u_0^B + \frac{\left(\sqrt{u_2^B - u_0^B} - \sqrt{u_1^B - u_0^B}\right)^2 \left(\sqrt{u_2^B - u_0^B} + \sqrt{u_1^B - u_0^B}\right)^2}{\left(\sqrt{u_2^B - u_0^B} \mp \sqrt{u_1^B - u_0^B}\right)^2}. \end{aligned} \tag{2.60}$$

In particular, some simplification gives

$$\begin{aligned} u_{\max} &= u_0^B + \left(\sqrt{u_2^B - u_0^B} + \sqrt{u_1^B - u_0^B}\right)^2 \\ &= u_0^B - u_1^B + u_2^B + 2\left(u_1^B - u_0^B\right) + 2\sqrt{(u_2^B - u_0^B)(u_1^B - u_0^B)} \end{aligned} \tag{2.61}$$

and

$$\begin{aligned} u_{\min} &= u_0^B + \left(\sqrt{u_2^B - u_0^B} - \sqrt{u_1^B - u_0^B}\right)^2 \\ &= u_0^B - u_1^B + u_2^B + 2\left(u_1^B - u_0^B\right) - 2\sqrt{(u_2^B - u_0^B)(u_1^B - u_0^B)}. \end{aligned} \tag{2.62}$$

We can thus quantify the size of the oscillations in the DSW region:

$$u_{\max} - u_{\min} = 2h = 4\sqrt{(u_2^B - u_0^B)(u_1^B - u_0^B)} \leq 4\left(u_2^B - u_0^B\right). \tag{2.63}$$

Lastly, at the left edge of the region of oscillations we have that $u_1^B(x_l, t) = u_0^B(x_l, t)$. This implies, not surprisingly, that

$$u_{\max}(x_l, t) = u_{\min}(x_l, t) = u_2^B(x_l, t). \tag{2.64}$$

Also, noting $u_2^B(x_r, t) = u_1^B(x_r, t)$ we have that

$$u_{\max}(x_r, t) = 4u_2^B(x_r, t) - 3u_0^B(x_r, t) \quad \text{and} \quad u_{\min}(x_r, t) = u_0^B(x_r, t). \tag{2.65}$$

In particular, for the entire region of oscillations $x_l \leq x \leq x_r$,

$$u_{\min}(x, t) \geq u_0^B(x, t). \tag{2.66}$$

This means that a solution $u(x, t)$ of the BO equation in the zero dispersion limit stays positive for all time if the initial condition is strictly positive. The equivalent statement for negative initial conditions is false.

2.2 Direct Scattering for Rational Potentials

In this section, we present our procedure to construct the scattering data for the BO equation with rational initial data. Specifically, we present an explicit construction procedure

for the Jost solutions and eigenfunctions of the associated Lax pair. We also derive an Evans function for the eigenvalues and formulas for the phase constants and reflection coefficient.

2.2.1 Rational Potentials

We consider real bounded rational potentials u_0 simple poles of the form:

$$u_0(x) = \sum_{p=1}^{P} \frac{c_p}{x - z_p} + \text{c.c.}, \tag{2.67}$$

where $\{c_p\}_{p=1}^P$ are nonzero complex numbers and the poles $\{z_p\}_{p=1}^P$ with $\text{Im}\{z_p\} > 0$ have distinct real parts increasing with p. We impose in addition the condition $\sum_{p=1}^P \left(c_p + c_p^*\right) = 0$, ensuring $u_0 \in L^1(\mathbb{R})$ and denote $I := \int_{-\infty}^{\infty} u_0(x)\, dx$.

Remark II.11. The requirement that the poles z_p have distinct real parts is imposed as a convenience. The procedure described here should be generalizable with some changes in notation.

Remark II.12. The number P of poles of u_0 in $\mathbb{C}^+$ *is not related in any simple way with the number* N *of discrete eigenvalues associated with* u_0. The number N of eigenvalues λ_j and phase constants γ_j is, however, dependent on the dispersion parameter ϵ. In fact,

$$N = N(\epsilon) \sim \frac{1}{2\pi\epsilon} \int\limits_{u_0>0} u_0(x)\, dx \tag{2.68}$$

as $\epsilon \to 0$ for P fixed; see Corollary II.52, [49], or [90, equation (2.139)].

In addition, let $f(x)$ be the particular anti-derivative of $u_0(x)$ given by

$$f(x) := \sum_{p=1}^{P} c_p \left(\text{Log}\left(i(x - z_p)\right) + \frac{\pi i}{2} \right) + \text{c.c.}, \quad f'(x) = u_0(x), \tag{2.69}$$

where $\text{Log}(\cdot)$ is the principal branch ($|\text{Im}\{\text{Log}(\cdot)\}| < \pi$).

Definition II.13. We denote by G the domain of analyticity of f as illustrated in Figure 2.2. Given $R > 0$ sufficiently large, we define subdomains G_p of $G \cap \{|x| > R\}$, $p = 1, \ldots, P-1$, consisting of points lying between the branch cuts emanating from the points z_p and z_{p+1}. Then we define G_0 as the half-plane $\mathrm{Re}\{x\} < \mathrm{Re}\{z_1\}$ and G_P as the half-plane $\mathrm{Re}\{x\} > \mathrm{Re}\{z_P\}$. The usage of an asterisk denotes the Schwarz reflection (reflection through the real axis) of the subdomain.

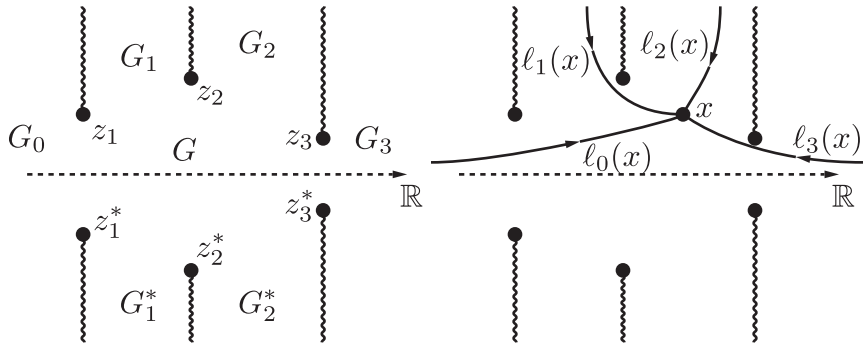

Figure 2.2: Left: a domain G with subdomains G_p and G_p^* in the x-plane for a potential u_0 of the form (2.67) with $P = 3$. The zig-zagged half-lines denote the (logarithmic) branch cuts of $f(x)$. Right: contours $\ell_m(x)$ in G originating at ∞ in the subdomains G_m and terminating at $x \in G$.

Lemma II.14. For any $m = 0, 1, \ldots, P$, we have

$$\lim_{x \to \infty \in G_m} f(x) = 2\pi i \sum_{p=1}^{m} c_p \quad \text{and} \quad \lim_{x \to \infty \in G_m^*} f(x) = -2\pi i \sum_{p=1}^{m} c_p^*. \tag{2.70}$$

Proof. This follows immediately from formula (2.69). ∎

Remark II.15. Note that $G_0 = G_0^*$ and $G_P = G_P^*$. In particular, if follows that $f \to 0$ as $x \to \infty$ in G_0 while $f(x) \to 2\pi i\,(c_1 + \cdots + c_P) = -2\pi i\,(c_1^* + \cdots + c_P^*) = I$ as $x \to \infty$ in G_P.

The advantage of considering potentials u_0 of the form (2.67), as first noted in [43], is that upon applying the Plemelj formula to the right-hand side of (2.12), the expression $\mathcal{C}^-[u_0 w^+]$

can be evaluated by residues because w^+ is analytic and bounded for $\mathrm{Im}\{x\} > 0$. Therefore,

$$i\epsilon \frac{dw^+}{dx} + \lambda w^+ + u_0 w^+ = \lambda w_0 + \sum_{p=1}^{P} \frac{c_p w^+(z_p;\lambda)}{x - z_p}. \tag{2.71}$$

As the x-dependence on the right-hand side is explicit and elementary this first-order equation can be solved in closed-form, at least assuming $w^+(z_p;\lambda)$ is known for all p.

2.2.2 The Jost Solutions and Reflection Coefficient

Let $\lambda > 0$. To obtain a formula for the Jost solutions $W_\pm(x;\lambda)$ we take $w_0 = 1$ in (2.71), apply the boundary condition $W_\pm(x;\lambda) \to 1$ as $x \to \mp\infty$, and integrate to obtain the convergent improper integral representations:

$$W_+(x;\lambda) = -\frac{i}{\epsilon} e^{ih(x;\lambda)/\epsilon} \int_{-\infty}^{x} e^{-ih(z;\lambda)/\epsilon} \left(\lambda + \sum_{p=1}^{P} \frac{v_p(\lambda)}{z - z_p} \right) dz \tag{2.72}$$

and

$$W_-(x;\lambda) = \frac{i}{\epsilon} e^{ih(x;\lambda)/\epsilon} \int_{x}^{\infty} e^{-ih(z;\lambda)/\epsilon} \left(\lambda + \sum_{p=1}^{P} \frac{w_p(\lambda)}{z - z_p} \right) dz, \tag{2.73}$$

where

$$h(x;\lambda) := \lambda x + f(x) \tag{2.74}$$

and

$$v_p(\lambda) := c_p W_+(z_p;\lambda), \quad w_p(\lambda) := c_p W_-(z_p;\lambda) \quad \text{for} \quad p = 1, \dots, P. \tag{2.75}$$

Note that by Lemma II.14, $h(x;\lambda)$ is dominated by λx for large $|x|$ provided $\lambda \neq 0$. Since $\lambda > 0$, the integral can therefore be made absolutely convergent by rotation of the contour at infinity into the lower-half x-plane.

Remark II.16. We will only work with the Jost solution W_+ in this dissertation. In fact, our results for W_+ hold with minor modifications for the Jost solution W_-.

It is important to note, and central to our theory, that while $W_\pm(x;\lambda)$ given by (2.72)–(2.73) satisfy the differential equation (2.71) and is analytic for all x near $\mathbb{R}$, extra conditions are required to ensure analyticity for all x with $\mathrm{Im}\{x\} > 0$. Indeed, the factor

$$e^{ih(x;\lambda)/\epsilon} = e^{i\lambda x/\epsilon} \prod_{p=1}^{P} (i(x-z_p))^{ic_p/\epsilon} \left(i^*(x-z_p^*)\right)^{ic_p^*/\epsilon} e^{\pi(c_p^*-c_p)/(2\epsilon)} \tag{2.76}$$

appearing in (2.72) may have singularities for $\mathrm{Im}\{x\} > 0$. Requiring analyticity of $W_+(x;\lambda)$ imposes constraints on the quantities $\{v_p(\lambda)\}_{p=1}^P$, thus eliminating them entirely from (2.72).

Definition II.17. $\ell_m(x)$ denotes any member of the equivalence class of contours in G originating at ∞ in the subdomain G_m and ending at the point $x \in G$; see Figure 2.2.

Definition II.18. $U_m^>$ denotes any member of the equivalence class of contours on the Riemann surface of f originating at $-i\infty$ in G_0 with value $f(\infty) = 0$ and ending at $-i\infty$ in G_0 having encircled the singularities $z_1, \ldots, z_m$ each exactly once in the positive sense; see Figure 2.3.

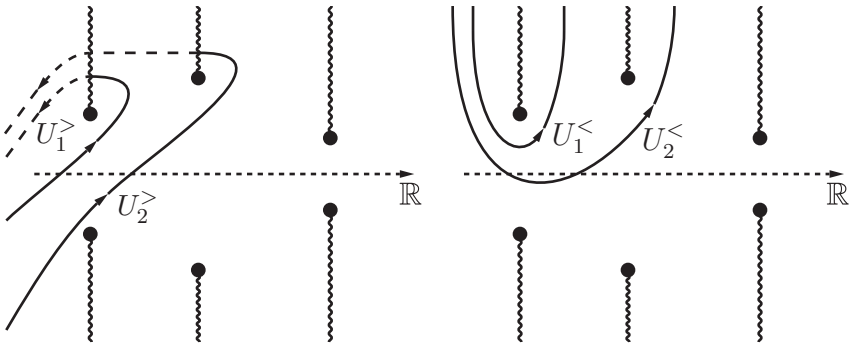

Figure 2.3: Left: contours $U_m^>$ for the configuration illustrated in Figure 2.2. Right: contours $U_m^<$ for the same configuration.

Remark II.19. We use the superscript $>$ ($<$) to denote an object defined specifically for use in the case $\lambda > 0$ ($\lambda < 0$).

Proposition II.20. Let $\lambda > 0$. For each $m = 1, \dots, P$, define the function $H_m = H_m(z;\lambda)$, $z \in G$, by the equation

$$(i(z - z_m))^{-ic_m/\epsilon - 1} H_m(z;\lambda) := -e^{-ih(z;\lambda)/\epsilon} \left(\lambda + \sum_{p=1}^{P} \frac{v_p(\lambda)}{z - z_p} \right). \tag{2.77}$$

(H_m is analytic not only on G, but also in a neighborhood of the point z_m.) Furthermore, let $ic_m = \epsilon(\mu_m + \omega_m)$, where $\mu_m = [\mathrm{Re}\{ic_m/\epsilon\}]$ denotes the integer (floor) part of $\mathrm{Re}\{ic_m/\epsilon\}$ and the remainder satisfies $0 \le \mathrm{Re}\{\omega_m\} < 1$. Then, the function $W_+(x;\lambda)$ given by (2.72) is analytic in the upper-half x-plane if and only if for each $m = 1, \dots, P$, either

$$\int_{\ell_0(z_m)} (i(z - z_m))^{-ic_m/\epsilon - 1} H_m(z;\lambda)\, dz = 0 \quad \text{when} \quad \mathrm{Re}\{ic_m/\epsilon\} < 0, \tag{2.78}$$

$$\int_{\ell_0(z_m)} (i(z - z_m))^{-\omega_m} \frac{d^{\mu_m+1}}{dz^{\mu_m+1}} H_m(z;\lambda)\, dz = 0 \quad \text{when} \quad \mathrm{Re}\{ic_m/\epsilon\} \ge 0 \text{ and } \omega_m \ne 0, \tag{2.79}$$

or

$$\operatorname*{Res}_{z=z_m} \frac{H_m(z;\lambda)}{(z - z_m)^{\mu_m+1}} = 0 \quad \text{when} \quad \mathrm{Re}\{ic_m/\epsilon\} \ge 0 \text{ and } \omega_m = 0, \tag{2.80}$$

where $\ell_0(z_m)$ (see Definition II.17) originates in at $-i\infty$ in G_0.

Proposition II.20 is proved in Appendix A.1. However, we describe the highlights here. Note that the singularity at $z = z_m$ arising from the exponential $e^{-ih(z;\lambda)/\epsilon}$ on the right-hand side of (2.77) is explicitly cancelled from $H_m(z;\lambda)$ by the factor $(i(z - z_m))^{-ic_m/\epsilon - 1}$ on the left-hand side. The essence of the proof is to deform the contour of integration in (2.72) to pass through the point z_m when the integrand is integrable there, i.e., $\mathrm{Re}\{ic_m/\epsilon\} < 0$. Thus, we deal separately with integrals on $\ell_0(z_m)$ and on the line from z_m to x. If the integrand is not integrable at z_m, we first integrate by parts repeatedly to achieve integrability, or invoke the Residue Theorem if the singularity is a pole.

It is useful to rewrite the conditions for analyticity of $W_+(x;\lambda)$ given in Proposition II.20 by eliminating the derivatives of H_m as follows.

Corollary II.21. Let $\lambda > 0$. The function $W_+(x;\lambda)$ defined by (2.72) is analytic in the upper-half x-plane if and only if $\{v_p(\lambda)\}_{p=1}^P$ satisfy

$$\int_{C_m^>} e^{-ih(z;\lambda)/\epsilon} \left(\sum_{p=1}^{P} \frac{v_p(\lambda)}{z - z_p} + \lambda \right) dz = 0, \quad m = 1, \dots, P, \tag{2.81}$$

where $C_m^> := U_m^>$ unless ic_m/ϵ is a strictly negative integer, in which case $C_m^> := \ell_0(z_m)$. (The contours $C_m^>$ extend to $-i\infty$ in G_0 making the integrals absolutely convergent.)

Remark II.22. In the general case that ic_m/ϵ is not a negative integer, the basic method of proof is to convert integrals over contours terminating at z_m for integrals over loops surrounding z_m, and this process relies on the existence of algebraic branching at z_m which is absent if ic_m/ϵ is a strictly negative integer. This explains the reason for requiring $C_m^> = \ell_0(z_m)$ valid in this exceptional case. While Corollary II.21 provides a nearly universal form for the analyticity conditions, the case that ic_m/ϵ is a nonnegative integer (or even better, if ic_m/ϵ is a nonnegative integer for all $m = 1, \dots, P$) is obviously exceptional in a different way, as the integral in (2.81) may be evaluated by residues. The resulting form of the condition (see (2.80), from which (2.81) originates in this special case) may be more useful, especially if a closed-form expression is preferable.

Proof. Let L_m be a contour on the Riemann surface of f beginning (with value $f(\infty) = 0$) and ending in at $-i\infty$ in G_0, and encircling only the singularity z_m once in the positive sense. Now, in the exceptional case when $C_m^> = \ell_0(z_m)$, (2.81) is just the condition (2.78) rewritten using (2.77). In (2.78) with $\omega_m \neq 0$ as well as in (2.79), the integrand has an integrable algebraic branch point at $z = z_m$ so up to a nonzero factor of the form $(1-\alpha)/2$ the integrals appearing in (2.78) and (2.79) can both be rewritten equivalently with the contour $\ell_0(z_m)$ replaced by L_m. In the case of (2.79) we then integrate by parts $\mu_m + 1$ times on L_m, which produces no boundary terms for $\lambda > 0$ due to exponential decay of the integrand at $-i\infty$. Eliminating H_m by means of (2.77) then gives (2.81) with $C_m^>$ replaced by L_m in these cases. The final condition of Proposition II.20, (2.80), can of course be written as an integral about

a small loop centered at $z = z_m$, and applying Cauchy's Theorem under the condition $\lambda > 0$ allows this loop to be deformed into the contour L_m (in the case of a locally single-valued integrand), so (2.80) also takes the form (2.81) with $C_m^>$ replaced by L_m after elimination of $H_m(z;\lambda)$. Finally, the replacement of L_m with $U_m^>$ in each case to arrive at (2.81) amounts to multiplication of the system (2.81) by a triangular nonsingular matrix. ■

Note that the conditions for analyticity of $W_+(x;\lambda)$ as expressed in Corollary II.21 take the form of a square linear system of equations on the unknowns $\{v_p(\lambda)\}_{p=1}^P$ assembled in a vector $\mathbf{v}(\lambda) := (v_1(\lambda), \dots, v_P(\lambda))^\mathsf{T}$:

$$\mathbf{A}^>(\lambda)\mathbf{v}(\lambda) = \mathbf{b}^>(\lambda), \quad \lambda > 0, \tag{2.82}$$

where $\mathbf{A}^>(\lambda) \in \mathbb{C}^{P\times P}$ and $\mathbf{b}^>(\lambda) \in \mathbb{C}^P$ have components

$$A_{mp}^>(\lambda) := \int_{C_m^>} \frac{e^{-ih(z;\lambda)/\epsilon}}{z - z_p}\, dz \quad \text{and} \quad b_m^>(\lambda) := -\lambda \int_{C_m^>} e^{-ih(z;\lambda)/\epsilon}\, dz. \tag{2.83}$$

The linear system (2.82) can then be solved using, say, Cramer's rule to determine $\{v_p(\lambda)\}_{p=1}^P$ uniquely. This procedure is guaranteed to succeed (i.e., the determinant of $\mathbf{A}^>(\lambda)$ is nonzero for all $\lambda > 0$) because according to the general theory described briefly in Section 2.1.2, the Jost solutions exist uniquely for all $\lambda > 0$. More to the point we have the following remark.

Remark II.23. Though the construction of the matrix $\mathbf{A}^>(\lambda)$ is carried out above strictly for $\lambda > 0$ it is natural to ask whether it admits analytic continuation into the upper-half complex λ-plane. There is indeed a matrix function $\mathbf{A} = \mathbf{A}(\lambda)$ defined for $\lambda \in \mathbb{C} \setminus \mathbb{R}^+$ whose boundary value for $\lambda > 0$ from the upper half-plane is $\mathbf{A}^>(\lambda)$ (the corresponding boundary value from the lower half-plane can be viewed as a matrix analogous to $\mathbf{A}^>(\lambda)$ that arises from conditions guaranteeing that $W_-(x;\lambda)$ is analytic in the upper-half z-plane). It is the fact that the eigenvalues are all negative real numbers that guarantees that $\mathbf{A}(\lambda)$ is singular

only when λ is a negative real number coinciding with an eigenvalue of the spectral problem.

The reflection coefficient $\beta(\lambda)$ is now easily calculated. Combining the scattering relation (2.15) with the boundary conditions (2.14) gives

$$\beta(\lambda) = \lim_{x\to\infty} e^{-i\lambda x/\epsilon}\left(W_+(x;\lambda) - 1\right) \quad \text{for} \quad \lambda > 0. \tag{2.84}$$

For a rational potential u_0 of the form (2.67), we may explicitly write

$$\beta(\lambda) = -\frac{i}{\epsilon} e^{iI/\epsilon} \int_{-\infty}^{\infty} e^{-ih(z;\lambda)/\epsilon} \left(\lambda + \sum_{p=1}^{P} \frac{v_p(\lambda)}{z - z_p}\right) dz, \tag{2.85}$$

using equation (2.72) for the Jost solution $W_+(x;\lambda)$ and Lemma II.14 (recall $I := \int_{-\infty}^{\infty} u_0(x)\, dx$). Similarly, we may use integration by parts to write

$$\beta(\lambda) = \frac{i}{\epsilon} e^{iI/\epsilon} \int_{-\infty}^{\infty} e^{-ih(z;\lambda)/\epsilon} \left(u_0(z) - \sum_{p=1}^{P} \frac{v_p(\lambda)}{z - z_p}\right) dz. \tag{2.86}$$

The coefficients $\{v_p(\lambda)\}_{p=1}^{P}$ are determined by the linear system (2.82).

2.2.3 Eigenfunctions and Eigenvalues

Eigenfunctions $w^+ = \Phi(x;\lambda)$ corresponding to eigenvalues $\lambda < 0$ are nontrivial solutions in $\mathbb{H}^+(\mathbb{R})$ of (2.12), or equivalently (2.71) for u_0 of the form (2.67), with $w_0 = 0$. Integrating (2.71) with $w_0 = 0$ and imposing $\Phi(x;\lambda) \to 0$ as $x \to -\infty$ shows that $\Phi(x;\lambda)$ necessarily has the form

$$\Phi(x;\lambda) = -\frac{i}{\epsilon} e^{ih(x;\lambda)/\epsilon} \int_{-\infty}^{x} e^{-ih(z;\lambda)/\epsilon} \sum_{p=1}^{P} \frac{\phi_p(\lambda)}{z - z_p}\, dz, \tag{2.87}$$

where $\phi_p(\lambda) := c_p \Phi(z_p;\lambda)$ for $p = 1, \ldots, P$. Since $\lambda < 0$, the convergent improper integral on the right-hand side of (2.87) becomes absolutely convergent if the contour is rotated at infinity into the upper-half x-plane.

Of course for (2.87) to indeed describe an eigenfunction of the Lax pair, the formula

must be evaluated at an eigenvalue λ of equation (2.12). This, evaluation however, does not apparently guarantee analyticity of $\Phi(x;\lambda)$ for $\mathrm{Im}\{x\} > 0$ in and of itself. In analogy to the construction of the Jost solutions, we must impose conditions on the values $\{\phi_p(\lambda)\}_{p=1}^P$ so that (2.87) describes an analytic function. Namely, while (2.87) satisfies (2.71) and is analytic for x in a neighborhood of $\mathbb{R}$, to obtain the required analyticity for all x with $\mathrm{Im}\{x\} > 0$ requires conditions on both $\lambda < 0$ and $\{\phi_p(\lambda)\}_{p=0}^P$. As we will see these conditions will also guarantee that $\Phi(\cdot;\lambda) \in \mathbb{H}^+(\mathbb{R})$, i.e., that it vanishes as $|x| \to \infty$ in all directions of the upper half-plane.

Definition II.24. $U_m^<$ denotes any member of the equivalence class of contours in G originating at $i\infty$ in G_0 and ending at $i\infty$ in G_m; see Figures 2.2 and 2.3.

Proposition II.25. Let $\lambda < 0$. The function $\Phi(x;\lambda)$ is analytic in the upper-half x-plane if and only if $\{\phi_p(\lambda)\}_{p=1}^P$ satisfy

$$\int_{C_m^<} e^{-ih(z;\lambda)/\epsilon} \sum_{p=1}^P \frac{\phi_p(\lambda)}{z - z_p}\, dz = 0, \quad m = 1, \dots, P, \tag{2.88}$$

where $C_m^< := U_m^<$ unless ic_m/ϵ is a strictly negative integer, in which case $C_m^< := \ell_0(z_m)$. (The contours $C_m^<$ extend to $i\infty$ in G_0 and G_m making the integrals absolutely convergent.)

Proof. The proof is analogous to those of Proposition II.20 and Corollary II.21 with the function H_m defined instead as

$$(i(z - z_m))^{-ic_m/\epsilon - 1} H_m(z;\lambda) := -e^{-ih(z;\lambda)/\epsilon} \sum_{p=1}^P \frac{\phi_p(\lambda)}{z - z_p}. \tag{2.89}$$

We omit the remaining details for the sake of brevity. ∎

Introducing the vector $\boldsymbol{\phi}(\lambda) = (\phi_1(\lambda), \dots, \phi_P(\lambda))^\mathsf{T}$ of unknowns, the conditions of Proposition II.25 take the form of a square homogeneous linear system $\mathbf{A}^<(\lambda)\boldsymbol{\phi}(\lambda) = \mathbf{0}$ where the

matrix $\mathbf{A}^<(\lambda) \in \mathbb{C}^{P\times P}$ has components

$$A^<_{mp}(\lambda) := \int_{C^<_m} \frac{e^{-ih(z;\lambda)/\epsilon}}{z - z_p}\, dz, \quad \lambda < 0. \tag{2.90}$$

Remark II.26. If the coefficient ic_m/ϵ is a positive integer N_m, the integral (2.90) can be calculated by residues, and thus $A^<_{mp}(\lambda) = e^{-i\lambda z_m/\epsilon}\mathcal{P}_p(\lambda)$, where $\mathcal{P}_p(\lambda)$ is a polynomial in λ of order at most N_m. If this holds for all $m = 1, \ldots, P$, then $D(\lambda)$ is proportional via a nonvanishing exponential factor to a polynomial of order at most $\prod_{m=1}^P N_m$. If $P = 1$, this polynomial is a scaled Laguerre polynomial; see [43].

Corollary II.27. The function $\Phi(x;\lambda)$ given by (2.87) is both nontrivial and analytic in the closed upper-half x-plane if and only if $D(\lambda) := \det(\mathbf{A}^<(\lambda)) = 0$ and $\boldsymbol{\phi}(\lambda)$ is a nontrivial nullvector of $\mathbf{A}^<(\lambda)$, unique up to a constant multiple.

Proof. This follows from Proposition II.25. It only remains to check that $\dim\left(\ker\left(\mathbf{A}^<(\lambda)\right)\right) = 1$ whenever $\det(\mathbf{A}^<(\lambda)) = 0$. We proceed by rewriting $\mathbf{A}^<(\lambda)$ using its *local Smith form* [29]. Namely, for λ in a neighborhood of a λ_j such that $D(\lambda_j) = 0$, an analytic matrix function $\mathbf{A}^<(\lambda)$ may be written in the form

$$\mathbf{A}^<(\lambda) = \mathbf{E}(\lambda)\mathrm{diag}\left((\lambda - \lambda_j)^{\kappa_1}, \ldots, (\lambda - \lambda_j)^{\kappa_P}\right)\mathbf{F}(\lambda), \tag{2.91}$$

where $\mathbf{E}(\lambda)$ is an analytic matrix function with $\det\left(\mathbf{E}(\lambda_j)\right) \neq 0$, $\mathbf{F}(\lambda)$ is a matrix polynomial with a constant non-zero determinant, and $\kappa_1 \geq \ldots \geq \kappa_P$ are non-negative integers. Now, from the general theory presented in Section 2.1.2 we know that the eigenvalues λ_j are simple, which implies that function $\det\left(\mathbf{A}^<(\lambda)\right)$ must have a simple zero at λ_j. Since $\kappa_1 \geq \ldots \geq \kappa_P$ are non-negative integers, using the multiplicative properties of the determinant, we find λ_j being a simple zero of $\det\left(\mathbf{A}^<(\lambda)\right)$ demands that $\kappa_1 = 1$ while $\kappa_2 = \cdots = \kappa_P = 0$. Thus, the

local Smith form (2.91) evaluated at $\lambda = \lambda_j$ allows us to deduce that

$$\operatorname{rank}\left(\mathbf{A}^{<}(\lambda_j)\right) = \operatorname{rank}\left(\operatorname{diag}\left(0, 1, \ldots, 1\right)\right) = P - 1, \tag{2.92}$$

since the matrices $\mathbf{E}(\lambda_j)$ and $\mathbf{F}(\lambda_j)$ are full rank. Lastly, by the Rank-Nullity Theorem we conclude that $\dim\left(\ker\left(\mathbf{A}^{<}(\lambda_j)\right)\right) = 1$ as stated. ■

Corollary II.28. Suppose $D(\lambda) = 0$, and that the components $\{\phi_p(\lambda)\}_{p=1}^{P}$ of the nullvector $\boldsymbol{\phi}(\lambda)$ are normalized to satisfy

$$\sum_{p=1}^{P} \phi_p(\lambda) = \lambda < 0, \tag{2.93}$$

then the function $\Phi(x;\lambda)$ satisfies the normalization condition (2.18), $\Phi(\cdot;\lambda) \in \mathbb{H}^+$, and therefore $\Phi(x;\lambda)$ is an eigenfunction of equation (2.12).

Proof. By Corollary II.27, the function $\Phi(x;\lambda)$ is analytic in the closed upper-half x-plane. From (2.87) one then invokes the Residue Theorem to see that (2.93) implies (2.18). The fact that $\boldsymbol{\phi}(\lambda)$ can be scaled to satisfy (2.93), i.e., that $\ker(\mathbf{A}^{<}(\lambda))$ is not orthogonal to $(1, 1, \ldots, 1)^{\mathsf{T}}$, is proven in Proposition II.34. Lastly, to show that $\Phi \in \mathbb{H}^+$, we check that $\Phi \to 0$ as $|x| \to \infty$ in the closed upper-half x-plane. First, applying Jordan's Lemma to (2.87) shows that $\Phi \to 0$ when $x \to \infty$ anywhere in G_0. To let $x \to \infty$ elsewhere in G, we rewrite the limiting contour integral in (2.87) as $\lim_{x\to\infty\in G_m} \int_{-\infty}^{x} = \int_{U_m^{<}}$ whenever ic_m/ϵ is not a negative integer. From Proposition II.25, this limit is identically zero for all $m = 1, \ldots, P$, since each integral on $U_m^{<}$ vanishes when $D(\lambda) = 0$. ■

Corollary II.29. Let $D(\lambda) = 0$ for some $\lambda = \lambda_j < 0$. The eigenfunction $\Phi_j(x) := \Phi(x;\lambda_j)$ given by (2.87) with the conditions from Corollary II.28 satisfies the asymptotic condition (2.17).

Proof. Applying L'Hôpital's rule to (2.87) and using $h'(x;\lambda) = \lambda + u_0(x)$,

$$\lim_{x\to\infty} x\Phi(x;\lambda) = \lim_{x\to\infty} \frac{ix^2}{\epsilon + i\lambda x + ixu_0(x)} \sum_{p=1}^{P} \frac{\phi_p(\lambda)}{x - z_p} = \frac{1}{\lambda} \sum_{p=1}^{P} \phi_p(\lambda), \tag{2.94}$$

and the desired result then follows from (2.93). ■

The stronger result $\Phi_j(x) = x^{-1} + O\left(x^{-2}\right)$ as $|x| \to \infty$ in the upper-half x-plane can be proven with a little more work.

Corollary II.30. The eigenfunction $\Phi_j(x) := \Phi(x; \lambda_j)$ with the conditions from Corollary II.28 satisfies the asymptotic condition $\Phi(x; \lambda) = x^{-1} + O\left(x^{-2}\right)$ as $|x| \to \infty$ uniformly in the closed upper-half x-plane.

Proof. Note that for initial conditions of the form (2.67) there exists a number L such that $|u_0(x)| < |u_0(L)|$ for $|x| > L$. Indeed, $|\lambda + u_0(x)| > ||\lambda| - |u_0(L)||$ and $u_0'(x) = O\left(x^{-3}\right)$ for $|x| > L$. Assuming that the conditions of Corollary II.28 hold, we may then proceed as in the proof of Property 1 in Proposition 5.3 in [54] with minor modifications; namely, we consider the asymptotic properties of Φ in each subdomain G_m separately. First, suppose that $x \in G_m$ and $|x| > L$ for some m, then using Proposition II.25 we may rewrite (2.87) in the form

$$\Phi(x; \lambda) = \frac{i}{\epsilon} e^{ih(x;\lambda)/\epsilon} \int_x^{\infty \in G_m} e^{-ih(z;\lambda)/\epsilon} \sum_{p=1}^{P} \frac{\phi_p(\lambda)}{z - z_p}\, dz. \tag{2.95}$$

Second, integrating (2.95) by parts using the identity $(\lambda + u_0(x))\, e^{-ih(x;\lambda)/\epsilon} = i\epsilon \partial_x e^{-ih(x;\lambda)/\epsilon}$ we obtain

$$\begin{aligned} \Phi(x; \lambda) =& \frac{1}{\lambda + u_0(x)} \sum_{p=1}^{P} \frac{\phi_p(\lambda)}{x - z_p}\, dz - e^{ih(x;\lambda)/\epsilon} \int_x^{\infty \in G_m} \frac{e^{-ih(z;\lambda)/\epsilon} u_0'(z)}{(\lambda + u_0(z))^2} \sum_{p=1}^{P} \frac{\phi_p(\lambda)}{z - z_p}\, dz \\ & - e^{ih(x;\lambda)/\epsilon} \int_x^{\infty \in G_m} \frac{e^{-ih(z;\lambda)/\epsilon}}{\lambda + u_0(z)} \sum_{p=1}^{P} \frac{\phi_p(\lambda)}{(z - z_p)^2}\, dz, \end{aligned} \tag{2.96}$$

where the contour of integration can, in general, be taken to be a straight line. Integrating

the last term of the right hand side of (2.96) once more gives

$$\int_x^{\infty\in G_m} \frac{e^{-ih(z;\lambda)/\epsilon}}{(\lambda+u_0(z))(z-z_p)^2}\,dz = i\epsilon\left[-\frac{e^{-ih(x;\lambda)/\epsilon}}{(\lambda+u_0(x))^2(x-z_p)^2}\right.$$
$$\left. + \int_x^{\infty\in G_m} \frac{2e^{-ih(z;\lambda)/\epsilon}u_0'(z)}{(\lambda+u_0(z))^3(z-z_p)^2}\,dz + \int_x^{\infty\in G_m} \frac{2e^{-ih(z;\lambda)/\epsilon}}{(\lambda+u_0(z))^2(z-z_p)^3}\,dz\right]. \tag{2.97}$$

Now, note that for the chosen contours of integration, it follows from (2.69) that $\mathrm{Im}\{h(z;\lambda)\} \leq \mathrm{Im}\{h(x;\lambda)\}$ since $\lambda < 0$ for large enough x. Therefore, combining (2.96) and (2.97) we see that the leading term of Φ as $x \to \infty$ in the closed upper-half x-plane arises from the first term on the right hand side of (2.96). To see this, consider the second term on the right hand side of (2.96). This term may be bounded as follows

$$\left|e^{ih(x;\lambda)/\epsilon}\int_x^{\infty\in G_m} \frac{e^{-ih(z;\lambda)/\epsilon}u_0'(z)}{(\lambda+u_0(z))^2}\sum_{p=1}^{P}\frac{\phi_p(\lambda)}{z-z_p}\,dz\right| \leq \int_x^{\infty\in G_m} \frac{|u_0'(z)|}{|\lambda+u_0(z)|^2}\left|\sum_{p=1}^{P}\frac{\phi_p(\lambda)}{z-z_p}\right||dz|$$
$$\leq \frac{K|x|^{-2}}{||\lambda|-|u_0(L)||^2}\int_L^{\infty\in G_m}\left|\frac{1}{z}\sum_{p=1}^{P}\frac{\phi_p(\lambda)}{z-z_p}\right||dz| \tag{2.98}$$

where K is a constant independent of x and ϵ. The remaining terms the sum may be treated similarly. Thus, for any m such that $x \in G_m$ and $|x| > L$, we have

$$\left|\Phi(x;\lambda_j) - \frac{1}{x}\right| \leq \frac{K}{x^2} \tag{2.99}$$

using the normalization (2.93), where K is independent of x and ϵ. ■

Corollaries II.27–II.30 show that the equation $D(\lambda) = 0$ involving the determinant of the $P \times P$ matrix $\mathbf{A}^<(\lambda)$ is exactly the condition that $\lambda = \lambda_j < 0$ is an eigenvalue for the rational potential u_0 of the form (2.67), and that $\Phi_j(x) := \Phi(x;\lambda_j)$ is the corresponding normalized eigenfunction.

2.2.4 Analytic Continuation, Evans Function, and Phase Constants

The matrix $\mathbf{A}^<(\lambda)$ defined for $\lambda < 0$ by (2.90) has an analytic continuation to the maximal domain $\mathbb{C} \setminus \mathbb{R}^+$. Recalling that $h(z;\lambda)$ is dominated for large $|z|$ by the term λz, it is clear that this continuation is afforded simply by rotating the infinite "tails" of the integration contours $C_m^<$ so that they tend to complex ∞ in the direction $\arg(z) = -\arg(i\lambda)$. In the process of rotating the contours from their initially upward vertical configuration when $\lambda < 0$, the function $f(z)$ appearing in the integrand via $h(z;\lambda)$ must be analytically continued through its vertical branch cuts as well. Therefore we observe that

$$D(\lambda) := \det\left(\mathbf{A}^<(\lambda)\right), \quad \lambda \in \mathbb{C} \setminus \mathbb{R}^+. \tag{2.100}$$

is an *Evans function* for the spectral problem with rational potential u_0 of the form (2.67); it is an analytic function in the domain $\mathbb{C} \setminus \mathbb{R}^+$ (complementary to the continuous spectrum for the problem) whose roots are precisely the eigenvalues.

Remark II.31. We note in passing that it may be potentially useful to rewrite the Evans function D as an P-fold integral using the multi-linearity properties of the determinant, though we don't pursue this fact further in this dissertation. Namely, D may be rewritten as

$$D(\lambda) = \int_{C_1^<} \cdots \int_{C_P^<} Q_P(\zeta_1, \ldots, \zeta_P) \exp\left(-\frac{i}{\epsilon}\sum_{p=1}^{P} h(\zeta_p;\lambda)\right) d\zeta_P \cdots d\zeta_1, \tag{2.101}$$

where $Q_P(\zeta_1, \ldots, \zeta_P) = \det \mathbf{C}$ and $\mathbf{C}$ is the $P \times P$ Cauchy matrix with elements $C_{mp} = 1/(\zeta_m - z_p)$. The determinant of a Cauchy matrix is well known [74] and we use it to define the function Q_M for future use:

$$Q_M(\zeta_1, \ldots, \zeta_M) = Q_M(\zeta_1, \ldots, \zeta_M; z_1, \ldots, z_P) := \frac{\displaystyle\prod_{m=2}^{M}\prod_{p=1}^{m-1}(\zeta_m - \zeta_p)(z_p - z_m)}{\displaystyle\prod_{m=1}^{M}\prod_{p=1}^{P}(\zeta_m - z_p)}. \tag{2.102}$$

for some number $M \leq P$.

If the analytic continuation of $\mathbf{A}^{<}(\lambda)$ is carried out through the upper-half λ-plane to the cut $\mathbb{R}^+$, the contour $C_m^{\leq}$ will be rotated through the left-half z-plane to coincide precisely with $C_m^{>}$; compare the two panels of Figure 2.3. Thus, the analytic continuation of the matrix $\mathbf{A}^{<}(\lambda)$ from $\lambda < 0$ through the upper half-plane to $\lambda > 0$ coincides with the matrix $\mathbf{A}^{>}(\lambda)$ defined in (2.83). In a similar way, the vector $\mathbf{b}^{<}(\lambda)$ defined for $\lambda < 0$ with components

$$b_m^{<}(\lambda) := -\lambda \int_{C_m^{\leq}} e^{-ih(z;\lambda)/\epsilon}\, dz, \quad \lambda < 0, \tag{2.103}$$

has an analytic continuation to $\mathbb{C} \setminus \mathbb{R}^+$, taking a boundary value on $\mathbb{R}^+$ from the upper half-plane that coincides with the vector $\mathbf{b}^{>}(\lambda)$ also defined in (2.83). This shows that indeed for each x with $\mathrm{Im}\{x\} \geq 0$, the Jost solution $W_+(x;\lambda)$ given by the formula (2.72) is the boundary value on $\mathbb{R}^+$ from $\mathrm{Im}\{\lambda\} > 0$ of a function $W(x;\lambda)$ analytic for $\lambda \in \mathbb{C} \setminus \mathbb{R}^+$ with the possible exception of the eigenvalues $\lambda < 0$ satisfying $D(\lambda) = 0$; only at these points do the quantities $\{v_p(\lambda)\}_{p=1}^P$ entering into (2.72) become indeterminate as the matrix of the analytic continuation of the system (2.82) through $0 \leq \arg(\lambda) < 2\pi$ becomes singular. This idea allows us to deduce the remaining scattering data corresponding to u_0 of the form (2.67), namely the phase constants $\{\gamma_j\}_{j=1}^N$ corresponding to the negative eigenvalues $\{\lambda_j\}_{j=1}^N$. The first step is to determine how analytic continuation of the system (2.82) fails near an eigenvalue $\lambda = \lambda_j$.

Proposition II.32. Let $\mathbf{v}(\lambda)$ be the unique solution of $\mathbf{A}^{<}(\lambda)\mathbf{v}(\lambda) = \mathbf{b}^{<}(\lambda)$ for each $\lambda \in \mathbb{C} \setminus \mathbb{R}^+$ for which $D(\lambda) \neq 0$. Then each component $v_p(\lambda)$ is analytic in $\mathbb{C} \setminus \mathbb{R}^+$ except at the eigenvalues $\{\lambda_j\}_{j=1}^N$ which are simple poles, with corresponding Laurent expansion

$$v_p(\lambda) = -i\epsilon \frac{\phi_p(\lambda_j)}{\lambda - \lambda_j} + (z_p + \Gamma_j)\phi_p(\lambda_j) + O(\lambda - \lambda_j) \quad \text{as} \quad \lambda \to \lambda_j \tag{2.104}$$

where Γ_j is a constant (independent of p) given by

$$\Gamma_j := -\frac{i\epsilon}{2\lambda_j} - \frac{1}{2\lambda_j}\sum_{p=1}^{P} z_p\phi_p(\lambda_j) - \frac{i\epsilon}{2}\frac{\mathbf{m}^\mathsf{T}\mathbf{b}^{<\prime}(\lambda_j)}{\mathbf{m}^\mathsf{T}\mathbf{b}^{<}(\lambda_j)}, \tag{2.105}$$

and where $\{\phi_p(\lambda_j)\}_{p=1}^P$ are the components of the right nullvector of $\mathbf{A}^<(\lambda_j)$ normalized by (2.93) and $\mathbf{m}$ is a nonzero left nullvector of $\mathbf{A}^<(\lambda_j)$.

We include the proof of Proposition II.32 in Appendix A.1. The following result then shows that the phase constant associated with each eigenvalue $\lambda_j < 0$ by means of the Laurent expansion (2.19) is given by $\gamma_j = \Gamma_j$.

Corollary II.33. The function $W(x;\lambda)$ whose boundary value from $\mathrm{Im}\{\lambda\} > 0$ on $\mathbb{R}^+$ is the Jost solution $W_+(x;\lambda)$ is meromorphic on $\mathbb{C}\setminus\mathbb{R}^+$ with simple poles only at the eigenvalues $\{\lambda_j\}_{j=1}^N$, at each of which it has a Laurent expansion of the form (2.19) with $\Phi_j(x) := \Phi(x;\lambda_j)$ defined as in Section 2.2.3 and with $\gamma_j = \Gamma_j$ defined as in (2.105).

Proof. The function $W(x;\lambda)$ is given by (2.72) with the coefficients $\{v_p(\lambda)\}_{p=1}^P$ obtained by solving $\mathbf{A}^<(\lambda)\mathbf{v}(\lambda) = \mathbf{b}^<(\lambda)$, and with the contour of integration rotated appropriately to ensure absolute convergence of the integral. Substitution of (2.104) into (2.72) we obtain a Laurent expansion for $W(x;\lambda)$ of the form

$$W(x;\lambda) = \frac{W_j^{-1}(x)}{\lambda-\lambda_j} + W_j^0(x) + O(\lambda-\lambda_j) \quad \text{as} \quad \lambda\to\lambda_j. \tag{2.106}$$

The coefficient $W_j^{-1}(x)$ can be read directly to be

$$W_j^{-1}(x) = -\frac{i}{\epsilon}e^{ih(x;\lambda_j)/\epsilon}\int_{-\infty}^{x} e^{-ih(z;\lambda_j)/\epsilon}\sum_{p=1}^{P}\frac{(-i\epsilon\phi_p(\lambda_j))}{z-z_p}\,dz = -i\epsilon\Phi(x;\lambda_j) \tag{2.107}$$

using formula (2.87) of the eigenfunction Φ. Similarly, we may read off the the coefficient

$W_j^0(x)$ from formula (2.72) to be

$$\begin{aligned} W_j^0(x) &= -\frac{i}{\epsilon} e^{ih(x;\lambda_j)/\epsilon} \int_{-\infty}^{x} e^{-ih(z;\lambda_j)/\epsilon} \left(\lambda_j + \sum_{p=1}^{P} \frac{(z_p + \Gamma_j)\phi_p(\lambda_j)}{z - z_p} \right) dz \\ &= -(z + \Gamma_j) \frac{i}{\epsilon} e^{ih(x;\lambda_j)/\epsilon} \int_{-\infty}^{x} e^{-ih(z;\lambda_j)/\epsilon} \sum_{p=1}^{P} \frac{\phi_p(\lambda_j)}{z - z_p} \, dz = (z + \Gamma_j)\Phi(x;\lambda_j) \end{aligned} \tag{2.108}$$

using the normalization condition (2.93). This yields (2.19) in which $\gamma_j = \Gamma_j$. ■

We finish the section by establishing some facts about the nullspace of $\mathbf{A}^<(\lambda_j)$.

Proposition II.34. The components of the only nontrivial nullvector $\mathbf{n}(\lambda_j)$ of $\mathbf{A}^<(\lambda_j)$ are given (up to a common constant) by the formula

$$n_p(\lambda) = D_p(\lambda), \tag{2.109}$$

where $D_p(\lambda) := \det\left(\mathbf{A}_p^<(\lambda)\right)$ and $\mathbf{A}_p^<(\lambda)$ denotes the matrix $\mathbf{A}^<(\lambda)$ with its p^{th} column replaced with $\mathbf{b}(\lambda)$. Moreover, the components satisfy the condition

$$\sum_{p=1}^{P} n_p(\lambda) \neq 0. \tag{2.110}$$

We include the proof of this fact in Appendix A.1

Remark II.35. The statement of Proposition II.34 is indeed correct if $\mathbf{b}^<(\lambda)$ is chosen to be any vector that is not in the column space of $\mathbf{A}^<(\lambda)$. The fact that $\sum_{p=1}^{P} n_p(\lambda_j) \neq 0$, however, is a special feature of our problem.

2.2.4.1 Summary

For ease of reference, we summarize the results on the scattering data for rational initial conditions presented in this section. For information concerning the scattering data derived in the general theory refer to Table 1.

Theorem II.36. For the rational initial condition (2.67), the scattering data for the Lax pair equation (2.12) is explicitly characterized from the integral formulas of the Jost solution (2.72) and eigenfunctions (2.87). In particular,

- Reflection coefficient $\beta(\lambda)$ for $\lambda > 0$: defined by equation (2.85) with coefficients $\{v_p(\lambda)\}_{p=1}^{P}$ satisfying the linear algebra problem (2.82).
- Negative eigenvalues $\{\lambda_j\}_{j=1}^{N}$: determined by the zeros of the Evans function D in (2.100), as shown in Corollary II.27.
- Phase constants $\{\gamma_j\}_{j=1}^{N}$: defined by (2.105) in Proposition II.32.

2.3 Small Dispersion Asymptotics

In this section we derive direct asymptotic expansions of the scattering data in the small dispersion limit $\epsilon \to 0$. In particular, we provide rigorous justification to previous results in the literature: a formal argument for the eigenvalue density [49], a leading term approximation of the phase constants obtained from the inverse problem [55], and a conjecture on the leading term of the reflection coefficient [90]. To simplify the analysis, we restrict our discussion to a simpler set of initial conditions: the so-called Klaus-Shaw potentials discussed in [42].

2.3.1 Klaus-Shaw Initial Conditions

Definition II.37. A rational Klaus-Shaw (rKS) potential u_0 is a rational function of the form (2.67) for which there exists $x_c \in \mathbb{R}$ such that u_0 is strictly monotone on the intervals $(-\infty, x_c)$ and $(x_c, +\infty)$ (making x_c the unique local maximum/minimum).

Rational Klaus-Shaw potentials are of one sign, and hence there are distinct classes of positive and negative rKS potentials. Colloquially one can describe a rKS potential as having a graph consisting of a single "lobe" or "bump", a property that was found by Klaus and

Shaw [42] to be useful in confining the spectrum of certain non-self-adjoint operators. For a rKS potential u_0 there are at most two real roots $\lambda \in \mathbb{R}$ of the equation $u_0(x) = -\lambda$, and hence the corresponding inverse function $x(\lambda)$ has at most two real branches.

The function h defined in (2.74) has critical points given by the equation

$$h'(x;\lambda) = \lambda + u_0(x) = 0 \quad \text{or} \quad u_0(x) = -\lambda; \tag{2.111}$$

here the prime denotes a derivative in x. The nature of the critical points is clearly dependent on the value of the parameter λ.

Definition II.38. The bulk $\mathcal{B} \subset \mathbb{R}$ is the set $\mathcal{B} = \{\lambda : -\sup\{u_0\} < \lambda < -\inf\{u_0\}\}$. We say that $\lambda \in \mathcal{B}$ lies in the bulk, while $\lambda \in \mathbb{R} \setminus \overline{\mathcal{B}}$ lies outside the bulk, where $\overline{\mathcal{B}}$ denotes the closure of $\mathcal{B}$.

Remark II.39. For a strictly positive (resp., negative) rKS potential u_0, $\mathcal{B} = (-\sup\{u_0\}, 0)$ (resp., $\mathcal{B} = (0, -\inf\{u_0\})$) so the bulk is a negative (resp., positive) interval abutting the origin. We refer to the origin as the "hard edge" of the bulk and to the nonzero endpoint of $\mathcal{B}$ as the "soft edge" of the bulk.

Definition II.40 (Critical points for λ in the bulk)**.** Let u_0 be a rKS potential with corresponding bulk $\mathcal{B}$. The functions $x_\pm : \mathcal{B} \longrightarrow \mathbb{R}$ represent the two real branches of the inverse of $-u_0$ with $x_-(\lambda) < x_+(\lambda)$ for $\lambda \in \mathcal{B}$ and are the only real critical points of the function h; see Figure 2.4. The remaining $2P-2$ critical points of h form complex-conjugate pairs for $\lambda \in \mathcal{B}$ and will be denoted by $x_p, x_p^* : \mathcal{B} \longrightarrow \mathbb{C} \setminus \mathbb{R}$ with $\mathrm{Im}\{x_p\} > 0$ for $p = 1, \ldots, P-1$ (not necessarily distinct for all $\lambda \in \mathcal{B}$).

Definition II.41 (Critical points for λ outside the bulk)**.** Let u_0 be a rKS potential with corresponding bulk $\mathcal{B}$. If $\lambda \in \mathbb{R} \setminus \overline{\mathcal{B}}$, then h has $2P$ critical points in complex-conjugate pairs denoted by $x_p, x_p^* : \mathbb{R} \setminus \overline{\mathcal{B}} \longrightarrow \mathbb{C} \setminus \mathbb{R}$ with $\mathrm{Im}\{x_p\} > 0$ for $p = 0, \ldots, P-1$ (not necessarily distinct for all $\lambda \in \mathbb{R} \setminus \overline{\mathcal{B}}$).

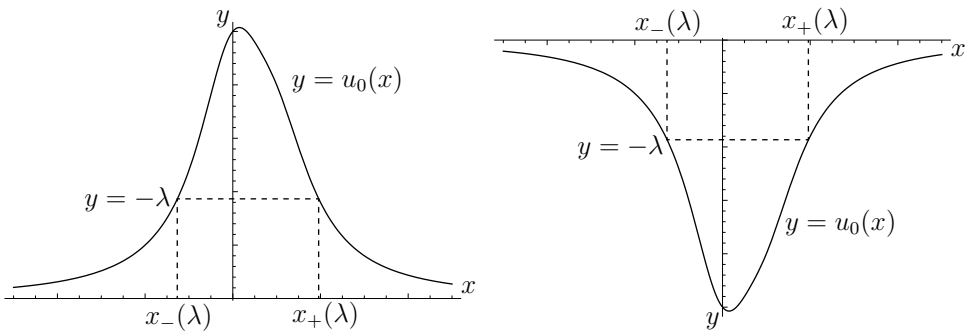

Figure 2.4: Strictly positive (left) and negative (right) rKS potentials u_0 illustrating the two real branches $x_\pm(\lambda)$ of the inverse function of $-u_0$ defined for suitable $\lambda < 0$.

In a neighborhood of $\lambda \in \mathcal{B}$ or $\lambda \in \mathbb{R} \setminus \mathcal{B}$ for which all $2P$ critical points of h are simple, each critical point is a locally analytic function $x = x(\lambda)$ of λ that satisfies (2.111). As λ exits the bulk through the soft edge, the two real critical points $x_\pm(\lambda)$ coalesce and bifurcate into the complex plane, where they become renamed as the conjugate pair $x_0(\lambda)$ and $x_0(\lambda)^*$. As λ exits the bulk through the hard edge, the two real critical points $x_\pm(\lambda)$ tend to $\pm\infty$ and then become finite again as a (large) conjugate pair $x_0(\lambda)$ and $x_0(\lambda)^*$.

According to (2.67) and (2.111) $h'(x;\lambda)$ is a rational function of x with simple poles at $\{z_p, z_p^*\}_{p=1}^P$ such that $h'(x;\lambda) \to \lambda$ as $x \to \infty$, using Definition II.40 we may explicitly write

$$h'(x;\lambda) = \lambda \Psi^-(x;\lambda)\Psi^+(x;\lambda) \quad \text{for} \quad \lambda \in \mathcal{B}, \tag{2.112}$$

where, assuming the complex critical points $x_p = x_p(\lambda)$ are distinct for $\lambda \in \mathcal{B}$,

$$\begin{aligned} \Psi^-(x;\lambda) &:= (x - x_-(\lambda)) \frac{\prod_{p=1}^{P-1} (x - x_p(\lambda))}{\prod_{p=1}^{P} (x - z_p)}, \\ \Psi^+(x;\lambda) &:= (x - x_+(\lambda)) \frac{\prod_{p=1}^{P-1} (x - x_p(\lambda)^*)}{\prod_{p=1}^{P} (x - z_p^*)}. \end{aligned} \tag{2.113}$$

Remark II.42. Note that Ψ^+ (Ψ^-) is analytic in the upper-half (lower-half) x-plane. For λ outside the bulk, we may also write (2.112) provided we extend the definitions of $\Psi^\pm$ by replacing x_- with x_0 and x_+ with x_0^* in (2.113). These extended definitions satisfy

$\Psi^-(x;\lambda)^* = \Psi^+(x^*;\lambda)$ for all $\lambda \in \mathbb{R} \setminus \overline{\mathcal{B}}$.

Differentiating (2.112) and using the definitions (2.113), we may express $h''(x_\pm(\lambda);\lambda)$ for $\lambda \in \mathcal{B}$ as

$$h''(x_\pm(\lambda);\lambda) = u_0'(x_\pm(\lambda)) = \pm\lambda \frac{|\Psi^\mp(x_\pm(\lambda);\lambda)|^2}{x_+(\lambda) - x_-(\lambda)} \quad \text{for} \quad \lambda \in \mathcal{B}. \tag{2.114}$$

From equation (2.114) we may immediately observe that

$$\operatorname{sign}\{h''(x_\pm(\lambda);\lambda)\} = \pm\operatorname{sign}\{\lambda\} \quad \text{for} \quad \lambda \in \mathcal{B}. \tag{2.115}$$

This fact can also be deduced from equation (2.111) and Figure 2.4. Namely, for a positive rKS potential u_0, we have $u_0'(x_-(\lambda)) > 0$ and $u_0'(x_+(\lambda)) < 0$ for a given $\lambda \in \mathcal{B}$, so $h''(x_\pm(\lambda);\lambda) = \mp|u_0'(x_\pm(\lambda))|$. Similarly, for strictly negative u_0, we have that $h''(x_\pm(\lambda);\lambda) = \pm|u_0'(x_\pm(\lambda))|$ for $\lambda \in \mathcal{B}$.

Lemma II.43. Let $\lambda \in \mathcal{B}$, and set $h_\pm(\lambda) := h(x_\pm(\lambda);\lambda)$. Then

$$h_+(\lambda) = \theta_+(\lambda) + I \quad \text{and} \quad h_-(\lambda) = \theta_-(\lambda), \tag{2.116}$$

where

$$I := \int_{-\infty}^{\infty} u_0(y)\,dy \quad \text{and} \quad \theta_\pm(\lambda) := \int_0^{\lambda} x_\pm(\eta)\,d\eta \tag{2.117}$$

are well-defined because $u_0 \in L^1(\mathbb{R})$.

Proof. Because $f(x)$ is an antiderivative of $u_0(x)$ that vanishes as $x \to -\infty$,

$$f(x_-(\lambda)) = \int_{-\infty}^{x_-(\lambda)} u_0(y)\,dy = -\lambda x_-(\lambda) - \int_{-\infty}^{x_-(\lambda)} y u_0'(y)\,dy, \tag{2.118}$$

by integration by parts. We rewrite the final integral on the right-hand side using the substitution $u_0(y) = -\eta$, where $y = x_-(\eta)$ for $y \in (-\infty, x_-(\lambda)]$. Hence equation (2.118)

becomes

$$f\left(x_{-}(\lambda)\right) = -\lambda x_{-}(\lambda) + \int_{0}^{\lambda} x_{-}(\eta)\, d\eta, \tag{2.119}$$

i.e., $h_{-}(\lambda) = \theta_{-}(\lambda)$. Starting instead from the formula

$$f\left(x_{+}(\lambda)\right) = \int_{-\infty}^{\infty} u_0(y)\, dy - \int_{x_{+}(\lambda)}^{\infty} u_0(y)\, dy \tag{2.120}$$

but making the substitution $y = x_{+}(\eta)$ after integrating by parts in the second term then gives the desired formula for $h_{+}(\lambda)$. This finishes the proof. ■

2.3.2 Reflection Coefficient

An asymptotic expansion for the reflection coefficient $\beta(\lambda)$, $\lambda > 0$, can be directly obtained from formula (2.86) and the linear system (2.82). We first prove that the solution $\mathbf{v}(\lambda)$ of the linear system (2.82) has a limit $\mathbf{v}^0(\lambda)$ as $\epsilon \to 0$. This, in turn, allows us to apply the method of stationary phase to the integral in (2.86) to determine the leading-order behavior of $\beta(\lambda)$.

To study the asymptotic behavior of the solution $\mathbf{v}(\lambda)$, we consider an analogous linear system to (2.82). Namely, we define the system written in the form $\mathbf{N}\mathbf{A}^{>}(\lambda)\mathbf{v}(\lambda) = \mathbf{N}\mathbf{b}^{>}(\lambda)$ for some square matrix $\mathbf{N}$. Since each row of the original system (2.82) shares the same contour of integration and the same integrand on each column, the system obtained through multiplication by $\mathbf{N}$ is simply the original system over new contours, say W_m, in each row. Indeed, we intend to choose the matrix $\mathbf{N}$ in such a way that asymptotic analysis of the solution $\mathbf{v}(\lambda)$ is possible. In particular, we consider a modified linear system as *suitable* provided (i) the contour W_m passes through only one critical point of $h(x;\lambda)$ and maximizes the value of the real part of h (W_m is locally a steepest descent curve), (ii) only one contour W_m is associated with each critical point (one-to-one correspondence), and (iii) the matrix $\mathbf{N}$ is invertible (the new and modified linear systems are equivalent).

We believe that in general a suitable linear system can be found. We present the following proposition as proof of principle.

Proposition II.44. Let u_0 be a negative rKS potential for which c_p is a positive imaginary number for all $p = 1, \ldots, P$, and let $\lambda > 0$ be such that the complex critical points of $h(\cdot; \lambda)$ are all simple and correspond to distinct nonzero values of $\mathrm{Re}\{-ih(z; \lambda)\}$. Then there exists a family of contours $\{W_m\}_{m=1}^P$ such that the modified system $\mathbf{N}\mathbf{A}^>(\lambda)\mathbf{v}(\lambda) = \mathbf{N}\mathbf{b}^>(\lambda)$ having integrals over contours $\{W_m\}_{m=1}^P$ in place of $\{C_m^>\}_{m=1}^P$ is suitable.

The restriction on $\{c_p\}_{p=1}^P$ in Proposition II.44 implies that u_0 is a linear combination of negative Lorentzian profiles. The proof of Proposition II.44 is given by Miller and Wetzel [54] so we refrain from presenting the, rather technical, details here. For illustration, we present two example potentials and deduce the respective matrices $\mathbf{N}$ (with associated contours $\{W_m\}_{p=1}^P$) that make the linear system suitable in Appendix A.2. The following proposition shows that for suitable systems, the solution $\mathbf{v}(\lambda)$ has a well-defined limit as $\epsilon \to 0$.

Proposition II.45. Let $\lambda > 0$ be given. Suppose that the system (2.82) is equivalent via a matrix $\mathbf{N}$ to a suitable system $\mathbf{N}\mathbf{A}^>(\lambda)\mathbf{v}(\lambda) = \mathbf{N}\mathbf{b}^>(\lambda)$ having integration contours $\{W_m\}_{m=1}^P$ in its P rows. Then

$$\lim_{\epsilon \to 0} \mathbf{v}(\lambda) = \mathbf{v}^0(\lambda), \tag{2.121}$$

where for each $p = 1, \ldots, P$ we have

$$v_p^0(\lambda) = \lambda \lim_{z \to z_p} (z - z_p)\Psi^-(z; \lambda) = \lambda \operatorname*{Res}_{z=z_p} \Psi^-(z; \lambda). \tag{2.122}$$

Proof. We consider the case that all critical points associated with contours $\{W_m\}_{m=1}^P$ are simple, which is generically true. Let x_m denote the unique critical point of $-ih(z; \lambda)$ on the contour W_m traversed at the local steepest descent angle θ_m. Now, define the diagonal matrix $\mathbf{D}$ whose elements are $D_{mm} := |h''(x_m; \lambda)|^{1/2} e^{ih(x_m;\lambda)/\epsilon} e^{-i\theta_m}/\sqrt{2\pi\epsilon}$ (reciprocal of the leading order contribution). Let $\hat{\mathbf{A}}^>(\lambda) := \mathbf{D}\mathbf{N}\mathbf{A}^>(\lambda)$ and $\hat{\mathbf{b}}^> := \mathbf{D}\mathbf{N}\mathbf{b}^>(\lambda)$ so that $\hat{\mathbf{A}}^>(\lambda)\mathbf{v}(\lambda) =$

$\hat{\mathbf{b}}^{>}(\lambda)$. Therefore, applying the method of steepest descent shows that $\hat{\mathbf{A}}^{>}(\lambda) \to \hat{\mathbf{A}}^{>0}(\lambda)$ and $\hat{\mathbf{b}}^{>}(\lambda) \to \hat{\mathbf{b}}^{>0}(\lambda)$ as $\epsilon \to 0$, where

$$\hat{A}^{>0}_{mp}(\lambda) := \frac{1}{x_m - z_p} \quad \text{and} \quad \hat{b}^{>0}_m(\lambda) := -\lambda. \tag{2.123}$$

Note that the limiting matrix $\mathbf{A}^{>0}(\lambda)$ is an invertible Cauchy matrix (all x_m are distinct). To show that $\mathbf{v}^0(\lambda)$ given by (2.122) is the unique solution of $\hat{\mathbf{A}}^{>0}(\lambda)\mathbf{v}^0(\lambda) = \hat{\mathbf{b}}^{>0}(\lambda)$, we use the Residue Theorem; multiplying (2.122) on the left by $\hat{\mathbf{A}}^{>0}(\lambda)$ and dividing by $-1/\lambda$ gives

$$-\frac{1}{\lambda}\sum_{p=1}^{P}\frac{v_p^0(\lambda)}{x_m - z_p} = \sum_{p=1}^{P}\frac{\operatorname{Res}_{z=z_p}\Psi^-(z;\lambda)}{z_p - x_m} = \frac{1}{2\pi i}\oint_C \frac{\Psi^-(z;\lambda)}{z - x_m}\,dz = 1, \tag{2.124}$$

where C is a counter-clockwise contour encircling the points $\{z_p\}_{p=1}^P$ (the poles of the integrand) and the last equality follows by taking a residue at $z = \infty$ where the integrand behaves like $1/z$; see [74] for an alternative approach. ■

Theorem II.46. Let u_0 be a rKS potential, and suppose that $\lambda > 0$ is such that there exists a suitable modification of the linear system (2.82). If λ lies outside the bulk (true for all $\lambda > 0$ if u_0 is a positive rKS potential) then $\beta(\lambda)$ is exponentially small as $\epsilon \to 0$. If λ lies in the bulk, then

$$\beta(\lambda) = -\sqrt{\frac{2\pi\lambda\,(x_+(\lambda) - x_-(\lambda))}{\epsilon}}\,e^{-\frac{i}{\epsilon}\theta_+(\lambda) - i\psi_+(\lambda)} + o(\epsilon^{-1/2}) \tag{2.125}$$

as $\epsilon \to 0$, where θ_+ is defined in (2.117) and ψ_+ is defined by

$$e^{\mp i\psi_\pm(\lambda) - i\pi/4} := \frac{\Psi^{\mp}\left(x_\pm(\lambda);\lambda\right)}{\left|\Psi^{\mp}\left(x_\pm(\lambda);\lambda\right)\right|} \tag{2.126}$$

with $\Psi^{\pm}$ given in (2.113). The error term in (2.125) can be written as $O(\epsilon^{1/2})$ if the critical points of $h(z;\lambda)$ are all simple.

Proof. Since the limiting value of the solution $\mathbf{v}(\lambda)$ is known, it remains to analyze the

integrand of (2.86) as $\epsilon \to 0$.

For $\lambda \in \mathbb{R}^+ \setminus \overline{\mathcal{B}}$ (outside the bulk), there are exist no real critical points of $h(z;\lambda)$, and hence the contour of integration in (2.85) may be deformed away from the real axis in the direction of decrease of $\mathrm{Re}\{-ih(z;\lambda)\}$ onto the nearest critical point or singular point, giving exponential decay of $\beta(\lambda)$ as $\epsilon \to 0$.

For $\lambda \in \mathcal{B}$ (in the bulk) there are exactly two simple stationary phase points at $x_\pm(\lambda)$. By the method of stationary phase, if $g(z) = u_0(z)$ or $g(z) = (z - z_p)^{-1}$,

$$\begin{aligned}\frac{1}{\sqrt{2\pi\epsilon}} \int_{-\infty}^{\infty} e^{-ih(z;\lambda)/\epsilon} g(z)\, dz =& \frac{e^{i\pi/4} e^{-ih_-(\lambda)/\epsilon} g(x_-(\lambda))}{\sqrt{|h''(x_-(\lambda);\lambda)|}} \\ &+ \frac{e^{-i\pi/4} e^{-ih_+(\lambda)/\epsilon} g(x_+(\lambda))}{\sqrt{|h''(x_+(\lambda);\lambda)|}} + O(\epsilon)\end{aligned} \tag{2.127}$$

where $h_\pm(\lambda) := h(x_\pm(\lambda);\lambda)$, and we used the fact that as u_0 is a negative rKS potential (to have positive $\lambda \in \mathcal{B}$ and $\mathrm{sign}(u_0'(x_\pm(\lambda))) = \pm 1$). Therefore,

$$-ie^{-iI/\epsilon}\sqrt{\frac{\epsilon}{2\pi}}\beta(\lambda) = \frac{e^{i\pi/4} e^{-ih_-(\lambda)/\epsilon} G_\epsilon^-(\lambda)}{\sqrt{|h''(x_-(\lambda);\lambda)|}} + \frac{e^{-i\pi/4} e^{-ih_+(\lambda)/\epsilon} G_\epsilon^+(\lambda)}{\sqrt{|h''(x_+(\lambda);\lambda)|}} + O(\epsilon), \tag{2.128}$$

where, using $u_0(x_\pm(\lambda)) = -\lambda$,

$$G_\epsilon^\pm(\lambda) := -\lambda - \sum_{p=1}^{P} \frac{v_p(\lambda)}{x_\pm(\lambda) - z_p}. \tag{2.129}$$

The ϵ-dependence in $G_\epsilon^\pm(\lambda)$ enters through the coefficients $\{v_p(\lambda)\}_{p=1}^P$; we may replace $G_\epsilon^\pm(\lambda)$ with $G_0^\pm(\lambda)$ (in which $v_p(\lambda)$ is replaced with $v_p^0(\lambda)$ given by (2.122)) in (2.128) provided the error term is replaced with $o(1)$. Now using $v_p^0(\lambda)$ given by (2.122), we apply the Residue Theorem to obtain

$$G_0^\pm(\lambda) + \lambda = -\lambda \sum_{p=1}^{P} \mathrm{Res}_{z=z_p} \frac{\Psi^-(z;\lambda)}{x_\pm(\lambda) - z} = -\frac{\lambda}{2\pi i} \oint_C \frac{\Psi^-(z;\lambda)}{x_\pm(\lambda) - z}\, dz \tag{2.130}$$

where C encircles the poles $z_1, \ldots, z_P$ positively but excludes the points $x_\pm(\lambda)$. Now $\Psi^-(z;\lambda)$ includes a factor of $z - x_-(\lambda)$, so to compute $G_0^-(\lambda) + \lambda$ the singularity of the integrand at $z = x_-(\lambda)$ outside of C is removable, so as the integrand behaves like $-1/z$ as $z \to \infty$ we obtain $G_0^-(\lambda)+\lambda = \lambda$ or simply $G_0^-(\lambda) = 0$. To compute $G_0^+(\lambda)+\lambda$ we have the same behavior of the integrand for large z but now there is a simple pole outside C at $z = x_+(\lambda)$, so taking it into account gives $G_0^+(\lambda) + \lambda = \lambda - \lambda\Psi^-(x_+(\lambda);\lambda)$ or simply $G_0^+(\lambda) = -\lambda\Psi^-(x_+(\lambda);\lambda)$. Therefore, using (2.126), (2.128) may be written as

$$\beta(\lambda) = -\sqrt{\frac{2\pi}{\epsilon}}\frac{\lambda|\Psi^-(x_+(\lambda);\lambda)|}{\sqrt{|h''(x_+(\lambda);\lambda)|}}e^{iI/\epsilon}e^{-ih_+(\lambda)/\epsilon}e^{-i\psi_+(\lambda)} + o(\epsilon^{-1/2}). \tag{2.131}$$

The phase is simplified using Lemma II.43, while the amplitude is simplified using (2.114).

■

Corollary II.47 (Matsuno's modulus formula [50])**.** Let u_0 be a negative rKS potential. If $\lambda \in \mathcal{B}$, then

$$\lim_{\epsilon\to 0} \epsilon|\beta(\lambda)|^2 = 2\pi\lambda\left(x_+(\lambda) - x_-(\lambda)\right). \tag{2.132}$$

2.3.3 Evans Function

We study the asymptotic properties of the Evans function (2.100) for rKS potentials u_0 in the limit $\epsilon \to 0$ for $\lambda < 0$. For our asymptotic calculations, it is sufficient to consider only *positive* rKS potentials since only in this case is the intersection between the bulk $\mathcal{B}$ and $\lambda < 0$ not trivial. The essence of our strategy is to calculate the determinant in (2.100) for $\epsilon > 0$ small by first applying the method of steepest descent to the individual entries (2.83) of the matrix $\mathbf{A}(\lambda)$, and then compute the leading term of the (finite) determinant $D(\lambda)$. As in Section 2.3.2, it is useful to apply this method not to $D(\lambda) = \det(\mathbf{A}(\lambda))$ but rather to $\tilde{D}(\lambda) = \det(\tilde{\mathbf{A}}(\lambda))$ where $\tilde{\mathbf{A}}(\lambda) = \mathbf{N}\mathbf{A}(\lambda)$ for some invertible $\mathbf{N}$. As before, the matrix $\mathbf{N}$ is chosen so that the entries of $\tilde{\mathbf{A}}(\lambda)$ are given by (2.83) but with the contours $\{C_m\}_{m=1}^P$ replaced with $\{W_m\}_{m=1}^P$ to make the system $\tilde{\mathbf{A}}(\lambda)$ suitable. In addition, the suitable system

$\tilde{\mathbf{A}}(\lambda)$ will not lead to cancellations beyond all orders in the computation of the determinant (Evans function) since different rows of the matrix $\tilde{\mathbf{A}}(\lambda)$ will not have identical asymptotic expansions.

Again, as proof of principle that in a general case a suitable matrix $\tilde{\mathbf{A}}(\lambda) = \mathbf{N}\mathbf{A}(\lambda)$ can be found, we offer the following special case.

Proposition II.48. Let u_0 be a positive rKS potential for which c_p is a negative imaginary number for all $p = 1, \ldots, P$, and let $\lambda < 0$ be such that the complex critical points of $h(\cdot; \lambda)$ are all simple and correspond to distinct nonzero values of $\mathrm{Re}\{-ih(z; \lambda)\}$. Then there exists a family of contours $\{W_m\}_{m=1}^P$ such that the modified matrix $\tilde{\mathbf{A}}(\lambda) := \mathbf{N}\mathbf{A}(\lambda)$ having integrals over contours $\{W_m\}_{m=1}^P$ in place of $\{C_m^<\}_{m=1}^P$ is suitable and at most one of the contours, say W_Q, may pass over either one complex critical point or at most two real critical points $x_\pm(\lambda)$ for which $\mathrm{Re}\{-ih(x_\pm(\lambda); \lambda)\} = 0$.

This proposition is proven by Miller and Wetzel [54], so we refrain from its exposition here. As previously stated, we refer the reader to Appendix A.2 for an illustrative example potential leading to nontrivial matrix $\mathbf{N}$ and associated contours $\{W_m\}_{p=1}^P$. Note that the condition on the coefficients $\{c_p\}_{p=1}^P$ in Proposition II.48 implies that u_0 is a linear combination of positive Lorentzian profiles. We now show the leading order of $\tilde{D}(\lambda)$ as $\epsilon \to 0$ for a suitable matrix $\tilde{\mathbf{A}}(\lambda)$.

Proposition II.49. Let u_0 be a positive rKS potential and let $\lambda < 0$ be fixed. Suppose that for some invertible $\mathbf{N}$, the matrix $\tilde{\mathbf{A}}(\lambda) := \mathbf{N}\mathbf{A}(\lambda)$ is suitable. If λ is outside the bulk, then

$$\tilde{D}(\lambda) = U_\epsilon(\lambda)(1 + O(\epsilon)), \tag{2.133}$$

where $U_\epsilon(\lambda) \neq 0$ is the function defined by (2.138), while for λ in the bulk,

$$\tilde{D}(\lambda) = V_\epsilon(\lambda)\left(\sin(\pi T_\epsilon(\lambda)) + O(\epsilon)\right) \tag{2.134}$$

where $V_\epsilon(\lambda) \neq 0$ is the function defined by (2.143), and

$$T_\epsilon(\lambda) := \frac{1}{2\pi\epsilon} \int_{-\max\{u_0\}}^{\lambda} (x_+(\eta) - x_-(\eta))\, d\eta + \frac{1}{2\pi}(\psi_+(\lambda) - \psi_-(\lambda)) \tag{2.135}$$

with $\psi_\pm(\lambda)$ being defined by (2.126). The error terms in (2.133) and (2.134) are uniform on compact subintervals of the indicated sets.

Proof. Let Q be the index of the row of $\tilde{\mathbf{A}}(\lambda)$ whose integration contour may pass through one or two real critical points. Since W_m contains exactly one dominant critical point $x_m = x_m(\lambda) \in \mathbb{C}^+$ for $m \neq Q$, from the method of steepest descent we obtain

$$\tilde{A}_{mp}(\lambda) = \sqrt{\frac{2\pi\epsilon}{|h''(x_m;\lambda)|}} e^{i\theta_m} e^{-ih_m(\lambda)/\epsilon} \left(\frac{1}{x_m - z_p} + O(\epsilon) \right) \tag{2.136}$$

as $\epsilon \to 0$, where $h_m(\lambda) := h(x_m(\lambda);\lambda)$, and θ_m is the steepest descent direction with which W_m traverses the critical point x_m. This same formula also holds for $m = Q$, provided $\lambda \in \mathbb{R}^- \setminus \overline{\mathcal{B}}$ (outside the bulk). Aside from the Cauchy factors $(x_m - z_p)^{-1}$ there is no dependence on the column index p in these leading terms. Therefore, by row-multilinearity of the determinant, if $\lambda \in \mathbb{R}^- \setminus \overline{\mathcal{B}}$,

$$\tilde{D}(\lambda) = (2\pi\epsilon)^{P/2} \left(\prod_{m=1}^{P} \frac{e^{i\theta_m} e^{-ih_m(\lambda)/\epsilon}}{\sqrt{|h''(x_m;\lambda)|}} \right) \det\left(\mathbf{C} + O(\epsilon)\right), \tag{2.137}$$

where $\mathbf{C}$ is the $P \times P$ Cauchy matrix $C_{mp} := (x_m - z_p)^{-1}$. This matrix is invertible because the complex critical points $\{x_m\}_{m=1}^P$ are distinct, as are the poles $\{z_p\}_{p=1}^P$. Hence if the nonvanishing function $U_\epsilon(\lambda)$ is defined by

$$U_\epsilon(\lambda) := (2\pi\epsilon)^{P/2} \left(\prod_{m=1}^{P} \frac{e^{i\theta_m} e^{-ih_m(\lambda)/\epsilon}}{\sqrt{|h''(x_m;\lambda)|}} \right) \det\left(\mathbf{C}\right), \tag{2.138}$$

the proof of (2.133) is complete.

On the other hand, if $\lambda \in \mathcal{B}$ (in the bulk), the contour W_Q traverses two simple real critical points $x_\pm = x_\pm(\lambda)$ by the Klaus-Shaw condition, both of which contribute at the leading order to $\tilde{A}_{Qp}(\lambda)$:

$$\begin{aligned}\tilde{A}_{Qp}(\lambda) =& \sqrt{\frac{2\pi\epsilon}{|h''(x_+(\lambda);\lambda)|}} e^{i\pi/4} e^{-ih_+(\lambda)/\epsilon} \left(\frac{1}{x_+(\lambda) - z_p} + O(\epsilon) \right) \\ &+ \sqrt{\frac{2\pi\epsilon}{|h''(x_-(\lambda);\lambda)|}} e^{-i\pi/4} e^{-ih_-(\lambda)/\epsilon} \left(\frac{1}{x_-(\lambda) - z_p} + O(\epsilon) \right),\end{aligned} \tag{2.139}$$

where $h_\pm(\lambda)$ are defined in Lemma II.43. Then at the cost of a factor $(-1)^{P-Q}$ we may split the last row before computing determinants to obtain

$$\begin{aligned}\tilde{D}(\lambda) =& (2\pi\epsilon)^{P/2} (-1)^{P-Q} \left(\prod_{m=1}^{P-1} \frac{e^{i\theta_m} e^{-ih_m(\lambda)/\epsilon}}{\sqrt{|h''(x_m;\lambda)|}} \right) \\ &\cdot \left(\frac{e^{i\pi/4} e^{-ih_+(\lambda)/\epsilon} \det(\mathbf{C}^+)}{\sqrt{|h''(x_+(\lambda);\lambda)|}} + \frac{e^{-i\pi/4} e^{-ih_-(\lambda)/\epsilon} \det(\mathbf{C}^-)}{\sqrt{|h''(x_-(\lambda);\lambda)|}} + O(\epsilon) \right)\end{aligned} \tag{2.140}$$

where the matrix $\mathbf{C}^\pm$ is the Cauchy matrix $\mathbf{C}$ with the last row replaced by $((x_\pm(\lambda) - z_1)^{-1}, \ldots, (x_\pm(\lambda) - z_P)^{-1})$, and we used the fact that $|e^{-ih_\pm(\lambda)/\epsilon}| = 1$. Next we observe that

$$\begin{aligned}\det\left(\mathbf{C}^+\right) &= Q_{P-1}(x_1, \ldots, x_{P-1}) \frac{\Psi^-(x_+(\lambda);\lambda)}{x_+(\lambda) - x_-(\lambda)} \\ \det\left(\mathbf{C}^-\right) &= -Q_{P-1}(x_1, \ldots, x_{P-1}) \frac{\Psi^+(x_-(\lambda);\lambda)^*}{x_+(\lambda) - x_-(\lambda)},\end{aligned} \tag{2.141}$$

where $\Psi^\pm$ are defined by (2.113) and where $Q_{P-1} \neq 0$ is given by (2.102). Using (2.114) and Lemma II.43, noting that because u_0 is a positive rKS potential

$$\begin{aligned}h_+(\lambda) - h_-(\lambda) &= \int_{-\infty}^{\infty} u_0(x)\, dx - \int_{\lambda}^{0} (x_+(\eta) - x_-(\eta))\, d\eta \\ &= \int_{-\max\{u_0\}}^{\lambda} (x_+(\eta) - x_-(\eta))\, d\eta,\end{aligned} \tag{2.142}$$

and defining

$$V_\epsilon(\lambda) := \frac{-2i(-1)^{P-Q}(2\pi\epsilon)^{P/2}P(\lambda)}{\sqrt{-\lambda(x_+(\lambda)-x_-(\lambda))}} \cdot \left(\prod_{m=1}^{P-1} \frac{e^{i\theta_m}e^{-ih_m(\lambda)/\epsilon}}{\sqrt{|h''(x_m;\lambda)|}}\right) e^{-i(h_+(\lambda)+h_-(\lambda))/(2\epsilon)}e^{-i(\psi_+-\psi_-)/2} \tag{2.143}$$

then establishes (2.134) and completes the proof. ∎

Proposition II.49 allows us to establish a number of important asymptotic properties of the discrete eigenvalues for positive rKS potentials.

Theorem II.50 (Uniform approximation of eigenvalues)**.** Let u_0 be a positive rKS potential and suppose that for each $\lambda \in K \subset \mathcal{B}$, K compact, a suitable modification $\tilde{\mathbf{A}}(\lambda)$ of $\mathbf{A}(\lambda)$ can be found. Then there is a constant C_K such that for each eigenvalue λ in K there exists $\lambda_0 < 0$ satisfying $T_\epsilon(\lambda_0) \in \mathbb{Z}$, such that $|\lambda - \lambda_0| \le C_K\epsilon^2$ holds for all sufficiently small $\epsilon > 0$. Likewise, for each $\lambda_0 \in K$ satisfying $T_\epsilon(\lambda_0) \in \mathbb{Z}$ there is an eigenvalue λ such that the same estimate holds true.

Proof. The eigenvalues are characterized exactly by $\tilde{D}(\lambda) = 0$, or using Proposition II.49 in the case $\lambda \in \mathcal{B}$, $\sin(\pi T_\epsilon(\lambda)) = O(\epsilon)$, the error term being uniform for $\lambda \in K \subset \mathcal{B}$. Solving this equation for $T_\epsilon(\lambda)$ and multiplying by ϵ gives, for some $n \in \mathbb{Z}$,

$$\frac{1}{2\pi}\int_{-\min\{u_0\}}^{\lambda} (x_+(\eta)-x_-(\eta))\,d\eta + \frac{1}{2\pi}\epsilon(\psi_+(\lambda)-\psi_-(\lambda)) = \epsilon n + O(\epsilon^2) \tag{2.144}$$

uniformly on K. Since the left-hand side is differentiable with respect to λ with a derivative that is strictly positive on K, the result follows from the Implicit Function Theorem. ∎

Corollary II.51 (Local approximation of eigenvalues)**.** Fix a closed interval $K \subset \mathcal{B}$ and a positive integer J. Under the hypotheses of Theorem II.50, for each $\Lambda \in K$ the $2J+1$

eigenvalues λ closest to Λ are given by

$$\lambda = \Lambda + 2\pi\epsilon \left(\frac{j + [T_\epsilon(\Lambda)] - T_\epsilon(\Lambda)}{x_+(\Lambda) - x_-(\Lambda)} \right) + O(\epsilon^2), \quad j = -J, -(J-1), \ldots, J-1, J, \tag{2.145}$$

in the limit $\epsilon \to 0$, where the error term depends on J and K, and where $[\cdot]$ denotes the nearest integer function. Note that $|[T_\epsilon(\Lambda)] - T_\epsilon(\Lambda)| \leq 1/2$.

Proof. We set $\lambda = \Lambda + \Delta$ and write the exact eigenvalue condition as $\epsilon T_\epsilon(\Lambda + \Delta) = \epsilon n + O(\epsilon^2)$ for some integer n. Writing the integer n in the form $n = [T_\epsilon(\Lambda)] + j$ for an integer $j \in [-J, J]$ and applying Taylor expansion,

$$\epsilon T_\epsilon'(\Lambda)\Delta + O(\Delta^2) = \epsilon \left(j + [T_\epsilon(\Lambda)] - T_\epsilon(\Lambda) \right) + O(\epsilon^2), \tag{2.146}$$

because derivatives of ϵT_ϵ are bounded uniformly for $\lambda \in K$. Now we seek solutions $\Delta = \lambda - \Lambda$ that are $O(\epsilon)$, so the error terms may be combined and we may solve for Δ:

$$\lambda - \Lambda = \frac{\epsilon(j + [T_\epsilon(\Lambda)] - T_\epsilon(\Lambda)) + O(\epsilon^2)}{\epsilon T_\epsilon'(\Lambda)} \tag{2.147}$$

Since $\epsilon T_\epsilon'(\Lambda) = (x_+(\Lambda) - x_-(\Lambda))/(2\pi) + O(\epsilon)$ uniformly for $\Lambda \in K$ the proof is complete. ■

This result shows that the eigenvalues are locally equally spaced with $O(\epsilon)$ spacing that depends on the point Λ of local expansion. Ignoring the details of the equal spacing and the offset of the grid given by the term $[T_\epsilon(\Lambda)] - T_\epsilon(\Lambda)$, we may reproduce a result obtained by Matsuno [49] by formal asymptotic analysis of trace formulas (conservation laws).

Corollary II.52 (Matsuno's density formula [49])**.** The asymptotic density of eigenvalues at a point λ in the bulk is $\rho_{\mathrm{M}}(\lambda)/\epsilon$, where Matsuno's density is

$$\rho_{\mathrm{M}}(\lambda) := \frac{1}{2\pi}(x_+(\lambda) - x_-(\lambda)). \tag{2.148}$$

In other words, the number $N[a,b]$ of eigenvalues in a subinterval $[a,b]$ of the bulk $\mathcal{B}$ satisfies

$$N[a,b] = \frac{1}{\epsilon}\int_a^b \rho_{\mathrm{M}}(\lambda)\,d\lambda + O(1), \quad \epsilon \to 0. \tag{2.149}$$

For all positive rKS potentials u_0, Matsuno's density ρ_{M} vanishes at the soft edge and blows up at the hard edge $\lambda = 0$.

2.3.4 Phase Constants

Fix a number $\Lambda \in \mathcal{B}$ and for each $\epsilon > 0$ sufficiently small let λ_ϵ be the eigenvalue closest to Λ. It follows from Corollary II.51 that there exists some constant $C > 0$ depending only on Λ such that $|\Lambda - \lambda_\epsilon| \le C\epsilon$ holds for all $\epsilon > 0$ sufficiently small. Denote by γ_ϵ and $\Phi_\epsilon(x)$ the phase constant and normalized eigenfunction, respectively, associated to the eigenvalue λ_ϵ. Then (2.21) takes the form

$$\begin{aligned}
\gamma_\epsilon &= \frac{\epsilon}{2\pi\lambda_\epsilon}\int_{-\infty}^{\infty} \Phi_\epsilon(x)^*\,(x\Phi_\epsilon(x) - 1)\,dx \\
&= \lim_{R\to+\infty}\left[\int_{-R}^{R} xI_\epsilon(x)\,dx - \frac{\epsilon}{2\pi\lambda_\epsilon}\int_{-R}^{R}\Phi_\epsilon(x)^*\,dx\right] \\
&= \lim_{R\to+\infty}\int_{-R}^{R} xI_\epsilon(x)\,dx - \frac{i\epsilon}{2\lambda_\epsilon}, \quad I_\epsilon(x) := \frac{\epsilon}{2\pi\lambda_\epsilon}|\Phi_\epsilon(x)|^2.
\end{aligned} \tag{2.150}$$

In the last step we deformed the contour $[-R,R]$ to a semicircle in the lower half-plane where $\Phi_\epsilon(z^*)^*$ is analytic and satisfies $\Phi_\epsilon(z^*)^* = z^{-1} + O(z^{-2})$ as $z \to \infty$; see Corollary II.30.

Proposition II.53. Let $\Lambda \in \mathcal{B}$ be such that the complex critical points of $h(\cdot;\Lambda)$ are simple, and set $L = L(\Lambda) := 1 + \max\{|x_+(\Lambda)|, |x_-(\Lambda)|\}$. The function $I_\epsilon(x)$ defined for $x \in \mathbb{R}$ in (2.150) has the following properties.

1. There is a constant $K_0 = K_0(\Lambda)$ such that

$$\left| I_\epsilon(x) - \frac{\epsilon}{2\pi\lambda_\epsilon x^2} \right| \leq \frac{K_0\epsilon}{|x|^3}, \quad |x| \geq L \tag{2.151}$$

holds for all $\epsilon > 0$ sufficiently small.

2. There exists a constant $K_1 = K_1(\Lambda)$ such that

$$|I_\epsilon(x)| \leq K_1, \quad |x| \leq L \tag{2.152}$$

holds for all $\epsilon > 0$ sufficiently small.

3. For each real $x \neq x_\pm(\Lambda)$,

$$\lim_{\epsilon\to 0} I_\epsilon(x) = I_0(x) := -\frac{\chi_{(x_-(\Lambda),x_+(\Lambda))}(x)}{x_+(\Lambda) - x_-(\Lambda)}, \tag{2.153}$$

where $\chi_{(a,b)}(x)$ denotes the characteristic function of (a, b).

Proposition II.53 is proved in [54]. The proof hinges on the fact that the method of stationary phase is essentially unchanged when we allow λ to depend mildly on ϵ as in the form above. We then use this fact to obtain estimates and boundedness results on the eigenfunction Φ_ϵ in the regions $|x| > L$ and $|x| < L$ separately. Its main purpose is to prove the following theorem.

Theorem II.54. For each $\Lambda \in \mathcal{B}$,

$$\lim_{\epsilon\to 0} \gamma_\epsilon = -\frac{1}{2}\left(x_+(\Lambda) + x_-(\Lambda)\right), \tag{2.154}$$

where γ_ϵ is the phase constant for the eigenvalue λ_ϵ nearest Λ.

Proof. Since $\lambda_\epsilon \to \Lambda < 0$ as $\epsilon \to 0$, from (2.150) we see that $\mathrm{Im}\{\gamma_\epsilon\} \to 0$ as $\epsilon \to 0$, so it remains to analyze the real part. Next, recalling $L = L(\Lambda)$ as defined in Proposition II.53

and using property 1 from the same yields, for each $R > L$,

$$\begin{aligned}\left|\int_{\mathcal{I}_L^R} xI_\epsilon(x)\,dx\right| &\le \left|\int_{\mathcal{I}_L^R} \frac{\epsilon\,dx}{2\pi\lambda_\epsilon x}\right| + \int_{\mathcal{I}_L^R} |x|\left|I_\epsilon(x) - \frac{\epsilon}{2\pi\lambda_\epsilon x^2}\right| dx \\ &= \int_{\mathcal{I}_L^R} |x|\left|I_\epsilon(x) - \frac{\epsilon}{2\pi\lambda_\epsilon x^2}\right| dx \le \int_{\mathcal{I}_L^R} \frac{K_0\epsilon}{x^2}\,dx \le \frac{2K_0\epsilon}{L},\end{aligned} \tag{2.155}$$

an upper bound that is independent of $R > L$ and that tends to zero with ϵ, where $\mathcal{I}_L^R := [-R,-L]\cup[L,R]$. Hence

$$\lim_{\epsilon\to 0}\gamma_\epsilon = \lim_{\epsilon\to 0}\int_{-L}^{L} xI_\epsilon(x)\,dx, \tag{2.156}$$

so applying the Lebesgue Dominated Convergence Theorem to calculate the latter limit, making use of properties 2 (integrable ϵ-independent domination) and 3 (pointwise limit), gives

$$\lim_{\epsilon\to 0}\gamma_\epsilon = \int_{-L}^{L} xI_0(x)\,dx = -\int_{x_-(\Lambda)}^{x_+(\Lambda)} \frac{x\,dx}{x_+(\Lambda) - x_-(\Lambda)}. \tag{2.157}$$

Evaluating the integral then proves (2.154). ■

Remark II.55. Theorem II.54 shows that the limit of γ_ϵ is a purely real number. For some applications, however, the imaginary part $\mathrm{Im}\{\gamma_\epsilon\} = -i\epsilon/(2\lambda_\epsilon)$ cannot be neglected outright and gives rise to important terms of the zero-dispersion analysis, e.g., see [55].

A version of the asymptotic formula (2.154) characterizing the real part of γ_ϵ for small ϵ was hypothesized in [55]. Theorem II.54 provides a careful restatement and rigorous proof of this result.

2.4 Numerical Verification

We illustrate the accuracy of our asymptotic formulas for the scattering data by comparing them to exact calculations in the case of the rKS potential:

$$u_0(x) = -\nu i \left(\frac{2}{x-i} + \frac{1}{x-(i+1)} \right) + c.c. \tag{2.158}$$

for $\nu = \pm 1$. The graph of u_0 is plotted in Figure 2.4 for $\nu = 1$ (left panel; strictly positive) and $\nu = -1$ (right panel; strictly negative), which confirms the Klaus-Shaw condition. The bulk $\mathcal{B}$ consists of the interval $(-\max\{u_0\}, 0)$ when $\nu = 1$ with $\max\{u_0\} \approx 5.07308$ and $(0, -\min\{u_0\})$ when $\nu = -1$ with $\min\{u_0\} \approx -5.07308$; see Section 2.3.1 and Definition II.38. We selected this rKS potential so that it is not even about any point, a property that makes the phase constants γ_j nontrivial to calculate and interesting to compare with small-dispersion asymptotics; see Remark II.4.

We compute the exact eigenvalues λ_j for the potential (2.158) with $\nu = 1$ using the Evans function (2.100). To simplify computations only values of $\epsilon > 0$ for which $ic_1/\epsilon \in \mathbb{Z}$ and $ic_2/\epsilon \in \mathbb{Z}$ are considered. Since ic_1 and ic_2 are integers, this requires $\epsilon = 1/m$ for some $m \in \mathbb{N}$. In this case, the integrals (2.82) defining the elements of the matrix $\mathbf{A}(\lambda)$ can be calculated explicitly by the Residue Theorem and the exact eigenvalues are thus obtained as the roots of a polynomial; see Remark II.26.

In Figure 2.5 we show the exact eigenvalues (black dots) and compare them (i) with their uniform approximations obtained from Theorem II.50 by solving $T_\epsilon(\lambda) = n$ for positive integers n (overlaid squares) and (ii) with their local equally-spaced approximations described by (2.145) in Corollary II.51 (overlaid circles), under the scaling $(\lambda - \Lambda)/\epsilon$. As expected, the uniform approximation's squares track the black dots very well, while the circles do so best near the point $\lambda = \Lambda$. In Figure 2.6 we present histograms of the exact eigenvalues to highlight their distribution on $\mathbb{R}^-$. The histograms clearly match better, in the limit $\epsilon \to 0$, the density $\rho_{\mathrm{M}}(\lambda)$ of Matsuno (Corollary II.52), here normalized to have integral equal to

unity.

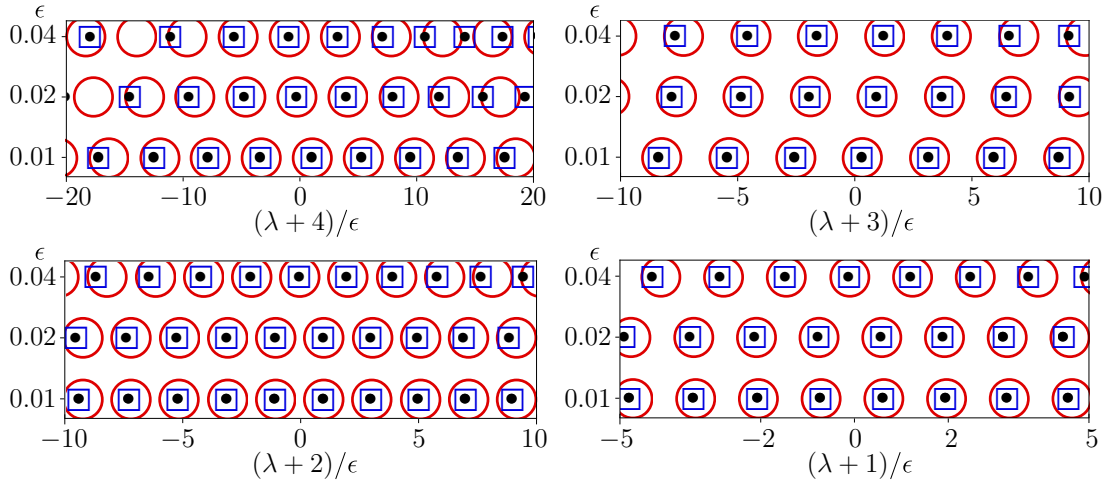

Figure 2.5: Exact eigenvalues for the rKS potential (2.158) with $\nu = 1$ as a function of ϵ (Black dots) with the uniform approximations (squares) and the local approximations based at various values of $\Lambda \in \mathcal{B}$ (circles) overlaid. Note that the horizontal axis is the rescaled local coordinate $(\lambda - \Lambda)/\epsilon$ with $\Lambda = -4$ (top left), $\Lambda = -3$ (top right), $\Lambda = -2$ (bottom left), and $\Lambda = -1$ (bottom right).

In Figure 2.7 we illustrate Theorem II.54 by plotting the real parts of the three exact phase constants γ_ϵ corresponding to eigenvalues λ_ϵ closest to $\Lambda = -1, -4, -5$ in the bulk $\mathcal{B}$. The limiting values predicted by Theorem II.54 are indicated with dashed lines. The exact values of $\mathrm{Re}\{\gamma_\epsilon\}$ were computed using the alternate formula for γ_ϵ presented in [53].

The reflection coefficient is calculated exactly from (2.85) with the coefficients $v_p(\lambda)$ computed by solving the linear system (2.82) for $\nu = \pm 1$. To aid in numerical computation we deform the path of integration in (2.85) into the complex plane to exploit the exponential decay of the integrand. We compare the exact reflection coefficient β with its asymptotic approximation (Theorem II.46) in Figures 2.8 and 2.9. In Figure 2.8 we plot the normalized magnitude $\sqrt{\epsilon}|\beta|$ as a function of the spectral parameter λ. The left panel ($\nu = -1$) shows that as $\epsilon \to 0$, $\sqrt{\epsilon}|\beta|$ indeed approaches the expected limit — whose support is the bulk — given by Matsuno's modulus formula (Corollary II.47). The right panel ($\nu = 1$) shows that $\sqrt{\epsilon}|\beta| \to 0$ for positive rKS potentials as predicted by Theorem II.46. In Figure 2.9 we plot the derivative of the phase of β (computed indirectly using $\epsilon \mathrm{Im}\{\beta'(\lambda)/\beta(\lambda)\}$)

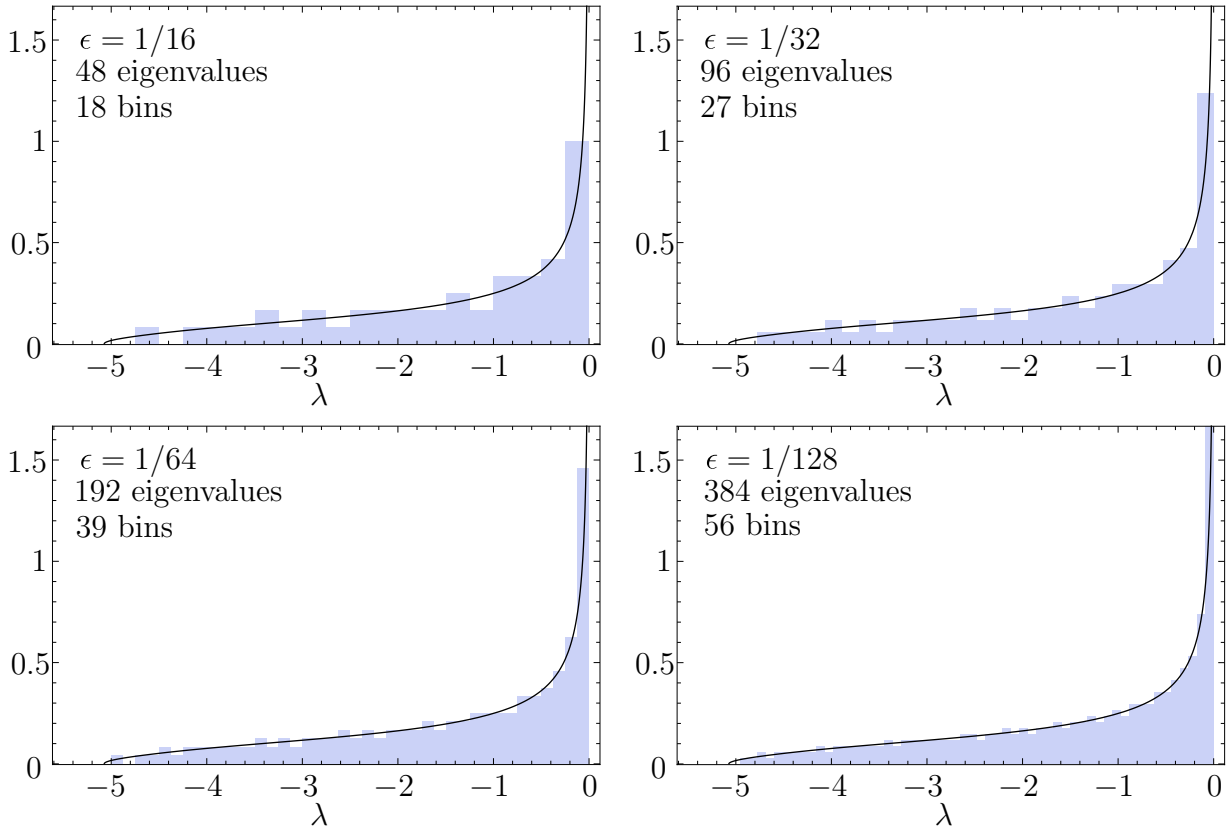

Figure 2.6: Histograms of exact eigenvalues for the rKS potential (2.158) with $\nu = 1$ illustrating the distribution of eigenvalues. The limiting curve (solid black) is Matsuno's density $\rho_{\mathrm{M}}(\lambda)$ normalized to unit mass.

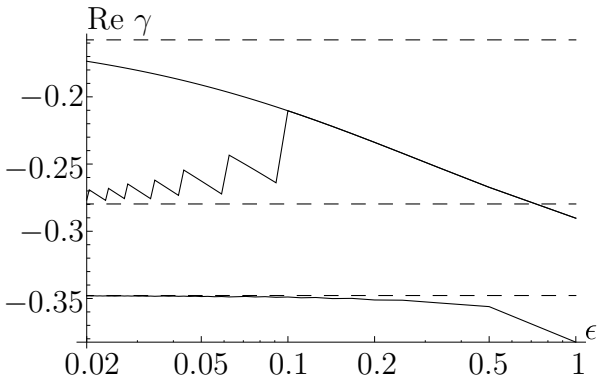

Figure 2.7: Three values of $\mathrm{Re}\{\gamma_\epsilon\}$ (solid lines) as a function of ϵ for the rKS potential (2.158) with $\nu = 1$. The limiting values (dashed lines) are from (2.154) evaluated for $\Lambda = -1, -4, -5$. The values $\mathrm{Re}\{\gamma_\epsilon\}$ are computed at each ϵ with the exact eigenvalue λ_ϵ chosen to be the closest to Λ.

and compare with the corresponding limiting curve predicted by Theorem II.46, namely $-\theta'_+(\lambda) = -x_+(\lambda)$.

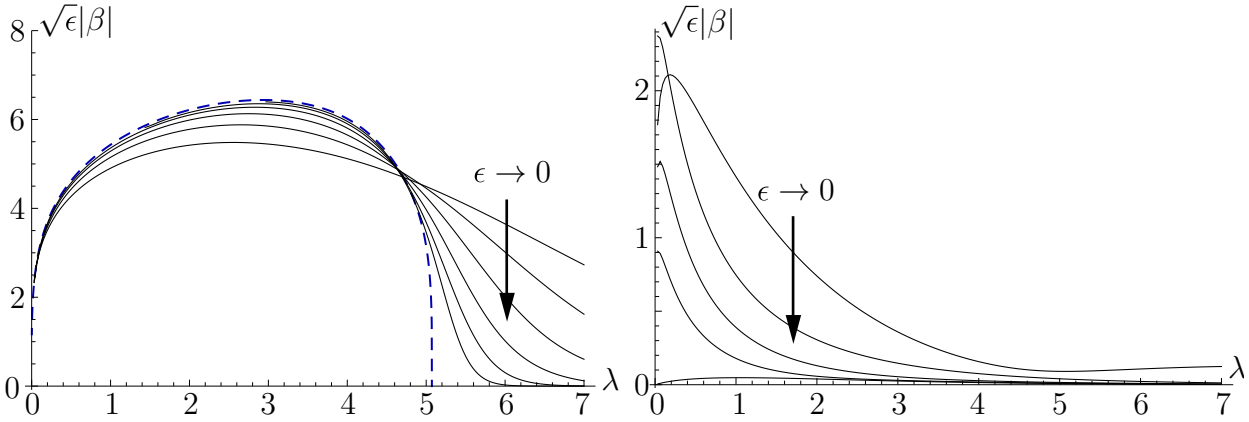

Figure 2.8: Left: $\sqrt{\epsilon}|\beta|$ as a function of λ for the rKS potential (2.158) with $\nu = -1$ for $\epsilon = 2,\ 1,\ 1/2,\ 1/4,\ 1/8,\ 1/16$ (solid black curves). The apparent limiting curve (dashed-blue) is obtained from Corollary II.47. For ϵ and λ both small the graphs become difficult to compute and are not plotted. Right: Same as the left panel but for $\nu = 1$ and $\epsilon = 4,\ 2,\ 7/4,\ 13/8,\ 3/2$, showing convergence to zero.

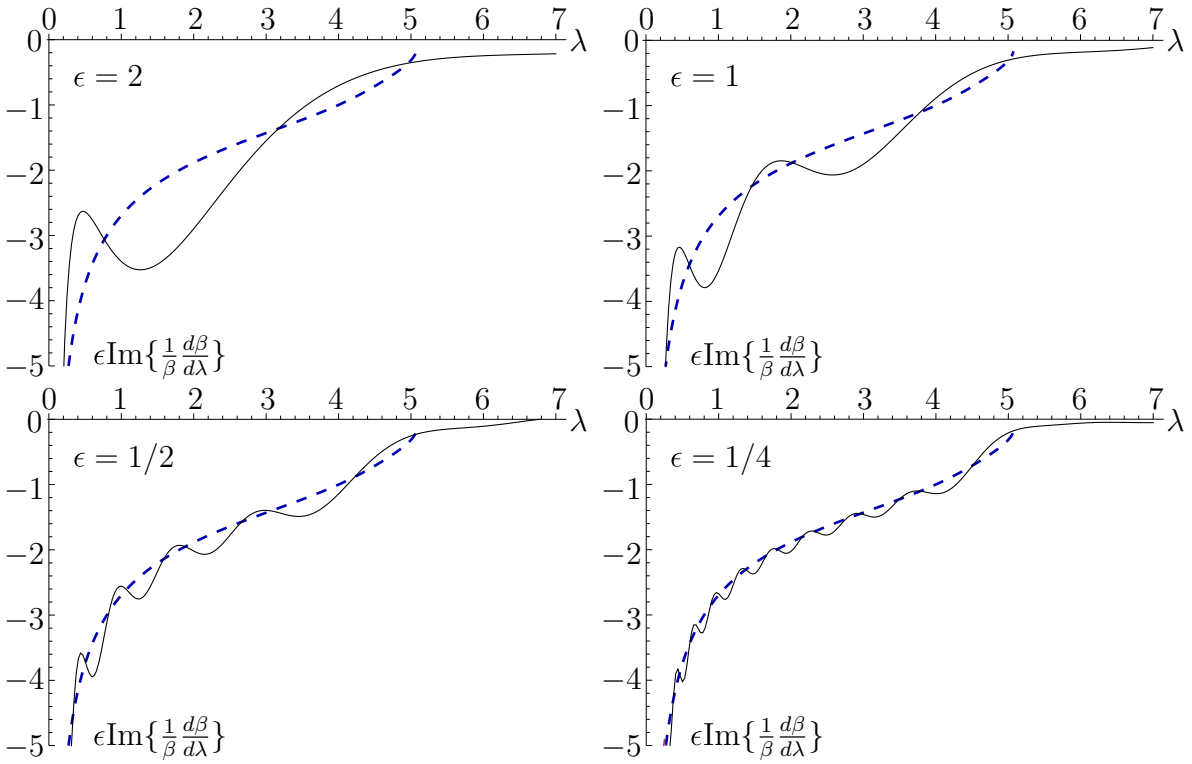

Figure 2.9: Plots of the rescaled derivative of the phase for the function β as a function λ. The limiting curve $-x_+(\lambda)$ (dashed-blue) is obtained from taking the derivative of the leading order phase $-\theta_+(\lambda)$ in (2.125).

CHAPTER III

Effects of Stratification on the Surface Tides in a Two-Layer Tidal Model

The contents of this chapter are currently in revision [83] and are presented here nearly verbatim with some modifications to improve readability and maintain a consistent format.

3.1 Illustrative Global Numerical Simulations

To illustrate the effects of stratification on large-scale tides, in somewhat greater detail than in [5] and [75], we briefly discuss numerical one- and two-layer tide simulations performed in a realistic near-global domain. The realistic-domain simulations are executed with the Hallberg Isopycnal Model (HIM) [31] on a 1/8° grid covering the latitudes from 86°S to 82°N. The realistic-domain model satisfies the Boussinesq approximation [81, for example], a point we will return to later. We apply the astronomical forcing of the principal lunar semidiurnal tide M_2 only. The self-attraction and loading term is computed iteratively using the full spherical harmonic treatment [33]. As in [5] and many other studies, we employ a parameterized topographic internal wave drag. As in [5], the wave drag acts only on the bottom flow in our two-layer simulations. Because the exact details of the three-dimensional geography of tidal dissipation are still a matter of active research, the assumed format of the damping employed here could be debated. Our assumption is that wave drag represents

damping due to breaking of high-vertical-mode internal waves (unresolved in a two-layer model) generated by bottom flow over topography. Here, we use a multiplicative topographic wave drag factor of 5, where the wave drag acts only in waters deeper than 1000 m. This topographic drag factor is chosen to minimize the discrepancy in deep-ocean tidal elevations between the tide model and the GOT99 satellite-altimeter constrained tide model [67]. This value of the drag factor yields a 6.69 cm Root Mean Square (RMS) elevation error, comparable to the errors seen in other tuned forward tide models, in the one-layer run. We refer the interested reader to [5] for further details on the model setup and parameters.

We evaluate the strength of the small-scale tides in the 1/8° HIM results utilized in this chapter via the same methods used to evaluate the 1/12° HYbrid Coordinate Ocean Model (HYCOM) results in [8]. The small-scale tides are separated out using a spatial high-pass filter applied to M_2 amplitudes and phases in a North Pacific region of strong internal tides. In the HIM "interface perturbation" run to be described shortly, the RMS M_2 internal tide perturbations to sea surface elevations, averaged over a large box around Hawai'i 35 degrees in latitude by 50 degrees in longitude, are 0.64 cm for amplitude and 3.56 degrees for phase. The 1/8° "g' perturbation" and "control" runs to be described shortly yield very similar internal tide strengths. These numbers compare well to the 0.87 cm and 4.35 degree perturbations estimated in the same box from the along-track satellite altimeter observations of [70, 71]. Therefore the 1/8° two-layer HIM simulations produce internal tides that are of comparable magnitudes to those in observations.

The impacts of stratification, and of climatic perturbations to stratification, on the modeled surface tidal elevations are displayed in Figures 3.1 and 3.2. Figure 3.1(a) displays the difference in surface elevation amplitudes between the two-layer control simulation, having an interface at 700 m and a reduced gravity of $g' = 1.64 \times 10^{-2}$ m s^{-2}, and the one-layer simulation. Differences of up to 5 cm (about 5% of a typical 1 m signal) between the two- and one-layer amplitudes are clearly seen. Notably, the amplitude differences are clearly in the large-scale tides, as well as in the small-scale tides that are introduced into the two-

layer case. Figure 3.1(b) demonstrates that phase changes of up to 5° take place when stratification is introduced into the model. Figure 3.2(a) displays a global map of the RMS (in time) M_2 surface elevation differences between the two-layer and one-layer simulations. The RMS computations account for differences in both amplitude and phase, in the manner usually assumed in tidal computations [5, amongst many]. Again, differences are seen in both large- and small-scales, and are up to 4 cm in some regions. A climatic perturbation to the stratification in the two-layer numerical model, either in the form of an interface lying at 800 m instead of 700 m, or in an increase in the value of g' to 1.78×10^{-2} m s^{-2}, leads to further alterations of order 1 cm (about 1% of a typical 1 m tidal signal) in the surface tides; see Figures 3.2(b) and (c). Our choices for the perturbed values of interface depth and reduced gravity (g') are meant to roughly mimic centennial-scale climate change and are explained in Section 3.4.2. Figures 3.1 and 3.2 demonstrate that the change in the tides due to stratification and climatically-perturbed stratification is a robust effect. The area-weighted RMS elevation difference over 66°S to 66°N (the range of latitudes covered by the TOPEX/POSEIDON and JASON class altimeters) is 2.70 cm for the two-layer vs. one-layer comparison (Figure 3.2(a)), 0.63 cm for the two-layer interface perturbation vs. two-layer control comparison (Figure 3.2(b)), and 0.66 cm for the two-layer g' perturbation vs. two-layer control comparison (Figure 3.2(c)). The patterns of change in Figures 3.2(b) and (c), and the corresponding area-weighted RMS values, are remarkably similar. Later in this chapter we will use our analytical model to put forth a suggestive explanation for these similarities.

3.2 Governing Equations

In the analytical model, we let $u_1(x,t)$, $u_2(x,t)$, $\eta_1(x,t)$, $\eta_2(x,t)$ be, respectively, the upper and lower layer velocities and the perturbation surface and internal tidal elevations (displacements), where x denotes the horizontal spatial coordinate, t denotes time, and subscripts 1 and 2 denote the upper and lower layers, respectively. We let the resting layer

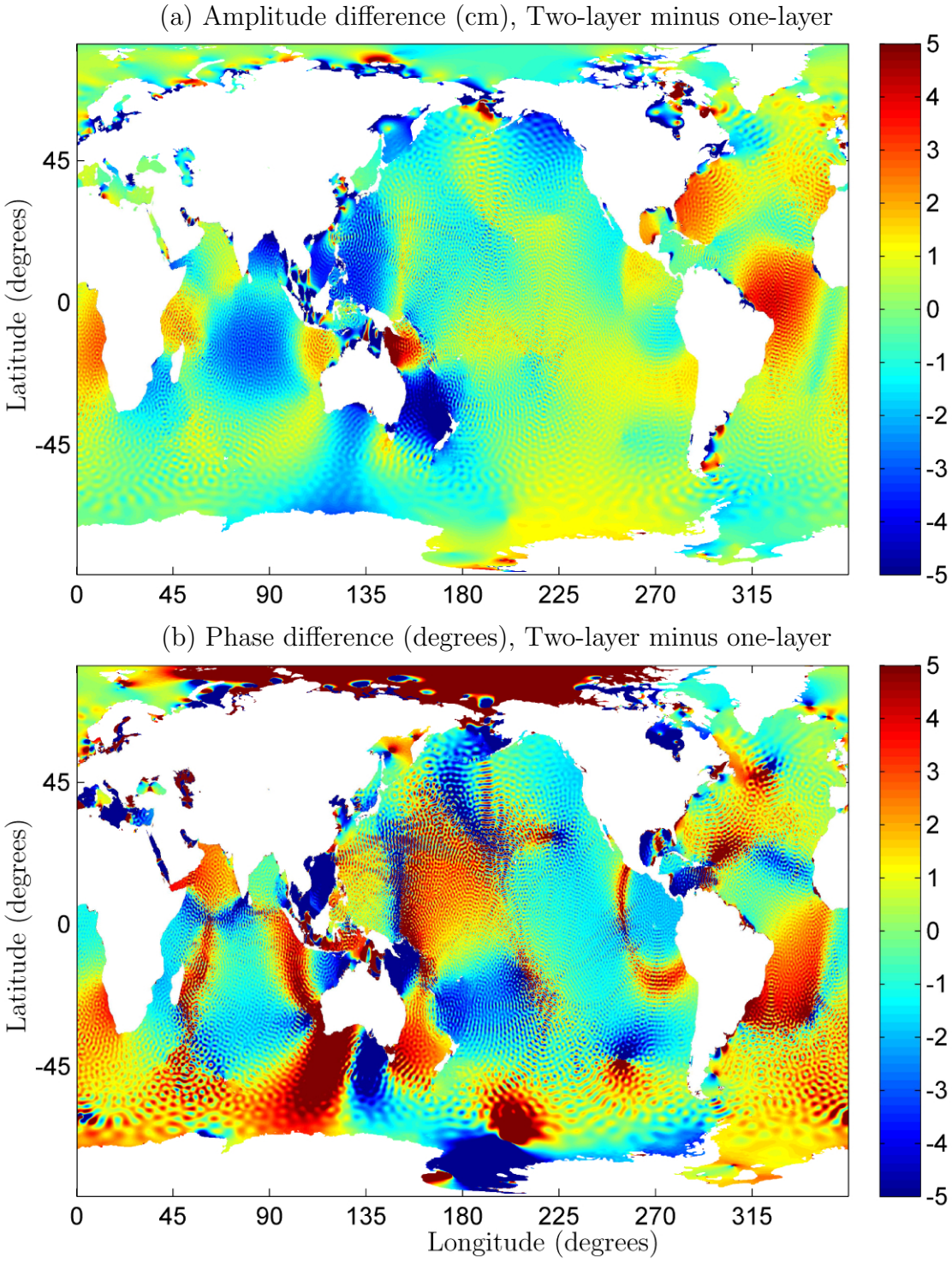

Figure 3.1: (a) Amplitude and (b) phase differences, in the surface elevations of the two-layer 1/8° M_2 control simulation of the global realistic-domain numerical model minus the one-layer 1/8° M_2 control simulation.

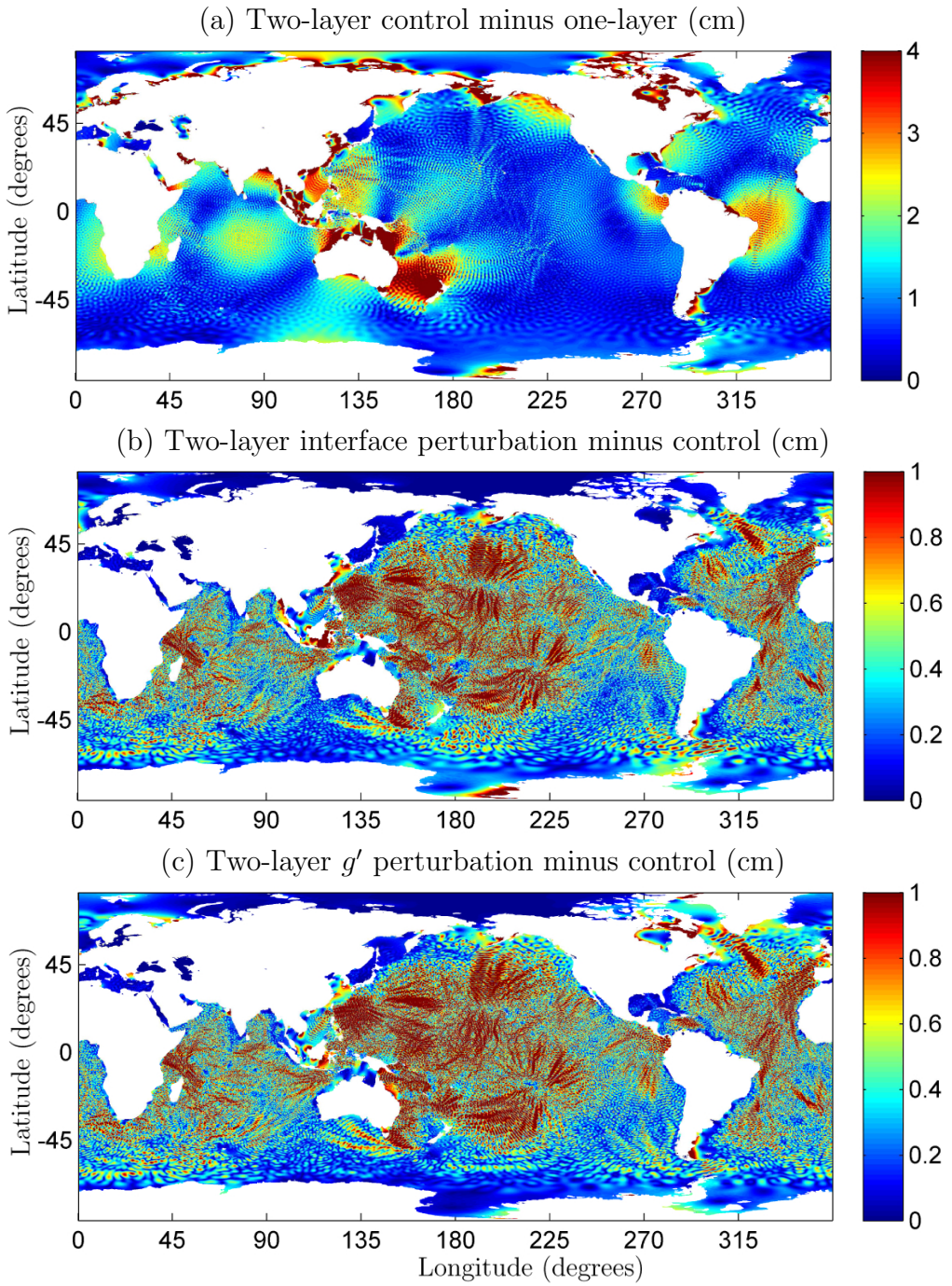

Figure 3.2: RMS in time of surface elevation differences (cm) in 1/8° M_2 global realistic-domain numerical simulations: (a) two-layer control simulation minus one-layer control simulation, (b) two-layer "interface perturbation" simulation (layer interface at 800 m) minus two-layer control simulation (layer interface at 700 m), (c) two-layer "g' perturbation" simulation (perturbed g' value of 1.78×10^{-2} m s^{-2}) minus two-layer control simulation (g' value of 1.64×10^{-2} m s^{-2}).

depths be H_1 and $h_2(x)$. See Figure 3.3 for a sketch of the model notation.

Remark III.1. Note that while H_1 is a constant, the resting depth $h_2(x)$ is variable due to variable bottom topography. We define $H_2 = \max\{h_2(x)\}$, the largest resting depth in the bottom layer.

The assumption of linearized, non-rotating, shallow-water dynamics in one horizontal dimension with astronomical forcing and linear damping yields the upper- and lower-layer mass conservation equations

$$\frac{\partial}{\partial t}(\eta_1 - \eta_2) + H_1 \frac{\partial u_1}{\partial x} = 0, \tag{3.1}$$

$$\frac{\partial \eta_2}{\partial t} + \frac{\partial}{\partial x}\left[h_2(x) u_2\right] = 0, \tag{3.2}$$

respectively, and the upper- and lower-layer momentum equations

$$\frac{\partial u_1}{\partial t} = -g\frac{\partial \eta_1}{\partial x} + g\frac{\partial F_A}{\partial x} - r_1 u_1, \tag{3.3}$$

$$\frac{\partial u_2}{\partial t} = (g' - g)\frac{\partial \eta_1}{\partial x} - g'\frac{\partial \eta_2}{\partial x} + g\frac{\partial F_A}{\partial x} - r_2 u_2, \tag{3.4}$$

respectively. Here g denotes gravitational acceleration, $g' = g(\rho_2 - \rho_1)/\rho_2$ is the reduced gravity with ρ_1 and ρ_2 being the upper and lower layer densities ($\rho_1 \leq \rho_2$), r_1 and r_2 are damping rates on the upper and lower layer flows, respectively, and the astronomical tidal forcing is given by $F_A(x,t) = \eta_0 \cos(Kx + \Omega t)$, where η_0, K, and Ω are the forcing amplitude, wavenumber, and frequency. We impose the no-normal flow boundary conditions

$$u_1(-L,t) = u_1(0,t) = u_2(-L,t) = u_2(0,t) = 0, \tag{3.5}$$

where $x = -L, 0$ are the locations of the basin boundaries.

Omission of the Coriolis force yields a simple, more analytically tractable, one-dimensional model. As shown in Section B.1, the addition of the Coriolis force does not change an essen-

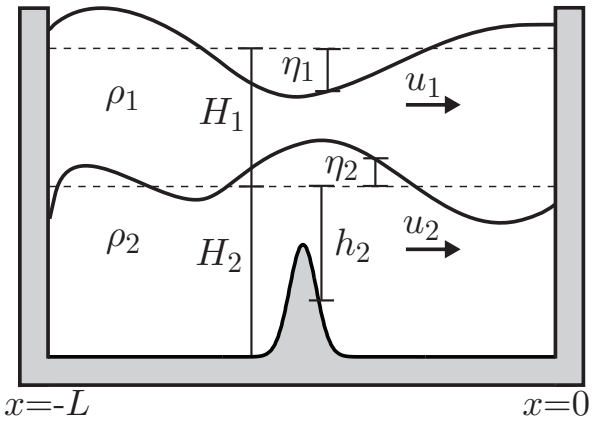

Figure 3.3: Sketch of analytical two-layer model with Gaussian bump topography.

tial argument of the chapter, that the introduction of stratification produces a shift in the phase speed of the large-scale tides. It may, however, lead to significant quantitative changes in the solution. We see the simplicity of the one-dimensional governing equations (3.1)–(3.4) as an asset to more easily highlight the key parameters at work.

We nondimensionalize the governing equations using the scaling:

$$x = Lx^*, \quad t = t^*/\Omega, \quad \eta_j = \eta_0\eta_j^*, \quad u_j = \eta_0\frac{\Omega L}{H_j}u_j^*, \quad h_2 = H_2[1 - \sigma\beta^*(x^*)], \tag{3.6}$$

where $j = 1, 2$, and the asterisks denote nondimensional variables. The parameter L is the basin length. The quantity $\sigma\beta^*(x^*)$ is a nondimensionalization of the bottom topography such that $0 < \beta^* < 1$ and the vertical topographical scaling is given by σ. Hence, H_2 is the maximum value attained by h_2 in the basin.

After dropping asterisks, the governing equations (3.1)–(3.4) can be written nondimensionally as

$$\frac{\partial}{\partial t}(\eta_1 - \eta_2) + \frac{\partial u_1}{\partial x} = 0, \tag{3.7}$$

$$\frac{\partial \eta_2}{\partial t} + \frac{\partial}{\partial x}\left[(1 - \sigma\beta)u_2\right] = 0, \tag{3.8}$$

$$\frac{1}{\epsilon^2\gamma}\left(\frac{\partial u_1}{\partial t} + \delta_1 u_1\right) + \frac{\partial \eta_1}{\partial x} = \frac{\partial}{\partial x}\cos(\phi x + t), \tag{3.9}$$

$$\frac{1}{\epsilon^2(1-\gamma)}\left(\frac{\partial u_2}{\partial t}+\delta_2 u_2\right)+(1-\alpha)\frac{\partial \eta_1}{\partial x}+\alpha\frac{\partial \eta_2}{\partial x}=\frac{\partial}{\partial x}\cos(\phi x+t), \tag{3.10}$$

where the governing nondimensional parameters are

$$\epsilon=\frac{\sqrt{g(H_1+H_2)}}{\Omega L},\quad \gamma=\frac{H_1}{H_1+H_2},\quad \alpha=\frac{g'}{g},\quad \delta_1=\frac{r_1}{\Omega},\quad \delta_2=\frac{r_2}{\Omega},\quad \phi=KL. \tag{3.11}$$

Thus, δ_1, δ_2 are the damping parameters and α, γ are the two stratification parameters. By definition $0 \leq \alpha < 1$, where $\alpha = 0$ represents unstratified flow. Similarly, by definition, $0 \leq \gamma \leq 1$. The parameter ϕ is the nondimensional forcing wavenumber, while ϵ is a tidal resonance parameter. The boundary conditions simply become

$$u_1(-1,t)=u_1(0,t)=u_2(-1,t)=u(0,t)=0. \tag{3.12}$$

Note that setting $\alpha = 0$ and $\delta_1 = \delta_2$ yields the governing equations in the one-layer shallow-water model used in [6]. This follows from the fact that if we set $\alpha = 0$ (so that the layer densities are equal) and $\delta_1 = \delta_2$ (equal damping rates in the two layers), then the momentum equations for both u_1 and u_2 coincide with the momentum equation in the one-layer case after redimensionalization.

In the open ocean, where tidal flows are relatively weak, the quadratic bottom boundary layer drag is also weak [60]. We therefore assume that the r_2 term in the open ocean is dominated by topographic internal wave drag [5, 20, 36, among others]. In our treatment we allow drag on the upper-layer flow to be a convenience that proves useful in examining limiting cases. We believe that the bottom drag-only case ($\delta_1 = 0$) is more relevant for the ocean. We note also that the presence of the coefficient α of $\partial\eta_1/\partial x$ in equation (3.10) arises from the fact that we do not make the Boussinesq approximation when deriving these equations. The effects of stratification on the large-scale and small-scale surface tidal elevations are qualitatively similar in both Boussinesq and non-Boussinesq approximations. For the sake of brevity, the differences between Boussinesq and non-Boussinesq model solutions will

not be studied further in this chapter.

Remark III.2. The particular choice of δ_1, δ_2 significantly affects the sensitivity of the system to the stratification parameter α as is shown in Section 3.4.

3.3 Solution Methods

To solve the two-layer system (3.7)–(3.10) it is enough to obtain a homogeneous and a particular solution of the model. However, since we are not primarily interested in transient behavior, we only seek a particular solution, which would be expected to dominate the large-time asymptotic behavior when damping is present. A particular solution can be obtained through the method of undetermined coefficients, which leads us to seek a harmonically oscillating solution of the form:

$$u_1 = \mathrm{Re}\{U_1(x)e^{it}\}, \quad u_2 = \mathrm{Re}\{U_2(x)e^{it}\}, \quad \eta_1 = \mathrm{Re}\{N_1(x)e^{it}\}, \quad \eta_2 = \mathrm{Re}\{N_2(x)e^{it}\}. \tag{3.13}$$

Substituting (3.13) into the mass equations (3.7)–(3.8) and solving for N_1, N_2 gives the system

$$N_1(x) = i\frac{dU_1}{dx} + i\frac{dV_2}{dx}, \tag{3.14}$$

$$N_2(x) = i\frac{dV_2}{dx}, \tag{3.15}$$

where $V_2 = (1 - \sigma\beta)\, U_2$. This system can be written concisely in vector form as

$$\mathbf{N}(x) = i\,(\mathbb{I} + \mathbf{e}_{12})\,\frac{d\mathbf{U}}{dx}(x), \tag{3.16}$$

where $\mathbf{N}(x) = [N_1 \;\; N_2]^\mathsf{T}$, $\mathbb{I} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ is the identity matrix, $\mathbf{e}_{12} = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$, and $\mathbf{U}(x) = [U_1 \;\; V_2]^\mathsf{T}$; the superscript T denotes the transpose. Similarly, with the use of equations (3.14) and

(3.15), substituting (3.13) into the momentum equations (3.9) and (3.10) gives

$$s_1 U_1 + \frac{d^2 U_1}{dx^2} + \frac{d^2 V_2}{dx^2} = \phi e^{i\phi x}, \tag{3.17}$$

$$s_2 U_2 + (1-\alpha)\frac{d^2 U_1}{dx^2} + \frac{d^2 V_2}{dx^2} = \phi e^{i\phi x}, \tag{3.18}$$

where $s_1 = (1 - i\delta_1)/(\epsilon^2 \gamma)$ and $s_2 = (1 - i\delta_2)/(\epsilon^2(1-\gamma))$. Multiplication of (3.17) by s_2 and (3.18) by s_1 allows us to write these equations in vector form as

$$\mathbf{A}\mathbf{U}(x) + \mathbf{A}\mathbf{e}_{22}\left(\frac{\sigma\beta(x)}{1-\sigma\beta(x)}\right)\mathbf{U}(x) + \mathbf{B}\frac{d^2\mathbf{U}}{dx^2}(x) = \mathbf{F}(x), \tag{3.19}$$

where

$$\mathbf{A} = s_1 s_2 \mathbb{I}, \quad \mathbf{e}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} s_2 & s_2 \\ s_1(1-\alpha) & s_1 \end{bmatrix}, \quad \text{and} \quad \mathbf{F}(x) = \phi \begin{bmatrix} s_2 \\ s_1 \end{bmatrix} e^{i\phi x}. \tag{3.20}$$

In the next few subsections we present two distinct methods, each having different strengths, to solve equations (3.17) and (3.18) or, equivalently, equation (3.19). First, we show the derivation of a Fourier series solution method (infinite series of sines). The Fourier series method is simple to implement and is numerically robust. Second, we introduce a Neumann series solution method (infinite series of operators). The Neumann series method, valid only for small-amplitude topography, suggests an analytically valuable decomposition into large-scale and small-scale modes. We use the Fourier series method for computing results from the analytical model.

In Section B.2 we introduce a scattering solution method, calculated for an infinite basin. We include the scattering method in this study because it bears similarities to the internal tide generation problem which has received much attention from the community [47, 78, 41, among others]. The Neumann series method highlighted shortly has similar algebraic structure to the scattering method. The scattering method is, however, not emphasized in

this discussion because, first, the infinite basin is less realistic than the finite basin, and second, the scattering method requires prior knowledge of the "incident" large-scale tidal velocity on the topography; whereas here we are mainly interested in obtaining, without prior assumptions, both the large- and small-scale tidal solutions.

3.3.1 Fourier Series Solution

We can construct a solution for (3.17) and (3.18) with general topography using Fourier series. To satisfy the boundary conditions the velocities must be expansions of sines only, i.e.,

$$U_1(x) = \sum_{n=1}^{\infty} c_n \sin(n\pi x), \quad U_2(x) = \sum_{n=1}^{\infty} d_n \sin(n\pi x), \quad V_2(x) = \sum_{n=1}^{\infty} e_n \sin(n\pi x), \tag{3.21}$$

and

$$e_n = 2\int_{-1}^{0} (1 - \sigma\beta(x))\, U_2(x) \sin(n\pi x)\, dx = d_n - \sum_{m=1}^{\infty} d_m \sigma_{mn} \tag{3.22}$$

with

$$\sigma_{mn} = 2\sigma \int_{-1}^{0} \beta(x) \sin(m\pi x) \sin(n\pi x)\, dx = \sigma\left(\beta_{m-n} - \beta_{m+n}\right), \tag{3.23}$$

where $\beta_n = \int_{-1}^{0} \beta(x) \cos(n\pi x)\ dx$. We may also write the forcing term in a Fourier series expansion:

$$e^{i\phi x} = \sum_{n=1}^{\infty} f_n \sin n\pi x, \tag{3.24}$$

where

$$f_n = 2\int_{-1}^{0} e^{i\phi x} \sin(n\pi x)\ dx = \frac{2n\pi[1 - (-1)^n e^{-i\phi}]}{\phi^2 - (n\pi)^2}. \tag{3.25}$$

Remark III.3. The convergence of the Fourier series (3.24) is not uniform. It suffers from the Gibbs phenomenon at the boundaries because the term $e^{i\phi x}$ is not zero at the boundaries $x = -1, 0$. This, however, does not affect our analysis since convergence in the $L^2(-1, 0)$ sense (RMS) is guaranteed.

Therefore, equations (3.17) and (3.18) can be written in terms of the coefficients c_n and d_n as

$$\left((n\pi)^2 - s_1\right) c_n + (n\pi)^2 d_n = (n\pi)^2 \sum_{m=1}^{\infty} d_m \sigma_{mn} - \phi f_n, \tag{3.26}$$

$$(n\pi)^2(1-\alpha)c_n + \left((n\pi)^2 - s_2\right) d_n = (n\pi)^2 \sum_{m=1}^{\infty} d_m \sigma_{mn} - \phi f_n. \tag{3.27}$$

Solving this system for the coefficients c_n and d_n gives

$$c_n = \frac{s_2}{D_n}\left(\phi f_n - (n\pi)^2 \sum_{m=1}^{\infty} d_m \sigma_{mn}\right), \tag{3.28}$$

$$d_n = \frac{s_1 - \alpha(n\pi)^2}{D_n}\left(\phi f_n - (n\pi)^2 \sum_{m=1}^{\infty} d_m \sigma_{mn}\right), \tag{3.29}$$

where

$$D_n = s_1 s_2 - (s_1 + s_2)\,(n\pi)^2 + \alpha(n\pi)^4. \tag{3.30}$$

Note that we have written the bottom topography as an extra "forcing" term. Indeed, equation (3.29) for d_n may be viewed as an infinite system of linear equations. To highlight this fact, (3.29) can be written as $\mathbf{Md} = \mathbf{f}$, where $\mathbf{d} = [d_1 \;\; d_2 \;\; \cdots]^{\mathsf{T}}$, the components of $\mathbf{M}$ are

$$M_{nm} = \begin{cases} D_n + (s_1 - \alpha(n\pi)^2)\,(n\pi)^2\sigma_{nn} & \text{for } n = m \\ (s_1 - \alpha(n\pi)^2)\,(n\pi)^2\sigma_{nm} & \text{for } n \neq m, \end{cases} \tag{3.31}$$

and the components of $\mathbf{f}$ are

$$\mathrm{f}_n = \left(s_1 - \alpha(n\pi)^2\right)\phi f_n. \tag{3.32}$$

For a finite number of modes the system can be solved for $\mathbf{d}$ by simply inverting the matrix $\mathbf{M}$. Then, we solve for c_n in (3.28) and using equations (3.14), (3.15) we can calculate the elevations

$$N_1(x) = \sum_{n=1}^{\infty} a_n \cos n\pi x, \quad N_2(x) = \sum_{n=1}^{\infty} b_n \cos n\pi x, \tag{3.33}$$

with

$$a_n = in\pi \left(c_n + d_n - \sum_{m=1}^{\infty} d_m \sigma_{mn} \right), \quad b_n = in\pi \left(d_n - \sum_{m=1}^{\infty} d_m \sigma_{mn} \right). \tag{3.34}$$

3.3.2 Neumann Series Solution

For sufficiently small topography, we can obtain a solution of the model equations using a Green's function derived from the no topography case. This method is thus limited by the vertical scale σ of the topography, but leads to a natural decomposition of the system into large and small horizontal scales, as is discussed further in Section 3.3.3. For the remainder of this subsection we assume $\alpha \neq 0$, ensuring that $\mathbf{B}$ is non-singular. Technical details aside, the one-layer solution is recovered by taking the limit $\alpha \to 0$.

We may nearly uncouple the system (3.19) by diagonalizing the matrix $\mathbf{B}$; $\mathbf{BV} = \mathbf{V\Lambda}$, where $\mathbf{\Lambda}$ is a diagonal matrix of the eigenvalues of $\mathbf{B}$, in decreasing magnitude, and $\mathbf{V}$ is a matrix with the respective eigenvectors. This allows us to change basis, left-multiplying the system by $\mathbf{V}^{-1}$, and rewriting (3.19) as

$$\mathbf{AW}(x) + \mathbf{\Lambda} \frac{d^2\mathbf{W}}{dx^2}(x) = \mathbf{V}^{-1}\mathbf{F}(x) - \sigma \mathbf{A}\mathbf{V}^{-1}\mathbf{e}_{22} \left(\frac{\beta(x)}{1 - \sigma\beta(x)} \right) \mathbf{VW}(x), \tag{3.35}$$

where the left-hand side terms of (3.35) are fully uncoupled. This is easily seen from the fact that $\mathbf{A}$ commutes with any other matrix. The new variable $\mathbf{W}(x) = \begin{bmatrix} U^L(x) & U^S(x) \end{bmatrix}^{\mathsf{T}} = \mathbf{V}^{-1}\mathbf{U}(x)$ represents a separation of the system into large- (L) and small-scale (S) motions. Namely, we associate the first element $U^L(x)$ with the motions arising from the larger of the two eigenvalues (larger scales) of $\mathbf{B}$ and the second element $U^S(x)$ with the motions arising from the smaller of the eigenvalues (smaller scales). As explained in Section 3.3.3, the elements U^L and U^S roughly equate to the barotropic and baroclinic modes of the system, respectively. However, we avoid the words "barotropic" and "baroclinic" in our description of the elements of $\mathbf{W}$, because our separation may not be consistent with all the definitions of these terms in the literature.

Remark III.4. It is not strictly necessary to uncouple the system (3.19) (as is shown in (3.35)) to solve it using the Neumann series approach, but we do so for convenience.

We rewrite (3.35) as

$$\mathbf{E}\mathbf{W}(x) + \frac{d^2\mathbf{W}}{dx^2}(x) = \mathbf{\Lambda}^{-1}\mathbf{V}^{-1}\mathbf{F}(x) - \sigma\mathbf{E}\mathbf{V}^{-1}\mathbf{e}_{22}\left(\frac{\beta(x)}{1-\sigma\beta(x)}\right)\mathbf{V}\mathbf{W}(x), \tag{3.36}$$

where $\mathbf{E} = \mathbf{\Lambda}^{-1}\mathbf{A} = \begin{bmatrix} (k^L)^2 & 0 \\ 0 & (k^S)^2 \end{bmatrix}$. Note that k^L and k^S arise from the eigenvalues of $\mathbf{B}$ and satisfy the polynomial equation

$$k^4 - \left(\frac{s_1+s_2}{\alpha}\right)k^2 + \frac{s_1 s_2}{\alpha} = 0. \tag{3.37}$$

To avoid ambiguity about which roots of (3.37) we refer to, we pick those roots such that $\mathrm{Im}(k^L), \mathrm{Im}(k^S) > 0$ with $|k^L| < |k^S|$. That is, the equation (3.37) has the four well-defined roots $\{\pm k^L, \pm k^S\}$. In the case $\alpha \ll 1$, typical for the ocean, it follows that $|k^L| \ll |k^S|$. In other words, k^L and k^S define two well separated scales (complex wavenumbers) for the problem.

The new system (3.36) can be readily solved using the Green's function of the problem

$$\mathbf{E}\mathbf{W}(x) + \frac{d^2\mathbf{W}}{dx^2}(x) = \mathbf{R}(x), \tag{3.38}$$

where $\mathbf{R}(x)$ stands for the right-hand side of (3.36). The solution of (3.38) is given by

$$\mathbf{W}(x) = \int_{-1}^{0} \mathbf{G}(x,y)\mathbf{R}(y)\,dy \quad \text{with} \quad \mathbf{G}(x,y) = \begin{bmatrix} g^L(x,y) & 0 \\ 0 & g^S(x,y) \end{bmatrix}, \tag{3.39}$$

where $\mathbf{G}(x,y)$, $-1 < x, y < 0$, satisfies the system

$$\mathbf{E}\mathbf{G}(x,y) + \frac{\partial^2\mathbf{G}}{\partial y^2}(x,y) = \delta(y-x)\mathbb{I} \quad \text{with} \quad \mathbf{G}(x,-1) = \mathbf{G}(x,0) = \mathbf{0}. \tag{3.40}$$

Note that $g^L(x,y)$ is the Green's function arising from the large-scale mode, top equation of system (3.38), and $g^S(x,y)$ is the Green's function arising from the small-scale mode, bottom equation of system (3.38). Moreover, all information about the boundary conditions is contained in the Green's functions. The Green's function for our finite basin is

$$g(x,y) = \begin{cases} \dfrac{\sin k(x+1)\sin ky}{k\sin k} & x < y, \\ \dfrac{\sin k(y+1)\sin kx}{k\sin k} & y < x; \end{cases} \tag{3.41}$$

see for example [30]. In this case, k is a placeholder for k^L or k^S depending on whether $g(x,y)$ is meant to denote $g^L(x,y)$ or $g^S(x,y)$, respectively.

Remark III.5. We merely need to change the given Green's function to treat the infinite basin case; see Section B.2, where the Green's function is of the form $e^{ik|x|}/(2ik)$. Thus, the separation of the model solutions into large- and small-scales is preserved independently of the basin geometry.

Formula (3.39) allows us to write

$$\mathbf{W}(x) = \int_{-1}^{0} \mathbf{D}(x,y)\mathbf{F}(y)\,dy - \sigma s_1 s_2 \int_{-1}^{0} \mathbf{D}(x,y)\mathbf{e}_{22}\left(\frac{\beta(y)}{1-\sigma\beta(y)}\right)\mathbf{V}\mathbf{W}(y)\,dy, \tag{3.42}$$

where $\mathbf{D}(x,y) = \mathbf{G}(x,y)\mathbf{\Lambda}^{-1}\mathbf{V}^{-1}$ and $\mathbf{F}(x)$ is defined in (3.20). Letting

$$\mathbf{F}_0(x) = \int_{-1}^{0} \mathbf{D}(x,y)\mathbf{F}(y)\,dy \tag{3.43}$$

and defining an operator K by

$$(K\mathbf{W})(x) = -s_1 s_2 \int_{-1}^{0} \mathbf{D}(x,y)\mathbf{e}_{22}\left(\frac{\beta(y)}{1-\sigma\beta(y)}\right)\mathbf{V}\mathbf{W}(y)\,dy \tag{3.44}$$

we can concisely write the equation (3.42) as

$$\mathbf{W}(x) = \mathbf{F}_0(x) + \sigma K \mathbf{W}(x). \tag{3.45}$$

Equation (3.45) may be solved by the formal Neumann series

$$\mathbf{W}(x) = \mathbf{F}_0(x) + \sigma K \mathbf{F}_0(x) + \sigma^2 K^2 \mathbf{F}_0(x) + \cdots \tag{3.46}$$

giving the full solution of $\mathbf{U}(x) = \mathbf{V}\mathbf{W}(x)$ in a finite basin when the topography scale σ is small enough that the series (3.46) converges.

Remark III.6. The solution (3.46) is a perturbation series expansion for small σ about $\sigma = 0$ (no topography).

3.3.3 Separation into Large- and Small-Scales

A separation for the model solutions into large- and small-scales is achieved through multiplication of the system (3.19) by the inverse of $\mathbf{V}$; $\mathbf{W}(x) = \mathbf{V}^{-1}\mathbf{U}(x)$. To make multiplication by $\mathbf{V}$ completely determined, we normalize the eigenvectors as presented in [22] and [59]. Essentially, for a given eigenvector $\mathbf{v} = [v_1, v_2]^\mathsf{T}$ we impose the condition $v_1^2/\gamma + v_2^2/(1-\gamma) = 1$; a depth weighted normalization consistent with the nondimensionalization (3.6) of the velocities. Under these assumptions we may write the matrix $\mathbf{V}$ of eigenvectors as

$$\mathbf{V} = \begin{bmatrix} Y^L & Y^S \\ Y^L\left(\frac{s_1}{(k^L)^2} - 1\right) & Y^S\left(\frac{s_1}{(k^S)^2} - 1\right) \end{bmatrix} \tag{3.47}$$

and

$$\mathbf{V}^{-1} = \frac{1}{s_1}\left(\frac{1}{(k^L)^2} - \frac{1}{(k^S)^2}\right)^{-1} \begin{bmatrix} -\frac{1}{Y^L}\left(\frac{s_1}{(k^S)^2} - 1\right) & \frac{1}{Y^L} \\ \frac{1}{Y^S}\left(\frac{s_1}{(k^L)^2} - 1\right) & -\frac{1}{Y^S} \end{bmatrix}, \tag{3.48}$$

where

$$Y^{L,S} = \left(\frac{1}{\gamma} + \frac{1}{1-\gamma} \left(\frac{s_1}{(k^{L,S})^2} - 1 \right)^2 \right)^{-1/2} \tag{3.49}$$

to satisfy the normalization conditions. Thus, the matrix $\mathbf{V}$ is invertible when α is small since the scales of the problem are well separated, which implies $(k^L)^2 \neq (k^S)^2$.

Remark III.7. We use the notation $Y^{L,S}$ to mean that Y^L can be obtained by only using the first superscript in every term of the equation and Y^S by only using the second superscript.

The separation of modes defined by $\mathbf{V}^{-1}$ is analogous to the common separation into barotropic and baroclinic modes. This fact is particularly clear when $\alpha, \delta_1, \delta_2 \ll 1$, in which case we have

$$\begin{bmatrix} U^L(x) \\ U^S(x) \end{bmatrix} = \mathbf{V}^{-1}\mathbf{U}(x) \approx \begin{bmatrix} U_1(x) + V_2(x) \\ \sqrt{\frac{1-\gamma}{\gamma}} U_1(x) - \sqrt{\frac{\gamma}{1-\gamma}} V_2(x) \end{bmatrix}. \tag{3.50}$$

Equation (3.50) is equivalent to a barotropic and baroclinic splitting when variables are redimensionalized using (3.6); depth average flow for the barotropic component (first entry) and difference in layer velocities for the baroclinic component (second entry).

Using the concrete eigenvector matrix $\mathbf{V}$, equation (3.42) for U^L and U^S can be written as

$$\begin{aligned} Y^{L,S}U^{L,S} =& \frac{1}{s_1} \left(\frac{1}{(k^{L,S})^2} - \frac{1}{(k^{S,L})^2} \right)^{-1} \left[\phi \int_{-1}^{0} g^{L,S} e^{i\phi y} \, dy \right. \\ &+ \sigma (k^{L,S})^2 \left(1 - \frac{s_1}{(k^L)^2} \right) \int_{-1}^{0} g^{L,S} Y^L U^L \left(\frac{\beta}{1-\sigma\beta} \right) \, dy \\ &\left. + \sigma (k^{L,S})^2 \left(1 - \frac{s_1}{(k^S)^2} \right) \int_{-1}^{0} g^{L,S} Y^S U^S \left(\frac{\beta}{1-\sigma\beta} \right) \, dy \right], \end{aligned} \tag{3.51}$$

where we have written g^L as opposed to $g^L(x,y)$, β as opposed to $\beta(y)$, etc., to avoid clutter. For the layer elevations, a separation can be achieved simply by rewriting (3.16) as $\mathbf{N}(x) = i\left(\mathbb{I} + \mathbf{e}_{12}\right)\mathbf{V}\mathbf{W}'(x)$. In this manner we can define the large- and small-scale components of

the surface and interfacial elevations as those arising from U^L and U^S, respectively. It follows that

$$N_1^{L,S} = i\left(\frac{s_1}{(k^{L,S})^2}\right) Y^{L,S} \frac{dU^{L,S}}{dx}, \tag{3.52}$$

$$N_2^{L,S} = i\left(\frac{s_1}{(k^{L,S})^2} - 1\right) Y^{L,S} \frac{dU^{L,S}}{dx}. \tag{3.53}$$

Equation (3.51) clearly shows that in the presence of topography there is mixing between the large- and small-scale modes $U^{L,S}$. The coupling between $U^{L,S}$ only occurs when the topography is non-zero. Importantly, (3.51) corroborates that our splitting into large- and small-scale modes is reasonable since the large-scale Green's function g^L appears in the equation for U^L while the small-scale Green's function g^S appears in the equation for U^S.

Remark III.8. The horizontal scale of $U^{L,S}$ is given by the wavenumber $k^{L,S}$ because the scale of the Green's function $g^{L,S}$ is determined by $k^{L,S}$ in the x variable; see (3.41).

3.4 Effects of Stratification on the Surface Elevation

In this section we examine the effects of stratification, combined with bottom topography and an assumed layer structure of damping, on surface tidal elevations. We utilize a suitably truncated version of the Fourier series expansion (3.33), as well as the decomposition into large- and small-scales (3.48) justified by the Neumann series (3.46), to study several physically motivated trials.

3.4.1 Analytical Description

Some understanding of the key parameters underlying the effect of stratification on the tidal elevations can be obtained analytically. In particular, in the low damping regime ($\delta_1, \delta_2 \ll 1$), the polynomial (3.37) — whose zeros determine $k^{L,S}$ — can be written in the

approximate form

$$k^4 - \left(\frac{s_1 + s_2}{\alpha}\right) k^2 + \frac{s_1 s_2}{\alpha} \approx k^4 - \frac{k^2}{\epsilon^2 \alpha\gamma(1-\gamma)} + \frac{1}{\epsilon^4 \alpha\gamma(1-\gamma)} \tag{3.54}$$

as $\alpha \to 0$. This is consistent with cases of oceanographic interest, where $\alpha, \delta_1, \delta_2 \ll 1$; see Section 3.4.2 for a discussion of realistic parameters. Writing the equation in this manner clearly shows that in the low damping regime we essentially have one stratification parameter dictating the horizontal scales of the problem: $\alpha\gamma(1-\gamma) = g'H_1H_2/(g(H_1+H_2)^2)$. This stratification parameter equals the ratio of the square of the leading order small-scale phase speed $\sqrt{g'H_1H_2/(H_1+H_2)}$ to the square of the leading order large-scale phase speed $\sqrt{g(H_1+H_2)}$. For this reason, in the low damping regime, perturbations in α or γ can lead to the same $k^{L,S}$ scales. In particular, in the oceanographically relevant case $\gamma < 1/2$ (a thin upper layer over a thicker lower layer), a positive perturbation in α (increased density contrast) may yield equivalent scales as a positive perturbation in γ (increased upper layer thickness). It is worth noting that in this low damping regime, under the substitution $c = \Omega L/k$, equation (3.54) is equivalent to the formula for the phase speed of linear waves in a two-layer fluid:

$$c^4 - g(H_1+H_2)c^2 + gg'H_1H_2 = 0, \tag{3.55}$$

as given by Gill [28]. Therefore, the scales $k^{L,S}$ of the problem are determined by the large-scale and small-scale phase speeds of the system; as given in (3.55).

Remark III.9. Note that in the low damping regime (3.55), the large-scale phase speed c^L is given by $(c^L)^2 = g(H_1+H_2) - (c^S)^2$, where c^S is the small-scale phase speed. Under the Boussinesq approximation, an additional $O(g')$ term enters this relation.

The analytical model's dependence on the length scales can be seen directly in (3.51) and (3.52), where the only parameters determining the model solutions are k^L, k^S, and s_1 (or s_2 with some substitutions). The system has its strongest dependence on the scales k^L, k^S as we show below in a limiting case.

For ease of reference, we define the RMS value of a function φ as

$$(\varphi)_{RMS} = \sqrt{\int_{-1}^{0} |\varphi(x)|^2 \, dx}, \tag{3.56}$$

i.e., the $L^2(-1,0)$ norm. Now, the RMS of the large-scale surface tidal elevation, N_1^L, may be approximated by the RMS of the first term (largest scale) in the Fourier series expansion of N_1:

$$(N_1^L)_{RMS} = \sqrt{\int_{-1}^{0} |N_1^L(x)|^2 \, dx} \approx \sqrt{\frac{1}{2}|a_1|^2} = \frac{|a_1|}{\sqrt{2}}. \tag{3.57}$$

This approximation is accurate if the scales of the problem are well separated; as is expected when $\alpha \ll 1$. The coefficient a_1 can be further simplified in the absence of topography ($\sigma = 0$). Without topography the coefficient a_1 can be written exactly as

$$a_1 = \frac{\phi f_1}{i\pi}\left(1 - \frac{(k^L)^2(k^S)^2}{((k^L)^2 - \pi^2)\,((k^S)^2 - \pi^2)}\right), \tag{3.58}$$

where we have used the expression

$$D_n = \alpha\left((k^L)^2 - (n\pi)^2\right)\left((k^S)^2 - (n\pi)^2\right) \tag{3.59}$$

on (3.28), (3.29), (3.34), and f_n given in equation (3.25). Equation (3.58) demonstrates that, in the case with no topography, the large-scale surface tidal elevation is dependent only on the scales k^L, k^S and the forcing parameter ϕ. That is, the effects of damping and stratification are only accounted for through the scales k^L and k^S; equivalently, the phase speeds in the basin.

Remark III.10. Expression (3.59) is obtained by using the conditions $k^L + k^S = (s_1 + s_2)/\alpha$ and $k^L k^S = s_1 s_2/\alpha$, derived from (3.37), on (3.30). Recall that the scales k^L, k^S are in general complex numbers when $\delta_1, \delta_2 \neq 0$.

The dependence of the solution on the parameter $\alpha\gamma(1-\gamma)$ is particularly salient if we examine equations (3.28) and (3.29) of the Fourier coefficients c_n and d_n. We expect the Fourier coefficients of c_n and d_n to be large when the denominator D_n is small. In the low damping regime ($\delta_1, \delta_2 \ll 1$), D_n in (3.59) is zero when $k^S = n\pi$ for an integer $n > 1$. Similarly, using equation (3.30), the condition $D_n = 0$ follows the curves

$$\alpha\gamma(1-\gamma) = \frac{1}{(\epsilon n\pi)^2}\left(1 - \frac{1}{(\epsilon n\pi)^2}\right) \tag{3.60}$$

for any integer n. Thus, the curves given by equation (3.60) are a rough predictor of the resonance peaks seen in upcoming Figures 3.5–3.8, but not for all values of n. This follows from the fact that not all coefficients in the Fourier expansion of N_1 and N_1^L contribute equally to their RMS values.

More rigorously, we note the existence of a resonance that is induced by the topography on the system. The matrix $\mathbf{M}$ of the linear system $\mathbf{Md} = \boldsymbol{f}$ given by (3.31) and (3.25) may become nearly singular given realistic oceanic parameters. By Cramer's rule, the solution $\mathbf{d}$ can become large when $\det \mathbf{M}$ becomes small, and vice versa. This effect leads to the oscillations observed in Figures 3.5–3.8 as the matrix $\mathbf{M}$ changes from nearly singular to not singular as parameters are varied. Note that this oscillation is not present in the absence of topography; in that case, the denominator of (3.58) does not change between nearly singular and not singular as parameters are varied (it only has poles at π^2).

3.4.2 Description of Trials and Figures

For all computations utilizing the analytical model, as in [6], we take $g = 9.81$ m s^{-2}, $K = \frac{2}{6{,}371\times 10^3}$ m^{-1} (2π over the zonal wavelength of the semidiurnal tidal potential at the equator), $\Omega = 1.405189 \times 10^{-4}$ s^{-1} (the M_2 frequency), $L = 4{,}086 \times 10^3$ m (a typical ocean basin scale), and $H_1 + H_2 = 4000$ m (an average ocean basin depth). This implies $\epsilon = 0.345$ and $\phi = 1.2827$; see (3.11). In analogy to the global realistic-domain numerical simulations of

Section 3.1, we denote as the control case the trial with a reduced gravity of $g' = 1.64 \times 10^{-2}$ m s^{-2}, layer resting depths $H_1 = 700$ m, $H_2 = 3300$ m, and canonical values of open ocean damping $\delta_1 = 0$, $\delta_2 = 0.0412$, where the δ_2 value is taken from [6]. Thus, for the control case, the stratification parameters are $\alpha = 0.0017$ and $\gamma = 0.175$. Unless specified otherwise, all parameters in the numerical trials are as in the control case.

We choose a nominal topography with a Gaussian form

$$\beta(x) = \frac{H_0}{H_2} e^{-(x+0.5)^2/q^2} \tag{3.61}$$

so that it is centered in the basin. The parameter q is chosen to be $q = \frac{\pi}{2q_1} = 8.49 \times 10^{-3}$ using values of the control case, where $q_1 = \Omega L \sqrt{(H_1 + H_2)/(g' H_1 H_2)}$ is the topographic wavenumber that excites the first baroclinic mode in the absence of damping. In addition, we set $H_0 = 2,350$ m and allow σ to vary in the numerical trials. As an aside, even though in the different trials the values of g', H_1, and H_2 change, we keep the value of q — obtained using g', H_1, and H_2 from the control case above — constant throughout. For simplicity, in order to focus on the changes in other parameters, we do not allow the shape of the topography to vary in our results section.

Remark III.11. The chosen horizontal length scale q of the topography ensures that the characteristic length of the topography is significantly smaller than the basin scale. The value of H_0 was chosen so that the ratio of the spatially and temporally averaged square of the upper-layer to lower-layer velocity in the control case (ratio ≈ 2.5) is approximately equal to the value obtained for the $1/8^\circ$ global realistic-domain two-layer control simulation of Section 3.1.

We highlight four significant parameter cases coinciding to the trials shown in Figure 3.2 of the global realistic-domain numerical simulations: one-layer solution $(\alpha, \delta_1, \delta_2) = (0, 0.0412, 0.0412)$, two-layer control solution with $(\alpha, \gamma, \delta_1, \delta_2) = (0.0017, 0.175, 0, 0.0412)$, two-layer g' perturbation solution with $(\alpha, \gamma, \delta_1, \delta_2) = (0.0018, 0.175, 0, 0.0412)$, and two-

layer interface perturbation solution with $(\alpha, \gamma, \delta_1, \delta_2) = (0.0017, 0.2, 0, 0.0412)$. Figure 3.4, which displays results from the analytical model for these four cases, is directly comparable with Figure 3.2, which displays the global realistic-domain numerical simulation trials for the same four cases. Note that in Figure 3.4, in contrast to Figures 3.5–3.9, the RMS in time is displayed at each grid point of the domain in order to make Figure 3.4 directly comparable to Figure 3.2.

The perturbation values of α and γ are obtained as follows. We obtain the g' perturbation value of $\alpha = 0.0018$ by assigning $g' = 1.78 \times 10^{-2}$ m s^{-2} and the interface perturbation value of $\gamma = 0.2$ by assigning $H_1 = 800$ m. These values of α and γ are intended to represent climatic perturbations to the values of g' and H_1. The interface perturbation is motivated by Figure 10 of [7], which shows ∼100 m displacements of isopycnals over decadal timescales in hydrographic observations of the North Atlantic. The g' perturbation is estimated from the 0.5°C century^{-1} nominal maximum warming trend found in intermediate depth waters in the same paper. We compute the change in g' (with potential densities referenced to 1780 db) that would take place if a water parcel at 100 db having salinity 37 psu and temperature of 20°C warmed by 0.5°C. The deep reference parcel has depth of 3000 db, salinity 34.5 psu and temperature of 4°C. We note that because the global realistic-domain numerical simulations were performed well before we had developed an analytical model, the choices of perturbed g' and interface depth values in the global realistic-domain numerical simulations were motivated by the Arbic and Owens [7] results and were not guided by the analytical model — a point we will return to shortly.

The relative simplicity of the analytical model allows for a more complete exploration of parameter space than is possible with the global realistic-domain numerical simulations. We survey parameter space by computing the RMS value, calculated over space and time (only in time for Figure 3.4), of relevant quantities. The RMS values of the large- and small-scale velocity and surface tidal elevation are presented in Figure 3.5 as a function of α for different damping structures δ_1, δ_2 (damping in the bottom layer only, damping in the top layer only,

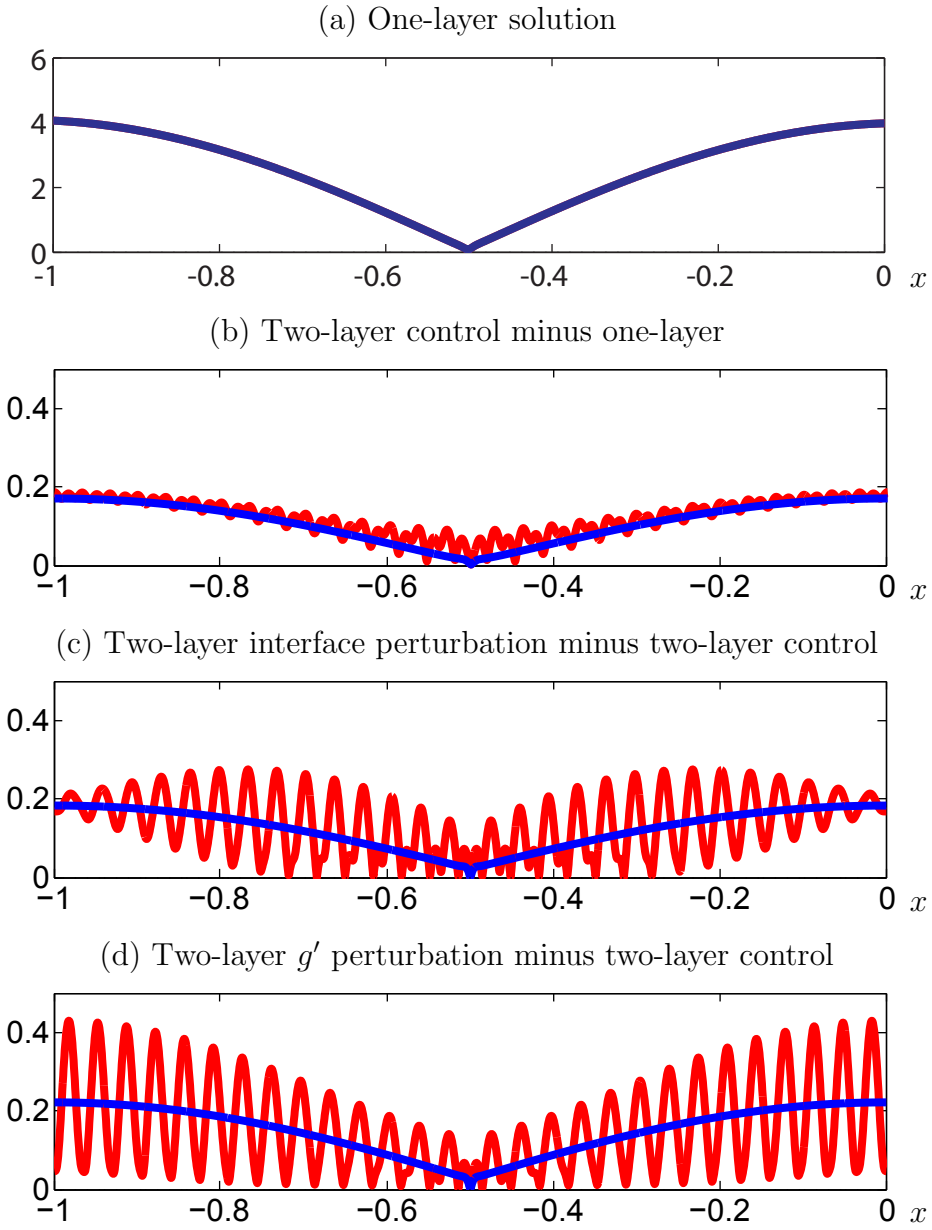

Figure 3.4: As in Figure 3.2 but for analytical model: RMS in time of surface elevation differences (red curve) and large-scale surface elevations (blue curve) magnitudes for analytical model. The range in the vertical scales of (b), (c), and (d) is identical. Topography consists of the centered Gaussian (3.61). For reference the one-layer solution is shown in (a).

and equal damping in the two layers). The variations of the RMS of the large-scale surface tidal elevation N_1^L with α and γ for different values of the topography scale σ are shown in Figure 3.6, while Figure 3.7 displays variations of the RMS of the large- and small-scale

elevations N_1^L and N_1^S as a function of α and δ_1 with δ_2 fixed at the control value of 0.0412. Similarly, Figure 3.8 displays variations of the RMS of N_1^L and N_1^S as a function of α and δ_2 with the control value $\delta_1 = 0$. Note that the curves of Figure 3.5(c) represent a transect of Figure 3.7(a) about the $\delta_1 = 0$ (black curve) and $\delta_1 = \delta_2 = 0.0412$ (red curve) lines. Similarly, Figure 3.5(d) is a transect of Figure 3.7(b) about the $\delta_1 = 0$ (black curve) and $\delta_1 = \delta_2 = 0.0412$ (red curve) lines.

Remark III.12. Recall that all results concerning the analytical model have been nondimensionalized as given by (3.6). Moreover, for a function of the form (3.13), say $u_1(x,t)$, taking the RMS in just time gives

$$(u_1(x,\cdot))_{RMS} = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |\mathrm{Re}\left\{U_1(x)e^{it}\right\}|^2 \, dt} = \frac{|U_1(x)|}{\sqrt{2}} \tag{3.62}$$

while taking the RMS in time and space gives $(u_1)_{RMS} = (U_1)_{RMS}/\sqrt{2}$. For the analytical results, we omit the extra $\sqrt{2}$ factor to focus on the x dependence of the analytic solution because the time dependence is of the nearly trivial form e^{it}.

Lastly, we briefly bring up a point which will not be pursued further. For the purposes of this chapter, all computational results of the analytical model were carried out using the Fourier series expansion method of Section 3.3.1. For the topography used, the Fourier series method was more numerically reliable than a computational implementation of the Neumann series method of Section 3.3.2. The numerical accuracy of the model solutions was confirmed by checking that they satisfy a discretized version of the original set of differential equations.

3.4.3 Discussion

Consistent with the 1–5% changes obtained using global realistic-domain numerical model results shown in Figures 3.1 and 3.2 and the observed secular changes of 1% [57], the analyti-

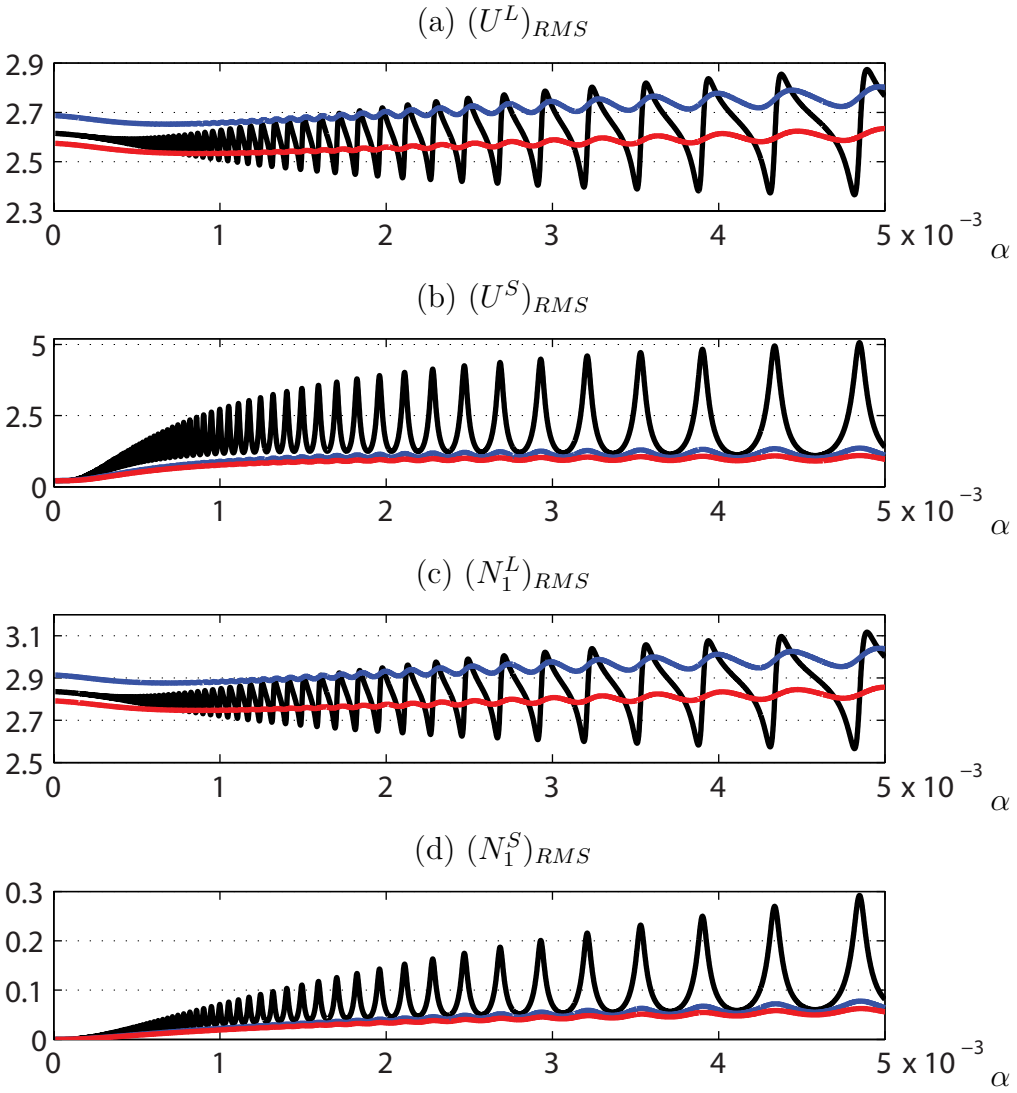

Figure 3.5: RMS in space of large- and small-scale components of velocities and surface elevations, versus stratification parameter α for two layers damped ($\delta_1 = \delta_2 = 0.0412$), in red, only the top layer damped ($\delta_1 = 0.0412, \delta_2 = 0$), in blue, and only the bottom layer damped ($\delta_1 = 0, \delta_2 = 0.0412$), in black. Topography consists of the centered Gaussian (3.61).

cal model shows that both the addition of stratification and of perturbations to stratification yield changes in both the large- and small-scale surface elevation amplitudes of up to 10%; see Figure 3.4 for a case by case comparison in the vein of Figure 3.2 and Figure 3.5 for a continuous exploration in the parameter α. We see that the large-scale RMS quantities may be perturbed as much as 10%; see black curves in Figures 3.5 (a) and (c). That is, the large-scale RMS values may change by as much as 10% from peak to trough in the parameter

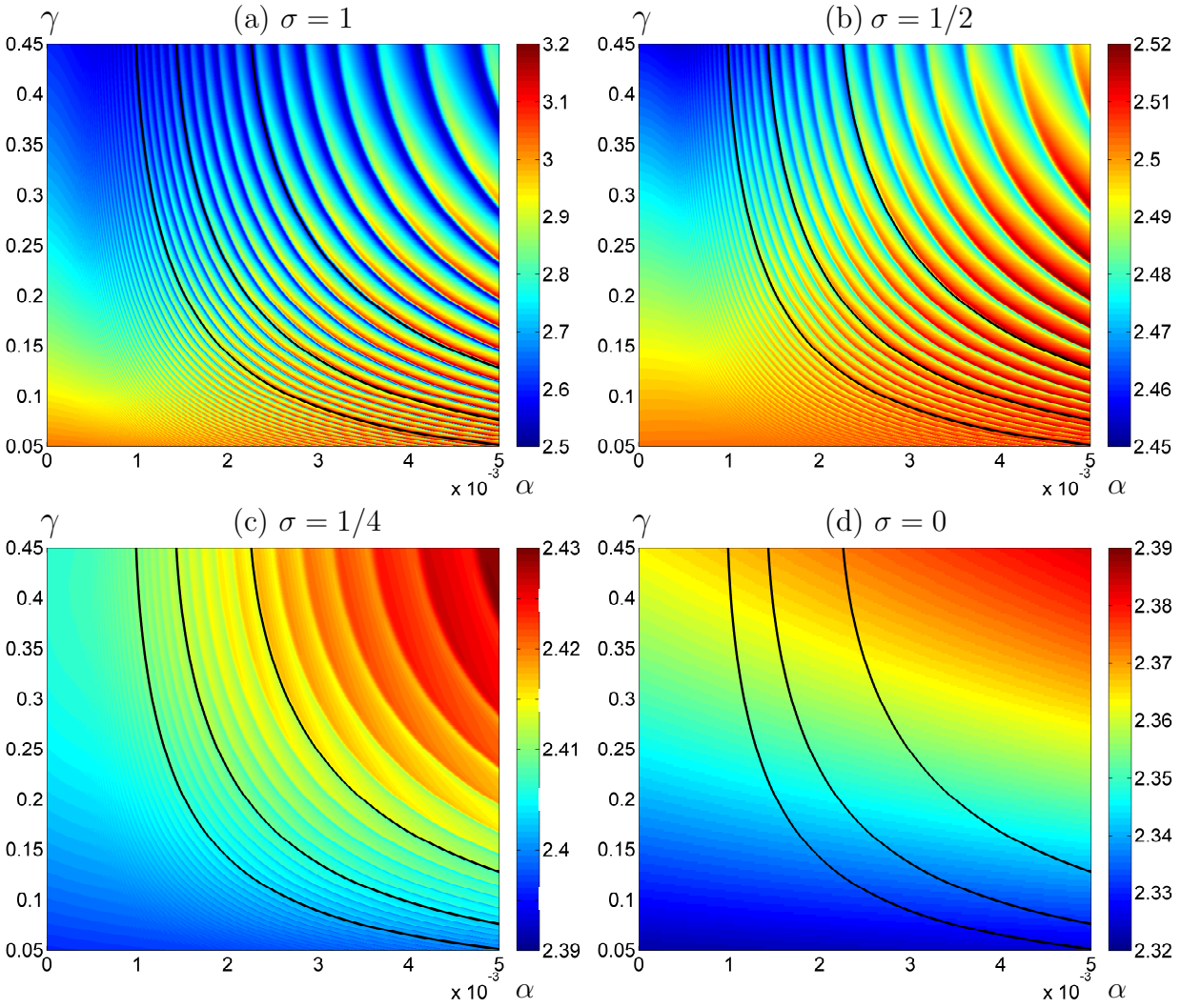

Figure 3.6: RMS in space of the large-scale surface tidal elevation N_1^L as a function of α and γ when $\sigma = 1,\ 1/2,\ 1/4,\ 0$. The black curves are the curves given by (3.60) for $\epsilon = 0.345$ and $n = 39,\ 49,\ 59$, i.e., $\alpha\gamma(1-\gamma) = 2.4 \times 10^{-4},\ 3.5 \times 10^{-4},\ 5.6 \times 10^{-4}$. Topography consists of the centered Gaussian (3.61).

range explored in this dissertation. Even when $\alpha = 0$, the addition of a second layer changes the large-scale surface tidal elevation amplitude by 2% compared to the one-layer case, due to the structure of damping as discussed in the introduction; see the differences between red, blue, and black curves at $\alpha = 0$ in Figure 3.5 (c). A climatic perturbation in α to the two-layer solution may induce a further change in the amplitude of both large- and small-scales by about 3%; see the oscillations in Figure 3.5 (c) and (d). Climatic perturbations

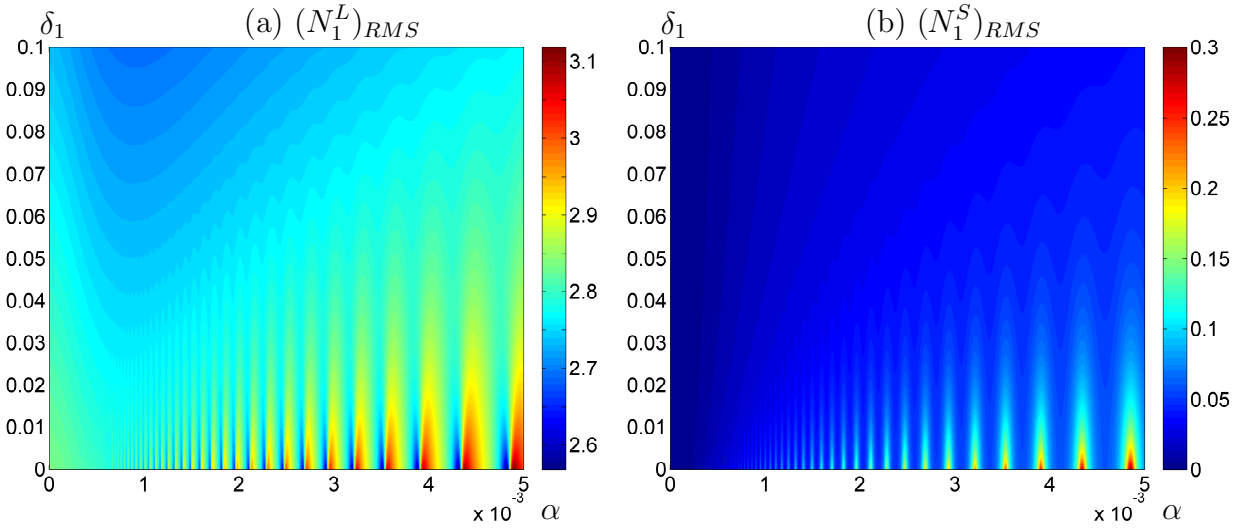

Figure 3.7: RMS in space of the large- and small-scale surface tidal elevation N_1^L and N_1^S, respectively, as a function of α and δ_1 when $\sigma = 1$. All other model parameters are as in the control case described in Section 3.4.2 and topography consists of the centered Gaussian (3.61).

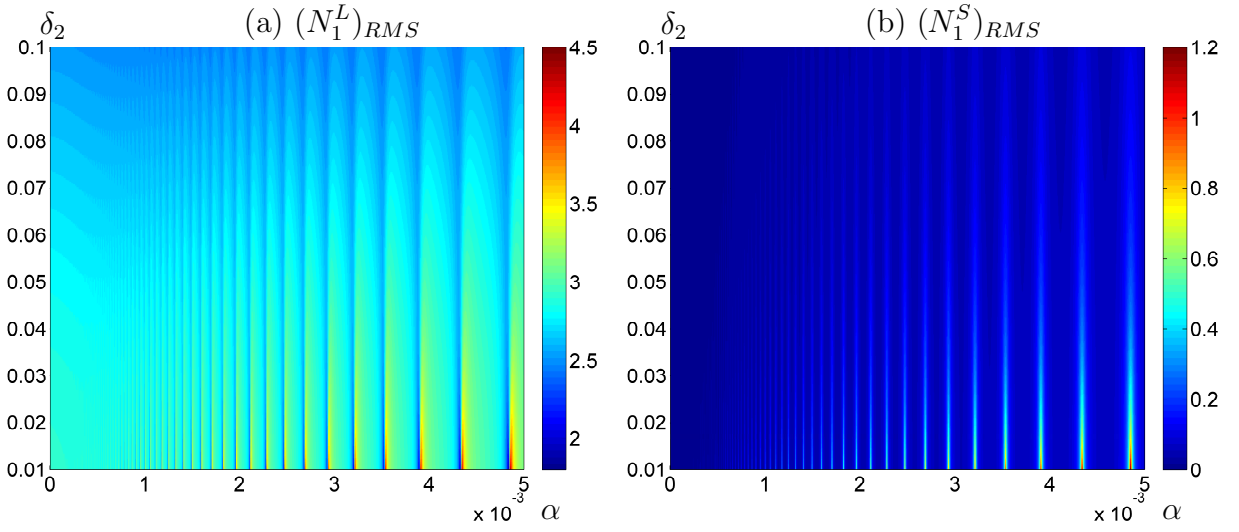

Figure 3.8: RMS in space of the large- and small-scale surface tidal elevation N_1^L and N_1^S, respectively, as a function of α and δ_2 when $\sigma = 1$. All other model parameters are as in the control case described in Section 3.4.2 and topography consists of the centered Gaussian (3.61).

in γ induce qualitatively similar responses in the system; see Figure 3.6 and note that the sensitivity of the system is strongest to changes on the $\alpha\gamma(1-\gamma)$ parameter as indicated

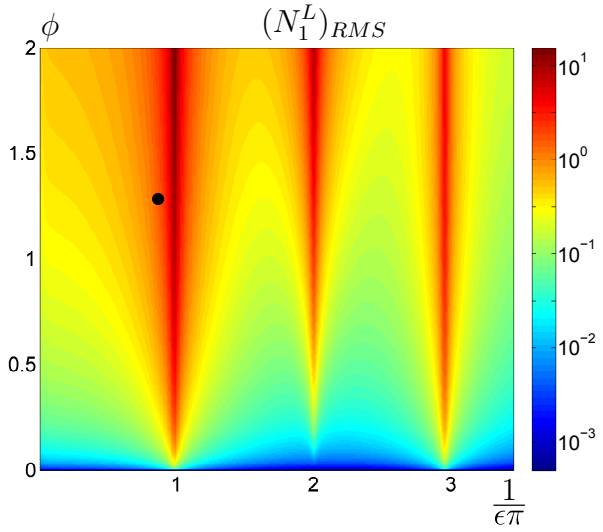

Figure 3.9: RMS in space of the large-scale surface elevation amplitude N_1^L as a function of $1/(\epsilon\pi)$ and ϕ. All other parameters are as in the control solution and topography consists of the centered Gaussian (3.61). Black dot indicates the location in parameter space for the two-layer control solution.

by the black lines. Perturbations in $\alpha\gamma(1-\gamma)$ may lead to an increase or a decrease in the large-scale mode's amplitude comparable in magnitude to the amplitude perturbation of the small-scale mode; compare the size of the oscillations in Figure 3.5 (c), (d) or those in Figure 3.7. The sensitivity of the system to bottom damping is significant when damping is absent in the top layer ($\delta_1 = 0$); see Figure 3.8.

The results presented here show strong sensitivity, in the form of oscillations of significant amplitude, of both the large- and small-scale tidal elevations to the stratification. It is worth noting that the observed oscillations are not about a constant mean; instead, the mean values in these plots vary with the stratification in an observable manner; see Figure 3.5. This implies that these are not simply oscillations about the limiting one-layer case. It is also observed that the sensitivity of the system to stratification is more significant when only the bottom layer is damped than when both the top and bottom layers are identically damped, or when only the top layer is damped; compare black, blue, red curves in Figure 3.5 or the sensitivity in Figures 3.7 and 3.8 to δ_1, δ_2. The greater sensitivity (oscillations) in the black

versus red or blue curves in Figure 3.5 implies that the tidal sensitivity to stratification is itself dependent on assumptions about how the tidal signal is damped in the layers. The amplitude of the topography also plays an important role in the sensitivity of tidal elevations to the stratification; see Figure 3.6 and note the response as σ varies. Note that $\sigma = 1$ (the most realistic case, as discussed in Section 3.4.2) yields greater sensitivity than the $\sigma = 1/2, 1/4$ cases and significantly greater sensitivity than the $\sigma = 0$ (flat bottom) case.

The sensitivity of the system to the quantity $1/(\epsilon\pi)$, which acts as a nondimensional measure of the forcing frequency of the system, and to ϕ, the nondimensional forcing wavenumber, is displayed in Figure 3.9. For our chosen parameters, the control solution (black dot) lies near a resonance peak of the system. That is, the plot shows that effects of stratification on the large-scale surface tidal elevation might be accentuated due to our chosen location in parameter space.

Finally, with our analytical model in place, we return to the global realistic-domain numerical model results shown in Figures 3.1 and 3.2. In our analytical model, when topography is present and damping is in the bottom layer only, as in the global realistic-domain numerical model, the quantity $\alpha\gamma(1-\gamma)$ is a strong predictor of model behavior. We note that for the control simulation of the global realistic-domain numerical model, $\alpha = 0.0017$, $\gamma = 0.175$, and $\alpha\gamma(1-\gamma) = 2.41 \times 10^{-4}$. For the interface and g' perturbation simulations of the global realistic-domain numerical model $\alpha\gamma(1-\gamma) = 2.67 \times 10^{-4}$ and 2.62×10^{-4}, respectively. The similarity of Figures 3.2(b) and 3.2(c) may be explained by the similarity of their $\alpha\gamma(1-\gamma)$ values, a similarity that is somewhat coincidental because the stratification values in the global realistic-domain numerical simulations were chosen well before a guiding analytical model was developed. Motivated by our analytical model results, we plan to write a follow-up paper in which a global realistic-domain numerical model, run in the non-Boussinesq limit is used to perform a suite of simulations with different $\alpha\gamma(1-\gamma)$ values.

CHAPTER IV

Energy Cascade in Forced-Dissipated Quasi-Geostrophy

4.1 Background

In this section we introduce some of the general theory concerning the Quasi-Geostrophic (QG) equations.

4.1.1 Governing Equations of Quasi-Geostrophy

The continuously stratified QG equation is given by the conservation of the *quasi-geostrophic potential vorticity* $q = q(x, y, z, t)$:

$$\frac{Dq}{Dt} = 0, \quad t \geq 0 \tag{4.1}$$

in a domain of interest, where

$$q = \nabla_2^2 \psi + \mathcal{A}_0 \psi + f \quad \text{with} \quad \mathcal{A}_0 := \frac{\partial}{\partial z} \frac{f_0^2}{N^2(z)} \frac{\partial}{\partial z}. \tag{4.2}$$

Here we use the standard convention for the variables (x, y, z, t): x represents the east-west direction, y the north-south direction, z the vertical direction, and t is time. The function $\psi = \psi(x, y, z, t)$ is the so-called *geostrophic streamfunction*, the *Brunt-Väisälä frequency*

$N(z)$ is the buoyancy frequency quantifying the vertical stability of a fluid parcel in a known reference state, and f is the so-called *Coriolis parameter* which quantifies the effects of the earth's rotation in the z direction; see (4.5). The operator D/Dt is the horizontal material derivative given by

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla}_2, \tag{4.3}$$

where $\boldsymbol{\nabla}_2 = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ is the horizontal gradient operator and $\mathbf{u} = \boldsymbol{\nabla}_2^{\perp}\psi = \left(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x}\right)$ is the horizontal fluid velocity at a given location.

The evolution equation (4.1) may be written in a concise manner using the fact that the horizontal fluid velocity $\mathbf{u}$ can be written in terms of a streamfunction ψ. Namely, we may simply write the material derivative operator as

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + J(\psi, \cdot), \tag{4.4}$$

where $J(f, g) = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial g}{\partial x}\frac{\partial f}{\partial y}$ is the Jacobian operator; see Appendix C.1 for a discussion on properties of the Jacobian. Thus, the QG equation (4.1) with the definitions (4.2) are given only in terms of the streamfunction ψ.

We may fully expect that the Coriolis "parameter" f will be a function of the latitude y since the vertical component of rotation increases (decreases) as we move northward (southward) from the equator in the northern hemisphere. For scales of the size of the Rossby deformation radius L_d (synoptic scale), it is common to approximate f linearly by using the *β-plane* approximation. Namely,

$$f(y) = f_0 + \beta y, \tag{4.5}$$

where β is the linear correction to the constant value f_0 at the latitude of interest. In this manner, we assume that the QG dynamics take place on a tilted plane (*beta*-plane) rather than the more complicated geometry of a sphere or more realistic ellipsoid.

For the purposes of this dissertation, boundary conditions for the QG equation (4.1) will consist of requiring double periodicity in the horizontal variables (x, y) and rigid surfaces at

$z = -H$ and $z = 0$ in a vertical domain of depth H. The vertical boundary conditions are equivalent to requiring

$$\frac{Db}{Dt} = 0 \quad \text{at} \quad z = -H,\ 0, \tag{4.6}$$

where the *buoyancy* $b = b(x, y, z, t)$ is given in terms of the streamfunction by

$$b = f_0 \frac{\partial \psi}{\partial z}. \tag{4.7}$$

The vertical boundary conditions (4.6) are indeed rigid lid approximations as they require the vertical velocity of the flow to be zero at the top and the bottom boundary, though this might not be immediate from (4.6); see, for instance, page 222 in Chapter 5 of [81].

Since relation (4.2) is linear in ψ, it is convenient to decompose the streamfunction into contributions arising from the boundaries and the interior [45, 80]. That is, we let

$$\psi = \psi^I + \psi^T + \psi^B, \tag{4.8}$$

where the superscript I stands for the "interior" solution, T for "top" boundary solution, and B for "bottom" boundary solution. We require the differing components of ψ to satisfy

$$\left(\nabla_2^2 + \mathcal{A}_0\right)\psi^I + f = q, \quad \left(\nabla_2^2 + \mathcal{A}_0\right)\psi^T = 0, \quad \left(\nabla_2^2 + \mathcal{A}_0\right)\psi^B = 0, \quad \text{for } -H < z < 0, \tag{4.9}$$

$$\frac{\partial \psi^I}{\partial z} = 0, \quad f_0 \frac{\partial \psi^T}{\partial z} = b, \quad \frac{\partial \psi^B}{\partial z} = 0, \quad \text{at} \quad z = 0, \tag{4.10}$$

$$\frac{\partial \psi^I}{\partial z} = 0, \quad \frac{\partial \psi^T}{\partial z} = 0, \quad f_0 \frac{\partial \psi^B}{\partial z} = b, \quad \text{at} \quad z = -H. \tag{4.11}$$

In this form, each component of ψ describes relevant limiting cases in geophysics. One of those limiting cases, surface Quasi-Geostrophy (SQG), consists of only using the evolution equation (4.6) for $z = 0$ coupled with equations (4.9)–(4.11) with the pared down streamfunction $\psi = \psi^T$ ($\psi^I = \psi^B = 0$) along with the condition $\psi \to 0$ as $z \to -\infty$ (infinite bottom depth). By neglecting the effects of the interior dynamics, SQG describes

flow driven by buoyancy advection at the surface boundary and provides a simplification to nearly two-dimensional dynamics [32, 45]. A complementary regime is that of "interior" QG. In this case, we use the evolution equation (4.1) coupled with equations (4.9)–(4.11) and the streamfunction $\psi = \psi^I$ with $\psi^T = \psi^B = 0$. Note that in this case, equation (4.6) for the evolution of the buoyancy is redundant since $b = 0$ at the boundaries. Many previous studies have indeed considered the interior QG regime [3, 4, 38, 76, to name a few], however, allowing the buoyancy b to be nontrivial at the boundaries (i.e., governed by (4.6)) provides a more realistic spectrum for the kinetic and potential energy consistent with observations in both the ocean and atmosphere [44, 80].

4.1.2 Forced-Dissipated Quasi-Geostrophy

For the purposes of this dissertation, we assume a forcing arising from a known stationary mean flow Ψ, Q, B satisfying the QG equations and consider the evolution of perturbation variables ψ', q', b'. Namely, we write the streamfunction, potential vorticity, and buoyancy as $\psi = \Psi(x,y,z) + \psi'(x,y,z,t)$, $q = Q(x,y,z) + q'(x,y,z,t)$, and $b = B(x,y,z,t) + b'(x,y,z,t)$, respectively. Since the mean flow is known, all subsequent discussion will center around the perturbation variables. This can be seen to follow in the same vein as the work of Arbic and Flierl [3] and Arbic, Flierl, and Scott [4]. Substituting ψ, q, b in this form into the QG equations given in the introduction and dropping primes gives the system

$$\frac{\partial q}{\partial t} + J(\psi, q) + J(\psi, Q) + J(\Psi, q) = 0 \quad \text{for } -H < z < 0, \tag{4.12}$$

$$\frac{\partial b}{\partial t} + J(\psi, b) + J(\psi, B) + J(\Psi, b) = 0 \quad \text{for } z = 0, \tag{4.13}$$

$$\frac{\partial b}{\partial t} + J(\psi, b) + J(\psi, B) + J(\Psi, b) = 0 \quad \text{for } z = -H, \tag{4.14}$$

where we have separated terms using the bilinearity of the Jacobian. For consistency we demand that the mean flow quantities satisfy

$$Q = f + \nabla_2^2 \Psi + \mathcal{A}_0 \Psi \quad \text{and} \quad B = f_0 \frac{\partial \Psi}{\partial z} \tag{4.15}$$

so that the perturbation variables satisfy

$$q = \nabla_2^2 \psi + \mathcal{A}_0 \psi \quad \text{and} \quad b = f_0 \frac{\partial \psi}{\partial z}. \tag{4.16}$$

As dissipative mechanisms, we may place thermal damping ν^T on the surface boundary and potential vorticity damping ν^I in the interior. These are easily included in the form

$$\frac{\partial q}{\partial t} + J(\psi, q) + J(\psi, Q) + J(\Psi, q) = -\nu^I q \quad \text{for } -H < z < 0, \tag{4.17}$$

$$\frac{\partial b}{\partial t} + J(\psi, b) + J(\psi, B) + J(\Psi, b) = -\nu^T b \quad \text{for } z = 0. \tag{4.18}$$

In addition, we may include the effects of bottom friction on the QG system simply by modifying the bottom boundary condition [80, 81]:

$$\frac{\partial b}{\partial t} + J(\psi, b) + J(\psi, B) + J(\Psi, b) = -r_{\text{Ek}} N^2 \nabla_2^2 \psi \quad \text{for } z = -H, \tag{4.19}$$

where r_{Ek} is proportional to the thickness of the Ekman layer and thus quantifies friction. In this manner, we consider the system (4.15)–(4.19) as being forced by an imposed mean streamfunction Ψ with a known stratification profile N and dissipated by Ekman friction at the bottom surface, thermal damping at the upper surface, and potential vorticity damping in the interior.

A significant (and customary) simplification of the system (4.15)–(4.19) arises if we limit the mean flow to only the zonal (x) direction. That is, suppose $(U(z), 0) = (-\frac{\partial \Psi}{\partial y}, \frac{\partial \Psi}{\partial x})$. Without loss of generality, the zonal assumption implies the form $\Psi = \Psi(y, z)$ on the mean

geostrophic streamfunction which in turn implies by (4.15) that $Q = Q(y,z)$ and $B = B(y,z)$. In addition, it follows immediately that $J(\Psi, Q) = J(\Psi, B) = 0$ from the definition of the Jacobian. Moreover, the relations (4.15) indeed imply that

$$\frac{\partial Q}{\partial y} = \beta - \mathcal{A}_0 U \quad \text{and} \quad \frac{\partial B}{\partial y} = -f_0 \frac{\partial U}{\partial z}, \tag{4.20}$$

where $\beta = \frac{\partial f}{\partial y}$, which may be used to simplify the form of (4.17)–(4.19) significantly. For example, (4.17) becomes

$$\frac{\partial q}{\partial t} + J(\psi, q) + (\beta - \mathcal{A}_0 U)\frac{\partial \psi}{\partial x} + U\frac{\partial q}{\partial x} = -\nu^I q \tag{4.21}$$

under the zonal assumption.

4.1.3 Nondimensionalization

A nondimensionalization of the forced-dissipated QG system (4.17)–(4.19) may be obtained by changing variables as: $x = L_d x'$, $y = L_d y'$, $z = Hz'$, $t = Tt'$, $q = T^{-1}q'$, $\psi = L_d^2 T^{-1}\psi'$, $b = L_d T^{-2} b'$, where primes denote nondimensional quantities. Additionally, we nondimensionalize the buoyancy frequency using $N = \overline{N}N'$, where $\overline{N}$ is the vertical mean of the buoyancy frequency [77]. This allows our nondimensional time scale to be $T = 1/\overline{N}$ for convenience. Equation (4.17) then becomes

$$\frac{\partial q'}{\partial t'} + J'(\psi', q') + J'(\psi', Q') + J'(\Psi', q') = -\frac{\nu^I}{\overline{N}} q' \tag{4.22}$$

on the domain $0 \le x', y' \le 2\pi n$ for some $n \in \mathbb{N}$ and $-1 < z' < 0$. Similarly, the boundary conditions (4.18)–(4.19) become

$$\frac{\partial b'}{\partial t'} + J'(\psi', b') + J'(\Psi', b') + J'(\psi', B') = -\frac{\nu^T}{\overline{N}} b' \quad \text{at } z' = 0, \tag{4.23}$$

$$\frac{\partial b'}{\partial t'} + J'(\psi', b') + J'(\Psi', b') + J'(\psi', B') = -\frac{r_{\text{Ek}}}{L_d}(N')^2(\nabla_2')^2\psi' \quad \text{at } z' = -1. \tag{4.24}$$

Lastly, equation (4.16) becomes

$$q' = (\nabla_2')^2\psi' + \frac{L_d^2 f_0^2}{H^2\overline{N}^2}\mathcal{A}_0'\psi' \quad \text{and} \quad b' = \frac{f_0}{\overline{N}}\frac{\partial\psi'}{\partial z'}, \tag{4.25}$$

where $\mathcal{A}_0' = \frac{\partial}{\partial z'}\frac{1}{(N')^2}\frac{\partial}{\partial z'}$. For consistency we allow the approximate deformation scale in a continuously stratified system to be given by $L_d = \overline{N}H/f_0$. Therefore, a nondimensionalized QG system (4.15)–(4.19) merely requires us to consider the original system on the domain $0 \leq x', y' \leq 2\pi n$ for some $n \in \mathbb{N}$ and $-1 < z' < 0$ with the substitutions $(v^I)' = v^I/\overline{N}$, $(v^T)' = v^T/\overline{N}$, $(r_{\text{Ek}})' = r_{\text{Ek}}/L_d$, and $(f_0)' = f_0/\overline{N}$, where primes denote nondimensional parameters. Here onward we will consider only the nondimensionalized system.

4.1.4 Potential Vorticity Inversion for Quasi-Geostrophy

To solve the QG system (4.15)–(4.19), it is useful to invert relation (4.16) for q in terms of ψ. This is due primarily to the fact that it is q and not ψ that is being evolved by equation (4.17). Many methods exist to accomplish the inversion of the left hand side of (4.16). In the case of a vertically discretized (multi-level system; see Section 4.1.5) we may write (4.16) as a matrix system. Then, inversion of the operator is done by simple linear algebra. In the fully continuous case this inversion may be accomplished through Sturm-Liouville theory as we proceed to show; see, for instance, [30] for a general discussion of Sturm-Liouville theory, or [44] and [80] for the theory applied in special cases to the QG equations at hand.

From the fact that we are using a periodic domain in the horizontal variables (x, y), we may freely transform (4.16) into its spectral domain using Fourier series. Here we consider the Fourier series defined as

$$\varphi(\mathbf{x}) = \sum_{\mathbf{k}} \varphi^\wedge(\mathbf{k}) e^{i\mathbf{x}\cdot\mathbf{k}} \quad \text{with} \quad \varphi^\wedge(\mathbf{k}) = \frac{1}{(2\pi n)^2}\int_0^{2\pi n}\int_0^{2\pi n} \varphi(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}\, d^2\mathbf{x}, \tag{4.26}$$

where the superscript $\wedge$ represents the Fourier coefficient associated with the wavenumber

$\mathbf{k} = (k_x, k_y)$, $\mathbf{k} = \mathbf{m}/n$ wtih $\mathbf{m} \in \mathbb{Z} \times \mathbb{Z}$, and $\mathbf{x} = (x, y)$. Equation (4.16) then becomes

$$q^\wedge(z;\mathbf{k},t) = \left(\mathcal{A}_0 - |\mathbf{k}|^2\right)\psi^\wedge(z;\mathbf{k},t) \quad \text{and} \quad b^\wedge(z;\mathbf{k},t) = f_0\frac{\partial\psi^\wedge}{\partial z}(z;\mathbf{k},t). \tag{4.27}$$

The inversion of operator $\mathcal{A}_0 - |\mathbf{k}|^2$ may accomplished using the Green's function of the problem or equivalently a modal representation. Here we will show this process using the modal representation arising from the decomposition (4.8).

We begin by separating the perturbation streamfunction into its constituents, as we did in the introduction for the streamfunction. Namely,

$$\psi^\wedge(z;\mathbf{k},t) = \psi^{I\wedge}(z;\mathbf{k},t) + \psi^{T\wedge}(z;\mathbf{k},t) + \psi^{B\wedge}(z;\mathbf{k},t), \tag{4.28}$$

where I, T, B stand for interior, top, and bottom, respectively. From equations (4.9)–(4.11), we note that $\psi^{T\wedge}$ and $\psi^{B\wedge}$ may be normalized as $\varphi^{T\wedge}$ and $\varphi^{B\wedge}$ given by

$$\psi^{T\wedge}(z;\mathbf{k},t) = \frac{b^\wedge(0;\mathbf{k},t)}{f_0}\varphi^{T\wedge}(z;\mathbf{k}) \quad \text{and} \quad \psi^{B\wedge}(z;\mathbf{k},t) = \frac{b^\wedge(-H;\mathbf{k},t)}{f_0}\varphi^{B\wedge}(z;\mathbf{k}). \tag{4.29}$$

We can accordingly rewrite equations (4.9)–(4.11) for T and B as

$$\left(\mathcal{A}_0 - |\mathbf{k}|^2\right)\varphi^{T\wedge} = 0 \quad \text{and} \quad \frac{\partial\varphi^{T\wedge}}{\partial z} = 1 \quad \text{at } z = 0, \quad \frac{\partial\varphi^{T\wedge}}{\partial z} = 0 \quad \text{at } z = -1, \tag{4.30}$$

$$\left(\mathcal{A}_0 - |\mathbf{k}|^2\right)\varphi^{B\wedge} = 0 \quad \text{and} \quad \frac{\partial\varphi^{B\wedge}}{\partial z} = 0 \quad \text{at } z = 0, \quad \frac{\partial\varphi^{B\wedge}}{\partial z} = 1 \quad \text{at } z = -1. \tag{4.31}$$

Similarly, the internal mode I satisfies

$$\left(\mathcal{A}_0 - |\mathbf{k}|^2\right)\psi^{I\wedge} = q^\wedge \quad \text{and} \quad \frac{\partial\psi^{I\wedge}}{\partial z} = 0 \text{ at } z = -1, 0 \tag{4.32}$$

and may be inverted using standard Sturm-Liouville theory.

Consider the eigenvalue problem

$$\left(\mathcal{A}_0 - |\mathbf{k}|^2\right)\varphi^\wedge = \lambda\varphi^\wedge \quad \text{with} \quad \frac{\partial\varphi^\wedge}{\partial z} = 0 \quad \text{at } z = -1, 0, \tag{4.33}$$

where $\lambda \in \mathbb{C}$ and $\varphi^\wedge \in L_2([-1,0])$. By Sturm-Liouville theory, the system has a complete set of orthogonal eigenfunctions $\varphi_n^{I\wedge}(z)$ in $L_2([-1,0])$. So, we may write the expansions

$$\psi^{I\wedge} = \sum_n a_n\varphi_n^{I\wedge} \quad \text{and} \quad q^\wedge = \sum_n b_n\varphi_n^{I\wedge}, \tag{4.34}$$

where the eigenfunctions satisfy the orthogonality conditions

$$\left(\varphi_n^{I\wedge}, \varphi_m^{I\wedge}\right) = \int_{-1}^{0} \varphi_n^{I\wedge}(z)\varphi_m^{I\wedge}(z)\, dz = 0 \quad \text{for} \quad n \neq m. \tag{4.35}$$

To find the coefficients a_n, we may formally substitute the expressions (4.34) into the equation (4.32) for the internal mode to obtain

$$\sum_n b_n\varphi_n^{I\wedge} = \left(\mathcal{A}_0 - |\mathbf{k}|^2\right)\left(\sum_n a_n\varphi_n^{I\wedge}\right) = \sum_n a_n\lambda_n\varphi_n^{I\wedge} \quad \text{so that} \quad a_n = \frac{b_n}{\lambda_n} \tag{4.36}$$

for $\lambda_n \neq 0$ by orthogonality. Note that the case $\lambda_n = 0$ is dealt with since we assume that the operator is invertible. Then, by (4.34), the coefficient b_n is then simply given by the inner products $b_n = (q^\wedge, \varphi_n^{I\wedge})/(\varphi_n^{I\wedge}, \varphi_n^{I\wedge})$ which implies

$$\psi^{I\wedge} = \sum_n \frac{1}{\lambda_n}\frac{(q^\wedge, \varphi_n^{I\wedge})}{(\varphi_n^{I\wedge}, \varphi_n^{I\wedge})}\varphi_n^{I\wedge} \tag{4.37}$$

for $\lambda_n \neq 0$. Finally, we may write the inversion of the potential vorticity as

$$\psi(x, y, z, t) = \sum_{\mathbf{k}} \psi^\wedge(z; \mathbf{k}, t)e^{i\mathbf{k}\cdot\mathbf{x}} \tag{4.38}$$

with $\mathbf{x} = (x, y)$, where

$$\begin{aligned}\psi^{\wedge}(z;\mathbf{k},t) = &\sum_n \frac{1}{\lambda_n}\frac{(q^{\wedge}(\cdot;\mathbf{k},t),\varphi_n^{I\wedge})}{(\varphi_n^{I\wedge},\varphi_n^{I\wedge})}\varphi_n^{I\wedge}(z) \\ &+ \frac{b^{\wedge}(0;\mathbf{k},t)}{f_0}\varphi^{T\wedge}(z;\mathbf{k}) + \frac{b^{\wedge}(-H;\mathbf{k},t)}{f_0}\varphi^{B\wedge}(z;\mathbf{k}).\end{aligned} \tag{4.39}$$

4.1.5 Multi-Level Quasi-Geostrophy

A discretized multi-level QG system can be obtained simply by discretizing the operator $\mathcal{A}_0$ defined in (4.2) [23, 77, 81]. We denote the discretized version of the continuous operator $\mathcal{A}_0$ as $\mathbf{A}_0$. At each level $i = 1, \ldots, m$, we write the equation for level potential vorticity as

$$q_i = \nabla_2^2\psi_i + \frac{1}{\delta_i}\left(b_{i-1}s_{i-1}^2 - b_i s_i^2\right), \tag{4.40}$$

where $\delta_i = H_i/H$ represents the i^{th} level thickness and b_i represents the buoyancy while $s_i^{-1} = N_i/\overline{N}$ represents the buoyancy frequency on the interface between level i and $i+1$. In addition, we discretize the buoyancy in terms of the streamfunction by taking a simple one-sided difference of (4.7) to obtain

$$b_i = \frac{\psi_i - \psi_{i+1}}{(\delta_i + \delta_{i+1})/2} \tag{4.41}$$

at each interface $i \neq 0, m$. To include boundary effects on the multi-level system we merely need to account for the top and bottom surface buoyancies b_0 and b_m, respectively. Again, we impose new interpretations on the quantities evaluated at the top most and bottom most interfaces. A simple argument gives

$$b_0 = \frac{\psi_0 - \psi_1}{\delta_1/2} \quad \text{and} \quad b_m = \frac{\psi_m - \psi_{m+1}}{\delta_m/2}, \tag{4.42}$$

where ψ_0 and ψ_{m+1} represent the streamfunction ψ evaluated at the top and bottom boundaries, respectively. We may thus, accept the equation (4.41) for $i = 0, \ldots, m+1$ with

$\delta_0 = \delta_{m+1} := 0$.

Therefore, the multi-level QG system maps the quantities ψ_0, ψ_{m+1}, ψ_1, …, ψ_m onto b_0, b_m, q_1, …, q_m. Namely, replacing (4.41) into (4.40) we obtain

$$q_i = \nabla_2^2 \psi_i + \frac{1}{\delta_i}\left(\frac{\psi_{i-1} - \psi_i}{L_{i-1}} - \frac{\psi_i - \psi_{i+1}}{L_i}\right) \tag{4.43}$$

or

$$q_i = \nabla_2^2 \psi_i + \frac{1}{\delta_i}\frac{\psi_{i+1}}{L_i} - \frac{1}{\delta_i}\left(\frac{1}{L_i} + \frac{1}{L_{i-1}}\right)\psi_i + \frac{1}{\delta_i}\frac{\psi_{i-1}}{L_{i-1}} \tag{4.44}$$

for all $i = 1, \ldots, m$ if we define

$$L_i := \frac{\delta_i + \delta_{i+1}}{2s_i^2}. \tag{4.45}$$

This may be concisely written in matrix form as

$$\boldsymbol{q} = \left(\nabla_2^2 \mathbb{I}^I + \mathbf{A}_0\right)\boldsymbol{\psi} \tag{4.46}$$

with vectors

$$\boldsymbol{q} = [b_0 \;\; q_1 \;\; \cdots \;\; q_m \;\; b_m]^{\mathsf{T}} \quad \text{and} \quad \boldsymbol{\psi} = [\psi_0 \;\; \psi_1 \;\; \cdots \;\; \psi_m \;\; \psi_{m+1}]^{\mathsf{T}}, \tag{4.47}$$

where $\mathbb{I}^I = \mathrm{diag}(0, 1, \ldots, 1, 0)$ and the matrix coefficients $\{(\mathbf{A}_0)_{i,j}\}_{i,j=0}^{m+1}$ of the $m+2 \times m+2$ tridiagonal matrix $\mathbf{A}_0$ are given by

$$\delta_i(\mathbf{A}_0)_{i,i-1} = \delta_{i-1}(\mathbf{A}_0)_{i-1,i} = \frac{1}{L_{i-1}}, \quad \delta_i(\mathbf{A}_0)_{i,i} = \frac{1}{L_i} + \frac{1}{L_{i-1}}, \tag{4.48}$$

for $i = 1, \ldots, m$,

$$\delta_1(\mathbf{A}_0)_{1,0} = -\frac{\delta_1}{2L_0}(\mathbf{A}_0)_{0,1} = \frac{\delta_1}{2L_0}(\mathbf{A}_0)_{0,0} = \frac{1}{L_0}, \tag{4.49}$$

and

$$\delta_m(\mathbf{A}_0)_{m,m+1} = \frac{\delta_m}{2L_m}(\mathbf{A}_0)_{m+1,m} = -\frac{\delta_m}{2L_m}(\mathbf{A}_0)_{m+1,m+1} = \frac{1}{L_m}. \tag{4.50}$$

Remark IV.1. Here we use the notation $\mathbb{I}^I = \text{diag}(0, 1, \ldots, 1, 0)$ to denote the identity matrix only for the values on the diagonal that coincide with the "interior" solution. Similarly, we define, for future reference, $\mathbb{I}^T = \text{diag}(1, 0, \ldots, 0)$ and $\mathbb{I}^B = \text{diag}(0, \ldots, 0, 1)$ representing the top and bottom boundaries, respectively. Note that $\mathbb{I}^I + \mathbb{I}^T + \mathbb{I}^B = \mathbb{I}$ with $\mathbb{I} = \text{diag}(1, \ldots, 1)$ being the identity matrix.

In particular, we have nearly shown the following.

Lemma IV.2. Let $\mathbf{B} = \text{diag}\left(\frac{\delta_1}{2L_0}, -\delta_1, \ldots, -\delta_m, -\frac{\delta_m}{2L_m}\right)$. The matrix $\mathbf{BA}_0$ for vertically discretized QG is real symmetric and singular.

Proof. The fact that $\mathbf{BA}_0$ is symmetric is immediate and follows directly from (4.48)–(4.50). Moreover, $\mathbf{BA}_0$ is singular since

$$\det \mathbf{BA}_0 = \det \begin{bmatrix} L_0^{-1} & -L_0^{-1} & 0 & \cdots & 0 & 0 \\ -L_0^{-1} & L_1^{-1} + L_0^{-1} & -L_1^{-1} & \cdots & 0 & 0 \\ 0 & -L_1^{-1} & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & L_{m-1}^{-1} + L_{m-2}^{-1} & -L_{m-1}^{-1} & 0 \\ 0 & 0 & \ldots & -L_{m-1}^{-1} & L_m^{-1} + L_{m-1}^{-1} & -L_m^{-1} \\ 0 & 0 & \ldots & 0 & -L_m^{-1} & L_m^{-1} \end{bmatrix} \tag{4.51}$$

$$= \det \begin{bmatrix} L_0^{-1} & -L_0^{-1} & 0 & \cdots & 0 & 0 \\ 0 & L_1^{-1} & -L_1^{-1} & \cdots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & L_{m-1}^{-1} & -L_{m-1}^{-1} & 0 \\ 0 & 0 & \ldots & 0 & L_m^{-1} & -L_m^{-1} \\ 0 & 0 & \ldots & 0 & 0 & 0 \end{bmatrix} = 0.$$

The last determinant in (4.51) is obtained by adding the first row to the second row, then the second row to the third row, and so on. ■

As one can see, equation (4.43) is thus applicable for all levels of the system, though some

care is needed to interpret the topmost and bottommost levels, i.e., $i = 1, m$. In particular, if we impose zero boundary buoyancy (interior QG), one must require a cancellation condition on ψ_0, ψ_{m+1} so that the potential vorticity of the top level is given solely by

$$q_1 = \nabla_2^2 \psi_1 + \frac{1}{\delta_1}\left(\frac{\psi_2 - \psi_1}{L_1}\right) \tag{4.52}$$

and the bottom level is given by

$$q_n = \nabla_2^2 \psi_n + \frac{1}{\delta_n}\left(\frac{\psi_{n-1} - \psi_n}{L_{n-1}}\right). \tag{4.53}$$

Note that this is equivalent to requiring for the buoyancy frequency s_i at the boundaries $i = 0, n+1$ to be identically zero. Now, interior QG is given by the relation

$$\boldsymbol{q} = \left(\nabla_2^2 + \mathbf{A}_0\right)\boldsymbol{\psi} \tag{4.54}$$

with vectors

$$\boldsymbol{q} = [q_1 \ \cdots \ q_m]^\mathsf{T} \quad \text{and} \quad \boldsymbol{\psi} = [\psi_1 \ \cdots \ \psi_m]^\mathsf{T}, \tag{4.55}$$

where

$$\mathbf{B}\mathbf{A}_0 = \begin{bmatrix} L_1^{-1} & -L_1^{-1} & 0 & \cdots & 0 & 0 \\ -L_1^{-1} & L_2^{-1} + L_1^{-1} & -L_2^{-1} & \cdots & 0 & 0 \\ 0 & -L_2^{-1} & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -L_{m-2}^{-1} & 0 \\ 0 & 0 & \dots & -L_{m-2}^{-1} & L_{m-1}^{-1} + L_{m-2}^{-1} & -L_{m-1}^{-1} \\ 0 & 0 & \dots & 0 & -L_{m-1}^{-1} & L_{m-1}^{-1} \end{bmatrix} \tag{4.56}$$

with $\mathbf{B} = \operatorname{diag}(-\delta_1, \dots, -\delta_m)$.

Corollary IV.3. Let $\mathbf{B} = \operatorname{diag}(-\delta_1, \dots, -\delta_m)$. The matrix $\mathbf{B}\mathbf{A}_0$ for vertically discretized

interior QG is real symmetric and singular.

Remark IV.4. The difference between interior QG and QG is discussed in Section 4.1.1. For our purposes, one crucial difference between interior QG and QG is highlighted by the absence of the matrix $\mathbb{I}^I$ in (4.54) (cf. (4.46)). For interior QG, $\mathbb{I}^I = \mathbb{I}$ as there are no boundary components.

4.2 Modal Decomposition

We proceed to construct a natural modal decomposition for the forced-dissipated QG system (4.15)–(4.19) using the matrix structure presented in Section 4.1.5. The decomposition presented here is analogous to that in the work of Smith and Vanneste [77] with some details fleshed out. Our primary aim is to use this decomposition to diagonalize not only the system (4.46), making inversion trivial, but also the significant conserved quantities of energy and enstrophy simultaneously.

Remark IV.5. Though conceptually useful, the decomposition of the QG system into interior, top, and bottom solutions shown in Section 4.1.1 does not diagonalize the energy and enstrophy when we include the boundary conditions (4.18)–(4.19).

Equation (4.46) may be written in terms of Fourier coefficients as

$$\boldsymbol{q}^{\wedge} = \mathbf{A}_{\kappa}\boldsymbol{\psi}^{\wedge}, \tag{4.57}$$

where $\mathbf{A}_{\kappa} = \mathbf{A}_0 - \kappa^2\mathbb{I}^I$ and $\kappa = |\mathbf{k}|$ is the magnitude of the wavenumber given by the Fourier series defined in (4.26). Now, the matrix $\mathbf{A}_{\kappa}$ of the transformed system (4.57) has very amenable properties as we proceed to show.

Proposition IV.6. Let $\mathbf{B} = \operatorname{diag}\left(\frac{\delta_1}{2L_0}, -\delta_1, \ldots, -\delta_m, -\frac{\delta_m}{2L_m}\right)$. The matrix $\mathbf{B}\mathbf{A}_{\kappa}$ is positive definite for $\kappa \neq 0$ and the matrix $\mathbf{B}\mathbf{A}_0$ is positive semidefinite.

We include the proof of Proposition IV.6 in Appendix C.2.

Note that Proposition IV.6 implies that equation (4.57) is invertible since $\mathbf{A}_\kappa$ cannot have a zero eigenvalue when $\kappa \neq 0$. We may show the equivalent statement and more specific details in the simpler interior QG case.

Corollary IV.7. Let $\mathbf{B} = \operatorname{diag}(-\delta_1, \ldots, -\delta_m)$ and suppose that (4.57) represents the interior QG system, i.e., $\mathbf{A}_\kappa = \mathbf{A}_0 - \kappa^2 \mathbb{I}$, $\boldsymbol{q} = [q_1 \ \cdots \ q_m]^\mathsf{T}$, and $\boldsymbol{\psi} = [\psi_1 \ \cdots \ \psi_m]^\mathsf{T}$. Then, the matrix $\mathbf{A}_\kappa$ is negative definite for $\kappa \neq 0$ and $\mathbf{A}_0$ is negative semidefinite.

Definition IV.8. For a matrix $\mathbf{N}$ with real eigenvalues, we define the matrix $\mathbf{I}(\mathbf{N})$ as the *inertia matrix* corresponding to $\mathbf{N}$, where the operation $\mathbf{I}(\mathbf{N})$ lists the sign of the eigenvalues of the matrix $\mathbf{N}$ on the main diagonal from positive to negative.

Example IV.9. Consider the matrices

$$\mathbf{E} = \begin{bmatrix} 2 & 7 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 0 & -1 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{bmatrix}, \quad \text{and} \quad \mathbf{G} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \tag{4.58}$$

then

$$\mathbf{I}(\mathbf{E}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{I}(\mathbf{F}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \text{and} \quad \mathbf{I}(\mathbf{G}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.59}$$

We prove the following proposition in a very general case as it will be useful for our discussion.

Proposition IV.10. Suppose $\mathbf{M}$ is a positive definite real symmetric matrix and $\mathbf{N}$ is real symmetric matrix, then there exists an invertible real matrix $\mathbf{V}$ that simultaneously diagonalizes $\mathbf{M}$ and $\mathbf{N}$ in the form

$$\mathbf{V}^\mathsf{T}\mathbf{M}\mathbf{V} = \mathbf{\Lambda} \tag{4.60}$$

and

$$\mathbf{V}^{\mathsf{T}}\mathbf{N}\mathbf{V} = \mathbf{I}(\mathbf{N}), \tag{4.61}$$

where $\mathbf{\Lambda}$ is diagonal with nonnegative entries and $\mathbf{I}(\mathbf{N})$ denotes the inertia matrix of $\mathbf{N}$.

We show the proof of Proposition IV.10 in Appendix C.2. Proposition IV.10 is, however, a variation of the well known (and used) result in Classical Mechanics upon which two symmetric forms need to be simultaneously diagonalized (change of basis) to ensure certain conserved quantities are diagonalized. By picking carefully our chosen matrices $\mathbf{M}$ and $\mathbf{N}$ we intend to use this very idea on the energy and enstrophy; as was done in similar fashion in [77].

Corollary IV.11. For the matrices $\mathbf{A}_\kappa$ and $\mathbf{B}$ defined above, there exists a real matrix $\mathbf{V}_\kappa$ that allows for a diagonalization of the form

$$\mathbf{I}(\mathbf{B})\mathbf{A}_\kappa\mathbf{V}_\kappa = \mathbf{V}_\kappa\mathbf{\Lambda}_\kappa \quad \text{with} \quad \mathbf{V}_\kappa^{\mathsf{T}}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{V}_\kappa = \mathbb{I} \tag{4.62}$$

when $\kappa \neq 0$, where $\mathbf{\Lambda}_\kappa$ is diagonal with nonnegative real entries and $\mathbf{I}(\mathbf{B})$ denotes the inertia matrix of $\mathbf{B}$.

Proof. Let $\mathbf{M} = \mathbf{B}\mathbf{A}_\kappa$ and $\mathbf{N} = \mathbf{I}(\mathbf{B})\mathbf{B}$. By Proposition IV.6 the matrix $\mathbf{M}$ is positive definite. So, using Proposition IV.10 we obtain the diagonalizations

$$\mathbf{V}_\kappa^{\mathsf{T}}\mathbf{B}\mathbf{A}_\kappa\mathbf{V}_\kappa = \mathbf{\Lambda}_\kappa \quad \text{and} \quad \mathbf{V}_\kappa^{\mathsf{T}}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{V}_\kappa = \mathbb{I}. \tag{4.63}$$

Lastly, we note

$$\mathbf{\Lambda}_\kappa = \mathbf{V}_\kappa^{\mathsf{T}}\mathbf{B}\mathbf{A}_\kappa\mathbf{V}_\kappa = \mathbf{V}_\kappa^{\mathsf{T}}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{V}_\kappa\mathbf{V}_\kappa^{-1}\mathbf{I}(\mathbf{B})\mathbf{A}_\kappa\mathbf{V}_\kappa = \mathbf{V}_\kappa^{-1}\mathbf{I}(\mathbf{B})\mathbf{A}_\kappa\mathbf{V}_\kappa, \tag{4.64}$$

which finishes the proof. ■

In this manner, Corollary IV.11 allows us to define the bases $\mathbf{V}_\kappa$ and $\mathbf{W}_\kappa = \mathbf{I}(\mathbf{B})\mathbf{V}_\kappa$ for

fixed $\kappa \neq 0$. We define the transformations

$$\boldsymbol{q}^{\wedge} = \mathbf{W}_{\kappa}\boldsymbol{q}^{\wedge}_{\mathbf{W}} \quad \text{and} \quad \boldsymbol{\psi}^{\wedge} = \mathbf{V}_{\kappa}\boldsymbol{\psi}^{\wedge}_{\mathbf{V}}. \tag{4.65}$$

Indeed, we immediately note that the QG system (4.57) is diagonalized by the change of basis:

$$\boldsymbol{q}^{\wedge}_{\mathbf{W}} = \mathbf{V}_{\kappa}^{\mathsf{T}}\mathbf{B}\boldsymbol{q}^{\wedge} = \mathbf{V}_{\kappa}^{\mathsf{T}}\mathbf{B}\mathbf{A}_{\kappa}\mathbf{V}_{\kappa}\boldsymbol{\psi}^{\wedge}_{\mathbf{V}} = \boldsymbol{\Lambda}_{\kappa}\boldsymbol{\psi}^{\wedge}_{\mathbf{V}}. \tag{4.66}$$

Note that Corollary IV.11, and therefore (4.66), is essentially the same whether we consider interior QG or QG. The true difference lies in the interpretation of the vectors $\boldsymbol{q}$ and $\boldsymbol{\psi}$ and the properties of the eigenvalues $\boldsymbol{\Lambda}_{\kappa}$.

Remark IV.12. In the case of interior QG, the bases $\mathbf{V}_{\kappa}$ and $\mathbf{W}_{\kappa}$ satisfy the relation $\mathbf{V}_{\kappa} = -\mathbf{W}_{\kappa}$.

4.2.1 Conserved Quantities in Quasi-Geostrophy

The QG system conserves an infinite number of quantities [10, 81, for example]. For the purposes of our discussion, we limit ourselves to the two most physical, and perhaps primary, quantities: the energy density

$$E_{\kappa} = \frac{1}{2}\boldsymbol{\psi}^{\wedge *}\mathbf{B}\boldsymbol{q}^{\wedge} \tag{4.67}$$

and "generalized" enstrophy [77] density

$$P_{\kappa} = \frac{1}{2}\boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}\boldsymbol{q}^{\wedge}. \tag{4.68}$$

Remark IV.13. We refer to P_{κ} as generalized enstrophy density due to the fact that its definition includes terms arising from the conservation of potential energy at the boundary of the flow [11] and not merely enstrophy in the interior of the domain.

From the change of basis (4.65) obtained in Corollary IV.11 and relation (4.66), is it easy to show that both the energy and enstrophy density given by (4.67) and (4.68) are

simultaneously diagonalized. Namely, for the energy density we have

$$E_\kappa = \frac{1}{2}\boldsymbol{\psi}^{\wedge *}\mathbf{B}\boldsymbol{q}^{\wedge} = \frac{1}{2}\boldsymbol{\psi}_{\mathbf{V}}^{\wedge *}\mathbf{V}_\kappa^{\mathsf{T}}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{V}_\kappa\boldsymbol{q}_{\mathbf{W}}^{\wedge} = \frac{1}{2}\boldsymbol{\psi}_{\mathbf{V}}^{\wedge *}\boldsymbol{\Lambda}_\kappa\boldsymbol{\psi}_{\mathbf{V}}^{\wedge} \tag{4.69}$$

and for the generalized enstrophy density we obtain

$$P_\kappa = \frac{1}{2}\boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}\boldsymbol{q}^{\wedge} = \frac{1}{2}\boldsymbol{q}_{\mathbf{W}}^{\wedge *}\mathbf{W}_\kappa^{\mathsf{T}}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{W}_\kappa\boldsymbol{q}_{\mathbf{W}}^{\wedge} = \frac{1}{2}\boldsymbol{\psi}_{\mathbf{V}}^{\wedge *}\boldsymbol{\Lambda}_\kappa^2\boldsymbol{\psi}_{\mathbf{V}}^{\wedge}. \tag{4.70}$$

The fact that (4.67) and (4.68) indeed define conserved quantities is readily seen from the vertically discretized forced-dissipated QG system (4.15)–(4.19). Writing (4.15)–(4.19) in multi-level form using vector notation we obtain

$$\frac{\partial \boldsymbol{q}}{\partial t} + J(\boldsymbol{\psi}, \boldsymbol{q}) + J(\boldsymbol{\psi}, \boldsymbol{Q}) + J(\boldsymbol{\Psi}, \boldsymbol{q}) = -\mathbf{R}\nabla_2^2\boldsymbol{\psi}, \tag{4.71}$$

where we use the definition of the Jacobian of two vectors as given in Appendix C.1 and limit ourselves to the cases where only Ekman friction provides dissipation through a diagonal positive semidefinite matrix $\mathbf{R}$. To obtain an energy density equation from (4.71) we compute its Fourier transform and then multiply by $\boldsymbol{\psi}^{\wedge *}\mathbf{B}$ to obtain

$$\boldsymbol{\psi}^{\wedge *}\mathbf{B}\frac{\partial \boldsymbol{q}^{\wedge}}{\partial t} + \boldsymbol{\psi}^{\wedge *}\mathbf{B}J^{\wedge}(\boldsymbol{\Psi}, \boldsymbol{q}) = \kappa^2\boldsymbol{\psi}^{\wedge *}\mathbf{B}\mathbf{R}\boldsymbol{\psi}^{\wedge}, \tag{4.72}$$

where we have dropped the Jacobian terms $J^{\wedge}(\boldsymbol{\psi}, \boldsymbol{q})+J^{\wedge}(\boldsymbol{\psi}, \boldsymbol{Q})$ since they are identically zero when we add over all wavenumbers; this follows from Parseval's identity and Corollary C.3. Here $J^{\wedge}$ denotes the Fourier transform of the Jacobian, i.e., $J^{\wedge}(f, g) = (J(f, g))^{\wedge}$. Finally, using the complex conjugate of (4.72) and the fact that $\mathbf{B}\mathbf{A}_\kappa$ is symmetric allows us to write an evolution equation for the energy density E_κ as

$$\frac{\partial E_\kappa}{\partial t} + \operatorname{Re}\left\{\boldsymbol{\psi}^{\wedge *}\mathbf{B}J^{\wedge}(\boldsymbol{\Psi}, \boldsymbol{q})\right\} = \kappa^2\boldsymbol{\psi}^{\wedge *}\mathbf{B}\mathbf{R}\boldsymbol{\psi}^{\wedge}. \tag{4.73}$$

It is then clear from equation (4.73) that in the absence of forcing ($\boldsymbol{\Psi} = \boldsymbol{Q} = \mathbf{0}$) and dissipation ($\mathbf{R} = \text{diag}(0, \ldots, 0)$) that total energy $E = \sum_{\mathbf{k}} E_\kappa$ is a conserved quantity.

Similarly, to show that the generalized enstrophy is conserved, we compute the Fourier transform of (4.71) and then multiply by $\boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}$ to obtain

$$\boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}\frac{\partial \boldsymbol{q}^{\wedge}}{\partial t} + \boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}J^{\wedge}(\boldsymbol{\psi}, \boldsymbol{Q}) = \kappa^2 \boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{R}\boldsymbol{\psi}^{\wedge}, \tag{4.74}$$

where we have dropped the Jacobian terms $J^{\wedge}(\boldsymbol{\psi}, \boldsymbol{q}) + J^{\wedge}(\boldsymbol{\Psi}, \boldsymbol{q})$ since they are identically zero when summed over the spectral domain. As before for the energy, using the complex conjugate of (4.74), we obtain the evolution equation

$$\frac{\partial P_\kappa}{\partial t} + \text{Re}\left\{\boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}J^{\wedge}(\boldsymbol{\psi}, \boldsymbol{Q})\right\} = \kappa^2 \text{Re}\left\{\boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{R}\boldsymbol{\psi}^{\wedge}\right\}. \tag{4.75}$$

As before, in the absence of forcing ($\boldsymbol{\Psi} = \boldsymbol{Q} = \mathbf{0}$) and dissipation ($\mathbf{R} = \text{diag}(0, \ldots, 0)$), equation (4.75) shows that total generalized enstrophy $P = \sum_{\mathbf{k}} P_\kappa$ is a conserved quantity.

4.2.2 Conditions for Cascade Inequality

We may use formulas (4.73) and (4.75) on the quantities E_κ and P_κ to quantify the direction of energy transfer in the QG system for large time. To accomplish this feat, we restrict the nature of the forcing to simplify the analytical work ahead. First, we assume that the forcing is that of zonal mean flow, i.e., mean flow acting only in the x-direction; see Section 4.1.2. Second, we assume a localized forcing acting uniquely at wavenumber $\kappa_f = |\mathbf{k}_f|$ that modifies equation (4.15) into the form

$$\frac{\partial \boldsymbol{Q}}{\partial y} = \kappa_f^2 \mathbf{I}(\mathbf{B})\frac{\partial \boldsymbol{\Psi}}{\partial y}. \tag{4.76}$$

Remark IV.14. The assumption of localized forcing (4.76) is primarily for analytical convenience and is not expected to hold in realistic settings. In practice, such an assumption

may be considered as a crude bound on a forcing consisting of several well-defined forcing scales.

These assumptions on the forcing allow us to write (4.73), after using Corollary C.3, as

$$\frac{\partial E_\kappa}{\partial t} + \mathrm{Re}\left\{\boldsymbol{q}^{\wedge *}\mathbf{B}J^{\wedge}(\boldsymbol{\psi},\boldsymbol{\Psi})\right\} = \kappa^2\boldsymbol{\psi}^{\wedge *}\mathbf{B}\mathbf{R}\boldsymbol{\psi}^{\wedge}. \tag{4.77}$$

Similarly, we may write (4.75) as

$$\frac{\partial P_\kappa}{\partial t} + \kappa_f^2\mathrm{Re}\left\{\boldsymbol{q}^{\wedge *}\mathbf{B}J^{\wedge}(\boldsymbol{\psi},\boldsymbol{\Psi})\right\} = \kappa^2\mathrm{Re}\left\{\boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{R}\boldsymbol{\psi}^{\wedge}\right\}. \tag{4.78}$$

We may combine (4.77) and (4.78), assuming statistically steady state (giving the large time solution), to obtain

$$\kappa_f^2\sum_{\mathbf{k}}\kappa^2\boldsymbol{\psi}^{\wedge *}\mathbf{B}\mathbf{R}\boldsymbol{\psi}^{\wedge} = \sum_{\mathbf{k}}\kappa^2\mathrm{Re}\left\{\boldsymbol{q}^{\wedge *}\mathbf{I}(\mathbf{B})\mathbf{B}\mathbf{R}\boldsymbol{\psi}^{\wedge}\right\}. \tag{4.79}$$

including the (previously neglected) notation of summation over the spectral domain.

We may now highlight several special cases of the dynamics in the case of interior QG. First, we quantify the scale at which the kinetic energy acts in the system.

Definition IV.15. The centroid of the kinetic energy $\mathbf{k}_E$ is defined by

$$\mathbf{k}_E\sum_{\mathbf{k}}K_\kappa = \sum_{\mathbf{k}}\mathbf{k}K_\kappa, \tag{4.80}$$

where $K_\kappa = \kappa^2\boldsymbol{\psi}^{\wedge *}\mathbf{B}\boldsymbol{\psi}$.

Lemma IV.16. Let $\mathbf{k}_E$ be the centroid of the kinetic energy. Then, for the magnitude $\kappa_E = |\mathbf{k}_E|$, we have

$$\kappa_E^2\sum_{\mathbf{k}}K_\kappa \leq \sum_{\mathbf{k}}\kappa^2K_\kappa. \tag{4.81}$$

Proof. Let $\mathbf{k}_E = (k_E^x, k_E^y)$. Using Cauchy-Schwarz we have

$$\left(k_E^x \sum_{\mathbf{k}} K_\kappa\right)^2 = \left(\sum_{\mathbf{k}} k^x K_\kappa\right)^2 \leq \left(\sum_{\mathbf{k}} (k^x)^2 K_\kappa\right)\left(\sum_{\mathbf{k}} K_\kappa\right). \tag{4.82}$$

So,

$$(k_E^x)^2 \sum_{\mathbf{k}} K_\kappa \leq \sum_{\mathbf{k}} (k^x)^2 K_\kappa. \tag{4.83}$$

A similar result applies to k_E^y, implying

$$\kappa_E^2 \sum_{\mathbf{k}} K_\kappa \leq \sum_{\mathbf{k}} \kappa^2 K_\kappa \tag{4.84}$$

as stated. ■

Proposition IV.17 was obtained in a different manner in [4].

Proposition IV.17. In the interior QG case, suppose that the Ekman friction is vertically uniform, i.e., $\mathbf{R} = r_{\text{Ek}}\mathbb{I}$. Then, the forcing scale κ_f satisfies the inequality

$$\kappa_f^2 \geq \kappa_E^2 + \frac{\sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\Lambda}_0 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge}}{\sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\psi}_{\mathbf{V}}^{\wedge}}, \tag{4.85}$$

where $K_\kappa = \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\psi}_{\mathbf{V}}^{\wedge}$ and $\boldsymbol{\Lambda}_0$ is given by the eigenvalue system (4.62) when $\kappa = 0$.

Proof. For vertically uniform Ekman friction, equation (4.79) becomes

$$\kappa_f^2 \sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\psi}_{\mathbf{V}}^{\wedge} = \sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\Lambda}_\kappa \boldsymbol{\psi}_{\mathbf{V}}^{\wedge} \tag{4.86}$$

after using the decomposition (4.65)–(4.66). Note that in the interior QG case we may indeed write $\boldsymbol{\Lambda}_\kappa = \kappa^2 \mathbb{I} + \boldsymbol{\Lambda}_0$, where $\boldsymbol{\Lambda}_0$ is given by eigenvalue system (4.62) when $\kappa = 0$. This implies

$$\kappa_f^2 \sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\psi}_{\mathbf{V}}^{\wedge} = \sum_{\mathbf{k}} \kappa^4 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\psi}_{\mathbf{V}}^{\wedge} + \sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\Lambda}_0 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge}. \tag{4.87}$$

Lastly, noting that $K_\kappa = \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\psi}_{\mathbf{V}}^{\wedge}$ and using Lemma IV.16 we obtain

$$\kappa_f^2 \geq \kappa_E^2 + \frac{\sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\Lambda}_0 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge}}{\sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}_{\mathbf{V}}^{\wedge *} \boldsymbol{\psi}_{\mathbf{V}}^{\wedge}} \tag{4.88}$$

as stated. ■

We finish the chapter by stating a result in the case that the dissipation is only acting on the bottommost level of the system. First we derive an auxiliary result.

Lemma IV.18. Let $\mathbf{C}$ be an $n \times n$ matrix such that

$$\mathbf{C} = \begin{bmatrix} 0 & \mathbf{a} \\ \mathbf{a}^\mathsf{T} & b \end{bmatrix}, \tag{4.89}$$

where $\mathbf{a}$ is a vector and b is a scalar. Then, $\mathbf{C}$ has $n-2$ eigenvalues which are 0 while the remaining two are given by the formula $\lambda_\pm = \frac{1}{2}\left(b \pm \sqrt{b^2 + 4\mathbf{a}^\mathsf{T}\mathbf{a}}\right)$ and are associated with eigenvectors $\mathbf{v}_\pm = [\mathbf{a}^\mathsf{T}\ \lambda_\pm]^\mathsf{T}$.

Proof. The fact that there are $n-2$ eigenvalues equal to 0 is given by the fact that there are $n-2$ linear combinations of the entries of $\mathbf{a}^\mathsf{T}$ which are equal to zero. For the other eigenvalues, consider the eigenvector $\mathbf{v} = [\mathbf{c}^\mathsf{T}\ d]^\mathsf{T}$, where $\mathbf{c}$ is a vector and d a scalar. Then,

$$\mathbf{C}\mathbf{v} = \lambda \mathbf{v} \quad \text{implies} \quad \begin{bmatrix} 0 & \mathbf{a} \\ \mathbf{a}^\mathsf{T} & b \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ d \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{c} \\ d \end{bmatrix} \tag{4.90}$$

which means that

$$d\mathbf{a} = \lambda \mathbf{c} \quad \text{and} \quad \mathbf{a}^\mathsf{T}\mathbf{c} + bd = \lambda d. \tag{4.91}$$

These conditions can be combined (multiply the second by λ) to obtain

$$\mathbf{a}^\mathsf{T}\mathbf{a} + b\lambda = \lambda^2. \tag{4.92}$$

which has solutions $\lambda_\pm = \frac{1}{2}\left(b \pm \sqrt{b^2 + 4\mathbf{a}^\mathsf{T}\mathbf{a}}\right)$. The fact that the eigenvector is given as stated, follows from the equation that $\lambda_+ + \lambda_- = b$ and $\lambda_+\lambda_- = -\mathbf{a}^\mathsf{T}\mathbf{a}$. Note that these eigenvectors are orthogonal since

$$\mathbf{v}_+^\mathsf{T}\mathbf{v}_- = \begin{bmatrix}\mathbf{a}^\mathsf{T} & \lambda_+\end{bmatrix}\begin{bmatrix}\mathbf{a} \\ \lambda_-\end{bmatrix} = \mathbf{a}^\mathsf{T}\mathbf{a} + \lambda_+\lambda_- = 0. \tag{4.93}$$

This finishes the proof. ■

Proposition IV.19. In the interior QG case, suppose that the Ekman friction acts only on the bottom most level, i.e., $R = r_{\text{Ek}}\mathbb{I}^B$. Then, the forcing scale κ_f satisfies the inequality

$$\kappa_f^2 \geq \kappa_n^2 + \frac{L_{m-1}^{-1}}{2\delta_m}(1-\sqrt{2})\frac{\sum_{\mathbf{k}} K_\kappa}{\sum_{\mathbf{k}} K_{\kappa,m}}, \tag{4.94}$$

where $K_{\kappa,m} = \kappa^2|\psi_m^\wedge|^2$ denotes the kinetic energy of the bottommost level and κ_n is the centroid of $K_{\kappa,m}$.

Proof. For Ekman friction acting only on the bottom level, (4.79) becomes

$$\kappa_f^2\delta_m\sum_{\mathbf{k}}\kappa^2\psi_m^{\wedge *}\psi_m^\wedge = \delta_m\sum_{\mathbf{k}}\kappa^4\psi_m^{\wedge *}\psi_m^\wedge + \sum_{\mathbf{k}}\kappa^2\text{Re}\left\{\boldsymbol{\psi}^{\wedge *}\mathbf{A}_0^\mathsf{T}\mathbf{B}\mathbb{I}^B\boldsymbol{\psi}^\wedge\right\}. \tag{4.95}$$

We may then re-write the rightmost term of (4.95) in the form $\boldsymbol{\psi}^{\wedge *}\mathbf{H}\boldsymbol{\psi}^\wedge$ with

$$\mathbf{H} = \frac{1}{2}\left(\mathbb{I}^B\mathbf{B}\mathbf{A}_0 + \mathbf{B}\mathbf{A}_0\mathbb{I}^B\right) = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -L_{m-1}^{-1}/2 \\ 0 & \cdots & -L_{m-1}^{-1}/2 & L_{m-1}^{-1}\end{bmatrix}. \tag{4.96}$$

From Proposition IV.18 we find that $\mathbf{H}$ has eigenvalues (other than zero)

$$\lambda_\pm = \frac{L_{m-1}^{-1}}{2}\left(1 \pm \sqrt{2}\right) \tag{4.97}$$

with $\lambda_+ \geq 0 \geq \lambda_-$. Therefore, we may note by the Rayleigh-Ritz Theorem [34] that

$$\boldsymbol{\psi}^{\wedge *}\mathbf{H}\boldsymbol{\psi}^{\wedge} \geq \lambda_- \boldsymbol{\psi}^{\wedge *}\boldsymbol{\psi}^{\wedge}. \tag{4.98}$$

Finally, using Lemma IV.16 we obtain,

$$\kappa_f^2 \sum_{\mathbf{k}} \kappa^2 \psi_m^{\wedge *}\psi_m^{\wedge} \geq \kappa_n^2 \sum_{\mathbf{k}} \kappa^2 \psi_m^{\wedge *}\psi_m^{\wedge} + \frac{\lambda_-}{\delta_m} \sum_{\mathbf{k}} \kappa^2 \boldsymbol{\psi}^{\wedge *}\boldsymbol{\psi}^{\wedge} \tag{4.99}$$

as stated. ■

CHAPTER V

Summary

In this dissertation, we study three models of stratified fluid dynamics: the BO equation in Chapter II, a two-layer tidal model in Chapter III, and the QG equations in Chapter IV. Our intent is to extend and elaborate on certain properties of each model and showcase their mathematical richness.

In Chapter II, we compute the scattering data of the Benjamin-Ono (BO) equation for arbitrary rational initial conditions with simple poles, under mild restrictions. We obtain explicit formulas for the Jost solutions and eigenfunctions of the associated spectral problem, yielding an Evans function for the eigenvalues and formulas for the phase constants and reflection coefficient. For this class of initial conditions, the recovery of the scattering data can be done by using the analyticity properties of the Jost functions. We proceed to use the exact formulas for the scattering data to analyze their behavior in the small-dispersion limit. In particular, we deduce precise asymptotic formulas for the reflection coefficient, the location of the eigenvalues and their density, and the asymptotic dependence of the phase constant on the eigenvalue. Our results give direct confirmation of conjectures in the literature that have been partly justified by means of inverse scattering, and they also provide new details to the spectral problem of the BO equation. This work can be seen as a significant extension of Kodama, Ablowitz, and Satsuma's [43] who only considered specific Lorentzian initial conditions to recover the location of the eigenvalues.

A natural generalization of the work in Chapter II entails extending the results to include rational potentials with multiple poles (cf. (2.67)). This extension may be viewed as a stepping stone to a general theory of rational potentials. In premise a general theory for rational potentials may then be used, through some density argument, to construct approximations of arbitrary potentials in some suitable class. It is again conceivable that errors arising from the approximations may be controlled so as to allow a characterization of the BO initial value problem through IST. In addition, the work in Chapter II would conceivably allow for the explicit reconstruction of the solution of the BO equation — using IST methodology — in the zero dispersion limit ($\epsilon \to 0$). Such reconstructions are well known for reflectionless potentials ($\beta = 0$) for both the KdV [46] and BO [55] equations. Our theory of rational initial conditions would perhaps allow for the reconstruction of solutions even when the reflection coefficient is non-negligible. In particular, having direct access to formulas for the scattering data allows for concrete analysis of the inverse problem, presumably simplifying the analysis.

In Chapter III, we present an analytical tide model that demonstrates the influence of stratification on both large- and small-scale surface tidal elevations in a qualitatively similar manner as in the global realistic-domain numerical simulations shown in [5], [75], and Section 3.1 of this dissertation. Our analytical results demonstrate the potential for the presence of stratification to alter the large-scale (barotropic) tide, and for climatic perturbations of oceanic stratifications to contribute to temporal changes in tides, for instance seasonal and secular changes seen in regional model studies and in tide gauge records. Our derived analytical formulas for the large- and small-scale contributions to surface tidal elevations contain explicit dependence on the stratification parameter $\alpha\gamma(1-\gamma) = g'H_1H_2/(g(H_1+H_2)^2)$. The dependence on stratification is accentuated if there is significant bottom topography and if the damping acts only on the bottom layer. We quantify these effects with plots of the RMS values of the surface elevation over a Gaussian topography using representative oceanic parameters. We find that changes in stratification may change large-scale tidal elevations

by as much as 10%, making these perturbations of a size comparable to the small-scale tidal elevations.

With this analytical model in place, it is then possible to re-interpret the global realistic-domain numerical results shown in Figures 3.1, 3.2 and guide additional more realistic simulations. Namely, since the quantity $\alpha\gamma(1-\gamma)$ is a strong predictor — subject to bottom topography and damping — of model behavior in our simple analytical model, it is plausible that such a result is observable — at least in some manner — in more realistic oceanographic settings. Motivated by our analytical model results, we plan to write a follow-up paper in which a global realistic-domain numerical model, run in the non-Boussinesq limit, is used to perform a suite of simulations with different $\alpha\gamma(1-\gamma)$ values.

In Chapter IV, we investigate the influence of bottom friction on Quasi-Geostrophic (QG) turbulence in a multi-level forced-dissipated model that includes surface boundary effects, i.e., SQG effects. Motivated by earlier studies on the influence of bottom friction in two-layer QG turbulence, we focus on the influence of bottom friction on the horizontal scales and vertical structure of eddy kinetic energy. We briefly examine whether the inverse cascade to large, barotropic, energetic eddies in the weak drag limit of two-layer QG turbulence also takes place in multi-level turbulence. We discuss the usual interior and surface QG modal decomposition [45, 80], and a more natural surface-aware decomposition [77] with slight extensions, in our analytical work. In particular, we analyze relevant conserved quantities, e.g., energy and enstrophy, to showcase the influence of bottom friction on energy cascades.

In their current form, the results presented in Chapter IV are somewhat cumbersome to interpret directly. It is conceivable that generalizations and simplifications of the arguments presented in Chapter IV relating to the inverse cascade for the QG model could be fruitful. However, the utility of our results is perhaps mostly reserved to aid numerical computations of this complex system. Using our cascade inequality we intend to interpret energy cascades in numerical simulations and guide the further exploration of the QG system numerically.

APPENDICES

APPENDIX A

Benjamin-Ono

A.1 Proofs

In this section provide certain proofs that we consider too cumbersome to have in the body of the dissertation. We restate the propositions to be proven for ease of reference.

Proposition II.20. Let $\lambda > 0$. For each $m = 1, \ldots, P$, define the function $H_m = H_m(z; \lambda)$, $z \in G$, by the equation

$$(i(z - z_m))^{-ic_m/\epsilon - 1} H_m(z; \lambda) := -e^{-ih(z;\lambda)/\epsilon} \left(\lambda + \sum_{p=1}^{P} \frac{v_p(\lambda)}{z - z_p} \right). \tag{A.1}$$

(H_m is analytic not only on G, but also in a neighborhood of the point z_m.) Furthermore, let $ic_m = \epsilon(\mu_m + \omega_m)$, where $\mu_m = [\mathrm{Re}\{ic_m/\epsilon\}]$ denotes the integer (floor) part of $\mathrm{Re}\{ic_m/\epsilon\}$ and the remainder satisfies $0 \leq \mathrm{Re}\{\omega_m\} < 1$. Then, the function $W_+(x; \lambda)$ given by (2.72) is analytic in the upper-half x-plane if and only if for each $m = 1, \ldots, P$, either

$$\int_{\ell_0(z_m)} (i(z - z_m))^{-ic_m/\epsilon - 1} H_m(z; \lambda)\, dz = 0 \quad \text{when} \quad \mathrm{Re}\{ic_m/\epsilon\} < 0, \tag{A.2}$$

$$\int_{\ell_0(z_m)} (i(z-z_m))^{-\omega_m} \frac{d^{\mu_m+1}}{dz^{\mu_m+1}} H_m(z;\lambda)\, dz = 0 \quad \text{when} \quad \mathrm{Re}\{ic_m/\epsilon\} \geq 0 \text{ and } \omega_m \neq 0, \tag{A.3}$$

or

$$\operatorname*{Res}_{z=z_m} \frac{H_m(z;\lambda)}{(z-z_m)^{\mu_m+1}} = 0 \quad \text{when} \quad \mathrm{Re}\{ic_m/\epsilon\} \geq 0 \text{ and } \omega_m = 0, \tag{A.4}$$

where $\ell_0(z_m)$ (see Definition II.17) originates in at $-i\infty$ in G_0.

Proof of Proposition II.20: Fix $\lambda > 0$. The conditions of the proposition are equivalent to $W_+(x) = W_+(x;\lambda)$ being analytic at each possible singularity $z_1, \ldots, z_P$ in the upper-half x-plane. For any $m = 1, \ldots, P$, we use (A.1) defining $H_m(x) = H_m(x;\lambda)$ in (2.72) to express $W_+(x)$ in the form

$$W_+(x) = g_m(x)(i(x-z_m))^{ic_m/\epsilon} \int_{\ell_0(x)} (i(z-z_m))^{-ic_m/\epsilon-1} H_m(z)\, dz, \tag{A.5}$$

where we have also introduced a function $g_m(x)$ analytic and nonvanishing at $x = z_m$ defined precisely by the identity $\epsilon g_m(x)(i(x-z_m))^{ic_m/\epsilon} = ie^{ih(x)/\epsilon}$. We will now use Cauchy's Theorem to deform the path $\ell_0(x)$ of integration so as to pass directly through the branch point z_m on the way to a nearby point $x \in G$. While $H_m(z)$ is analytic at $z = z_m$, the other factor in the integrand of (A.5) may not be integrable at $z = z_m$.

First suppose that $\mathrm{Re}\{ic_m/\epsilon\} < 0$. The factor $(i(z-z_m))^{-ic_m/\epsilon-1}$ is integrable at $z = z_m$ so we can apply Cauchy's Theorem to obtain

$$\begin{aligned} W_+(x) = g_m(x)(i(x-z_m))^{ic_m/\epsilon} &\int_{\ell_0(z_m)} (i(z-z_m))^{-ic_m/\epsilon-1} H_m(z)\, dz \\ &+ g_m(x)(i(x-z_m))^{ic_m/\epsilon} \int_{z_m}^{x} (i(z-z_m))^{-ic_m/\epsilon-1} H_m(z)\, dz. \end{aligned} \tag{A.6}$$

The second term on the right-hand side is analytic for x near z_m. Indeed, substituting for

$H_m(z)$ its power series in $i(z - z_m)$,

$$H_m(z) = \sum_{k=0}^{\infty} \eta_{mk} (i(z - z_m))^k, \tag{A.7}$$

for small enough $|z - z_m|$, term-by-term integration yields

$$g_m(x)(i(x - z_m))^{ic_m/\epsilon} \int_{z_m}^{x} (i(z - z_m))^{-ic_m/\epsilon - 1} H_m(z)\, dz$$
$$= -ig_m(x) \sum_{k=0}^{\infty} \frac{\eta_{mk}}{k - ic_m/\epsilon} (i(x - z_m))^k. \tag{A.8}$$

Now, (A.8) is analytic at $x = z_m$ as it is a product of an analytic function and a convergent power series. So, requiring that $W_+(x)$ be analytic at $x = z_m$ is equivalent to demanding that the first term on the right-hand side of (A.6) be analytic at $x = z_m$. But this term is the product of (i) $g_m(x)$ nonvanishing at $x = z_m$, (ii) $(i(x - z_m))^{ic_m/\epsilon}$ which blows up as $x \to z_m$ since $\mathrm{Re}\{ic_m/\epsilon\} < 0$, and (iii) the x-independent integral on $\ell_0(z_m)$. Hence, the analyticity of $W_+(x)$ at $x = z_m$ is equivalent to the vanishing of this integral (cf. (A.2)).

Next let $\mathrm{Re}\{ic_m/\epsilon\} \geq 0$ with $\omega_m \neq 0$. Since $(i(z - z_m))^{-ic_m/\epsilon - 1}$ is branched and nonintegrable at $z = z_m$, we integrate by parts $\mu_m + 1$ times (note $\mu_m \geq 0$) to restore integrability before deforming the contour $\ell_0(x)$. With $K_m := i^{\mu_m + 1}\Gamma(\omega_m)/\Gamma(\omega_m + \mu_m + 1) \neq 0$ we have

$$(i(z - z_m))^{-ic_m/\epsilon - 1} = K_m \frac{d^{\mu_m + 1}}{dz^{\mu_m + 1}} (i(z - z_m))^{-\omega_m}. \tag{A.9}$$

Integrating (A.5) by parts using (A.9), note that $(i(z - z_m))^{-\omega_m}$ is integrable at $z = z_m$

$(0 \leq \mathrm{Re}\{\omega_m\} < 1)$, so Cauchy's Theorem gives

$$W_+(x) = (-1)^{\mu_m+1} K_m g_m(x)(i(x-z_m))^{ic_m/\epsilon} \int_{\ell_0(z_m)} \frac{H_m^{(\mu_m+1)}(z)}{(i(z-z_m))^{\omega_m}}\, dz$$
$$+ (-1)^{\mu_m+1} K_m g_m(x)(i(x-z_m))^{ic_m/\epsilon} \int_{z_m}^{x} \frac{H_m^{(\mu_m+1)}(z)}{(i(z-z_m))^{\omega_m}}\, dz$$
$$+ K_m g_m(x)(i(x-z_m))^{ic_m/\epsilon} \sum_{k=0}^{\mu_m} (-1)^k H_m^{(k)}(x) \frac{d^{\mu_m-k}}{dx^{\mu_m-k}} \frac{1}{(i(x-z_m))^{\omega_m}}. \quad \text{(A.10)}$$

Now, recalling $ic_m/\epsilon = \mu_m + \omega_m$, note that the terms on the third line of (A.10) are analytic at $x = z_m$. Likewise, with the help of the uniformly convergent power series (A.7) and term-by-term differentiation and integration (assuming only that $|x - z_m|$ is sufficiently small), the second line of (A.10) is analytic at $x = z_m$. The first line of the right-hand side of (A.10) is the product of (i) $K_m g_m(x)$ analytic and nonzero at $x = z_m$, (ii) a branched factor $(i(x-z_m))^{ic_m/\epsilon}$, and (iii) the definite integral on $\ell_0(z_m)$. This term of $W_+(x)$ is therefore analytic if and only if the integral on $\ell_0(z_m)$ vanishes (cf. (A.3)).

If $\mathrm{Re}\{ic_m/\epsilon\} \geq 0$ but instead $\omega_m = 0$, then (A.5) becomes

$$W_+(x) = -i g_m(x)(x-z_m)^{\mu_m} \int_{\ell_0(x)} \frac{H_m(z)}{(z-z_m)^{\mu_m+1}}\, dz, \quad \mu_m = 0, 1, 2, 3, \ldots. \quad \text{(A.11)}$$

Since g_m and H_m are analytic at z_m, $W_+(x)$ will be analytic at z_m exactly when the integral factor is single-valued as a function of x (cf. (A.4)). ■

Proposition II.32. Let $\mathbf{v}(\lambda)$ be the unique solution of $\mathbf{A}^<(\lambda)\mathbf{v}(\lambda) = \mathbf{b}^<(\lambda)$ for each $\lambda \in \mathbb{C} \setminus \mathbb{R}^+$ for which $D(\lambda) \neq 0$. Then each component $v_p(\lambda)$ is analytic in $\mathbb{C} \setminus \mathbb{R}^+$ except at the eigenvalues $\{\lambda_j\}_{j=1}^N$ which are simple poles, with corresponding Laurent expansion

$$v_p(\lambda) = -i\epsilon \frac{\phi_p(\lambda_j)}{\lambda - \lambda_j} + (z_p + \Gamma_j)\phi_p(\lambda_j) + O(\lambda - \lambda_j) \quad \text{as} \quad \lambda \to \lambda_j \quad \text{(A.12)}$$

where Γ_j is a constant (independent of p) given by

$$\Gamma_j := -\frac{i\epsilon}{2\lambda_j} - \frac{1}{2\lambda_j}\sum_{p=1}^{P} z_p\phi_p(\lambda_j) - \frac{i\epsilon}{2}\frac{\mathbf{m}^{\mathsf{T}}\mathbf{b}^{<\prime}(\lambda_j)}{\mathbf{m}^{\mathsf{T}}\mathbf{b}^{<}(\lambda_j)}, \tag{A.13}$$

and where $\{\phi_p(\lambda_j)\}_{p=1}^{P}$ are the components of the right nullvector of $\mathbf{A}^{<}(\lambda_j)$ normalized by (2.93) and $\mathbf{m}$ is a nonzero left nullvector of $\mathbf{A}^{<}(\lambda_j)$.

Proof of Proposition II.32: Since $\mathbf{A}(\lambda) = \mathbf{A}^{<}(\lambda)$ given by (2.83) and $\mathbf{b}(\lambda) = \mathbf{b}^{<}(\lambda)$ given by (2.103) are analytic for $\lambda \in \mathbb{C} \setminus \mathbb{R}^{+}$, the solution $\mathbf{v}(\lambda)$ of the system $\mathbf{A}(\lambda)\mathbf{v}(\lambda) = \mathbf{b}(\lambda)$ will be analytic in the same domain with the possible exception of the eigenvalues at which $D(\lambda) = 0$. As discussed for the general theory of Section 2.1.2, these isolated singular points are simple poles of $\mathbf{v}(\lambda)$, which therefore has a Laurent expansion about $\lambda = \lambda_j$ of the general form

$$\mathbf{v}(\lambda) = \frac{\mathbf{v}_j^{[-1]}}{\lambda - \lambda_j} + \mathbf{v}_j^{[0]} + (\lambda - \lambda_j)\mathbf{v}_j^{[1]} + O\left((\lambda - \lambda_j)^2\right) \quad \text{as} \quad \lambda \to \lambda_j. \tag{A.14}$$

Substitution into $\mathbf{A}(\lambda)\mathbf{v}(\lambda) = \mathbf{b}(\lambda)$ implies:

$$\mathbf{A}(\lambda_j)\mathbf{v}_j^{[-1]} = \mathbf{0}, \tag{A.15}$$

$$\mathbf{A}(\lambda_j)\mathbf{v}_j^{[0]} = \mathbf{b}(\lambda_j) - \mathbf{A}'(\lambda_j)\mathbf{v}_j^{[-1]}, \tag{A.16}$$

$$\mathbf{A}(\lambda_j)\mathbf{v}_j^{[1]} = \mathbf{b}'(\lambda_j) - \frac{1}{2}\mathbf{A}''(\lambda_j)\mathbf{v}_j^{[-1]} - \mathbf{A}'(\lambda_j)\mathbf{v}_j^{[0]}. \tag{A.17}$$

Since $\mathbf{A}(\lambda_j)$ is singular and has maximal rank $P - 1$, the first two equations (A.15) and (A.16) have the general solutions

$$\mathbf{v}_j^{[-1]} = \alpha_j^{[-1]}\boldsymbol{\phi}(\lambda_j) \quad \text{and} \quad \mathbf{v}_j^{[0]} = \mathbf{p} + \alpha_j^{[0]}\boldsymbol{\phi}(\lambda_j), \tag{A.18}$$

where $\mathbf{p}$ is a particular solution of (A.16) and $\boldsymbol{\phi}(\lambda_j)$ is the nullvector of $\mathbf{A}(\lambda_j)$ normalized according to (2.93). Here $\alpha_j^{[-1]}$ and $\alpha_j^{[0]}$ are constants determined by the fact that the right-hand sides of the equations (A.16) and (A.17) lie in the range of the singular matrix $\mathbf{A}(\lambda_j)$.

Therefore, multiplying these latter two equations on the left by $\mathbf{m}^\mathsf{T}$ where $\mathbf{m}$ is any nonzero left nullvector of $\mathbf{A}(\lambda_j)$ (unique up to scaling) and using (A.18) gives

$$\alpha_j^{[-1]} = \frac{\mathbf{m}^\mathsf{T}\mathbf{b}(\lambda_j)}{\mathbf{m}^\mathsf{T}\mathbf{A}'(\lambda_j)\boldsymbol{\phi}(\lambda_j)} \tag{A.19}$$

and

$$\alpha_j^{[0]} = \frac{\mathbf{m}^\mathsf{T}\left(\mathbf{b}'(\lambda_j) - \frac{1}{2}\alpha_j^{[-1]}\mathbf{A}''(\lambda_j)\boldsymbol{\phi}(\lambda_j) - \mathbf{A}'(\lambda_j)\mathbf{p}\right)}{\mathbf{m}^\mathsf{T}\mathbf{A}'(\lambda_j)\boldsymbol{\phi}(\lambda_j)}. \tag{A.20}$$

To simplify these formulas, we differentiate (2.90) to obtain the identity

$$\mathbf{A}'(\lambda) = \frac{i}{\epsilon\lambda}\mathbf{B}(\lambda) - \frac{i}{\epsilon}\mathbf{A}(\lambda)\mathbf{Z}, \tag{A.21}$$

where $\mathbf{B}(\lambda) := [\mathbf{b}(\lambda)\cdots\mathbf{b}(\lambda)]$ and $\mathbf{Z} := \operatorname{diag}(z_1,\ldots,z_N)$. Since $\mathbf{m}^\mathsf{T}\mathbf{A}(\lambda_j) = \mathbf{0}^\mathsf{T}$, (A.19) becomes

$$\alpha_j^{[-1]} = -i\epsilon\frac{\lambda_j\mathbf{m}^\mathsf{T}\mathbf{b}(\lambda_j)}{\mathbf{m}^\mathsf{T}\mathbf{B}(\lambda_j)\boldsymbol{\phi}(\lambda_j)} = -i\epsilon\frac{\mathbf{m}^\mathsf{T}\mathbf{b}(\lambda_j)}{\mathbf{m}^\mathsf{T}\mathbf{b}(\lambda_j)} = -i\epsilon, \tag{A.22}$$

due to (2.93). (That $\mathbf{m}^\mathsf{T}\mathbf{b}(\lambda_j) \neq 0$ is proven in Proposition II.34.) Now, substituting (A.18), (A.21), and (A.22) into (A.16) and using (2.93) again yields

$$\mathbf{A}(\lambda_j)\mathbf{p} = \mathbf{A}(\lambda_j)\mathbf{Z}\boldsymbol{\phi}(\lambda_j), \tag{A.23}$$

and hence the particular solution $\mathbf{p}$ may be taken as $\mathbf{p} = \mathbf{Z}\boldsymbol{\phi}(\lambda_j)$. Finally, substituting (A.21) and its derivative into (A.20) shows that $\alpha_j^{[0]} = \Gamma_j$ as defined by (A.13). ■

Proposition II.34. The components of the only nontrivial nullvector $\mathbf{n}(\lambda_j)$ of $\mathbf{A}^<(\lambda_j)$ are given (up to a common constant) by the formula

$$n_p(\lambda) = D_p(\lambda), \tag{A.24}$$

where $D_p(\lambda) := \det\left(\mathbf{A}_p^<(\lambda)\right)$ and $\mathbf{A}_p^<(\lambda)$ denotes the matrix $\mathbf{A}^<(\lambda)$ with its p^{th} column

replaced with $\mathbf{b}(\lambda)$. Moreover, the components satisfy the condition

$$\sum_{p=1}^{P} n_p(\lambda) \neq 0. \tag{A.25}$$

Proof of Proposition II.34. Consider the linear system

$$\mathbf{A}^{<}(\lambda)\mathbf{n}(\lambda) = \mathbf{b}^{<}(\lambda)D\left(\lambda\right), \tag{A.26}$$

where $\mathbf{n}(\lambda)$ is a vector in $\mathbb{C}^P$. By Corollary II.27, for an eigenvalue λ_j, we have that $D\left(\lambda_j\right) = 0$ and so $\mathbf{n}(\lambda_j)$ is a nullvector of $\mathbf{A}^{<}(\lambda_j)$. From Cramer's rule, for $\lambda \neq \lambda_j$, we have that each component of $\mathbf{n}(\lambda)$ satisfies

$$n_p(\lambda) = D_p\left(\lambda\right) \quad \text{for} \quad p = 1, \ldots, P. \tag{A.27}$$

From the analyticity of $\mathbf{A}_p^{<}(\lambda)$ we may deduce that

$$n_p(\lambda_j) = D_p\left(\lambda_j\right) \quad \text{for} \quad p = 1, \ldots, P. \tag{A.28}$$

The fact that $n_p(\lambda_j) \neq 0$ for each p follows from the fact that $\mathbf{b}^{<}(\lambda_j)$ is not in the column space of $\mathbf{A}^{<}(\lambda_j)$. Otherwise, the zero of the Evans function $D\left(\lambda\right)$ at $\lambda = \lambda_j$ is not simple, as we proceed to show. By contradiction, suppose that there exists a vector $\mathbf{x}$ such that

$$\mathbf{A}^{<}(\lambda_j)\mathbf{x} = \mathbf{b}^{<}(\lambda_j). \tag{A.29}$$

From formula (A.21) we may deduce that the derivative of the function $\det\left(\mathbf{A}^{<}(\lambda)\right)$ is given by

$$\frac{dD}{d\lambda}\left(\lambda\right) = \sum_{p=1}^{P}\left(\frac{i}{\epsilon\lambda}D_p\left(\lambda\right) - \frac{iz_p}{\epsilon}D\left(\lambda\right)\right). \tag{A.30}$$

At the point $\lambda = \lambda_j$, using (A.29) to rewrite $D_p(\lambda)$, we may obtain

$$\frac{dD}{d\lambda}(\lambda)\bigg|_{\lambda=\lambda_j} = D(\lambda_j)\sum_{p=1}^{P}\left(\frac{ix_p}{\epsilon\lambda_j} - \frac{iz_p}{\epsilon}\right) = 0 \tag{A.31}$$

from the fact that $D(\lambda_j) = 0$. This, however, is a contradiction with the fact that the zeros of $D(\lambda)$ are simple. So, $n_p(\lambda_j) \neq 0$ for at least some p. Lastly, $\sum_{p=1}^{P} n_p(\lambda_j) \neq 0$ follows from (A.30) and the fact that the zeros of $D(\lambda)$ are simple. ■

A.2 Elementary Examples of the Matrix N

Consider the positive rKS potential u_0 (cf. (2.67)) with $P = 3$ and data

$$(z_1, z_2, z_3) = \tfrac{1}{2}(2i, 2+4i, -2+i), \quad (c_1, c_2, c_3) = -\tfrac{1}{2}i(2, 1, 4), \tag{A.32}$$

and let $\lambda = -1$. While it is not obvious from the data (A.32), from a graph of u_0 one can see that it satisfies the Klaus-Shaw condition. The scheme for selection of appropriate contours $\{W_m\}_{m=1}^{3}$ in this case is illustrated in Figure A.1. Obviously W_j is homotopic on the domain of analyticity of f to $C_j = U_j^<$ (see Figure 2.3), so here $\mathbf{N} = \mathbb{I}$ and hence $\tilde{\mathbf{A}}(\lambda) = \mathbf{A}(\lambda)$ and $\tilde{D}(\lambda) = D(\lambda)$. Note that as desired, only $W_{P=3}$ traverses more than one critical point at the same level of $\mathrm{Re}\{-ih(z;-1)\}$.

To see that such a trivial outcome for $\mathbf{N}$ is not generally the case, consider the potential

$$(z_1, z_2, z_3) = \tfrac{1}{2}(-4+6i, 2i, 2+i), \quad (c_1, c_2, c_3) = -\tfrac{1}{3}i(3, 1, 3), \tag{A.33}$$

again yielding a positive rKS potential with $P = 3$. It follows from Figure A.2 that a possible matrix $\mathbf{N}$ is

$$\mathbf{N} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}. \tag{A.34}$$

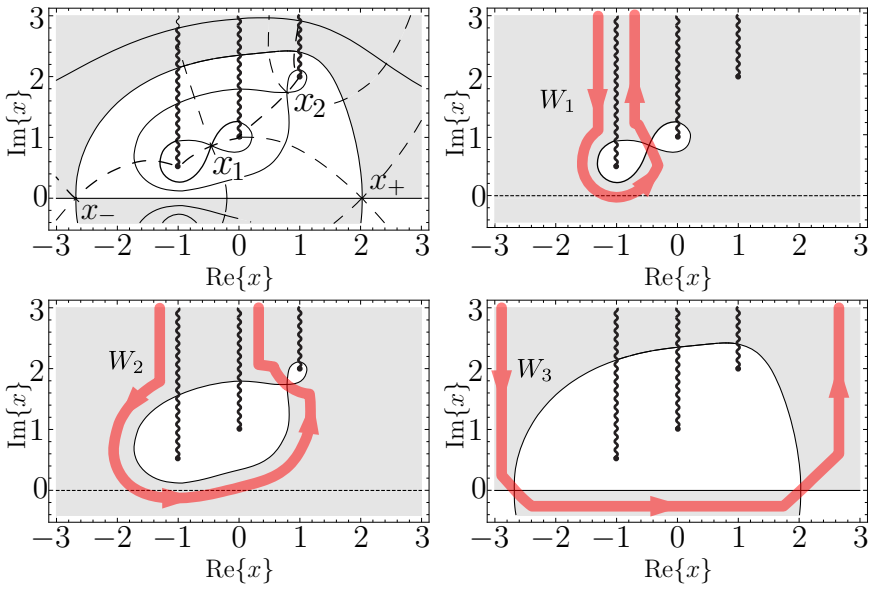

Figure A.1: Top left panel: for the $P = 3$ positive rKS potential u_0 with data given by (A.32), branch cuts of f emanating from the branch points $\{z_1, z_2, z_3\}$ are shown with zigzagged lines, the solid curves are the levels of $\mathrm{Re}\{-ih(x;-1)\}$, and the dashed curves are the levels of $\mathrm{Im}\{-ih(x;-1)\}$ (steepest descent/ascent). The intersection points mark the critical points of $h(x;-1)$ which are numbered according to the values $\mathrm{Re}\{-ih(x_j;-1)\}$ except for those $(x_\pm)$ on the real axis. The domain $\mathrm{Re}\{-ih(z;-1)\} < 0$ is shaded. Remaining panels: the contours $W_m \equiv U_m^< = C_m$ for which the integrals $\tilde{A}_{mp}(-1)$ are exponentially dominated by a contribution from neighborhoods of the critical point(s) over which the contour passes. The level curve $\mathrm{Re}\{-ih(z;-1)\} = \text{const.}$ containing the traversed critical point(s) is plotted, and the domain $\mathrm{Re}\{-ih(z;-1)\} < \text{const.}$ is shaded in each case.

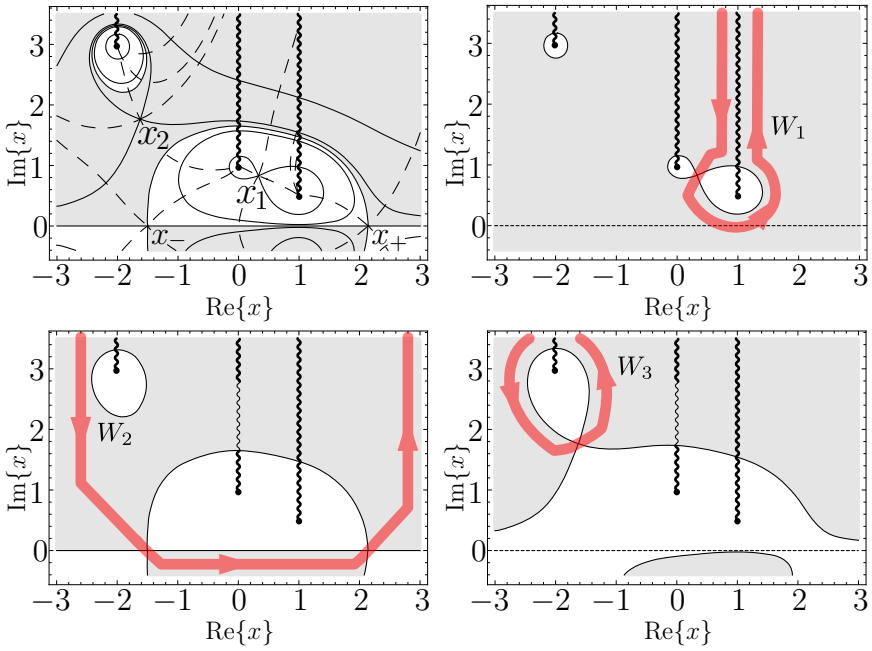

Figure A.2: Same as Figure A.1 but for the positive rKS potential with (A.33).

APPENDIX B

Tidal Model

B.1 Effects of Coriolis Force

To examine the effects of the Coriolis force on the system (3.1)–(3.4) we refer to the following two dimensional system:

$$\frac{\partial}{\partial t}(\eta_1 - \eta_2) + H_1 \left(\frac{\partial u_1}{\partial x} + \frac{\partial v_1}{\partial y} \right) = 0, \tag{B.1}$$

$$\frac{\partial \eta_2}{\partial t} + H_2 \left(\frac{\partial u_2}{\partial x} + \frac{\partial v_2}{\partial y} \right) = 0, \tag{B.2}$$

$$\frac{\partial u_1}{\partial t} - f v_1 + g \frac{\partial \eta_1}{\partial x} = 0, \tag{B.3}$$

$$\frac{\partial u_2}{\partial t} - f v_2 + (g - g') \frac{\partial \eta_1}{\partial x} + g' \frac{\partial \eta_2}{\partial x} = 0, \tag{B.4}$$

$$\frac{\partial v_1}{\partial t} + f u_1 + g \frac{\partial \eta_1}{\partial y} = 0, \tag{B.5}$$

$$\frac{\partial v_2}{\partial t} + f u_1 + (g - g') \frac{\partial \eta_1}{\partial y} + g' \frac{\partial \eta_2}{\partial y} = 0, \tag{B.6}$$

where x and y, respectively, denote east-west and north-south coordinates, u and v, respectively, denote zonal and meridional components of the velocity, f is the Coriolis parameter,

and for simplification we have neglected topography, astronomical forcing, and damping.

We may derive the dispersion relation of the system (B.1)–(B.5) by assuming harmonic solutions of the form $e^{i(kx+ly)-i\omega t}$. This yields the matrix equation

$$\begin{bmatrix} -i\omega & -f & ikg & 0 & 0 & 0 \\ f & -i\omega & ilg & 0 & 0 & 0 \\ ikH_1 & ilH_1 & -i\omega & 0 & 0 & i\omega \\ 0 & 0 & ik(g-g') & -i\omega & -f & ikg' \\ 0 & 0 & il(g-g') & f & -i\omega & ilg' \\ 0 & 0 & 0 & ikH_2 & ilH_2 & -i\omega \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ \eta_1 \\ u_2 \\ v_2 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{B.7}$$

The system (B.7) has a nontrivial solution only if the matrix has a zero determinant. If $\omega \neq 0$, computing the determinant yields the dispersion relation

$$\left(\frac{\omega^2 - f^2}{k^2 + l^2}\right)^2 - g(H_1 + H_2)\left(\frac{\omega^2 - f^2}{k^2 + l^2}\right) + gg'H_1H_2 = 0. \tag{B.8}$$

Remark B.1. A similar result to (B.8) can be seen, with some work, in the paper of Veronis and Stommel [82]: their equations 3.4 and 3.5.

Note that equation (B.8) is immediately comparable to equation (3.55) with the square of the phase speed translated by $f^2/(k^2 + l^2)$. In other words, the square of the phase speed for the two-dimensional system with the Coriolis effect present is, except for this translation, affected by stratification in a similar manner as the system without Coriolis effects. This suggests to us that the addition of the Coriolis force would not fundamentally change the effects studied in this dissertation. The fact that our analytical model without Coriolis effects behaves similarly as the global realistic-domain numerical model, which of course does include the Coriolis effects, lends further credence to this assumption of the relevance of our simplified analytical model.

B.2 Scattering Solution Method

A perturbation expansion with respect to the topographical parameter σ in equations (3.7)–(3.10) shows that the equations obtained at order σ for an infinite basin give the classical topographical scattering solution in terms of an incident velocity. That is, the order σ terms of the solution correspond to the scattering solution of the large-scale tide on the bottom roughness. The problem of internal wave generation from a bump that is impinged upon by a large-scale tidal velocity has received some attention in the oceanographic community; see for example [47], [78], [41]. We briefly derive the scattering solution in this subsection to highlight its similarities with our full solution.

Consider expanding all variables in the problem in terms of the topographical parameter σ, i.e., $\varphi(x) \sim \varphi_0(x) + \sigma\varphi_1(x) + \cdots$ for $\sigma \to 0$, where $\varphi(x)$ is the variable of interest. Gathering the order σ terms in equations (3.7)–(3.10), after using the separation (3.13), gives

$$i\left(N_1 - N_2\right) + \frac{dU_1}{dx} = 0, \tag{B.9}$$

$$iN_2 + \frac{dU_2}{dx} = F, \tag{B.10}$$

$$iU_1 + \gamma\frac{dN_1}{dx} + \delta_1 U_1 = 0, \tag{B.11}$$

$$iU_2 + (1-\gamma)\left((1-\alpha)\frac{dN_1}{dx} + \alpha\frac{dN_2}{dx}\right) + \delta_2 U_2 = 0, \tag{B.12}$$

where $F = \frac{\partial}{\partial x}\left(U_{2,0}b\right)$ with $U_{2,0}$ representing the incident tidal velocity on the topography or equivalently the $O(1)$ term in the expansion of U_2. For the scattering problem, we assume that this background state $U_{2,0}$ is known and focus on solving for the higher order terms. Note that all variables in this equation correspond to the order σ terms, but we have dropped any indicative subscripts to avoid cumbersome notation. That is, in this appendix the notation N_1 really stands for $N_{1,1}$ (the order σ term in the expansion of N_1).

Since the system obtained is linear and has constant coefficients, we may readily solve the system above in the infinite basin case using the Fourier transform, e.g., $\hat{u}(k) = \int_{-\infty}^{\infty} u(x)e^{-ikx}\,dx$.

Manipulating the equations yields

$$\left(s_1 - k^2\right)\hat{N}_1 = s_1\hat{N}_2, \tag{B.13}$$

$$\left(s_2 - \alpha k^2\right)\hat{N}_2 - (1-\alpha)k^2\hat{N}_1 = -is_2\hat{F}, \tag{B.14}$$

where s_1 and s_2 are as introduced previously. These equations may be combined to obtain

$$\left(k^4 - \left(\frac{s_1+s_2}{\alpha}\right)k^2 + \frac{s_1 s_2}{\alpha}\right)\hat{N}_1 = -i\frac{s_1 s_2}{\alpha}\hat{F}. \tag{B.15}$$

It is worth noting that the left-hand side of this equation is (3.37), making the separation of scales derived here identical to that in Section 3.3.3. This allows us to write equation (B.15) as

$$\hat{N}_1 = -i(k^L)^2(k^S)^2\left(\frac{1}{(k-k^L)(k+k^L)(k-k^S)(k+k^S)}\right)\hat{F} = \hat{R}_1\hat{F}, \tag{B.16}$$

where k^L and k^S are the roots of the polynomial (3.37) as defined previously; the left hand side polynomial in (B.15). Therefore, we obtain a closed form solution for $N_1(x)$ — the order σ term in the surface elevation — by applying the inverse Fourier transform to (B.16). The problem of inverting $\hat{N}_1(x)$ in equation (B.16) to obtain $N_1(x)$ is equivalent to taking a convolution of the functions $R_1(x)$ and $F(x)$. For this reason we seek a closed form solution of $R_1(x)$ because $F(x)$ is the known forcing function.

The integral $R_1(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{R}_1(k)e^{ikx}\,dk$ is readily done by summing residues separately for $x \geq 0$ and for $x < 0$ because k^L and k^S lie in the upper complex plane. In the case of $x \geq 0$ we close the contour of integration around the upper half plane leading to a positively oriented curve and in the case $x < 0$ we close the contour of integration around the lower half plane leading to a negatively oriented curve. After some manipulation we find $R_1(x) = R_1^L(x) + R_1^S(x)$, where

$$R_1^{L,S}(x) = \frac{e^{ik^{L,S}|x|}}{2k^{L,S}}\left(\frac{1}{(k^{S,L})^2} - \frac{1}{(k^{L,S})^2}\right)^{-1} \tag{B.17}$$

with the equation read only with the first or the second superscript in mind. This separation of R_1 leads to a very natural way of writing the upper layer solution in terms of its large- and small-scale components. Explicitly, we write

$$N_1^{L,S}(x) = \int_{-\infty}^{\infty} R_1^{L,S}(x-y)F(y)\,dy \tag{B.18}$$

for the order σ large- and small-scale components of the surface elevation, respectively.

APPENDIX C

Quasi-Geostrophy

C.1 Properties of the Jacobian

The Jacobian operator J of two differentiable functions $f = f(x, y)$ and $g = g(x, y)$ is given by

$$J(f, g) := \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial g}{\partial x}\frac{\partial f}{\partial y} \tag{C.1}$$

is bilinear, i.e., linear in each entry. It is also known as the Jacobian determinant as it arises from the determinant of the Jacobian matrix:

$$J(f, g) = \det \begin{bmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{bmatrix}. \tag{C.2}$$

Proposition C.1. Suppose that the functions $\varphi_i = \varphi_i(x, y)$ for $i = 1, 2, 3$ are doubly periodic and possess continuous second partial derivatives in the rectangular domain $\mathcal{D}$, then

$$\int_{\mathcal{D}} \varphi_1 J(\varphi_2, \varphi_3)\, dA = \int_{\mathcal{D}} \varphi_3 J(\varphi_1, \varphi_2)\, dA. \tag{C.3}$$

Proof. The assertion follows directly from the definition of the Jacobian operator (C.1) and integration by parts. Namely,

$$\begin{aligned}\int_{\mathcal{D}} \varphi_1 J(\varphi_2,\varphi_3)\,dA &= \int_{\mathcal{D}} \varphi_1 \left(\frac{\partial\varphi_2}{\partial x}\frac{\partial\varphi_3}{\partial y} - \frac{\partial\varphi_2}{\partial y}\frac{\partial\varphi_3}{\partial x}\right) dA \\ &= \int_{\mathcal{D}} \varphi_3 \left(\frac{\partial}{\partial x}\left(\varphi_1\frac{\partial\varphi_2}{\partial y}\right) - \frac{\partial}{\partial y}\left(\varphi_1\frac{\partial\varphi_2}{\partial x}\right)\right) dA = \int_{\mathcal{D}} \varphi_3 J(\varphi_1,\varphi_2)\,dA,\end{aligned} \tag{C.4}$$

as stated by means of Clairaut's Theorem for equality of mixed partial derivatives. ■

In particular, from Proposition C.1 and the definition of Jacobian operator (C.1) the corollary below easily follows.

Corollary C.2. Suppose φ_i for $i = 1, 2, 3$ are doubly periodic and possess continuous second partial derivatives in the rectangular domain $\mathcal{D}$ and let f be differentiable, then

$$\int_{\mathcal{D}} f(\varphi_1) J(\varphi_2,\varphi_3)\,dA = \int_{\mathcal{D}} \varphi_3 f'(\varphi_1) J(\varphi_1,\varphi_2)\,dA. \tag{C.5}$$

Note that as special cases of Corollary C.2 we have

$$\int_{\mathcal{D}} f(\psi) J(\psi,\varphi)\,dA = 0 \quad \text{and} \quad \int_{\mathcal{D}} J(\psi,\varphi)\,dA = 0. \tag{C.6}$$

We may additionally define the Jacobian operator of two vector valued functions $\boldsymbol{\varphi} = [\varphi_1 \cdots \varphi_n]^\mathsf{T}$ and $\boldsymbol{\psi} = [\psi_1 \cdots \psi_n]^\mathsf{T}$ as follows:

$$J(\boldsymbol{\varphi},\boldsymbol{\psi}) = \operatorname{diag}\left(\frac{\partial\boldsymbol{\varphi}}{\partial x}\right)\frac{\partial\boldsymbol{\psi}}{\partial y} - \operatorname{diag}\left(\frac{\partial\boldsymbol{\psi}}{\partial x}\right)\frac{\partial\boldsymbol{\varphi}}{\partial y} = \begin{bmatrix} J(\varphi_1,\psi_1) \\ \vdots \\ J(\varphi_n,\psi_n)\end{bmatrix}. \tag{C.7}$$

Corollary C.3. Suppose the vector valued functions $\boldsymbol{\varphi}_i$ for $i = 1, 2, 3$ are doubly periodic and possess continuous second partial derivatives in the rectangular domain $\mathcal{D}$ and let f be

differentiable, then

$$\int_{\mathcal{D}} \boldsymbol{\varphi}_1^{\mathsf{T}} J(\boldsymbol{\varphi}_2, \boldsymbol{\varphi}_3)\, dA = \int_{\mathcal{D}} \boldsymbol{\varphi}_3^{\mathsf{T}} J(\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2)\, dA. \tag{C.8}$$

C.2 Proofs

In this section provide certain proofs that have been neglected from the body of the dissertation. We restate the propositions to be proven for ease of reference.

Proposition IV.6. Let $\mathbf{B} = \operatorname{diag}\left(\frac{\delta_1}{2L_0}, -\delta_1, \ldots, -\delta_m, -\frac{\delta_m}{2L_m}\right)$. The matrix $\mathbf{BA}_\kappa$ is positive definite for $\kappa \neq 0$ and the matrix $\mathbf{BA}_0$ is positive semidefinite.

Proof of Proposition IV.6. Note that for a hermitian matrix $\mathbf{A}$, $\mathbf{A}$ is positive definite if and only if $\det \mathbf{A}^i > 0$ for $i = 0, \ldots m+1$, where $\mathbf{A}^i$ denotes the principal minor; see [34, Theorem 7.2.5]. Now define the $i+2 \times i+2$ matrix

$$\mathbf{C}_\kappa^{i+1} := \begin{bmatrix} \delta_0\kappa^2 + L_0^{-1} + L_{-1}^{-1} & -L_0^{-1} & \cdots & 0 \\ -L_0^{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \delta_i\kappa^2 + L_i^{-1} + L_{i-1}^{-1} & -L_i^{-1} \\ 0 & \ldots & -L_i^{-1} & \delta_{i+1}\kappa^2 + L_{i+1}^{-1} + L_i^{-1} \end{bmatrix} \tag{C.9}$$

Expanding the determinant of $\mathbf{C}_\kappa^{i+1}$ from the bottom right corner we have the formula

$$\det \mathbf{C}_\kappa^i = \left(\delta_i\kappa^2 + L_i^{-1} + L_{i-1}^{-1}\right) \det \mathbf{C}_\kappa^{i-1} - L_{i-1}^{-2} \det \mathbf{C}_\kappa^{i-2}. \tag{C.10}$$

which is valid for $i = 0, \ldots, m+1$ as long as we define $\det \mathbf{C}_\kappa^{-2} := 0$ and $\det \mathbf{C}_\kappa^{-1} := 1$. Now, it follows easily that

$$\det \mathbf{C}_\kappa^0 = \delta_0\kappa^2 + L_0^{-1} + L_{-1}^{-1} > \delta_0\kappa^2 + L_0^{-1} > 0 \tag{C.11}$$

and

$$\begin{aligned}\det \mathbf{C}_\kappa^1 &= \det \begin{bmatrix} \delta_0\kappa^2 + L_0^{-1} + L_{-1}^{-1} & -L_0^{-1} \\ -L_0^{-1} & \delta_1\kappa^2 + L_1^{-1} + L_0^{-1} \end{bmatrix} \\ &= \left(\delta_0\kappa^2 + L_0^{-1} + L_{-1}^{-1}\right)\left(\delta_1\kappa^2 + L_1^{-1} + L_0^{-1}\right) - L_0^{-2} \\ &> \det \mathbf{C}_\kappa^0 \left(\delta_1\kappa^2 + L_1^{-1}\right) > 0.\end{aligned} \tag{C.12}$$

To proceed by induction, suppose that for i we have

$$\det \mathbf{C}_\kappa^i > \det \mathbf{C}_\kappa^{i-1}\left(\delta_i\kappa^2 + L_i^{-1}\right) > 0 \quad \text{and} \quad \det \mathbf{C}_\kappa^{i-1} > \det \mathbf{C}_\kappa^{i-2}\left(\delta_{i-1}\kappa^2 + L_{i-1}^{-1}\right) > 0. \tag{C.13}$$

Then, by (C.10) at $i+1$ we have

$$\begin{aligned}\det \mathbf{C}_\kappa^{i+1} &= \left(\delta_{i+1}\kappa^2 + L_{i+1}^{-1} + L_i^{-1}\right)\det \mathbf{C}_\kappa^i - L_i^{-2}\det \mathbf{C}_\kappa^{i-1} \\ &= \left(\delta_{i+1}\kappa^2 + L_{i+1}^{-1}\right)\det \mathbf{C}_\kappa^i + L_i^{-1}\det \mathbf{C}_\kappa^i - L_i^{-2}\det \mathbf{C}_\kappa^{i-1} \\ &> \left(\delta_{i+1}\kappa^2 + L_{i+1}^{-1}\right)\det \mathbf{C}_\kappa^i + \left(L_i^{-1}\left(\delta_i\kappa^2 + L_i^{-1}\right) - L_i^{-2}\right)\det \mathbf{C}_\kappa^{i-1} \\ &> \left(\delta_{i+1}\kappa^2 + L_{i+1}^{-1}\right)\det \mathbf{C}_\kappa^i > 0.\end{aligned} \tag{C.14}$$

By induction, $\det \mathbf{C}_\kappa^{i+1}$ (principal minors of $\mathbf{C}_\kappa^{m+1}$) is positive for $i = 0, \ldots, m+1$. If we let $\delta_0 = \delta_{m+1} = L_{-1}^{-1} = L_{m+1}^{-1} = 0$ we note that $\mathbf{C}_\kappa^{m+1} = \mathbf{B}\mathbf{A}_\kappa$, which finishes the proof for $\kappa \neq 0$. Lastly, by continuity with respect to κ as $\kappa \to 0$ and Lemma IV.2 it follows that $\mathbf{B}\mathbf{A}_0$ must be positive semidefinite (at least one eigenvalue is equal to zero). ■

Proposition IV.10. Suppose $\mathbf{M}$ is a positive definite real symmetric matrix and $\mathbf{N}$ is real symmetric matrix, then there exists an invertible real matrix $\mathbf{V}$ that simultaneously diagonalizes $\mathbf{M}$ and $\mathbf{N}$ in the form

$$\mathbf{V}^\mathsf{T}\mathbf{M}\mathbf{V} = \mathbf{\Lambda} \tag{C.15}$$

and

$$\mathbf{V}^{\mathsf{T}}\mathbf{N}\mathbf{V} = \mathbf{I}(\mathbf{N}), \tag{C.16}$$

where $\mathbf{\Lambda}$ is diagonal with nonnegative entries and $\mathbf{I}(\mathbf{N})$ denotes the inertia matrix of $\mathbf{N}$.

Proof of Proposition IV.10. This proof essentially follows the proof of Theorem 7.6.4 of Horn and Johnson [34]. Since $\mathbf{M}$ is positive definite with real entries, there exists an invertible matrix $\mathbf{C}$ with real entries such that $\mathbf{M} = \mathbf{C}^{-\mathsf{T}}\mathbf{C}^{-1}$ (Cholesky decomposition). Now, the matrix $\mathbf{C}^{\mathsf{T}}\mathbf{N}\mathbf{C}$, which is symmetric, can be diagonalized in the form $\mathbf{C}^{\mathsf{T}}\mathbf{N}\mathbf{C} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathsf{T}}$, where $\mathbf{D}$ is diagonal and $\mathbf{U}$ is real orthogonal. In particular, by Theorem 4.5.8 in Horn and Johnson we know that $\mathbf{N}$ and $\mathbf{D}$ must have same inertia (number of positive and negative eigenvalues). Therefore, there exists a decomposition for $\mathbf{D}$ of the form $\mathbf{D} = \mathbf{S}\mathbf{I}(\mathbf{N})\mathbf{S}$ where $\mathbf{S}$ is the matrix given by $\mathbf{S} = \operatorname{diag}(\sqrt{|s_1|}, \ldots, \sqrt{|s_m|})$, where for each entry we let $s_i = d_i$ if $d_i \neq 0$ and $s_i = 1$ otherwise (here d_i denote the entries of $\mathbf{D}$). So, define the matrix with real entries $\mathbf{V} = \mathbf{C}\mathbf{U}\mathbf{S}^{-1}$, then

$$\mathbf{V}^{\mathsf{T}}\mathbf{N}\mathbf{V} = \mathbf{S}^{-1}\mathbf{U}^{\mathsf{T}}\mathbf{C}^{\mathsf{T}}\mathbf{N}\mathbf{C}\mathbf{U}\mathbf{S}^{-1} = \mathbf{S}^{-1}\mathbf{U}^{\mathsf{T}}\mathbf{U}\mathbf{D}\mathbf{U}^{\mathsf{T}}\mathbf{U}\mathbf{S}^{-1} = \mathbf{I}(\mathbf{N}) \tag{C.17}$$

and

$$\mathbf{V}^{\mathsf{T}}\mathbf{M}\mathbf{V} = \mathbf{S}^{-1}\mathbf{U}^{\mathsf{T}}\mathbf{C}^{\mathsf{T}}\mathbf{C}^{-\mathsf{T}}\mathbf{C}^{-1}\mathbf{C}\mathbf{U}\mathbf{S}^{-1} = \mathbf{S}^{-1}\mathbf{S}^{-1}. \tag{C.18}$$

The proof is finished by letting $\mathbf{\Lambda} := \mathbf{S}^{-2}$. Note that the sign of the entries in $\mathbf{\Lambda}$ is nonnegative by definition. ■

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] M. J. Ablowitz, A. S. Fokas, and R. L. Anderson, *The direct linearizing transform and the Benjamin-Ono equation*, Phys. Lett. A **93** (1983), no. 8, 375–378, doi:10.1016/0375-9601(83)90463-2.

[2] R. L. Anderson and E. Taflin, *The Benjamin-Ono equation — Recursivity of linearization maps — Lax pairs*, Lett. Math. Phys. **9** (1985), no. 4, 299–311, doi:10.1007/BF00397756.

[3] B. K. Arbic and G. R. Flierl, *Baroclinically unstable geostrophic turbulence in the limits of strong and weak bottom Ekman friction: Application to midocean eddies*, J. Phys. Ocean. **34** (2004), 2257–2273, doi:10.1175/1520-0485(2004)034<2257:BUGTIT>2.0.CO;2.

[4] B. K. Arbic, G. R. Flierl, and R. B. Scott, *Cascade inequalities for forced-dissipated geostrophic turbulence*, J. Phys. Ocean. **37** (2007), 1470–1487, doi:10.1175/JPO3067.1.

[5] B. K. Arbic, S. T. Garner, R. W. Hallberg, and H. L. Simmons, *The accuracy of surface elevations in forward global barotropic and baroclinic tide models*, Deep-Sea Res. II **51** (2004), no. 2526, 3069–3101, doi:10.1016/j.dsr2.2004.09.014.

[6] B. K. Arbic, R. H. Karsten, and C. Garrett, *On tidal resonance in the global ocean and the back-effect of coastal tides upon open-ocean tides*, Atmos.–Ocean **47** (2009), no. 4, 239–266, doi:10.3137/OC311.2009.

[7] B. K. Arbic and W. B. Owens, *Climatic warming of Atlantic intermediate waters*, J. Climate **14** (2001), 4091–4108, doi:10.1175/1520-0442(2001)014<4091:CWOAIW>2.0.CO;2.

[8] B. K. Arbic, A. J. Wallcraft, and E. J. Metzger, *Concurrent simulation of the eddying general circulation and tides in a global ocean model*, Ocean Modelling **32** (2010), no. 3-4, 175–187, doi:10.1016/j.ocemod.2010.01.007.

[9] T. B. Benjamin, *Internal waves of permanent form in fluids of great depth*, J. Fluid Mech. **29** (1967), 559–592, doi:10.1017/S002211206700103X.

[10] W. Blumen, *On the stability of quasi-geostrophic flow*, J. Atmos. Sci. **25** (1968), 929–931, doi:10.1175/1520-0469(1968)025<0929:OTSOQG>2.0.CO;2.

[11] _____, *Uniform potential vorticity flow: Part I. theory of wave interactions and two-dimensional turbulence*, J. Atmos. Sci. **35** (1978), 774–783, doi:10.1175/1520-0469(1978)035<0774:UPVFPI>2.0.CO;2.

[12] T. L. Bock and M. D. Kruskal, *A two-parameter Miura transformation of the Benjamin-Ono equation*, Phys. Lett. A **74** (1979), no. 3-4, 173–176, doi:10.1016/0375-9601(79)90762-X.

[13] D. E. Cartwright, *Tides and waves in the vicinity of Saint Helena*, Phil. Trans. R. Soc. Lond. A **270** (1971), no. 1210, 603–646, doi:10.1098/rsta.1971.0091.

[14] K. M. Case, *Benjamin-Ono-related equations and their solutions*, Proc. Natl. Acad. Sci. (USA) **76** (1979), no. 1, 1–3, http://www.jstor.org/stable/69418.

[15] W. Choi and R. Camassa, *Weakly nonlinear internal waves in a two-fluid system*, J. Fluid Mech. **313** (1996), 83–103, doi:10.1017/S0022112096002133.

[16] J. A. Colosi and W. Munk, *Tales of the venerable Honolulu tide gauge*, J. Phys. Ocean. **36** (2006), 967–996, doi:10.1175/JPO2876.1.

[17] R. E. Davis and A. Acrivos, *Solitary internal waves in deep water*, J. Fluid Mech. **29** (1967), 593–607, doi:10.1017/S0022112067001041.

[18] A. Defant, *Physical oceanography: Vol. 2.*, Pergamon Press, 1961.

[19] S. Y. Dobrokhotov and I. M. Krichever, *Multi-phase solutions of the Benjamin-Ono equation and their averaging*, Mathematical notes of the Academy of Sciences of the USSR **49** (1991), no. 6, 583–594, doi:10.1007/BF01156581.

[20] G. D. Egbert, R. D. Ray, and B. G. Bills, *Numerical modeling of the global semidiurnal tide in the present day and in the last glacial maximum*, J. Geophys. Res. **109** (2004), no. C3, doi:10.1029/2003JC001973.

[21] R. E. Flick, J. F. Murray, and L. C. Ewing, *Trends in United States tidal datum statistics and tide range*, J. Waterway, Port, Coastal and Ocean Engin. **129** (2003), no. 4, 155–164, doi:10.1061/(ASCE)0733-950X(2003)129:4(155).

[22] G. R. Flierl, *Models of vertical structure and the calibration of two-layer models*, Dyn. Atmos. Oceans **2** (1978), no. 4, 341–381, doi:10.1016/0377-0265(78)90002-7.

[23] _____, *Vertical structure*, personal communication, 2012.

[24] A. S. Fokas and M. J. Ablowitz, *The inverse scattering transform for the Benjamin-Ono equation: A pivot to multidimensional problems*, Stud. Appl. Math. **68** (1983), 1–10.

[25] A. S. Fokas and B. Fuchssteiner, *The hierarchy of the Benjamin-Ono equation*, Phys. Lett. A **86** (1981), no. 6-7, 341–345, doi:10.1016/0375-9601(81)90551-X.

[26] F. D. Gakhov, *Boundary value problems*, Dover Publications, 1990.

[27] J. Garnier, G. Xu, S. Trillo, and A. Picozzi, *Incoherent dispersive shocks in the spectral evolution of random waves*, Phys. Rev. Lett. **111** (2013), 113902, doi:10.1103/PhysRevLett.111.113902.

[28] A. E. Gill, *Atmosphere-ocean dynamics*, Academic Press, 1982.

[29] I. Gohberg and L. Rodman, *Analytic matrix functions with prescribed local data*, J. Anal. Math. **40** (1981), no. 1, 90–128, doi:10.1007/BF02790157.

[30] M. D. Greenberg, *Application of Green's functions in science and engineering*, Prentice-Hall Inc., 1971.

[31] R. Hallberg and P. Rhines, *Boundary-driven circulation in an ocean basin with isopycnals intersecting the sloping boundary*, J. Phys. Ocean. **26** (1996), 913–940, doi:10.1175/1520-0485(1996)026<0913:BDCIAO>2.0.CO;2.

[32] I. M. Held, R. T. Pierrehumbert, S. T. Garner, and K. L. Swanson, *Surface quasi-geostrophic dynamics*, J. Fluid Mech. **282** (1995), 1–20, doi:10.1017/S0022112095000012.

[33] M. C. Hendershott, *The effects of solid earth deformation on global ocean tides*, Geophys. J. Int. **29** (1972), no. 4, 389–402, doi:10.1111/j.1365-246X.1972.tb06167.x.

[34] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 1985.

[35] D. A. Jay, *Evolution of tidal amplitudes in the Eastern Pacific Ocean*, Geophys. Res. Lett. **36** (2009), no. 4, L04603, doi:10.1029/2008GL036185.

[36] S. R. Jayne and L. C. St. Laurent, *Parametrizing tidal dissipation over rough topography*, Geophys. Res. Lett. **28** (2001), no. 5, 811–814, doi:10.1029/2000GL012044.

[37] M. C. Jorge, A. A. Minzoni, and N. F. Smyth, *Modulation solutions for the Benjamin-Ono equation*, Physica D: Nonlinear Phenomena **132** (1999), no. 1-2, 1–18, doi:10.1016/S0167-2789(99)00039-1.

[38] M. Juckes, *Quasigeostrophic dynamics of the tropopause*, J. Atmos. Sci. **51** (1994), 2756–2768, doi:10.1175/1520-0469(1994)051<2756:QDOTT>2.0.CO;2.

[39] S. K. Kang, M. G. G. Foreman, H.-J. Lie, J.-H. Lee, J. Cherniawsky, and K.-D. Yum, *Two-layer tide modeling of the Yellow and East China seas with application to seasonal variability of the M_2 tide*, J. Geophys. Res. **107** (2002), no. C3, 6–1–6–18, doi:10.1029/2001JC000838.

[40] D. J. Kaup and Y. Matsuno, *The inverse scattering transform for the Benjamin-Ono equation*, Stud. Appl. Math. **101** (1998), no. 1, 73–98, doi:10.1111/1467-9590.00086.

[41] S. Khatiwala, *Generation of internal tides in an ocean of finite depth: analytical and numerical calculations*, Dee-Sea Res. I **50** (2003), no. 1, 3–21, doi:10.1016/S0967-0637(02)00132-2.

[42] M. Klaus and J. K. Shaw, *Purely imaginary eigenvalues of Zakharov-Shabat systems*, Phys. Rev. E **65** (2002), 036607, doi:10.1103/PhysRevE.65.036607.

[43] Y. Kodama, M. J. Ablowitz, and J. Satsuma, *Direct and inverse scattering problems of the nonlinear intermediate long wave equation*, J. Math. Phys. **23** (1982), no. 4, 564–576, doi:10.1063/1.525393.

[44] G. Lapeyre, *What vertical mode does the altimeter reflect? On the decomposition in baroclinic modes and on a surface-trapped mode*, J. Phys. Ocean. **39** (2009), 2857–2874, doi:10.1175/2009JPO3968.1.

[45] G. Lapeyre and P. Klein, *Dynamics of the upper oceanic layers in terms of surface quasigeostrphy theory*, J. Phys. Ocean. **36** (2006), 165–176, doi:10.1175/JPO2840.1.

[46] P. D. Lax and C. D. Levermore, *The small dispersion limit of the Korteweg-de Vries equation. I, II, III*, Comm. Pure Appl. Math. **36** (1983), no. 3, 5, 6, 253–290, 571–593, 809–829, doi:10.1002/cpa.3160360302, 10.1002/cpa.3160360503, 10.1002/cpa.3160360606.

[47] S. G. Llewellyn Smith and W. R. Young, *Conversion of the barotropic tide*, J. Phys. Ocean. **32** (2001), 1554–1566, doi:10.1175/1520-0485(2002)032<1554:COTBT>2.0.CO;2.

[48] Y. Matsuno, *Exact multi-soliton solution of the Benjamin-Ono equation*, J. Phys. A: Math. Gen. **12** (1979), no. 4, 619, doi:10.1088/0305-4470/12/4/019.

[49] ———, *Number density function of Benjamin-Ono solitons*, Phys. Lett. A **87** (1981), no. 1, 15–17, doi:10.1016/0375-9601(81)90603-4.

[50] ———, *Asymptotic properties of the Benjamin-Ono equation*, J. Phys. Soc. Jpn. **51** (1982), no. 2.

[51] ———, *Nonlinear modulation of periodic waves in the small dispersion limit of the Benjamin-Ono equation*, Phys. Rev. E **58** (1998), 7934–7940, doi:10.1103/PhysRevE.58.7934.

[52] ———, *The small dispersion limit of the Benjamin-Ono equation and the evolution of a step initial condition*, J. Phys. Soc. Jpn. **67** (1998), no. 6, 1814–1817, doi:10.1143/JPSJ.67.1814.

[53] P. D. Miller and A. N. Wetzel, *Direct scattering for the Benjamin-Ono equation with rational initial data*, submitted. arXiv:1504.07610, 2015.

[54] ———, *The scattering transform for the Benjamin-Ono equation in the small-dispersion limit*, (in press). arXiv:1505.02700, 10.1016/j.physd.2015.07.012, 2015.

[55] P. D. Miller and Z. Xu, *On the zero-dispersion limit of the Benjamin-Ono Cauchy problem for positive initial data*, Comm. Pure Appl. Math. **64** (2011), no. 2, 205–270, doi:10.1002/cpa.20345.

[56] M. Müller, *The influence of changing stratification conditions on barotropic tidal transport and its implications for seasonal and secular changes of tides*, Cont. Shelf Res. **47** (2012), 107–118, doi:10.1016/j.csr.2012.07.003.

[57] M. Müller, B. K. Arbic, and J. X. Mitrovica, *Secular changes in ocean tides: Observations and model results*, J. Geophys. Res. **116** (2011), no. C5, doi:10.1029/2010JC006387, C05013.

[58] M. Müller, J. Y. Cherniawsky, M. G. G. Foreman, and J.-S. von Storch, *Global M_2 internal tide and its seasonal variability from high resolution ocean circulation and tide modeling*, Geophys. Res. Lett. **39** (2012), no. 19, doi:10.1029/2012GL053320, L19607.

[59] P. Müller, *The equations of oceanic motions*, Cambridge University Press, 2006.

[60] W. Munk, *Once again: Once again — tidal friction*, Prog. Oceanog. **40** (1997), no. 1-4, 7–35, doi:10.1016/S0079-6611(97)00021-9.

[61] N. I. Muskhelishvili, *Singular integral equations: boundary problems of function theory and their application to mathematical physics*, Dover Publications, 2008.

[62] A. Nakamura, *Bäcklund transform and conservation laws of the Benjamin-Ono equation*, J. Phys. Soc. Jpn. **47** (1979), no. 4, 1335–1340, doi:10.1143/JPSJ.47.1335.

[63] H. Ono, *Algebraic solitary waves in stratified fluids*, J. Phys. Soc. Jpn. **39** (1975), no. 4, 1082–1091, doi:10.1143/JPSJ.39.1082.

[64] ———, *Algebraic rossby wave soliton*, J. Phys. Soc. Jpn. **50** (1981), no. 8, 2757–2761, doi:10.1143/JPSJ.50.2757.

[65] A. Porter and N. F. Smyth, *Modelling the morning glory of the Gulf of Carpentaria*, J. Fluid Mech. **454** (2002), 1–20, doi:10.1017/S0022112001007455.

[66] J. Proudman, *Dynamical oceanography*, London: Methuen & Co. Ltd., 1953.

[67] R. D. Ray, *A global ocean tide model from TOPEX/POSEIDON altimetry: GOT99.2, National Aeronautics and Space Administration technical memorandum*, Tech. report, 1999, NASA/TM-1999-209478.

[68] ———, *Secular changes of the M_2 tide in the Gulf of Maine*, Cont. Shelf Res. **26** (2006), no. 3, 422–427, doi:10.1016/j.csr.2005.12.005.

[69] ———, *Secular changes in the solar semidiurnal tide of the western North Atlantic Ocean*, Geophys. Res. Lett. **36** (2009), no. 19, doi:10.1029/2009GL040217, L19601.

[70] R. D. Ray and G. T. Mitchum, *Surface manifestation of internal tides generated near Hawaii*, Geophys. Res. Lett. **23** (1996), no. 16, 2101–2104, doi:10.1029/96GL02050.

[71] ———, *Surface manifestation of internal tides in the deep ocean: observations from altimetry and island gauges*, Prog. Oceanog. **40** (1997), no. 1-4, 135–162, doi:10.1016/S0079-6611(97)00025-6.

[72] R. Salmon, *Two-layer quasi-geostrophic turbulence in a simple special case*, Geophys. Astrophys. Fluid Dyn. **10** (1978), no. 1, 25–52, doi:10.1080/03091927808242628.

[73] P. Santini, M. J. Ablowitz, and A. S. Fokas, *On the limit from the intermediate long-wave equation to the Benjamin-Ono equation*, J. Math. Phys. **25** (1984), 892–899, doi:10.1063/1.526243.

[74] S. Schechter, *On the inversion of certain matrices*, Math. Tables Other Aids Comput. **13** (1959), no. 66, 73–77, doi:10.2307/2001955.

[75] J. F. Shriver, J. G. Richman, and B. K. Arbic, *How stationary are the internal tides in a high-resolution global ocean circulation model?*, J. Geophys. Res. **119** (2014), no. 5, 2769–2787, doi:10.1002/2013JC009423.

[76] K. S. Smith and G. K. Vallis, *The scales and equilibration of midocean eddies: freely evolving flow*, J. Phys. Ocean. **31** (2001), 554–571, doi:10.1175/1520-0485(2001)031<0554:TSAEOM>2.0.CO;2.

[77] K. S. Smith and J. Vanneste, *A surface-aware projection basis for quasigrostrophic flow*, J. Phys. Ocean. **43** (2013), 548–562, doi:10.1175/JPO-D-12-0107.1.

[78] L. C. St. Laurent, S. Stringer, C. Garrett, and D. Perrault-Joncas, *The generation of internal tides at abrupt topography*, Deep-Sea Res. I **50** (2003), no. 8, 987–1003, doi:10.1016/S0967-0637(03)00096-7.

[79] T. Tao, *Global well-posedness of the Benjamin-Ono equation in* $H^1(\mathbb{R})$, J. Hyp. Diff. Eq. **01** (2004), no. 01, 27–49, doi:10.1142/S0219891604000032.

[80] R. Tulloch and K. S. Smith, *Quasigeostrophic turbulence with explicit surface dynamics: application to the atmospheric energy spectrum*, J. Atmos. Sci. **66** (2009), 450–467, doi:10.1175/2008JAS2653.1.

[81] G. K. Vallis, *Atmospheric and oceanic fluid dyanmics: Fundamentals of large-scale circulation*, Cambridge University Press, Cambridge, 2006.

[82] G. Veronis and H. Stommel, *The action of variable wind stresses on a stratified ocean*, J. Mar. Res. **15** (1956), 43–75.

[83] A. N. Wetzel, B. K. Arbic, I. Cerovecki, M. C. Hendershott, R. H. Karsten, P. D. Miller, and J. F. Molinari, *On stratification, large-scale tides, and temporal changes in surface tidal elevations: Two-layer analytical model*, arXiv:1311.6349, 2015.

[84] G. B. Whitham, *Non-linear dispersive waves*, Proc. R. Soc. A **283** (1965), no. 1393, 238–261, doi:10.1098/rspa.1965.0019.

[85] ———, *Linear and nonlinear waves*, Wiley-Interscience, New York, 1974.

[86] P. L. Woodworth, *A survey of recent changes in the main components of the ocean tide*, Cont. Shelf Res. **30** (2010), no. 15, 1680–1691, doi:10.1016/j.csr.2010.07.002.

[87] P. L. Woodworth, S. M. Shaw, and D. L. Blackman, *Secular trends in mean tidal range around the British Isles and along the adjacent European coastline*, Geophys. J. Int. **104** (1991), no. 3, 593–609, doi:10.1111/j.1365-246X.1991.tb05704.x.

[88] Y. Wu, *Simplicity and finiteness of discrete spectrum of the Benjamin-Ono scattering operator*, arXiv:1506.06110, 2015.

[89] C. Wunsch, *Internal tides in the ocean*, Rev. Geosphys. **13** (1975), no. 1, 167–182, doi:10.1029/RG013i001p00167.

[90] Z. Xu, *Asymptotic analysis and numerical analysis of the Benjamin-Ono equation*, Ph.D. thesis, The University of Michigan, 2010, http://hdl.handle.net/2027.42/78806.

[91] E. D. Zaron and D. A. Jay, *An analysis of secular change in tides at open-ocean sites in the Pacific*, J. Phys. Ocean. **44** (2014), 1704–1726, doi:10.1175/JPO-D-13-0266.1.