Some Nonlinear Equations Arising in the Theory of Water Waves

by

Jianke Yang

B.S., Applied Mathematics, Tsinghua University
July, 1989

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Applied Mathematics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1994

© Massachusetts Institute of Technology 1994. All rights reserved.

Author ..........................................................................................................

Department of Mathematics
May 12, 1994

Certified by ........................................................................................................

David J. Benney
Professor of Applied Mathematics
Thesis Supervisor

Accepted by ........................................................................................................

David Vogan
Chairman, Departmental Graduate Committee

ARCHIVES

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

AUG 11 1994
Some Nonlinear Equations Arising in the Theory of Water Waves
by
Jianke Yang

Submitted to the Department of Mathematics
on May 12, 1994, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Applied Mathematics

Abstract

In this thesis, three different but related aspects of water waves, namely, cross-waves, edge waves and coupled nonlinear wave packets are studied. For cross-waves, the closed form solution for the linear cross-wave equations is derived and the neutral curve is obtained. The experimentally observed sloshing phenomenon is numerically reproduced and found to be irrelevant to the Benjamin-Feir instability of a uniform wavetrain. As for edge waves, the previous work on edge wave stability is improved and the entire stability structure is laid out. Stable edge-wave modes are found. These modes strongly affect edge-wave nonlinear evolution. An explanation is also given for the edge wave anti-node number alternation phenomenon which has been observed in experiments. Coupled nonlinear wave packets are generally governed by coupled nonlinear Schrödinger type equations. These equations are analytically and numerically studied. The interesting non-elastic solitary wave interactions are observed and explained. The mixed-mode solitary waves are also described and their stability is determined. The methods developed for these coupled nonlinear Schrödinger equations seem to be quite general.

Thesis Supervisor: David J. Benney
Title: Professor of Applied Mathematics
Acknowledgment

I would like to take this opportunity to thank my thesis advisor, Prof. D.J. Benney, for all the support he gave me throughout my years at MIT. I first met him five years ago in Beijing, where he was giving three-week lectures on fluid mechanics. I knew back then what a knowledgeable and friendly person he was. When I came to MIT, I chose him as my thesis advisor, and I began to benefit from his unique perspective, new ideas and diverse knowledge. Over the years, his guidance has made this thesis work possible, his generous summer support has been of great help and his kindness, encouragement and sense of humor have been heart-warming. To him I owe deep gratitude.

I am also grateful to Prof. R.R. Rosales and Prof. F. Waleffe for being on my thesis committee and spending time to review my thesis, Prof. C.C. Lin for helping me to come to MIT and for his kind general advice, and Prof. H. Cheng for many inspiring and interesting discussions on mathematics as well as on literature. They have always been very supportive.

I would also like to thank the Wang Foundation (Hong Kong) for providing me the opportunity to study in the U.S., and the mathematics department for the financial support. The help from the department’s staff members is also appreciated.

My stay at MIT would not have been so joyful without the friendship of my fellow graduate students, past and present. My special thanks go to Carlos, Dmitry, Esteban, Glenn, Lu, Paul, Peter, Pusun, Rodney, Shinya, Yiqun and Yuan.

I am lucky to have a loving family behind me. The emotional support from my parents, brother and sisters over all the years since my childhood is invaluable. I can not thank them enough for their love and support. I am deeply indebted to my wife, Huimin, for her love, patience and support over the years at MIT. This thesis is dedicated to her.
Contents

General Introduction 3

I The Sloshing Motion of Cross-waves 5
I.1 Introduction 5
I.2 The analytical solution of the linearized 2-D cross-wave equations for an arbitrary initial condition 8
I.2.1 Formulation 8
I.2.2 Derivation of the analytical solution 9
I.2.3 The asymptotic behavior of the analytical solution $A(z,t)$ for large time $t$ 11
I.2.4 The neutral curve 14
I.3 Sloshing motion and the nonlinear 3-D cross-wave equations 16
I.3.1 Formulation 16
I.3.2 The analytical solution of the linearized 3-D cross-wave equations for an arbitrary initial condition 17
I.3.3 Numerical methods for the nonlinear 3-D cross-wave equations 18
I.3.4 Numerical results and sloshing motion 18
I.4 Summary 23

II The Stability and Nonlinear Evolution of Edge Waves 24
II.1 Introduction 24
II.2 Formulation 26
II.2.1 Edge waves on an open beach 26
II.2.2 Edge waves on a wide bounded beach 27
II.3 Steady state standing edge waves and their stability 29
II.3.1 Steady state standing edge waves 29
II.3.2 Linear stability analysis 30
II.3.3 Stability results 32
II.4 Nonlinear evolution of edge waves on a wide beach 36
II.5 Summary 39

III Solitary Wave Interactions for Coupled Nonlinear Schrödinger Equa-
### III.1 Introduction

- III.1.1 The derivation of CPNLS in water waves: 44
- III.1.2 The derivation of CPNLS in nonlinear fiber optics: 46
  - III.1.2.1 Coupling between waves of different frequencies: 46
  - III.1.2.2 Coupling between polarization components of the same wave: 47
- III.1.3 System to be studied: 47

### III.2 The non-elastic solitary wave interactions

- III.2.1 The motion of a solitary wave in a slowly varying potential field: 48
- III.2.2 The relations (2.18), (2.19) and the conservation laws: 54
- III.2.3 Discussion: 55
- III.2.4 The determination of daughter wave profiles: 56

### III.3 Stability of equally mixed solitary wave mode

### III.4 Summary
General Introduction

Water waves are a very familiar physical phenomenon. When wind blows over the water surface, ripples are excited; when a ship passes by, ship waves trailing behind are observed; and when water waves travel toward a beach, their breaking and splashing are seen. Solitary waves in a shallow canal, edge waves on a coast and cross-waves in a wave tank all present interesting mathematical problems which are critical to understanding the physics.

The study of water waves can be traced back to the ancient times, and has been intense in the last one hundred years. First of all, there are numerous applications to such practical problems as ship construction, dam design, the laying of cable, etc. Secondly, the obvious mathematical difficulty of the subject poses a great challenge to the applied mathematicians. The difficulty comes not only from the nonlinear governing equations, but also from the nonlinear free surface boundary conditions and complex boundary geometries. Finally, the nonlinearity and dispersion, which are very important for water waves, are also of central interest in nonlinear optics, plasma physics and many other subjects. Mathematical techniques developed for water waves have been of general applicability.

Much of the development of the water wave theory is described in the books by Stoker (1957), Whitham (1974) and Craik (1985). An important assumption, which was often made, is that the motion is irrotational. If viscosity is negligible, and the motion is initially irrotational, then this assumption is fully justified. Under this assumption, the governing nonlinear equations are reduced to the linear Laplace equation. The difficulty now comes mainly from the nonlinear free surface boundary conditions and geometrical boundary complexity. If the wave amplitude is small, the free surface conditions can be linearized and linear wave modes are obtained.

Nonlinear effects are important in water waves. One such effect is the transfer of energy between wave modes through resonance mechanisms. Suppose we have a set of discrete wave modes with wave numbers \( k_j \) and frequencies \( \omega_j = \omega(k_j) \). If the resonance conditions:

\[
\sum_j k_j = 0
\]

\[
\sum_j \omega_j = 0
\]

are simultaneously satisfied, then energy will be transferred from some modes to others. This so-called weakly nonlinear wave interaction theory was proposed and developed in the early 1960’s by Phillips (1960), Longuet-Higgins (1962), Benney (1962) and Hasselmann (1962).

The finite amplitude aspects of periodic water waves were first considered by Stokes (1847), who showed that the surface elevation for a plane wave train on deep water could be expanded in powers of the amplitude. The evolution of a finite amplitude water wave train was intensively studied in the 1960’s. Benjamin and Feir (1967) first noticed the tendency of waves in a laboratory channel to disintegrate and in a resulting theoretical investigation discovered that this was due to side-band instability. An easy explanation of this so-called Benjamin-Feir instability lies in the fact that the modulation of a water wave train is governed by the nonlinear Schrödinger equation

\[
i A_T + \beta A_{XX} + \gamma A^2 A^* = 0
\]
where $A(X,T)$ is the envelope of the wave train. The Stokes solution can be easily shown to be unstable according to this equation.

This thesis studies three different but related aspects of water waves, namely, cross-waves, edge waves and coupled nonlinear wave packets. Cross-waves and edge waves, as we shall show later, are generated by a subharmonic resonance mechanism. Much work has been done on these two subjects, but some important questions still remain to be answered. Coupled nonlinear Schrödinger equations arise when two wave trains are present and interact with each other. New phenomena appear due to this coupling. These aspects have not been studied previously.

The format of this thesis is as follows. Part one studies cross-waves; part two studies edge waves; part three studies coupled nonlinear wave packets. Each part is self-contained and can be read independently.
Part I

The Sloshing Motion of Cross-waves

1.1 Introduction

Cross-waves are standing waves with their crests perpendicular to a wavemaker. They have a frequency equal to half that of the wavemaker. They can be generated in a wave tank when the wavemaker's frequency is near a subharmonic resonance (twice the natural frequency of a cross-wave mode) provided that the forcing is strong enough to overcome dissipative effects. These waves were first reported by Faraday in 1831. In his diary for July 1, 1831, Michael Faraday reported his observation that when a vibrating vertical plate was dipped into a basin of water, "Elevations, waves or crispations immediately formed but of a peculiar character ... beginning at the plate and projecting directly out from it ... like the teeth of a very short comb". He also remarked that these waves had frequency half that of the excitation. The generating mechanism of these waves remained a mystery until 1970 when Garrett (1970) found that, for a rectangular channel of finite length, it can be described by Mathieu's equation. Mahony (1972) extended the analysis to a rectangular channel of infinite length. He found that the resonant bandwidth is an order of magnitude smaller than that in a channel of finite length. He also indicated that nonlinear effects may control the decay of the cross-waves down the channel. Using the multiple-scale perturbation methods, Jones (1984) first derived the governing nonlinear equations for the longitudinal modulation of a cross-wave mode in a rectangular channel of infinite length. These equations were later identified by Miles (1985) as the nonlinear Schrödinger equation with a homogeneous wavemaker boundary condition and a null condition at infinity (referred to as the 2-D cross-wave equations hereafter). Using an averaged variational principle, Miles & Becker (1988) obtained the same results as those of Jones. When the channel is wide enough to allow several adjacent cross-wave modes to be simultaneously generated, the cross-wave will also be modulated in the spanwise direction. Under the assumption that the spanwise modulation is much milder than the longitudinal one, so that the spanwise dispersion is negligible, the governing equations for both modulations were derived by Ayanle, Bernoff & Lichter (1990) and will be referred to as the (spanwisely non-dispersive) 3-D cross-wave equations hereafter.

Experiments on cross-waves have been conducted by Barnard & Pritchard (1972), Lichter & Shemer (1986), Ayanle, Bernoff & Lichter (1990), Underhill, Lichter & Bernoff (1990) etc. Barnard & Pritchard (1972)'s experiments demonstrated the generation of cross-waves in a long water channel. They also observed that "the cross-waves never reach a true state of equilibrium, and after the cross-wave amplitude has passed through a maximum, a wave detaches itself from the wavemaker, propagates along the channel" and eventually decays. These results were in agreement with Lichter & Chen (1987)'s numerical calculations of the 2-D cross-wave equations (when damping was incorporated). Miles & Becker (1988) determined those stationary envelopes that are evanescent at large distances from the wavemaker through analytical approximations and numerical integration. They compared their envelopes with Barnard & Pritchard (1972)'s experiments and Lichter & Chen (1987)’s numerical calculations, and suggested that stationary envelopes with either no or one maximum are stable for sufficiently small amplitudes and evolve into limit cycles for somewhat larger amplitudes. Ayanle, Bernoff & Lichter (1990) observed a mixed-mode state. They then used a cen-
ter manifold analysis on the 3-D cross-wave equations to reduce the PDEs to a system of coupled Landau equations in the neighborhood of a codimension-two point where two adjacent cross-wave modes are marginally stable. They found four possible steady states of the system, one of which is a mixed-mode superposition of two cross-wave modes. They predicted a Hopf bifurcation from the mixed mode for some parameters. Their experiments showed good agreement with the theoretical predictions. Underhill, Lichter & Bernoff (1990)'s experiments revealed richer structures partially due to the presence of sloshing motion. They observed modulated, frequency-locked and chaotic cross-waves in different regions of the parameter space. In particular, they observed that "at large cross-wave amplitudes, the spanwise wave structure apparently breaks up, because of modulational instability, into coherent soliton-like structures that propagate in the spanwise direction and are reflected by the sidewalls."

The mechanism of cross-wave generation lies in the linearized 2-D cross-wave equations. The neutral curve, which sets the criterion as to when a cross-wave will be excited, should come from these linear equations. The cross-wave field characteristics along the channel at the initial stage of generation is also dictated by them. Therefore, the knowledge of the solution behaviors of these equations for an arbitrary initial condition is important. Previous research by Mahony (1972) and Jones (1984) found a special eigensolution of these linear equations. That solution was later used to establish the neutral curve. Such a neutral curve is apparently doubtful and a resolution has to be obtained from the solution for an arbitrary initial condition. In this paper, the analytical solution of the linearized 2-D cross-wave equations for an arbitrary initial condition is derived. The asymptotic behavior of this solution as time becomes large is given by simple formulas. These results fully describe the cross-wave field at its generating stage. The neutral curve based on this solution is also obtained. It is found that this neutral curve turns out to be the same as that based on the special eigensolution mentioned earlier.

When the water channel is wide, and if the spanwise dispersion is negligible, the mechanism of cross-wave generation lies in the linearized 3-D cross-wave equations. Making use of the fact that the solution of these linear 3-D equations can be decomposed into a sum of the 2-D cross-wave modes, whose amplitudes are governed by the corresponding linear 2-D equations, the analytical solution of the linear 3-D equations for an arbitrary initial condition is obtained. If the forcing of the wavemaker is strong, it is easy to show from this analytical solution that several adjacent cross-wave modes will be excited, and the spanwise modulation will therefore arise.

Sloshing motion is an intriguing nonlinear phenomenon. It has been experimentally observed by Underhill, Lichter & Bernoff (1990), but its dynamics is not yet known. Underhill, Lichter & Bernoff (1990) speculated that it "may be related to the Benjamin-Feir instability of a uniform wave train." Since the Benjamin-Feir instability relies crucially on dispersion as well as on nonlinearity, this speculation emphasizes the importance of the spanwise dispersion, which is neglected in the 3-D cross-wave equations mentioned above. Doubts naturally arise at first sight, since in the present situation, two oppositely-propagating wavetrains are involved, and the Benjamin-Feir instability of one uniform wavetrain may be irrelevant. This question is resolved because it can be proven that in deep water, each wavetrain still experiences the Benjamin-Feir instability in spite of the existence of the other. Serious doubts remain for the following reasons. First of all, the longitudinal modulation is coupled with the spanwise one, and will affect it in a non-trivial way. Secondly and more importantly, the cross-wave field is non-conservative. The wavemaker transfers energy into this field and strongly affects its dynamics. Due to the nature of this problem, it is proposed that the
spanwise dispersion does not play a major role; sloshing has a different nature from the Benjamin-Feir instability; and the dynamics of sloshing can be described by these (spanwisely non-dispersive) 3-D cross-wave equations. To support this proposition, the solutions of these 3-D equations were studied. In view of the complexity of these nonlinear equations, numerical approaches were taken. The numerical results show various features of the cross-wave field. In particular, they show that, for a certain range of parameters, sloshing motion really appears. These sloshing waves propagate in the spanwise direction, are reflected by the sidewalls, and interact with each other in a persistent way. Due to the interaction of the sloshing and longitudinal waves, the motion of fluid particles is very complicated. All these results qualitatively agree well with Underhill, Lichter & Bernoff (1990)'s observations on sloshing motion, and gives strong evidence of the connection between the sloshing motion and the 3-D cross-wave equations as well as the fact that the dynamics of sloshing is well described by these (spanwisely non-dispersive) 3-D equations. Spanwise dispersion may be important to suppress very short waves, but it plays only a minor role in the dynamics of sloshing.
I.2 The analytical solution of the linearized 2-D cross-wave equations for an arbitrary initial condition

I.2.1 Formulation

Consider a rectangular semi-infinite water channel. The coordinates are such that \( x(z > 0) \) measures distances along the channel, \( y(0 < y < b) \) is the spanwise coordinate and \( z(z < 0) \) is depth below the free surface. Its driving mechanism is a wavemaker at \( z = 0 \) with the prescribed motion

\[
x = af(z) \sin 2\sigma t \quad (z < 0).
\]

This motion will expectedly generate a progressive wavetrain along the channel and a local standing disturbance around the wavemaker, both with the frequency \( 2\sigma \). This generated motion is two dimensional and is described by the velocity potential

\[
\phi_g = \frac{ga}{\sigma} Pe^{4kz} \sin(4kz - 2\sigma t) - \int_0^\infty Q(\mu)(\cos \mu k z + \frac{4}{\mu} \sin \mu k z) e^{-\mu k z} d\mu \cos 2\sigma t,
\]

where

\[
k = \frac{\sigma^2}{g}, \quad P = 4 \int_{-\infty}^0 f(s)e^{4s}ds,
\]

and

\[
Q(\mu) = \frac{4}{\pi} \frac{\mu}{16 + \mu^2} \int_{-\infty}^0 f(s)(\cos \mu s + \frac{4}{\mu} \sin \mu s)ds
\]

(see Jones 1984). What is surprising is that the wavemaker may also excite cross-waves of frequency \( \sigma \). What happens is that this local standing disturbance may feed energy into a free cross-wave mode through a kind of subharmonic resonance mechanism when resonance conditions are approximately satisfied.

A free cross-wave mode with frequency \( \sigma \) and \( n \) transverse nodes has the velocity potential of the form

\[
\phi_c = \frac{ga}{\sigma}[A(X,T)e^{-i\sigma t} + \text{c.c.}] \cos k_0 y e^{kz},
\]

where \( A(X,T) \) is the slowly varying complex amplitude,

\[
X = ekz, \quad T = \epsilon^2 \sigma t, \quad \epsilon = ka, \quad k_0 = \frac{n\pi}{b}.
\]

This potential satisfies the linearized free surface boundary conditions and the two sidewall boundary conditions. For it to satisfy the Laplace equation \( \Delta \phi = 0 \), the cross-wave wavenumber \( k_0 \) and frequency \( \sigma \) need to approximately satisfy the dispersion relation \( \sigma^2 \approx gk_0 \), i.e. \( k \approx k_0 \).

The generated wave represented by the velocity potential \( \phi_g \) can resonantly interact and transfer energy into the cross-wave \( \phi_c \). As a result, the cross-wave can be excited. A multiple-scale perturbation method is used to determine the evolution equation of the cross-wave amplitude \( A(X,T) \).
It is found that $A$ satisfies the following nonlinear 2-D cross-wave equations (see Jones 1984, Miles 1985):

\begin{align}
    iA_T + \frac{1}{4} A_{XX} + \lambda A + \frac{1}{2} A^2 A^* &= 0 \\
    A_X = iRA^* &\quad X = 0 \\
    A &\rightarrow 0 \quad X \rightarrow \infty
\end{align}

where

\begin{align}
    \lambda &= \Lambda - 0.202G^2, \quad G = 4k \int_{-\infty}^{0} f(z)e^{4kz}dz, \quad \Lambda = \frac{1 - \frac{k^2}{\sigma^2}}{2\epsilon^2} = \frac{1 - (\sigma_0 \epsilon)^2}{2\epsilon^2}, \\
    \sigma_0 &= (k_0g)^{\frac{1}{2}} \text{ is the resonant (or cutoff) frequency of the cross-wave, and} \\
    R &= \int_{-\infty}^{0} [4kf(z) + f'(z)e^{2kz}]dz - 2f(0).
\end{align}

The linearized 2-D cross-wave equations are:

\begin{align}
    iA_T + \frac{1}{4} A_{XX} + \lambda A &= 0 \\
    A_X = iRA^* &\quad X = 0 \\
    A &\rightarrow 0 \quad X \rightarrow \infty
\end{align}

These linear equations govern the cross-wave field at its initial stage of generation.

As has been noted by Mahony (1972) and Jones (1984), the function

\[ A(X, T) = e^{-2\sqrt{\lambda + \sqrt{\frac{R^4}{16} - \lambda^2}} X + \sqrt{\frac{R^4}{16} - \lambda^2} T} \]

is a special eigensolution of the above linear equations (2.12), (2.13) and (2.14) with the eigenvalue $\sqrt{\frac{R^4}{16} - \lambda^2}$. Since this solution grows if $\frac{R^4}{16} - \lambda^2 > 0$ and is bounded if $\frac{R^4}{16} - \lambda^2 < 0$, it has been conjectured that the neutral curve is $\frac{R^4}{16} - \lambda^2 = 0$.

To clarify the neutral curve and to fully understand the cross-wave field behavior at its initial stage of generation, the solution of the equations (2.12), (2.13) and (2.14) for a general initial condition is required. In the next section, we will set out to derive the analytical solution of those equations for an arbitrary initial condition

\[ A|_{T=0} = A_0(X) \quad (X \geq 0). \]

**I.2.2 Derivation of the analytical solution**

For convenience, the variables $z$ and $t$ will be used instead of $X$ and $T$ in the remaining of the section I.2.
Rather than working with the real and imaginary parts of the complex amplitude $A$, here we choose to directly work with $A$ and $A^*$. Denoting $A^* = B$, it is easy to find from the equations (2.12), (2.13) and (2.14) that $A$ and $B$ satisfy the following equations:

\[
\begin{align*}
    iA_t + \frac{1}{4} A_{xx} + \lambda A &= 0 \\
    -iB_t + \frac{1}{4} B_{xx} + \lambda B &= 0 \\
    A_x = iRB & \quad z = 0 \\
    B_x = -iRA & \quad z = 0 \\
    A \rightarrow 0 & \quad z \rightarrow \infty \\
    B \rightarrow 0 & \quad z \rightarrow \infty
\end{align*}
\]  

(2.17)

together with the initial conditions:

\[
\begin{align*}
    A_{t=0} &= A_0(z) \quad (z \geq 0), \\
    B_{t=0} &= B_0(z) \quad (z \geq 0).
\end{align*}
\]  

(2.18)

Note: $B_0(z) = A^*_0(z)$.

When Laplace transforms are taken for the equations (2.17) and (2.18) with respect to the time $t$, $A$ and $B$’s Laplace transforms $\tilde{A}, \tilde{B}$ are found to satisfy the following equations:

\[
\begin{align*}
    \frac{1}{4} \tilde{A}_{xx} + (\lambda + is)\tilde{A} &= iA_0(z) \\
    \frac{1}{4} \tilde{B}_{xx} + (\lambda - is)\tilde{B} &= -iB_0(z) \\
    \tilde{A}_x = iR\tilde{B} & \quad z = 0 \\
    \tilde{B}_x = -iR\tilde{A} & \quad z = 0 \\
    \tilde{A} \rightarrow 0 & \quad z \rightarrow \infty \\
    \tilde{B} \rightarrow 0 & \quad z \rightarrow \infty
\end{align*}
\]  

(2.19) \hspace{1cm} (2.20) \hspace{1cm} (2.21) \hspace{1cm} (2.22) \hspace{1cm} (2.23) \hspace{1cm} (2.24)

The general solution of the equation (2.19) is:

\[
\tilde{A}(z, s) = c_1e^{2\sqrt{-\lambda - is}z} + d_1e^{2\sqrt{-\lambda - is}z} - \frac{i}{\sqrt{-\lambda - is}} \int_{0}^{\infty} A_0(z')e^{-2\sqrt{-\lambda - is}|z - z'|}dz',
\]  

(2.25)

where the complex function $\sqrt{-\lambda - is}$ has positive real part for $s$ on the line $\text{Re}(s) = \alpha_0 > 0$, $\alpha_0$ is sufficiently large, and $c_1, d_1$ are complex constants.

Similarly, the general solution of (2.20) is:

\[
\tilde{B}(z, s) = c_2e^{2\sqrt{-\lambda + is}z} + d_2e^{2\sqrt{-\lambda + is}z} + \frac{i}{\sqrt{-\lambda + is}} \int_{0}^{\infty} B_0(z')e^{-2\sqrt{-\lambda + is}|z - z'|}dz',
\]  

(2.26)

where the complex function $\sqrt{-\lambda + is}$ has positive real part for $s$ on the line $\text{Re}(s) = \alpha_0 > 0$, and $c_2, d_2$ are complex constants.

When the boundary conditions (2.23) and (2.24) are applied to the solutions (2.25) and (2.26), it is concluded that

\[
d_1 = d_2 = 0.
\]  

(2.27)
The wavemaker boundary conditions (2.21) and (2.22) are then applied to the solutions (2.25) and (2.26), and two equations to determine $c_1$ and $c_2$ are obtained,

$$2\sqrt{-\lambda - is} \ c_1 + i R \ c_2 = \frac{R}{\sqrt{-\lambda + is}} \int_0^\infty B_0(z') e^{-2\sqrt{-\lambda - is}z'} \ dz' - 2i \int_0^\infty A_0(z') e^{-2\sqrt{-\lambda - is}z'} \ dz', \tag{2.28}$$

$$i R \ c_1 - 2\sqrt{-\lambda + is} \ c_2 = -\frac{R}{\sqrt{-\lambda - is}} \int_0^\infty A_0(z') e^{-2\sqrt{-\lambda - is}z'} \ dz' - 2i \int_0^\infty B_0(z') e^{-2\sqrt{-\lambda + is}z'} \ dz'. \tag{2.29}$$

$c_1$ and $c_2$ are found to be

$$c_1 = \frac{-4R \int_0^\infty B_0(z') e^{-2\sqrt{-\lambda + is}z'} \ dz' + i(4\sqrt{-\lambda + is}\sqrt{-\lambda - is} + R^2) \int_0^\infty A_0(z') e^{-2\sqrt{-\lambda - is}z'} \ dz'}{R^2 - 4\sqrt{-\lambda - is}\sqrt{-\lambda + is}}, \tag{2.30}$$

$$c_2 = \frac{-4R \int_0^\infty A_0(z') e^{-2\sqrt{-\lambda - is}z'} \ dz' - i(4\sqrt{-\lambda - is}\sqrt{-\lambda + is} + R^2) \int_0^\infty B_0(z') e^{-2\sqrt{-\lambda + is}z'} \ dz'}{R^2 - 4\sqrt{-\lambda - is}\sqrt{-\lambda + is}}. \tag{2.31}$$

It is clear that $\tilde{A}$ now is fully determined and is in the form

$$\tilde{A}(z, s) = c_1 e^{-2\sqrt{-\lambda - is}z} - \frac{i}{\sqrt{-\lambda - is}} \int_0^\infty A_0(z') e^{-2\sqrt{-\lambda - is}z' - z} \ dz', \tag{2.32}$$

where $c_1$ is given by (2.30). The original solution $A(z, t)$ is the inverse Laplace transform of $\tilde{A}(z, s)$

$$A(z, t) = \frac{1}{2\pi i} \int_L \tilde{A}(z, s)e^{st} \ ds, \tag{2.33}$$

where the integration path $L$ is to the right hand of all the singularities of $\tilde{A}(z, s)$.

I.2.3 The asymptotic behavior of the analytical solution $A(z, t)$ for large time $t$

In order to determine $A(z, t)$'s large-time asymptotic behavior, $\tilde{A}(z, s)$'s singularities need to be examined.

First consider the $\lambda < 0$ case.

The solution (2.32) has two branch points at $s = \pm i\lambda$. To guarantee that the complex functions $\sqrt{-\lambda - is}$ and $\sqrt{-\lambda + is}$ have positive real parts on the integration path $L$, write

$$\sqrt{-\lambda - is} = e^{-\frac{i}{4}\pi i} \sqrt{s - i\lambda}, \tag{2.34}$$

$$\sqrt{-\lambda + is} = e^{\frac{i}{4}\pi i} \sqrt{s + i\lambda}. \tag{2.35}$$

The branch cuts of $\sqrt{s - i\lambda}$ and $\sqrt{s + i\lambda}$ are located as shown in Figure 1, and define

$$\text{arg}(s^+ - i\lambda) = \pi \quad \text{for } s^+ \text{ on the upper side of } \sqrt{s - i\lambda} \text{ cut;}$$
$$\text{arg}(s^- - i\lambda) = -\pi \quad \text{for } s^- \text{ on the lower side of } \sqrt{s - i\lambda} \text{ cut;}$$
$$\text{arg}(s^+ + i\lambda) = \pi \quad \text{for } s^+ \text{ on the upper side of } \sqrt{s + i\lambda} \text{ cut;}$$
$$\text{arg}(s^- + i\lambda) = -\pi \quad \text{for } s^- \text{ on the lower side of } \sqrt{s + i\lambda} \text{ cut.} \tag{2.36}$$
When (2.34) and (2.35) are introduced into the equations (2.30) and (2.32), $\hat{A}$ and $c_1$ can be rewritten in the following forms:

$$\hat{A}(x, s) = c_1 e^{-2e^{-\frac{xi}{4} \sqrt{s - i\lambda}} + \frac{e^{-\frac{xi}{4}}}{\sqrt{s - i\lambda}} \int_0^\infty A_0(x') e^{-2e^{-\frac{xi}{4} \sqrt{s - i\lambda}} x'} dx', \quad (2.37)$$

$$c_1 = \frac{-4R \int_0^\infty A_0^*(x') e^{-2e^{-\frac{xi}{4} \sqrt{s + i\lambda} x'}} dx' - \frac{e^{-\frac{xi}{4}} (4\sqrt{-\lambda + is\sqrt{-\lambda - is + R^2})}{\sqrt{s - i\lambda}} \int_0^\infty A_0(x') e^{-2e^{-\frac{xi}{4} \sqrt{s - i\lambda} x'}} dx'}{R^2 - 4\sqrt{s - i\lambda} \sqrt{s + i\lambda}}. \quad (2.38)$$

Since

$$\frac{1}{R^2 - 4\sqrt{s - i\lambda} \sqrt{s + i\lambda}} = \frac{R^2 + 4\sqrt{s - i\lambda} \sqrt{s + i\lambda}}{16(R^4 - \lambda^2 - s^2)}, \quad (2.39)$$

it is clear that $c_1(s)$ has poles at points

$$s = \pm \sqrt{\frac{R^4}{16} - \lambda^2}. \quad (2.40)$$

Making use of (2.39), $c_1$ can be expressed as

$$c_1(s) \equiv \frac{F(s)}{s^2 - (\frac{R^4}{16} - \lambda^2)} = \frac{F(s)}{(s - \sqrt{\frac{R^4}{16} - \lambda^2})(s + \sqrt{\frac{R^4}{16} - \lambda^2})}, \quad (2.41)$$

where

$$F(s) = \frac{R^2 + 4\sqrt{s - i\lambda} \sqrt{s + i\lambda}}{16} \{4R \int_0^\infty A_0^*(x') e^{-2e^{-\frac{xi}{4} \sqrt{s - i\lambda} x'}} dx'$$

$$+ \frac{e^{-\frac{xi}{4}} (4\sqrt{-\lambda + is\sqrt{-\lambda - is + R^2})}{\sqrt{s - i\lambda}} \int_0^\infty A_0(x') e^{-2e^{-\frac{xi}{4} \sqrt{s - i\lambda} x'}} dx' \}. \quad (2.42)$$

Note that integrals such as $\int_0^\infty A_0(x') e^{-2e^{-\frac{xi}{4} \sqrt{s - i\lambda} x'}} dx'$ do not contribute new singularities other than the branch points of $\sqrt{s - i\lambda}$ and $\sqrt{s + i\lambda}$, and they are less singular than $\frac{1}{\sqrt{s - i\lambda}}$ or $\frac{1}{\sqrt{s + i\lambda}}$ because $A_0(x)$ tends to zero as $x \to \infty$, so $F(s)$ has no singularities other than the branch points of $\sqrt{s - i\lambda}$ and $\sqrt{s + i\lambda}$, and these singularities are weaker than $\frac{1}{s - i\lambda}$ or $\frac{1}{s + i\lambda}$. Similarly, the second term in (2.37) has no singularities other than the branch point of $\sqrt{s - i\lambda}$ and the singularity is weaker than $\frac{1}{s - i\lambda}$.

The above results on $\hat{A}(x, s)$ are now used to determine the large-time asymptotic behaviors of $A(x, t)$. They are distinctively different depending on the sign of $\frac{R^4}{16} - \lambda^2$.

![Figure 1: branch cuts of $\sqrt{s - i\lambda}$ and $\sqrt{s + i\lambda}$](image-url)
Figure 2: the alternative integration path $L'$ for the case $\frac{R^4}{16} - \lambda^2 > 0$.

1. $\frac{R^4}{16} - \lambda^2 > 0$:

In this case, the two simple poles $s = \pm \sqrt{\frac{R^4}{16} - \lambda^2}$ are on the real axis. We choose an alternative integration path $L'$ as in Figure 2, and then use the residue theorem. After some simple asymptotic analysis, it is found that

$$A(x, t) \rightarrow \frac{F(\sqrt{\frac{R^4}{16} - \lambda^2})}{2\sqrt{\frac{R^4}{16} - \lambda^2}} e^{-2e^{-4t}\sqrt{\frac{R^4}{16} - \lambda^2 - i\lambda x + \sqrt{\frac{R^4}{16} - \lambda^2} t}}, \quad t \rightarrow \infty. \quad (2.43)$$

Note that this solution is exactly the same as the eigensolution noted by Mahony (1972) and Jones (1984). It grows exponentially in time with the growth rate $\sqrt{\frac{R^4}{16} - \lambda^2}$ and decays exponentially in the $x$ direction.

2. $\frac{R^4}{16} - \lambda^2 = 0$:

In this case, $s = 0$ is a double pole. Choosing an alternative integration path $L'$ as in Figure 3 and performing a similar analysis, it is found that

$$A(x, t) \rightarrow [F(0)t + F'(0) + iF(0)|\lambda|^{-\frac{1}{2}}x]e^{-2|\lambda|^{\frac{1}{2}}x}, \quad t \rightarrow \infty. \quad (2.44)$$

Note that this solution grows linearly in time and decays exponentially in the $x$ direction.

Figure 3: the alternative integration path $L'$ for the case $\frac{R^4}{16} - \lambda^2 = 0$. 

3. $\frac{R^4}{16} - \lambda^2 < 0$:

In this case, the two simple poles $s = \pm i\sqrt{\lambda^2 - \frac{R^4}{16}}$ are on the imaginary axis. When an alternative integration path $L'$ shown in Figure 4 is taken, a similar analysis reveals that

$$A(z, t) \rightarrow \frac{F(i\sqrt{\lambda^2 - \frac{R^4}{16}})}{2i\sqrt{\lambda^2 - \frac{R^4}{16}}} e^{-2e^{-\frac{R^4}{16}}(i\sqrt{\lambda^2 - \frac{R^4}{16}} - z + i\sqrt{\lambda^2 - \frac{R^4}{16}} t)}$$

$$- \frac{F(-i\sqrt{\lambda^2 - \frac{R^4}{16}})}{2i\sqrt{\lambda^2 - \frac{R^4}{16}}} e^{-2e^{-\frac{R^4}{16}}(-i\sqrt{\lambda^2 - \frac{R^4}{16}} - z - i\sqrt{\lambda^2 - \frac{R^4}{16}} t)}, \quad t \rightarrow \infty.$$  \hspace{1cm} (2.45)

Notice that this solution is bounded for all time and decays exponentially in the $z$ direction.

Now, the case $\lambda \geq 0$ is discussed briefly.

Using previous arguments, it is found that the asymptotic solution always decays exponentially in the $z$ direction.

1. When $\frac{R^4}{16} - \lambda^2 > 0$, it grows exponentially in time and is the same as (2.43).
2. When $\frac{R^4}{16} - \lambda^2 < 0$, it is bounded for all time and is the same as (2.45).
3. When $\frac{R^4}{16} - \lambda^2 = 0$, the analysis is more complicated, but no new features appear.

### I.2.4 The neutral curve

It is now clear that the neutral curve is $\frac{R^4}{16} - \lambda^2 = 0$. It is re-assured that the result is the same as that obtained based only on the special eigensolution (2.15) of the linearized 2-D cross-wave equations (2.12), (2.13) and (2.14). If $\frac{R^4}{16} - \lambda^2 > 0$, the growth rate of the cross-wave is $\sigma = \sqrt{\frac{R^4}{16} - \lambda^2}$. 
When dissipation is modelled into the linearized 2-D cross-wave equation (2.12), it becomes

\[ iA_t + \frac{1}{4} A_{xx} + (\lambda + iL)A = 0, \tag{2.46} \]

where \( L > 0 \) is linear damping constant. The boundary conditions (2.13) and (2.14) remain the same.

In this case, it is easy to show that the analytical solution and its asymptotic form are just the undamped ones (as previously given) multiplied by \( e^{-Lt} \). The neutral curve now becomes

\[ \sqrt{\frac{R^4}{16} - \lambda^2} - L = 0. \tag{2.47} \]
I.3 Sloshing motion and the nonlinear 3-D cross-wave equations

I.3.1 Formulation

Consider a rectangular water channel of semi-infinite length \((z \geq 0)\), infinite depth \((z \leq 0)\), and finite width \((0 \leq y \leq b)\). It is driven by a wavemaker at \(z = 0\) at frequency \(2\sigma\). The generated cross-wave has the primary wavenumber \(k_0 = \frac{N\pi}{b}\), where \(N\) is the number of transverse nodes. We assume that the channel is wide and \(N\) is large. This cross-wave is modulated in both longitudinal and transverse directions.

Introduce the perturbation parameter \(\epsilon = N^{-\frac{1}{2}}\), and suppose that the motion of the wavemaker is prescribed by:

\[
z = \frac{eg}{\sigma^2} f(z) \sin 2\sigma t \quad (z < 0).
\]

The velocity potential for this cross-wave is of the form:

\[
\phi = \frac{eg}{\sigma k} [A(X, Y, T)e^{i(k_0 y + \sigma t)} + B(X, Y, T)e^{i(k_0 y - \sigma t)} + c.c.]e^{kz},
\]

where

\[
X = \epsilon k z, \quad Y = \epsilon^2 k_0 y \quad T = \epsilon^2 \sigma t, \quad k = \frac{\sigma^2}{g}.
\]

A multiple-scale perturbation analysis results in the following equations for \(A\) and \(B\) (see Ayanle, Bernoff & Lichter 1990)

\[
\begin{align*}
-iA_T + \frac{1}{4}A_{XX} + \frac{1}{4}iA_Y + (J - iL)A - 2|A|^2 A + 4|B|^2 A &= 0, \\
iB_T + \frac{1}{4}B_{XX} + \frac{1}{4}iB_Y + (J + iL)B - 2|B|^2 B + 4|A|^2 B &= 0, \\
A_X = iRB & \quad X = 0, \\
B_X = -iRA & \quad X = 0, \\
A = B^* & \quad Y = 0, \pi, \\
A \to 0 & \quad X \to \infty, \\
B \to 0 & \quad X \to \infty.
\end{align*}
\]

where

\[
J = \Lambda - 0.202G^2, \quad G = 4k \int_{-\infty}^{0} f(z)e^{4kz} dz, \quad \Lambda = \frac{1 - \frac{k_0}{k}}{2\epsilon^2} = \frac{1 - \frac{\sigma^2}{\epsilon^2}}{2},
\]

\[
\sigma_0 = (kg)^{\frac{1}{2}}, \quad R = \int_{-\infty}^{0} [4k f(z) + f'(z)e^{2kz}] dz - 2f(0),
\]

and \(L > 0\) is a linear damping constant. With the introduction of new scalings

\[
\bar{X} = \sqrt{2}X, \quad \bar{Y} = Y, \quad \bar{T} = \frac{1}{2}T, \\
\bar{A} = 2A, \quad \bar{B} = 2B, \quad \bar{R} = \frac{1}{\sqrt{2}} R,
\]

and the bars dropped, the following system of equations for the complex amplitudes \(A\) and \(B\) of the cross-waves is obtained:

\[
-iA_T + A_{XX} + iA_Y + (J - iL)A - |A|^2 A + 2|B|^2 A = 0
\]
\[ iB_T + B_{XX} + iB_Y + (J + iL)B - |B|^2B + 2|A|^2B = 0 \]  
\[ A_X = iRB \quad X = 0 \]  
\[ B_X = -iRA \quad X = 0 \]  
\[ A = B^* \quad Y = 0, \pi \]  
\[ A \rightarrow 0 \quad X \rightarrow \infty \]  
\[ B \rightarrow 0 \quad X \rightarrow \infty \]

Note that in the above derivation, spanwise modulations are assumed to be an order of magnitude weaker than longitudinal ones so that spanwise dispersion is negligible.

**I.3.2 The analytical solution of the linearized 3-D cross-wave equations for an arbitrary initial condition**

At the initial stage of cross-wave generation, \( A \) and \( B \) are both very small. So they are governed by the linearized 3-D cross-wave equations, namely

\[ -iA_T + A_{XX} + iA_Y + (J - iL)A = 0 \]  
\[ iB_T + B_{XX} + iB_Y + (J + iL)B = 0 \]  
\[ A_X = iRB \quad X = 0 \]  
\[ B_X = -iRA \quad X = 0 \]  
\[ A = B^* \quad Y = 0, \pi \]  
\[ A \rightarrow 0 \quad X \rightarrow \infty \]  
\[ B \rightarrow 0 \quad X \rightarrow \infty \]

It has been shown that in this case, \( A \) and \( B \) can be written as a linear eigenmode expansion of the form (Ayanle, Bernoff & Lichter 1990)

\[ A = \sum_{k=-\infty}^{\infty} A_k(X,T)e^{ikY}, \]  
\[ B = \sum_{k=-\infty}^{\infty} A_k^*(X,T)e^{ikY}, \]

and \( A_k^* \) is governed by the equations

\[ iA_{kT}^* + A_{kXX}^* + (J - k + iL)A_k^* = 0 \]  
\[ A_{kX}^* = -iRA_k \quad X = 0 \]  
\[ A_k^* \rightarrow 0 \quad X \rightarrow \infty \]

It is clear that these equations are the same as the equations (2.46), (2.13) and (2.13), so the analytical solution for \( A_k^* \) and its asymptotic behavior are obtained as before. Therefore the analytical solutions for \( A \) and \( B \) can be constructed from the relations (3.22) and (3.23). The \( k \)-th cross-wave mode will be excited if \( \sqrt{R^2 - (J - k)^2} - L > 0 \). If the forcing is strong, i.e. \( R \) is sufficiently large, it is apparent that several adjacent cross-wave modes will be generated, and large-scale spanwise modulations are to be expected.
I.3.3 Numerical methods for the nonlinear 3-D cross-wave equations

Sloshing motion is, without doubt, a nonlinear phenomenon, but the role of spanwise dispersion is unclear. If this dispersion is crucial, sloshing appears to be related to the Benjamin-Feir instability of a uniform wavetrain. But this may not be the case. In view of the special nature of the cross-waves, it is proposed that the spanwise dispersion plays a minor role and that the dynamics of sloshing can be described by the (spanwisely non-dispersive) 3-D cross-wave equations (3.8)–(3.14). To verify this conjecture, numerical methods are developed for these equations and numerical calculations are subsequently carried out.

The nonlinear 3-D cross-wave equations (3.8)–(3.14) are first divided into real and imaginary parts and equations with respect to \( \text{Re}(A) \), \( \text{Im}(A) \), \( \text{Re}(B) \) and \( \text{Im}(B) \) are derived. These equations are solved numerically by an explicit finite difference scheme which uses second-order central difference operators to approximate spatial derivatives

\[
\partial_{XX} \rightarrow \delta_{XX}, \quad \partial_T \rightarrow \delta_T^2, \quad (3.25)
\]

and the fourth-order Runge-Kutta scheme to advance in time. The boundary conditions at the wavemaker \( X = 0 \) and at the sidewalls \( Y = 0, \pi \) are also approximated by second-order difference operators which use the boundary point and two adjacent points in the domain. A sufficiently large interval for \( X \) is used so that the outer boundary condition for \( X \) can be taken as \( A = B = 0 \).

The adopted scheme is consistent, and its truncation error is second-order in space and fourth order in time. The stability condition restricts the size of the time step. The determination of the computational step-sizes is based on the experience with testing the code as well as on stability and accuracy concerns.

The initial conditions were chosen to be

\[
A_0(X,Y) = e^{-RX}[a_1(1 - i) + a_2(1 + i)Y(Y - \pi)], \quad (3.26)
\]

and

\[
B_0(X,Y) = e^{-RX}[a_1(1 + i) - a_2(1 - i)Y(Y - \pi)], \quad (3.27)
\]

which satisfy the boundary conditions (3.10)–(3.14). The real constants \( a_1 \) and \( a_2 \) were chosen as \( a_1 = a_2 = 0.05 \) in the actual runs and \( a_1 = a_2 = 0.005 \) in the code testing.

The code was tested against the linear analytical solution for the case with parameters \( R = 1.6, J = 0, \) and \( L = 2 \). The domain was bounded by \( 0 \leq X \leq 7, 0 \leq Y \leq \pi, \) and \( 0 \leq T \leq 6 \). The step-sizes were chosen as \( \Delta X = 0.05, \Delta Y = \frac{\pi}{80} \approx 0.04, \) and \( \Delta T = 0.0003. \) \( a_1 = a_2 = 0.005 \) were taken in the initial conditions (3.26) and (3.27). The comparison showed excellent agreement between the numerical and analytical solutions.

I.3.4 Numerical results and sloshing motion

Many runs with different parameters have been carried out in detail. The results are basically similar to the ones shown in this paper with parameters \( R = 1.6, J = -0.5 \) and \( L = 1.6 \). The
domain was defined by \( 0 \leq X \leq 7, 0 \leq Y \leq \pi \), and \( 0 \leq T \leq 12 \). The step-sizes were chosen to be \( \Delta X = 0.07, \Delta Y = \frac{\pi}{80} \approx 0.04 \), and \( \Delta T = 0.0003 \). \( a_1 = a_2 = 0.05 \) were taken in the initial conditions (3.26) and (3.27).

The results are shown in Figures 5 to 8.

The following interesting features are to be emphasized:

1. Sloshing motion really appears, as can be seen in Figure 5. The cross-waves are excited first. They then become spanwisely more localized and sloshing motion begins to appear. These sloshing waves travel in the transverse direction, are reflected by the side walls, and interact with each other in a persistent way.

2. Due to the interaction of the sloshing and longitudinal waves, the motion of fluid particles is much more complicated than in cases without spanwise modulations (see Lichter & Chen 1987’s computations of the 2-D cross-wave equations, etc.). This feature is clear in Figure 6.

Comparison with the relevant experimental results is interesting. Underhill, Lichter & Bernoff (1990) observed that “at large cross-wave amplitudes, the noisy periodic states, consisting of two or more waves traversing the span of the tank, appeared. These waves became progressively more localized and appeared soliton-like as the forcing amplitude was increased.” They were also “reflected by the sidewalls.” These observations qualitatively agree well with the numerical calculations.

This agreement sheds some light on the dynamics of the sloshing motion. As the numerical results have shown, the sloshing waves can arise without including spanwise dispersion, and are well described by the 3-D cross-wave equations (3.8)–(3.14). Therefore, it would appear that sloshing motion bears little relationship to the Benjamin-Feir instability of a uniform wavetrain. Spanwisely uniform, standing cross-waves break up, because of modulational instability, into transversely localized sloshing waves. This instability is due to the forcing of the wavemaker and the nature of the problem, not to the Benjamin-Feir instability.
Figure 5: the time-evolution of the cross-waves at the wavemaker ($X = 0$). $0 \leq Y \leq \pi$, $0 \leq T \leq 12$. The upper figure is the $|A|$ plot, and the lower figure is the $|B|$ plot. Note that the sloshing waves are generated and they propagate across the span of the channel.
Figure 6: the time-evolution of the cross-waves at the following positions: \((X,Y) = (0,0), (0,\frac{\pi}{4}), (0,\frac{\pi}{2}), (0,\frac{3\pi}{4})\) and \((0,\pi)\). Only \(|A|\) is plotted. \(|B|\) is quite similar. \(0 \leq T \leq 12\).

Figure 7: the time-evolution of the cross-waves at one sidewall \((Y = 0)\). \(0 \leq X \leq 7, 0 \leq T \leq 12\). Only \(|A|\) is plotted. \(|B|\) is quite similar.
Figure 8: the cross-wave field at the time $T = 12$. $0 \leq X \leq 7$, $0 \leq Y \leq \pi$. The upper figure is the $|A|$ plot, and the lower figure is the $|B|$ plot.
I.4 Summary

In this part of the thesis, the analytical solution was obtained for the linearized cross-wave equations with arbitrary initial conditions. This solution conclusively established the neutral curve, and fully described the cross-wave field at the initial stages of its generation. The dynamics of the sloshing motion was explored, and the connection between the sloshing motion and the (spanwisely non-dispersive) 3-D cross-wave equations (3.8)–(3.14) has been established. The sloshing waves are due to a kind of modulational instability which is not related to the Benjamin-Feir instability of uniform wavetrains.
Part II

The Stability and Nonlinear Evolution of Edge Waves

II.1 Introduction

The classical linear water-wave problem on a wedge-shaped beach can be formulated in terms of a velocity potential $\phi$, which satisfies the Laplace equation, a free surface boundary condition, and a sea-bed condition, namely

$$\Delta \phi = 0,$$

$$g \phi_y + \phi_t = 0, \quad y = 0,$$

$$\phi_y = -\phi_x \tan \alpha, \quad y = -x \tan \alpha,$$

where $x$ is out to sea, $y$ is vertical, $z$ is along the beach, and $\alpha$ is the angle of the beach. Stokes (1846) first noted a solution of these equations that represents edge waves, which are propagating along the beach with their crests perpendicular to the shoreline and have an amplitude that decays exponentially off the coast. This solution for the velocity potential $\phi$ is

$$\phi = \frac{ga}{\omega} e^{-kx \cos \alpha + kx \sin \alpha} \sin(kz - \omega t),$$

where $a$ is the amplitude of the edge wave, and $k$ and $\omega$ are the wavenumber and frequency. The dispersion relation between $\omega$ and $k$ is

$$\omega^2 = gk \sin \alpha.$$  

Ursell (1952) further discovered that the Stokes solution is only one of many possible edge-wave modes and that successively more possible modes arise as $\alpha$ decreases. A second one is possible for $\alpha < \frac{1}{6}\pi$, a third for $\alpha < \frac{1}{10}\pi$ and so on. The $n$th mode appears as $\alpha$ drops below $\frac{\pi}{2(2n+1)}$ and its dispersion relation is

$$\omega^2 = gk \sin(2n + 1)\alpha.$$  

There is also a continuous spectrum of solutions with $\omega^2 > gk$ to complete the representation of general disturbances.

Edge waves are very distinctive on a beach because their maximum amplitudes are on the shoreline. It is now believed that they are responsible for the formation of beach cusps (Guza & Inman 1975) and the generation of rip currents and periodic circulation cells in the nearshore region (Bowen & Inman 1969).

The generation of edge waves has been intensively studied, both experimentally and theoretically, in the last forty years. Greenspan (1956) first demonstrated that large-scale edge waves can be excited by atmospheric forcing due to storms moving along the coastline. For smaller-scale edge waves, in an attempt to explain the experimental observations of Galvin (1965) and Bowen & Inman (1969), Guza & Davis (1974) proposed the nonlinear interaction mechanism of edge waves with incoming
wavetrains. Using the shallow-water approximation, they showed that a monochromatic harmonic wavetrain of frequency $\omega$, normally incident and strongly reflected on a beach, is unstable to subharmonic standing edge-wave perturbations of frequency $\frac{1}{2} \omega$. Guza & Inman (1975)'s experiments on a bounded beach indicated that this subharmonic resonance was the strongest, and a subharmonic standing edge-wave was preferentially excited. A synchronous edge wave (same period as the incident wave) was sometimes also excited, but the generation was a higher-order, weaker resonance, and was evident only when the subharmonic resonance was excluded by the beach geometry. If the edge-wave coastline antinode number was low, edge waves would reach a steady state. More interesting was the fact that if the wavenumber was high, the number of edge-wave antinodes sometimes alternated between adjacent integers. In a further development of the theory, Whitham (1976) calculated the leading-order nonlinear corrections to the linear dispersion relation of travelling Stokes edge waves and thus deduced that propagating finite-amplitude edge waves are always unstable to large-scale modulations. Later, Minzoni & Whitham (1977) studied the excitation of standing subharmonic edge waves by a normally incident, strongly reflected wavetrain. They formulated the problem in the full water-wave theory without making the shallow-water approximation and solved it for beach angles $\alpha = \pi/2N$, where $N$ is an integer. Their work confirms the results from the shallow-water theory in the small-beach-angle limit. Akylas (1983) studied the large-scale temporal and spatial modulations of subharmonic edge waves excited by resonant interactions with normally incident, strongly reflected wavetrains, and derived equations governing the modulations of edge-wave envelopes. He then re-examined the modulational stability of a propagating edge-wave train and confirmed that the instability, predicted by Whitham (1976), indeed leads to a series of envelope solitons. He also found that the steady state standing subharmonic edge wave with the wavenumber at exact subharmonic resonance is unstable to large-scale modulations.

Although much work has been done as mentioned above, some important questions remain open. Firstly, the effect of the beach geometry on edge waves has not been analytically studied. All the previous analytical work was done on an open beach. But if the beach is bounded by two sidewalls, which is always the case in experiments, this beach geometry will affect the edge-wave dynamics, sometimes even exclude the excitation of edge waves. This effect shows clearly in Guza & Inman (1975)'s experiments. Secondly, the nonlinear evolution of subharmonic edge waves on a wide beach is still not clear. Since in this situation the spatial large-scale modulation arises as well as the temporal one, the evolution equations of these modulations have been derived by Akylas (1983). But what these equations imply about the edge-wave evolution is not known. One steady state standing subharmonic edge wave was found by Akylas (1983) to be unstable to large-scale modulations. The significance of this finding is not clear. It is worth noting here that the relevant experiments conducted by Guza & Inman (1975) show that the number of edge-wave antinodes sometimes alternated between adjacent integers. This phenomenon has yet to be explained.

In this paper, the above two problems are studied. As to the first one, the beach geometry is found to introduce an additional detuning term to the governing equations which affects the edge-wave dynamics. As to the second problem, the stability of all possible steady state standing edge-wave modes to large-scale disturbances is first examined. Regions of stable and unstable modes are analytically specified. The unstable mode found by Akylas (1983) is shown to fall in the unstable-mode region. The significance of these stable and unstable modes is also discussed. Finally, numerical calculations of the equations governing the large-scale edge-wave modulations are carried out. The antinode-number alternation phenomenon is found in the numerical results. The nonlinear evolution of edge waves on a wide beach is commented on at the end of this paper.
II.2 Formulation

Equations governing the edge-wave amplitudes are a little different on a wide bounded beach and on an open beach. We treat them separately.

II.2.1 Edge waves on an open beach

Consider a normally incident and strongly-reflected wavetrain of frequency $\omega$ interacting with two Stokes edge-wave packets of frequency $\frac{1}{2}\omega$ propagating in opposite directions along an open beach of angle $\alpha$. The undisturbed incident wave and its reflexion is described by a potential

$$\phi_{\text{inc}} = \frac{ga}{\omega} S_1(z, y)e^{-i\omega t} + \text{c.c.},$$

(2.1)

where $g$ is the gravitational acceleration, $a$ is the amplitude scale of the incident wave, and $S_1(z, y)$ is a real-valued function which satisfies the Laplace equation and corresponding boundary conditions (see Minzoni & Whitham 1977 for details). This incident-wave field will be modified when edge waves are excited.

Following Akylas (1983), a suitable expansion of the velocity potential for the two Stokes edge-wave packets and the incident wavetrain is of the form (dimensions have been restored except as noted)

$$\phi = \frac{ga}{\omega} \left\{ e^{-\frac{i}{2}kz \cos \alpha + ky \sin \alpha} [A(X, Y, Z, T)e^{i(kz - \frac{1}{2}\omega t)} + B(X, Y, Z, T)e^{i(-kz - \frac{1}{2}\omega t)} + \text{c.c.}] ight\} + [S(x, y; X, Y, Z, T)e^{-i\omega t} + \text{c.c.}] + \text{higher order terms,} \right\}$$

(2.2)

where

$$k = \frac{\omega^2}{4g \sin \alpha}, \quad \epsilon = \frac{4ka}{\sin \alpha} = \frac{a \omega^2}{g \sin^2 \alpha}, \quad \text{which is assumed small,}$$

(2.3)

$$X = \mu 4kz, \quad Y = \mu 4ky, \quad Z = \mu 4kz, \quad T = \mu \omega t,$$

(2.4)

and $\mu^{-1} \gg 1$ is the dimensionless modulation scale.

A multiple-scale perturbation method is used to determine the evolution of the edge-wave amplitudes $A$ and $B$. It is found that $A, B$ satisfy the following equations on the shoreline:

$$A_T + A_Z = -i\mu A_{ZZ} + \epsilon \left\{ \frac{1}{2} \cos \alpha \chi(\alpha) S_0 B^* - \frac{i}{128} A^2 A^* + \delta BB^* A \right\} \quad (X = 0, Y = 0)$$

(2.5)

$$B_T - B_Z = -i\mu B_{ZZ} + \epsilon \left\{ \frac{1}{2} \cos \alpha \chi(\alpha) S_0 A^* - \frac{i}{128} B^2 B^* + \delta AA^* B \right\} \quad (X = 0, Y = 0)$$

(2.6)

where

$$S_0 = S_1(0, 0),$$

(2.7)

$$\chi(\alpha) S_0 = k \int_{0}^{\infty} S_1(z, 0)e^{-2kz \cos \alpha} dz,$$

(2.8)

and when $\alpha = 2\pi/N$,

$$\delta = \frac{1}{64} \left\{ -32N \chi^2 \sin 2\alpha + i(3 + \frac{32N \sin 2\alpha}{\pi} \int_{0}^{\infty} C_i^{2} \frac{dl}{l - \omega^2} \right\},$$

(2.9)
where the integral in (2.9) is to be interpreted as a principle value. (See Minzoni & Whitham 1977)

If the beach-angle $\alpha$ is small, $\chi$ and $\delta$ can be evaluated asymptotically (see Minzoni & Whitham 1977, Akylas 1983):

$$
\chi(\alpha) \sim \frac{1}{2e^2}, \quad \delta(\alpha) \sim \frac{1}{256}(-1.8413 + 0.4942i) \quad (2.10)
$$
$$
= -0.72 \times 10^{-2} + 0.19 \times 10^{-2}i \quad (\alpha \to 0).
$$

When the balance $\epsilon = \mu$ is chosen, the dispersive terms in equations (2.5) and (2.6) are relatively small and can be neglected. Further, if we simply denote $\frac{1}{2} \cos \alpha \chi(\alpha)S_0$ as $S_0$, equations for $A$ and $B$ on the shoreline ($X = 0, Y = 0$) will reduce to

$$
A_T + Cg A_Z = S_0 B^* + i\gamma A^2 A^* + \delta BB^* A \quad (2.11)
$$
$$
B_T - Cg B_Z = S_0 A^* + i\gamma B^2 B^* + \delta AA^* B \quad (2.12)
$$

where

$$
Cg = 1, \quad \gamma = -\frac{1}{128} \quad (2.13)
$$

and $\delta$ is as given by (2.9), or (2.10) if the beach angle $\alpha$ is small.

II.2.2 Edge waves on a wide bounded beach

When the beach is wide but bounded by two sidewalls normal to the shoreline, the forced edge-wave wavelength and the free edge-wave wavelength mismatch will also affect the dynamics of edge waves.

Consider a beach of angle $\alpha$, which is bounded by two sidewalls at $z = 0$ and $z = b$. A normally incident wave of frequency $\omega$ comes to the shore and is strongly reflected. The generated edge wave has the primary wavenumber $k_0 = \frac{N\pi}{b}$, where $N$ is the number of the coastline edge-wave antinodes and is assumed to be large. This edge wave is both temporally and spatially modulated.

Suppose the undisturbed incident wave field is described by a potential

$$
\phi_{\text{inc}} = \frac{g\alpha}{\omega} S_1(z, y) e^{-i\omega t} + c.c., \quad (2.14)
$$

where $S_1(z, y)$ is a real-valued function, and $\alpha$ is the incident-wave amplitude scale which is assumed small. Introduce the small perturbation parameter

$$
\epsilon = \frac{a\omega^2}{g \sin^2 \alpha}, \quad (2.15)
$$

and assume that

$$
\epsilon N = h \sim O(1), \quad (2.16)
$$
where $h$ is a dimensionless measure of the beach width. The suitable expansion of the velocity potential for the edge-wave packets and the incident wave is of the following dimensional form:

$$
\phi = \frac{g \alpha}{\omega} \left\{ e^{-\frac{1}{2} \epsilon - kz \cos \alpha + ky \sin \alpha} [A(X, Y, Z, T)e^{i(k_0z - \frac{1}{2}\omega t)} + B(X, Y, Z, T)e^{i(-k_0z - \frac{1}{2}\omega t)} + c.c.] 
+ [S(x, y; X, Y, Z, T)e^{-i\omega t} + c.c.] + \text{higher order terms}, \right\}
$$

(2.17)

where

$$
k = \frac{\omega^2}{4g \sin \alpha}, \quad k_0 = \frac{N\pi}{b},
$$

(2.18)

$$X = \epsilon 4k \epsilon, \quad Y = \epsilon 4k_y, \quad Z = \epsilon 4k_0z, \quad T = \epsilon \omega t.
$$

(2.19)

A perturbation analysis results in the following equations for $A$ and $B$ on the shoreline $(X = 0, Y = 0)$:

$$A_T + A_Z = iJA + S_0B^* + i\gamma A^2A^* + \delta BB^* A
$$

(2.20)

$$B_T - B_Z = iJB + S_0A^* + i\gamma B^2B^* + \delta AA^*B
$$

(2.21)

$$A = B, \quad Z = 0,4\pi h
$$

(2.22)

where

$$J = \frac{1 - \frac{hk}{4\epsilon}}{4\epsilon} \text{ is the detuning parameter,}
$$

(2.23)

$$\gamma = -\frac{1}{128}, \quad \delta = \frac{1}{64} \left\{ -32N \chi^2 \sin 2\alpha + i(3 + \frac{32N \sin 2\alpha}{\pi} \int_0^\infty \frac{C_l^2}{l - \omega^2} dl) \right\}.
$$

(2.24)

With a stretching of the $Z$ coordinate, the above equations can be rewritten as

$$A_T + CgA_Z = iJA + S_0B^* + i\gamma A^2A^* + \delta BB^* A
$$

(2.25)

$$B_T - CgB_Z = iJB + S_0A^* + i\gamma B^2B^* + \delta AA^*B
$$

(2.26)

$$A = B, \quad Z = 0, \pi
$$

(2.27)

with $C_g = \frac{1}{4h}$.

(2.28)

It needs to be pointed out that the value of $C_g$ depends on the actual width of the beach. If the beach is very wide, $C_g$ will be quite small.

When the beach-angle $\alpha$ is small,

$$\gamma = -\frac{1}{128}, \quad \delta = -0.72 \times 10^{-2} + 0.19 \times 10^{-2}i.
$$

(2.29)

Dissipation may be introduced by adding a linear damping term to the equations (2.25) and (2.26).
II.3 Steady state standing edge waves and their stability

On a bounded beach, since the possible free edge-wave frequencies are far apart, an incident wave will only be able to excite a single subharmonic standing edge wave, if any. When it does, this edge wave will reach a steady final state (see Guza & Inman 1975, Minzoni & Whitham 1977). But if the beach is open or bounded but wide, since the possible resonant frequencies are so close together, an incident wave usually can excite several adjacent standing edge-wave modes, so that large-scale amplitude modulations will arise. If this is the case, there are serious and important questions as to how edge waves evolve and what their final states may be.

A first step toward answering these questions is to study the stability of steady state standing edge-wave modes to large-scale disturbances. The equations governing these large-scale disturbances on the shoreline are as derived earlier. If dissipation is also included, they take the form

\[ A_T + C g A_Z = i (J + iL) A + S_0 B^* + i \gamma A^2 A^* + \delta B B^* A \]  
\[ B_T - C g B_Z = i (J + iL) B + S_0 A^* + i \gamma B^2 B^* + \delta A A^* B \]  
\[(A = B, \ Z = 0, \pi \text{ on a wide bounded beach})\]  

where \( J \) is the detuning parameter, (on an open beach, \( J = 0, \)) and \( L > 0 \) is the linear damping coefficient. Physical arguments show that \( \text{Re}(\delta) < 0 \) (Guza & Bowen 1976).

II.3.1 Steady state standing edge waves

The steady state standing edge-waves are described by

\[ A = a_K e^{-iKZ}, \]
\[ B = a_K e^{iKZ}, \]  

where \( K \) is any integer if the beach is wide but bounded, any real number if it is unbounded (or open). Each \( K \) represents one steady state standing edge-wave mode. \( a_K \) is a constant and is determined by the equation

\[ i (J + KC_g + iL) a_K + S_0 a_K^* + (\delta + i\gamma) a_K^2 a_K^* = 0. \]  

If \( a_K = re^{i\theta} \), it is a simple matter to show that \( r \) is given by

\[ r^2 = \frac{L \delta_r - (J + KC_g)(\delta_i + \gamma) + \sqrt{S_0^2 |\delta + i\gamma|^2 - [L(\delta_i + \gamma) + (J + KC_g)\delta_r]^2}}{|\delta + i\gamma|^2}, \]  

and \( \theta \) is given by the equation

\[ -i (J + KC_g) + L - (\delta + i\gamma)r^2 = S_0 e^{-2i\theta}, \]  

where \( \delta_r \) and \( \delta_i \) are the real and imaginary parts of \( \delta \).
Clearly such steady edge waves exist only for a limited range of the parameter combination \( J + KC_g \).

If \( \delta_i + \gamma < 0 \), which is the case for small-angle beaches, such steady states exist only when

\[
- \sqrt{S_0^2 - L^2} \equiv J_{\min} < J + KC_g < J_{\max} \equiv \frac{|S_0(\delta + i\gamma)| - |L(\delta_i + \gamma)|}{|\delta_r|}.
\]  

(3.7)

If \( \delta_i + \gamma > 0 \), such steady states exist when

\[
- \frac{|S_0(\delta + i\gamma)| - |L(\delta_i + \gamma)|}{|\delta_r|} \equiv J_{\min} < J + KC_g < J_{\max} \equiv \sqrt{S_0^2 - L^2}.
\]

(3.8)

II.3.2 Linear stability analysis

The stability of the steady mode (3.3) with \( K = 0 \) on an open beach was studied by Akylas (1983) and was found to be unstable to large-scale disturbances. Now a study is made of the stability of all possible steady modes of the form (3.3) on both wide bounded and open beaches.

Assume that \( A, B \) as given by (3.3) are slightly disturbed and are written as

\[
A = (a_K + \tilde{a}(Z,T))e^{-iKZ}
\]

\[
B = (a_K + \tilde{b}(Z,T))e^{iKZ}
\]

(3.9)

where \( \tilde{a}, \tilde{b} \) are infinitesimal disturbances. After (3.10) is substituted into the equations (3.1), (3.2) and higher order terms neglected, \( \tilde{a}, \tilde{b} \) are found to satisfy the following linear equations

\[
\dot{\tilde{a}} + CG\tilde{a}_Z = [i(J + KCg + iL) + (\delta + 2i\gamma)a_Ka_K^*\tilde{a} + i\gamma a_K^2\tilde{a}^* + \delta a_K a_K^* \tilde{b} + (S_0 + \delta a_K^2)\tilde{b}^*]
\]

(3.10)

\[
\dot{\tilde{b}} - CG\tilde{b}_Z = [i(J + KCg + iL) + (\delta + 2i\gamma)a_Ka_K^*\tilde{b} + i\gamma a_K^2\tilde{b}^* + \delta a_K a_K^* \tilde{a} + (S_0 + \delta a_K^2)\tilde{a}^*]
\]

(3.11)

\[
(\tilde{a} = \tilde{b}, \ Z = 0, \pi \quad \text{on a wide bounded beach})
\]

With the notation \( \tilde{c} \equiv \tilde{a}^*, \tilde{d} \equiv \tilde{b}^* \), equations for \( \tilde{a}, \tilde{b}, \tilde{c} \) and \( \tilde{d} \) are easily obtained from (3.10) and (3.11) to be

\[
\dot{\tilde{a}} + CG\tilde{a}_Z = [i(J + KCg + iL) + (\delta + 2i\gamma)a_Ka_K^*\tilde{a} + \delta a_K a_K^* \tilde{b} + i\gamma a_K^2\tilde{c} + (S_0 + \delta a_K^2)\tilde{d}]
\]

\[
\dot{\tilde{b}} - CG\tilde{b}_Z = [i(J + KCg + iL) + (\delta + 2i\gamma)a_Ka_K^*\tilde{b} + \delta a_K a_K^* \tilde{a} + i\gamma a_K^2\tilde{d} + (S_0 + \delta a_K^2)\tilde{c}]
\]

\[
\dot{\tilde{c}} + CG\tilde{c}_Z = [-i\gamma a_K^2\tilde{a} + (S_0 + \delta a_K^2)\tilde{b} + [i(J + KCg + iL) + (\delta + 2i\gamma)a_Ka_K^*]\tilde{c} + \delta^*a_K a_K^* \tilde{d}]
\]

\[
\dot{\tilde{d}} - CG\tilde{d}_Z = [-i\gamma a_K^2\tilde{b} + (S_0 + \delta a_K^2)\tilde{a} + [i(J + KCg + iL) + (\delta + 2i\gamma)a_Ka_K^*]\tilde{d} + \delta^*a_K a_K^* \tilde{c}]
\]

(3.12)

\[
(\tilde{a} = \tilde{b}, \ Z = 0, \pi \quad \text{on a wide bounded beach})
\]

(\( \tilde{c} = \tilde{d}, \ Z = 0, \pi \quad \text{on a wide bounded beach} \))

With the following change of variables

\[
W_1 = \tilde{a} + \tilde{b},
\]

\[
W_2 = \tilde{a} - \tilde{b},
\]

\[
W_3 = \tilde{c} + \tilde{d},
\]

\[
W_4 = \tilde{c} - \tilde{d},
\]

(3.13)
\( W_1, W_2, W_3, W_4 \) are found to satisfy the equations

\[
\begin{align*}
W_{1T} + C_g W_{2Z} &= [i(J + KC_g + iL) + (2\delta + 2i\gamma)a_K a_K^*]W_1 + (S_0 + (\delta + i\gamma)a_K^2)W_3 \\
W_{2T} + C_g W_{1Z} &= [i(J + KC_g + iL) + 2i\gamma a_K a_K^*]W_2 + ((i\gamma - \delta)a_K^2 - S_0)W_4 \\
W_{3T} + C_g W_{4Z} &= (S_0 + (\delta + i\gamma)a_K^2)^*W_1 + [i(J + KC_g + iL) + (2\delta + 2i\gamma)a_K a_K^*]^*W_3 \\
W_{4T} + C_g W_{3Z} &= ((i\gamma - \delta)a_K^2 - S_0)^*W_2 + [i(J + KC_g + iL) + 2i\gamma a_K a_K^*]^*W_4
\end{align*}
\]

(3.14)  
(\( W_2 = 0, \quad Z = 0, \pi \) on a wide bounded beach)  
(\( W_4 = 0, \quad Z = 0, \pi \) on a wide bounded beach)

Due to the boundary conditions, the normal mode analysis takes slightly different forms for an open beach and for a wide bounded beach.

1. On an open beach, the conventional normal mode analysis assumes that

\[
\begin{align*}
W_1 &= \overline{W}_1 e^{imZ + \sigma T} \\
W_2 &= \overline{W}_2 e^{imZ + \sigma T} \\
W_3 &= \overline{W}_3 e^{imZ + \sigma T} \\
W_4 &= \overline{W}_4 e^{imZ + \sigma T}
\end{align*}
\]

(3.15)

where the disturbance wavenumber \( m \) takes on any real value.

The eigenvalue \( \sigma \) is related to the wavenumber \( m \) by the equation

\[
\begin{vmatrix}
P_1 - \sigma & P_2 & -imC_g & 0 \\
P_2^* & P_1^* - \sigma & 0 & -imC_g \\
-imC_g & 0 & P_3 - \sigma & P_4 \\
0 & -imC_g & P_4^* & P_3^* - \sigma
\end{vmatrix} = 0, \quad (3.16)
\]

Where

\[
\begin{align*}
P_1 &= P_1(K) = i(J + KC_g) - L + 2(\delta + i\gamma)a_K a_K^* \\
P_2 &= P_2(K) = S_0 + (\delta + i\gamma)a_K^2 \\
P_3 &= P_3(K) = i(J + KC_g) - L + 2i\gamma a_K a_K^* \\
P_4 &= P_4(K) = (i\gamma - \delta)a_K^2 - S_0
\end{align*}
\]

(3.17)

and \( a_K \) is as given by the equation (3.4).

Notice that

\[
P_2 = -\frac{(iJ - L)a_K}{a_K^*}, \quad P_4 = \frac{a_K}{a_K^*} P_3 \quad (3.18)
\]
due to the equation (3.4).

2. On a wide bounded beach, due to the sidewall boundary conditions, the normal mode analysis assumes that

\[
\begin{align*}
W_1 &= \overline{W}_1 \cos mZ e^{\sigma T} \\
W_2 &= \overline{W}_2 \sin mZ e^{\sigma T} \\
W_3 &= \overline{W}_3 \cos mZ e^{\sigma T} \\
W_4 &= \overline{W}_4 \sin mZ e^{\sigma T}
\end{align*}
\]

(3.19)

where the disturbance wavenumber \( m \) takes on any integer value.

The eigenvalue \( \sigma \) is related to the wavenumber \( m \) by the equation

\[
\begin{vmatrix}
P_1 - \sigma & P_2 & -mC_g & 0 \\
P_2^* & P_1^* - \sigma & 0 & -mC_g \\
mC_g & 0 & P_3 - \sigma & P_4 \\
0 & mC_g & P_4^* & P_3^* - \sigma
\end{vmatrix} = 0, \quad (3.20)
\]
and \( P_1, P_2, P_3, P_4 \) are as given by (3.17).

After some algebra, it is found that (3.16) and (3.20) lead to the same quartic equation for \( \sigma \):

\[
\sigma^4 - [P_1 + P_1^* - 2L]\sigma^3 + [P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_1^*) + 2(mCg)^2]\sigma^2 \\
-[(mCg)^2(P_1 + P_1^* - 2L) - 2L(P_1 P_1^* - P_2 P_2^*)] \sigma \\
+(mCg)^2[P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 + (mCg)^2] = 0.
\]  

(3.21)

From (3.4) and (3.17), it is readily shown that

\[
P_1 + P_1^* = -2L + 4\delta a_K a_K^* (\leq 0),
\]

\[
P_1 P_1^* - P_2 P_2^* = 4a_K a_K^* \sqrt{S_0^2|\delta + i\gamma|^2 - [L(\delta + \gamma) + (J + KCg)\delta_r]^2} (\geq 0),
\]

\[
P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 = -4[L\delta a_K a_K^* + (J + KCg) + (\delta_i + \gamma)a_K a_K^*](J + KCg + 2\gamma a_K a_K^*).
\]  

(3.22)

In the next section, we will use these equations to determine the stability of steady standing edge waves.

II.3.3 Stability results

A steady edge-wave mode (3.3) is possible if \( \bar{J}_{\text{min}} < J + KCg < \bar{J}_{\text{max}} \), with \( \bar{J}_{\text{min}} \) and \( \bar{J}_{\text{max}} \) given by (3.7) or (3.8). It is unstable if some normal-mode disturbances are unstable, and vice versa. A normal-mode disturbance is unstable if its eigenvalue \( \sigma \) has a positive real part; it is stable otherwise. In the present situation, \( \sigma \) is a root of the quartic equation (3.21). To solve (3.21) for general values of \( K \) and \( m \) looks complicated, but actually it is not difficult. For our purpose, it is not necessary to solve (3.21). The signs of Re(\( \sigma \)) can be determined by simply examining the coefficients of the equation (3.21). The results are given below. (The derivation is given in Appendix 1 at the end of Part two.)

I. If the edge-wave mode (3.3) is such that \( P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 < 0 \), it is unstable. The unstable disturbance wavenumbers \( m \) are confined to the interval

\[
0 < m^2C_g^2 < m_2^2C_g^2 \equiv -(P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4).
\]  

(3.23)

These unstable disturbances are standing waves of growing amplitudes because they have positive real eigenvalues \( \sigma \).

II. If the edge-wave mode (3.3) is such that \( 0 < P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 < P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_2) \), it is stable.

III. If the edge-wave mode (3.3) is such that \( P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 > P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_2) \), it is unstable. The unstable disturbances have wavenumbers \( m \) such that

\[
m^2C_g^2 < m_2^2C_g^2
\]  

(3.24)

where

\[
n_2^2C_g^2 \equiv \frac{2L(P_1 P_1^* - P_2 P_2^*)(P_1 + P_2)(P_1^* P_1^* - P_2^* P_2^* - 2L)}{(P_1 + P_2 - 2L)((P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4) - (P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_2)))}.
\]  

(3.25)
These unstable disturbances are travelling waves of growing amplitudes since their eigenvalues \( \sigma \) are complex.

We can further show from the equation (3.21) that the eigenvalue

\[
\sigma \longrightarrow \pm mC_g i + \sigma^{(0)} \quad \text{as} \quad mC_g \longrightarrow \infty, \tag{3.26}
\]

where \( \sigma^{(0)} \) is the root of the quadratic equation

\[
\sigma^{(0)}^2 \frac{P_1 + P_1^* - 2L}{2} - \frac{P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 - [P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_2)]}{4} = 0. \tag{3.27}
\]

The derivation of (3.26) and (3.27) is given in Appendix 2 at the end of Part two.

In this third case, because \( P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 > P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_2) \), one root of the equation (3.27) is real and positive and the other one is real and negative. The unstable disturbance takes the positive \( \sigma^{(0)} \). Since \( \sigma^{(0)} \) is independent of \( mC_g \), very short disturbance waves tend to have the same growth rate \( \sigma^{(0)} \).

It should be noticed that edge-wave modes of this kind are very unusual, because they are unstable to small-scale disturbances. Interpretation of existence of these modes requires caution and will be discussed later.

The above stability results are not affected to any degree by a proportional change of \( \gamma \) and \( \delta \), and it is easy to show that

\[
\sigma(\beta \gamma, \beta \delta) = \sigma(\gamma, \delta). \tag{3.28}
\]

A change in \( S_0 \) causes \( \sigma \) to change in a simple way:

\[
\sigma(\beta m, \beta S_0, \beta L, \beta (J + K C_g)) = \beta \sigma(m, S_0, L, (J + K C_g)). \tag{3.29}
\]

These facts are helpful to determine the stability structure when different values of \( \gamma, \delta \) and \( S_0 \) are taken.

The above general stability results immediately give the precise stability structure of steady state standing edge-wave modes on a given beach. As an example, we determine this structure on a mildly sloping beach.

On a mildly sloping beach,

\[
\gamma = -\frac{1}{128}, \quad \delta = -0.72 \times 10^{-2} + 0.19 \times 10^{-2} i. \tag{3.30}
\]

The damping coefficient \( L \) is small but hard to determine. It is usually set to be zero for simplicity.

In this case, since \( \delta + \gamma < 0 \), \( \bar{J}_{\min} \) and \( \bar{J}_{\max} \) are given by the equation (3.7) as

\[
\bar{J}_{\min} = -\sqrt{S_0^2} = -S_0, \quad \bar{J}_{\max} = \frac{|S_0(\delta + i\gamma)|}{|\delta_r|} = 1.29S_0. \tag{3.31}
\]

Steady edge-wave modes (3.3) are possible if

\[
-S_0 < J + K C_g < 1.29S_0. \tag{3.32}
\]

The stability structure is as shown in Figure 9.
Figure 9: intervals of stable and unstable steady edge-wave modes (3.3) and regions of their stable and unstable disturbances on a mildly sloping beach. The unstable steady modes (3.3) are in the interval I and III. The unstable disturbances are in the dotted area.

Interval I: $-S_0 < J + KC_g < 0.82S_0$.

The steady edge-wave modes in this interval are unstable. From the equations (3.22) and (3.23) we get

$$m_1^2 C_g^2 = 4(J + KC_g + (\delta_i + \gamma)a_K^*a_K)(J + KC_g + 2\gamma a_K a_K^*),$$

(3.33)

where

$$a_K a_K^* = \frac{-(J + KC_g)(\delta_i + \gamma) + \sqrt{S_0^2|\delta + i\gamma|^2 - (J + KC_g)^2\delta_i^2}}{|\delta + i\gamma|^2}. (3.34)$$

Values of $m_1 C_g$ are plotted against $J + KC_g$ in Figure 9. The unstable disturbance wavenumbers $m$ are confined in the interval:

$$0 < m^2 C_g^2 < m_1^2 C_g^2,$$

(3.35)

These disturbances are standing waves of growing amplitudes.

Interval II: $0.82S_0 < J + KC_g < 1.24S_0$.

The steady edge-wave modes in this interval are stable.

Interval III: $1.24S_0 < J + KC_g < 1.29S_0$.

The steady edge-wave modes in this interval are unstable. $m_2^2 C_g^2$ is given by the equations (3.22) and (3.25). Since $L = 0$, $m_2 C_g = 0$. Any disturbances with $m_2^2 C_g^2 > 0$ are unstable, and they are travelling waves of growing amplitudes.

The unstable edge-wave mode Akylas (1983) found on an open, mildly sloping beach corresponds to the one in Interval I with $J = 0$, $K = 0$ and $L = 0$. With $S_0 = \frac{1}{4e^2}$ taken, he numerically determined that

$$m_1 C_g \approx 0.052,$$

and found that the eigenvalues $\sigma$ of the unstable disturbances have very small imaginary parts.

Actually, for this mode, (3.33) and (3.34) give

$$a_0 a_0^* = \frac{S_0}{|\delta + i\gamma|},$$

(3.36)
\[ m_1^2 C_g^2 = 8\gamma(\delta_i + \gamma)(a_K^* a_K)^2 \]
\[ = \frac{8\gamma(\delta_i + \gamma)S_0^2}{|\delta + i\gamma|^2} \]
\[ = 0.49 \times 10^{-2}. \]  

(3.37)

Therefore, the exact value for \( m_1 C_g \) is
\[ m_1 C_g = 0.070. \]  

(3.38)

This value can be checked in Figure 9. Moreover, the eigenvalues \( \sigma \) of the unstable disturbances are exactly real and positive, thus they represent standing waves of growing amplitudes.

Different values of \( \gamma, \delta, L \) will slightly change the parameters in Figure 9 and \( m_1 C_g \) and \( m_2 C_g \), but the basic stability structure (as in Figure 9) does not change.

The above stability structure has two distinctive features:

1. Although many steady state standing edge-wave modes of the form (3.3) are unstable to large-scale modulations (interval I in Figure 9), some of them are stable (interval II in Figure 9). These stable ones are an attracting set and will strongly affect the dynamics of edge-wave evolutions.

2. There exists a small interval of steady state standing edge-wave modes which are unstable to very short modulational disturbances (interval III in Figure 9). This may seem to suggest a mechanism of short-wave excitation by a long wave. If it really occurs in edge waves, it will invalidate the edge-wave modulational equations we derived before. But we need to be cautious here. For very short waves, surface tension and dispersion become important. But these effects are neglected in our formulation. Obviously more studies are needed to clarify the issue.
II.4 Nonlinear evolution of edge waves on a wide beach

The above results on the stability of steady state standing edge waves provide some insight on the nonlinear evolution of edge waves. On a wide bounded beach, the steady edge-wave modes of the form (3.3) are discrete. The number of stable and unstable such modes depends on $C_g$ and $J$ in the equations (2.25) and (2.26). Since stable steady modes form an attracting set, if they exist and are excited, the edge wave is likely to be attracted to one of these modes and settle down there. If they do not exist, the edge wave can not settle down to any steady mode, and its evolution will be quite different. In this case, one possibility is that the energy will mostly exchange among a few adjacent discrete modes, and the edge wave will evolve into a limit cycle. This corresponds to the antinode-number alternation phenomenon observed in Guza & Inman (1975)'s experiments.

To further investigate the edge-wave evolution on a wide bounded beach, numerical calculations are carried out for the governing equations (2.25), (2.26) and (2.27). To facilitate the computation, $A, B$ are decomposed in the following form:

$$A = \sum_{K=-\infty}^{\infty} a_K(T) e^{-iKZ}, \quad (4.1)$$  

$$B = \sum_{K=-\infty}^{\infty} a_K(T) e^{iKZ}. \quad (4.2)$$

When this decomposition is substituted into the equations (2.25), (2.26) and (2.27), an infinite-dimensional dynamical system for $a_K(T)$ are obtained. Truncation of this system is necessary for any numerical calculations. The choice for the number of $a_K$'s depends on the accuracy required. A fourth order Runge-Kutta scheme is used for this dynamical system in all our computations. Eleven $a_K$ modes are chosen in most cases, including the two cases shown later in this paper.

For easy comparison with the previous stability results, we consider a beach of small angle $\alpha$, where

$$\gamma = -\frac{1}{128}, \quad \delta = -0.72 \times 10^{-2} + 0.19 \times 10^{-2}i. \quad (4.3)$$

We also take

$$L = 0, \quad S_0 = \frac{1}{4e^2}. \quad (4.4)$$

Depending on the actual width of the beach, $C_g$ and $J$ may take different values.

For convenience, we normalize $S_0 = 1$, and other parameters change to

$$\gamma = -\frac{1}{128} \times 4e^2 = -0.23, \quad \delta = (-0.72 \times 10^{-2} + 0.19 \times 10^{-2}i) \times 4e^2 = -0.21 + 0.056i. \quad (4.5)$$

$L$ is still zero. $J$ and $C_g$ are multiplied by $4e^2$ and are still denoted as $J$ and $C_g$. A stretch of the $T$ coordinate is also needed.

Initially, the edge waves are very small. They are excited by the incident wave. In our computation, we take small "white noise" initial conditions with

$$a_K(0) = 0.01 + 0.01i, \quad K = 0, \pm 1, \pm 2, \ldots \quad (4.6)$$
Many runs with different values of $J$ and $C_g$ have been carried out. When $C_g$ is not very small, i.e. the beach is not very wide, the two solution behaviors for the following two sets of values of $J$ and $C_g$ are found to be typical.

1. $C_g = 1, J = 0$:

In this case, (3.32) shows that steady modes (3.3) exist for $K = 0$ and 1. The steady mode with $K = 0$ is unstable and the one with $K = 1$ is stable. At the initial stage of edge-wave generation, since the $K = 0$ mode is at exact subharmonic resonance, it quickly grows and reaches its steady state amplitude. But its steady state is unstable. It then gradually loses its energy to its side-band modes with $K = \pm 1$ and excites them. Notice that the steady mode with $K = 1$ is stable. When it is excited, it absorbs energy and attracts the edge wave to reach its own steady state. The edge wave finally settles down to this steady mode. The time-evolution of $a_K$'s is plotted in Figure 10.

2. $C_g = 1, J = 0.5$:

In this case, (3.32) shows that steady modes (3.3) exist for $K = -1$ and 0. These two steady modes are both unstable. Therefore, the edge wave can not settle down to any steady mode. Instead, it goes to a periodic state with energy largely confined to a few adjacent modes and exchanging among them, as is illustrated in Figure 11. Notice that this behavior corresponds to the antinode-number alternation phenomenon which was observed in Guza & Inman (1975)'s experiments.

The above two types of edge-wave evolution are very distinctive. They are both possible on a wide bounded beach. The actual beach geometry and the incident wave dictate which one should occur.

For some runs, rapidly oscillating disturbances appear, which is expected from the previous stability results since some important effects, like surface tension and dispersion, are neglected. When this happens, such effects need to be included and more studies is needed. But nevertheless, the qualitative behaviors are still the same as those without the appearance of rapidly oscillating disturbances.

On a wider beach, $C_g$ is smaller, and thus more stable and unstable steady modes (3.3) exist. Expectedly, in this situation, the dynamical system of $a_K$'s will show richer behaviors. For instance, the edge wave not only may go to a steady (stable mode) state or a limit cycle, but also may evolve into a quasi-periodic or even chaotic state. These aspects still remain to be studied.
Figure 10: the time-evolution of $a_K$, $K = 0, \pm 1, \pm 2, \pm 3$. Parameters $J = 0, C_0 = 1$. The solid line: Re($a_K$); the dotted line: Im($a_K$). This edge wave settles down to the stable steady mode (3.3) with $K = 1$.

Figure 11: the time-evolution of $a_K$, $K = 0, \pm 1, \pm 2, \pm 3$. Parameters $J = 0.5, C_0 = 1$. The solid line: Re($a_K$); the dotted line: Im($a_K$). This edge wave evolves into a limit cycle.
II.5 Summary

The stability of steady state standing edge waves to large-scale disturbances has been studied. Regions of stable and unstable edge-wave modes have been determined precisely, and the stability structure obtained analytically. The nonlinear evolution of edge waves on a wide beach has also been considered. An explanation has been found for the edge-wave antinode-number alternation phenomenon observed in Guza & Inman (1975)'s experiments.
Appendix 1. Root analysis of the quartic equation (3.21)

The quartic equation (3.21) is

\[
\sigma^4 - [P_1 + P_1^* - 2L] \sigma^3 + [P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_1^*) + 2(mCg)^2] \sigma^2 \\
- [(mCg)^2(P_1 + P_1^* - 2L) - 2L(P_1 P_1^* - P_2 P_2^*)] \sigma \\
+(mCg)^2[P_1 P_3 + (P_1 P_3)^* + P_2 P_4 + P_2^* P_4 + (mCg)^2] = 0, 
\]  
(A1.1)

Suppose the four roots are \(\sigma_i, i = 1, 2, 3, 4\). Since the equation (A1.1) has real coefficients, complex roots only appear in conjugate pairs. It is well known that these four roots satisfy the following relations:

\[
\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4 = P_1 + P_1^* - 2L \\
(\sigma_1 + \sigma_2)(\sigma_3 + \sigma_4) + \sigma_1 \sigma_2 + \sigma_3 \sigma_4 = P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_1^*) + 2m^2C_g^2 \\
\sigma_1 \sigma_2(\sigma_3 + \sigma_4) + \sigma_3 \sigma_4(\sigma_1 + \sigma_2) = m^2C_g^2(P_1 + P_1^* - 2L) - 2L(P_1 P_1^* - P_2 P_2^*) \\
\sigma_1 \sigma_2 \sigma_3 \sigma_4 = m^2C_g^2(P_1 P_3 + (P_1 P_3)^* + P_2 P_4 + P_2^* P_4 + m^2C_g^2) \]  
(A1.2) (A1.3) (A1.4) (A1.5)

Our purpose is to decide how the signs of the real parts of the roots depend on the coefficients in equation (A1.1).

Before doing the algebra, it should be noted from the equations (3.22) that

\[
P_1 + P_1^* < 0, \\
P_1 + P_1^* - 2L < 0, \\
P_1 P_1^* - P_2 P_2^* > 0, \\
P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_2) > 0. 
\]  
(A1.6)

These facts will be used in the following analysis.

1. If \(P_1 P_3 + (P_1 P_3)^* + P_2 P_4 + P_2^* P_4 < 0\):

In this case, the root relation (A1.5) immediately tells us that at least one root is real positive when \(m^2C_g^2 < -(P_1 P_3 + (P_1 P_3)^* + P_2 P_4 + P_2^* P_4)\). A further analysis will obtain more detailed information about these four roots.

When \(m^2C_g^2 = 0\), it is easy to show that

\[
\sigma_1 = 0, \quad \sigma_2 = -2L, 
\]  
(A1.7)

and \(\sigma_3, \sigma_4\) are given by the quadratic equation

\[
\sigma^2 - (P_1 + P_1^*) \sigma + P_1 P_1^* - P_2 P_2^* = 0. 
\]  
(A1.8)

Since

\[
P_1 P_1^* - P_2 P_2^* > 0, 
\]

and

\[
P_1 + P_1^* < 0, 
\]

it is clear that \(\text{Re}(\sigma_3) < 0, \text{Re}(\sigma_4) < 0\).
When \( m^2 C^2_\sigma \) gets larger but less than \(- (P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4)\), \( \sigma_1 \) will first move away from the origin along the positive \( \text{Re}(\sigma) \) axis, then it will change direction and move back to the origin along the positive \( \text{Re}(\sigma) \) axis. \( \sigma_2 \) always remains real and negative. \( \sigma_3 \) and \( \sigma_4 \) are always in the left half of the complex \( \sigma \) plane.

When \( m^2 C^2_\sigma = -(P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4) \), \( \sigma_1 = 0 \) again.

When \( m^2 C^2_\sigma > -(P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4) \), \( \sigma_1 \) moves onto the negative \( \text{Re}(\sigma) \) axis, and all four roots will stay in the left half of the complex \( \sigma \) plane.

To fully justify the above analysis, we need to prove that, as \( m^2 C^2_\sigma \) changes, no roots will ever cross the \( \text{Im}(\sigma) \) axis onto the right half of the complex \( \sigma \) plane from the left half. This is proved by a contradiction argument. Suppose they do, then they either cross the \( \text{Im}(\sigma) \) axis at the origin or not at the origin.

- If they cross at the origin, the equation (A1.4) tells us that at most one root does that and this root remains on the real axis when it crosses from the left half plane to the right half. This will mean that \( \sigma_1 \sigma_2 \sigma_3 \sigma_4 \) will change sign from positive to negative, which does not happen for any \( m^2 C^2_\sigma \).

- If they do not cross at the origin, since complex roots appear in conjugate pairs, two roots of conjugate pair, say \( \sigma_3 \) and \( \sigma_4 \), will cross the line simultaneously and \( \sigma_3 = -\sigma_4 \). The root relations now become:

\[
\begin{align*}
\sigma_1 + \sigma_2 &= P_1 + P_1^* - 2L \\
\sigma_1 \sigma_2 + \sigma_3 \sigma_4 &= P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_1^*) + 2m^2 C^2_\sigma \\
\sigma_3 \sigma_4 (\sigma_1 + \sigma_2) &= m^2 C^2_\sigma(P_1 + P_1^* - 2L) - 2L(P_1 P_1^* - P_2 P_2^*) \\
\sigma_1 \sigma_2 \sigma_3 \sigma_4 &= m^2 C^2_\sigma(P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 + m^2 C^2_\sigma) 
\end{align*}
\]

From equation (A1.9) and (A1.11) we obtain

\[
\sigma_3 \sigma_4 = m^2 C^2_\sigma - 2L(P_1 P_1^* - P_2 P_2^*) \\
\frac{P_1 + P_1^* - 2L}{P_1 + P_1^* - 2L}
\]

Therefore, from (A1.10) and (A1.12) we get

\[
\begin{align*}
\sigma_1 \sigma_2 &= m^2 C^2_\sigma + P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_1^*) + \frac{2L(P_1 P_1^* - P_2 P_2^*)}{P_1 + P_1^* - 2L} \\
\sigma_1 \sigma_2 &= m^2 C^2_\sigma + \frac{m^2 C^2_\sigma(P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 + \frac{2L(P_1 P_1^* - P_2 P_2^*)}{P_1 + P_1^* - 2L}}{m^2 C^2_\sigma - \frac{2L(P_1 P_1^* - P_2 P_2^*)}{P_1 + P_1^* - 2L}} 
\end{align*}
\]

The above two equations are consistent only if there exists \( m^2 C^2_\sigma \) such that

\[
m^2 C^2_\sigma = \frac{2L(P_1 P_1^* - P_2 P_2^*)}{(P_1 + P_2 - 2L)((P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4) - (P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_2)))}
\]

But such a value for \( m^2 C^2_\sigma \) does not exist since the right hand side is negative.

The above analysis concludes that if \( P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 < 0 \),

- when \( m^2 C^2_\sigma < m^2 C^2_\sigma \equiv -(P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4) \), one root is real positive. The other three roots have negative real parts.
• When \( m^2C_g^2 > m_1^2C_g^2 \), all four roots have negative real parts.

2. If \( P_1P_3 + (P_1P_3)^* + P_2P_4^* + P_2^*P_4 > 0 \):

In this case, when \( m^2C_g^2 = 0 \),
\[ \sigma_1 = 0, \quad \sigma_2 = -2L, \]
and \( \sigma_3, \sigma_4 \) are given by
\[ \sigma^2 - (P_1 + P_1^*)\sigma + P_1P_1^* - P_2P_2^* = 0 \]
and \( \text{Re}(\sigma_3) < 0, \text{Re}(\sigma_4) < 0 \) as discussed before.

When \( m^2C_g^2 > 0 \) and small, since \( \sigma_1, \sigma_2, \sigma_3, \sigma_4 > 0 \), \( \sigma_1 \) will move onto the negative \( \text{Re}(\sigma) \) axis, while \( \sigma_2 \) remains real negative. \( \sigma_3 \) and \( \sigma_4 \) are somewhere in the left half of the complex \( \sigma \) plane. So the four roots all have negative real parts. As \( m^2C_g^2 \) gets larger, the main interest is whether some roots could cross the \( \text{Im}(\sigma) \) axis into the right half plane. If they do, since \( \sigma_1 \sigma_2 \sigma_3 \sigma_4 \) is now always positive, they could not cross through the origin and have to cross over from somewhere else on the \( \text{Im}(\sigma) \) axis. The analysis for the case \( P_1P_3 + (P_1P_3)^* + P_2P_4^* + P_2^*P_4 < 0 \) shows that when the roots cross the imaginary axis, \( m^2C_g^2 \) has to be

\[
m^2C_g^2 = -\frac{2L(P_1P_1^* - P_2P_2^*)(P_1 + P_2)(P_1P_1^* - P_2P_2^* - 2L)}{(P_1 + P_2 - 2L)((P_1P_3 + (P_1P_3)^* + P_2P_4^* + P_2^*P_4) - (P_1P_1^* - P_2P_2^* - 2L(P_1 + P_2)))}
\]

(\text{A1.17})

• When \( P_1P_3 + (P_1P_3)^* + P_2P_4^* + P_2^*P_4 > P_1P_1^* - P_2P_2^* - 2L(P_1 + P_2) \), the right hand side of (\text{A1.17}) is positive, and such a value for \( m^2C_g^2 \) exists. When \( m^2C_g^2 \) gets larger than this value, the conjugate pair of roots will cross the imaginary axis onto the right half plane from somewhere other than the origin. As \( m^2C_g^2 \) gets even larger, this root pair will stay in the right half plane and will not move back onto the left half plane. The equations (\text{A1.2}) and (\text{A1.5}) show that the other two roots stays in the left half plane for all values of \( m^2C_g^2 \).

• When \( P_1P_3 + (P_1P_3)^* + P_2P_4^* + P_2^*P_4 < P_1P_1^* - P_2P_2^* - 2L(P_1 + P_2) \), the right hand side of (\text{A1.17}) is negative, and such a value for \( m^2C_g^2 \) does not exist. Thus the roots can not cross the imaginary axis and they always stay in the left half plane.

The above analysis concludes that

• If \( P_1P_3 + (P_1P_3)^* + P_2P_4^* + P_2^*P_4 > P_1P_1^* - P_2P_2^* - 2L(P_1 + P_2) \), denoting
\[
m^2C_g^2 \equiv -\frac{2L(P_1P_1^* - P_2P_2^*)(P_1 + P_2)(P_1P_1^* - P_2P_2^* - 2L)}{(P_1 + P_2 - 2L)((P_1P_3 + (P_1P_3)^* + P_2P_4^* + P_2^*P_4) - (P_1P_1^* - P_2P_2^* - 2L(P_1 + P_2)))}
\]

(\text{A1.18})

then, when
\[
m^2C_g^2 < m^2C_g^2,
\]
all four roots have negative real parts. When
\[
m^2C_g^2 > m^2C_g^2,
\]
two complex roots of conjugate pair have positive real parts, and the other two roots have negative real parts.

• If \( 0 < P_1P_3 + (P_1P_3)^* + P_2P_4^* + P_2^*P_4 < P_1P_1^* - P_2P_2^* - 2L(P_1 + P_2) \), all four roots have negative real parts.
Appendix 2. The asymptotic behavior of the roots $\sigma$ of the quartic equation (3.21) for large values of $mC_g$

The quartic equation (3.21) is:

$$\begin{align*}
\sigma^4 - [P_1 + P_1^* - 2L]\sigma^3 &+ [P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_1^*) + 2(mC_g)^2]\sigma^2 \\
-[(mC_g)^2(P_1 + P_1^* - 2L) - 2L(P_1 P_1^* - P_2 P_2^*)]\sigma &+ (mC_g)^2[P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4 + (mC_g)^2] = 0,
\end{align*}$$

(A2.1)

When $mC_g$ is large, the following expansion for $\sigma$ is suitable:

$$\sigma = \sigma^{(1)} mC_g + \sigma^{(0)} + \frac{\sigma^{(-1)}}{mC_g} + \ldots$$

(A2.2)

When this expansion is substituted into (A2.1) and terms of the same order in $mC_g$ are collected, we obtain the relations

$$\begin{align*}
(\sigma^{(1)} - 1)^2 & = 0 \quad \text{(A2.3)} \\
(4\sigma^{(0)} - (P_1 + P_1^* - 2L))(\sigma^{(1)} + 1)\sigma^{(1)} & = 0 \quad \text{(A2.4)} \\
6\sigma^{(1)} \sigma^{(0)^2} + 2\sigma^{(0)^2} - 3(P_1 + P_1^* - 2L)\sigma^{(1)^2} \sigma^{(0)} - (P_1 + P_1^* - 2L)\sigma^{(0)} & + [P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_1^*)]\sigma^{(1)^2} + (P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4) \\
+ 4\sigma^{(1)}(\sigma^{(1)} + 1)\sigma^{(-1)} & = 0 \quad \text{(A2.5)}
\end{align*}$$

$\sigma^{(1)}$, solved from the equation (A2.3), is found to be:

$$\sigma^{(1)} = \pm i$$

(A2.6)

(A2.4) is satisfied automatically because of (A2.6).

From (A2.5) and (A2.6), $\sigma^{(0)}$ is found to satisfy the quadratic equation:

$$\sigma^{(0)^2} - \frac{P_1 + P_1^* - 2L}{2}\sigma^{(0)} - \left[\frac{P_1 P_3 + (P_1 P_3)^* + P_2 P_4^* + P_2^* P_4}{4}\right] - \left[\frac{P_1 P_1^* - P_2 P_2^* - 2L(P_1 + P_2)}{4}\right] = 0.$$  

(A2.7)
Part III

Solitary Wave Interactions for Coupled Nonlinear Schrödinger Equations

III.1 Introduction

The large scale modulations of a wave packet in a dispersive weakly nonlinear medium are governed by the nonlinear Schrödinger equation

\[ iA_T + \gamma A_{\xi\xi} + \beta A^2 A^* = 0, \]

where \( \xi = x - C_g T, \) \( C_g \) is the group velocity of the wave packet, \( x \) and \( T \) are the large-scale space and time variables, and \( A(\xi, T) \) is the slowly varying amplitude. If \( \omega = \omega(k) \) is the linear dispersion relation, and \( k_0 \) is the dominant wave number, then \( C_g = \omega'(k_0) \) and \( \gamma = \frac{1}{2} \omega''(k_0). \) \( \beta \) is a real parameter. This nonlinear Schrödinger equation is generic and naturally arises in many physical situations, such as water waves, nonlinear optics, plasma physics, etc. Many studies have been made on this equation in recent years. One discovery is that when \( \beta \gamma > 0, \) the spatially independent solution \( A = r_0 e^{i\beta t^2 T} \) is unstable to large-scale modulations. In the deep water case, it naturally explains the famous Benjamin-Feir instability of a Stokes wavetrain. Another discovery is solitons. Solitons are solitary waves which interact with each other "elastically". This finding has led to much insight into the physical wave interactions and has wide practical applications.

When two wave packets are present in a dispersive, weakly nonlinear medium, their large scale modulations are usually governed by the coupled nonlinear Schrödinger type equations (CPNLS). A schematic derivation of these equations for water waves and nonlinear optics is given in the next two subsections.

III.1.1 The derivation of CPNLS in water waves

Consider the evolution of two water wave packets in deep water. The velocity potential of these two waves is given by

\[ \phi = \epsilon\{A(X_1, X_2, T) e^{ik^{(1)}x - i\omega^{(1)}t + |k^{(1)}|y} + B(X_1, X_2, T) e^{ik^{(2)}x - i\omega^{(2)}t + |k^{(2)}|y} + c.c.\} + \text{higher order terms}, \]

where \( x_1, x_2 \) are the two horizontal axis, \( x = (x_1, x_2), \) and \( y \) is the vertical axis. The wavenumber \( k^{(l)} \) and the frequency \( \omega^{(l)} \) for \( l = 1, 2 \) are related by the dispersion relation \( \omega = \omega(k). \) \( A(X_1, X_2, T) \) and \( B(X_1, X_2, T) \) are the slowly varying amplitudes of the two wave packets with

\[ X_1 = \mu x_1, \quad X_2 = \mu x_2, \quad T = \mu t, \]

where \( \mu \) is the scale of the slow space and time variations of the waves. \( \epsilon \) is a measure of nonlinearity. It is assumed that \( \mu \) and \( \epsilon \) are both small parameters.
A multiple scale perturbation analysis shows that the amplitudes \( A(X_1, X_2, T) \) and \( B(X_1, X_2, T) \) are governed by the following equations (Benney & Newell 1967):

\[
\begin{align*}
&i\mu(A_T + C^{(1)}_g \cdot \nabla A) + \mu^2 \sum_{i,j} \gamma^{(1)}_{ij} A_{X_i X_j} + \epsilon^2(\beta_{11}|A|^2 + \beta_{12}|B|^2)A = 0, \quad (1.4) \\
&i\mu(B_T + C^{(2)}_g \cdot \nabla B) + \mu^2 \sum_{i,j} \gamma^{(2)}_{ij} B_{X_i X_j} + \epsilon^2(\beta_{21}|A|^2 + \beta_{22}|B|^2)B = 0. \quad (1.5)
\end{align*}
\]

Here,

\[
\nabla = i_1 \frac{\partial}{\partial X_1} + i_2 \frac{\partial}{\partial X_2},
\]

and \( i_1, i_2 \) are the \( x_1, x_2 \) direction unit vectors.

Eqs. (1.4) and (1.5) are quite complicated. But simplifications can be made in the following situations.

1. In the 3-D case, if we assume that variations of \( A \) and \( B \) only take place in the direction \( u \) which is perpendicular to \( C^{(1)}_g - C^{(2)}_g \), then the envelope equations (1.4) and (1.5) will reduce to

\[
\begin{align*}
&iA_\tau + \gamma_1 A_{\xi \xi} + (\beta_{11}|A|^2 + \beta_{12}|B|^2)A = 0, \quad (1.7) \\
&iB_\tau + \gamma_2 B_{\xi \xi} + (\beta_{21}|A|^2 + \beta_{22}|B|^2)B = 0. \quad (1.8)
\end{align*}
\]

where

\[
\begin{align*}
\xi = u \cdot X - (C^{(l)}_g \cdot u)T, \quad \tau = \epsilon T, \\
\gamma_l = \sum_{i,j} \gamma^{(l)}_{ij} u_i u_j, \quad l = 1, 2, \quad (1.9)
\end{align*}
\]

and the balance of \( \mu = \epsilon \) has been taken (Roskes 1976).

If we further assume that \(|k^{(1)}| = |k^{(2)}|\), it is readily found that

\[
\gamma_1 = \gamma_2, \quad \beta_{11} = \beta_{22}, \quad \beta_{12} = \beta_{21}. \quad (1.10)
\]

So, Eqs. (1.7) and (1.8) become symmetric.

For gravity waves,

\[
\gamma_1(= \gamma_2) < 0, \quad \beta_{11}(= \beta_{22}) < 0, \quad (1.11)
\]

while \( \beta_{12}(= \beta_{21}) \) depends on the angle of the two wavetrains and can be both positive and negative. A rescaling of \( \tau \) and \( \xi \) coordinates reduces Eqs. (1.7) and (1.8) to the form

\[
\begin{align*}
&iA_\tau = A_{\xi \xi} + (|A|^2 + \beta|B|^2)A, \quad (1.12) \\
&iB_\tau = B_{\xi \xi} + (|B|^2 + \beta|A|^2)B \quad (1.13)
\end{align*}
\]

with \( \beta = \beta_{12}/\beta_{11} \).

2. In the 2-D case, the dispersion relation of the capillary-gravity waves shows that, for any gravity wave with the wavenumber \( k^{(1)} \), there is a capillary wave with the wavenumber \( k^{(2)} \) such that

\[
C_g^{(1)} = C_g^{(2)}. \quad (1.14)
\]
When these two waves come together and interact, the envelope equations (1.4) and (1.5) will reduce to

\[ iA_r + \gamma_1 A_{\xi \xi} + (\beta_{11} |A|^2 + \beta_{12} |B|^2)A = 0, \]  
\[ iB_r + \gamma_2 B_{\xi \xi} + (\beta_{21} |A|^2 + \beta_{22} |B|^2)B = 0, \]  

where

\[ \xi = X - C_{\phi}^{(1)} T, \quad \tau = \epsilon T. \]  

The parameters \( \gamma_1, \gamma_2, \beta_{11}, \beta_{12}, \beta_{21} \) and \( \beta_{22} \) are functions of the wavenumbers \( k^{(1)} \) and \( k^{(2)} \).

### III.1.2 The derivation of CPNLS in nonlinear fiber optics

An optical wave field is characterized by both its frequency and its polarization. The coupling between waves of different frequencies and the coupling between polarization components of the same wave are both described by the coupled nonlinear Schrödinger type equations.

#### III.1.2.1 Coupling between waves of different frequencies

Consider the propagation of two optical waves of different frequencies inside an optical fiber. Suppose the fiber is isotropic and that the induced polarization is parallel to the electric field. This electric field then can be written in the form

\[ E(x, t) = \frac{1}{2} e \{ F_1(x, y) A(z, t) e^{i(k_1 z - \omega_1 t)} + F_2(x, y) B(z, t) e^{i(k_2 z - \omega_2 t)} \} + \text{c.c.}, \]  

where \( e \) is the polarization unit vector, \( z \) is the coordinate along the direction of wave propagation, and \( x, y \) are the transverse coordinates. \( \omega_1 \) and \( \omega_2 \) are the center frequencies of the two pulses, and they are related to the wavenumbers \( k_1 \) and \( k_2 \) by the dispersion relation \( k = k(\omega) \). \( F_1(x, y) \) and \( F_2(x, y) \) are the fiber-mode cross-section distributions. \( A(z, t) \) and \( B(z, t) \) are the slowly varying amplitudes of the two waves.

A multiple scale perturbation analysis reveals that \( A(z, t) \) and \( B(z, t) \) are governed by the following equations (Agrawal 1989)

\[ A_z + \delta_1 A_t + \frac{i}{2} \gamma_1 A_{tt} = i \beta_1 (|A|^2 + 2|B|^2) A, \]  
\[ B_z + \delta_2 B_t + \frac{i}{2} \gamma_2 B_{tt} = i \beta_2 (|B|^2 + 2|A|^2) B, \]

where the coefficients \( \delta_1, \gamma_1 \) and \( \beta_i (i = 1, 2) \) are functions of the pulse frequencies \( \omega_i (i = 1, 2) \). With the change of variables

\[ A = e^{i(\frac{\epsilon}{\gamma_1} - \frac{\epsilon^2}{2\gamma_1^2} t)} \tilde{A}, \quad B = e^{i(\frac{\epsilon}{\gamma_2} - \frac{\epsilon^2}{2\gamma_2^2} t)} \tilde{B}, \]

and the bars dropped, Eqs. (1.20) and (1.21) become

\[ A_z + \frac{i}{2} \gamma_1 A_{tt} = i \beta_1 (|A|^2 + 2|B|^2) A, \]  
\[ B_z + \frac{i}{2} \gamma_2 B_{tt} = i \beta_2 (|B|^2 + 2|A|^2) B. \]
III.1.2.2 Coupling between polarization components of the same wave

Consider the propagation of an optical wave inside a birefringent optical fiber. The electric field associated with the elliptical polarized optical wave can be written in the form

\[ E(r, t) = \frac{1}{2} \{ e_1 F_1(z, y) A(z, t) e^{i k_1 z} + e_2 F_2(z, y) B(z, t) e^{i k_2 z} \} e^{-i \omega t} + \text{c.c.}, \]  

(1.25)

where \( e_1 \) and \( e_2 \) are two orthogonal polarization unit vectors. \( A(z, t) \) and \( B(z, t) \) are the slowly varying complex amplitudes of the polarization components of the wave. Similar perturbation analysis results in the following equations for \( A \) and \( B \) (Agrawal 1989)

\[ A_z + \delta_1 A_t + \frac{i}{2} \gamma_1 A_{tt} = i\beta(|A|^2 + \frac{2}{3}|B|^2) A, \]  

(1.26)

\[ B_z + \delta_2 B_t + \frac{i}{2} \gamma_2 B_{tt} = i\beta(|B|^2 + \frac{2}{3}|A|^2) B, \]  

(1.27)

where \( \delta_i, \gamma_i (i = 1, 2) \) and \( \beta \) are functions of the pulse frequencies \( \omega_i(i = 1, 2) \). The change of variables (1.22) will then reduce Eqs. (1.26) and (1.27) to the form

\[ A_z + \frac{i}{2} \gamma_1 A_{tt} = i\beta(|A|^2 + \frac{2}{3}|B|^2) A, \]  

(1.28)

\[ B_z + \frac{i}{2} \gamma_2 B_{tt} = i\beta(|B|^2 + \frac{2}{3}|A|^2) B. \]  

(1.29)

In practice, \( \gamma_1 \) and \( \gamma_2 \) can be assumed to be the same, so that Eqs. (1.28) and (1.29) are symmetric. A rescaling of the coordinates \( z \) and \( t \) can make \( \gamma_1(= \gamma_2) = 2 \) and \( \beta = 1 \) or \(-1\).

III.1.3 System to be studied

In this part of the thesis we study the solitary wave interactions governed by the following symmetric CPNLS:

\[ i A_t = A_{xx} + (A A^* + \beta B B^*) A, \]  

(1.30)

\[ i B_t = B_{xx} + (\beta A A^* + B B^*) B, \]  

(1.31)

where the real valued coupling coefficient \( \beta \) is considered as arbitrary. When \( \beta = 0 \), these two equations are decoupled. \( A \) and \( B \) are each governed by a nonlinear Schrödinger equation, which can be solved exactly by the inverse scattering method. When \( \beta = 1 \), they are called Manakov equations and can also be solved exactly by the inverse scattering method. In both cases, the solitary waves are solitons and interact with each other elastically. When \( \beta \neq 0 \) and \( 1 \), these equations are not exactly solvable, and the interactions between solitary waves are generally non-elastic. Some interesting new interactions arise and will be studied below.
III.2 The non-elastic solitary wave interactions

The solutions of Eqs. (1.30) and (1.31) for general initial conditions are very complicated, and no attempt is made to address this question. Rather, attention is focused on the interaction of two solitary waves, i.e., initially, we take

\[
A(z, 0) = \sqrt{2} r_{A0} \text{sech} r_{A0}(z - z_{A0}) e^{-i \frac{U_{A0}}{2}z}, \tag{2.1}
\]

\[
B(z, 0) = \sqrt{2} r_{B0} \text{sech} r_{B0}(z - z_{B0}) e^{-i \frac{U_{B0}}{2}z}. \tag{2.2}
\]

Initially, \(A\) is a solitary wave centered at \(z = z_{A0}\) with the speed \(U_{A0}\) and amplitude \(\sqrt{2}r_{A0}\), and \(B\) is another solitary wave centered at \(z = z_{B0}\) with the speed \(U_{B0}\) and amplitude \(\sqrt{2}r_{B0}\). The interaction between these two solitary waves is generally non-elastic. A numerical study of this interaction was first made for various values of \(\beta\) and different initial conditions of the form (2.1) and (2.2). The typical behaviors are shown in Fig. 12 to Fig. 18.

The interaction in Fig. 12 is typical for initially well separated solitary waves when \(\beta > 0\). After the interaction, the two waves pass through each other with some reshaping and radiation shedding, and daughter waves are generated. These daughter waves are small pulses that split off from a solitary wave and propagate along beside it in the other mode. The amount of radiation and the sizes of daughter waves are dependent on the initial conditions and the value of \(\beta\), but the qualitative structures remain the same (see Fig. 13 and Fig. 14). When \(\beta < 0\), if the approaching velocity \(U_{A0} - U_{B0}\) is not very large, the two initially well separated solitary waves are reflected off of each other after the interaction, as shown in Fig. 15 and Fig. 16. There is also some radiation, but the amount is very small. If the approaching velocity \(U_{A0} - U_{B0}\) is very large, these two waves simply pass through each other. The interaction scenario is quite different if initially these two solitary waves are overlapped. In this case, when \(\beta > 0\), they tend to trap each other and form a bound, oscillatory state as shown in Fig. 17. When \(\beta < 0\), they escape from each other as shown in Fig. 18.

III.2.1 The motion of a solitary wave in a slowly varying potential field

A useful quantitative as well as qualitative understanding of those various interaction behaviors may be reached by studying the motion of a solitary wave in a slowly varying potential field. This motion is governed by the equation

\[
iA_t = A_{xx} + (AA^* - f(X))A, \tag{2.3}
\]

where \(X = \varepsilon z, (\varepsilon \ll 1)\) and \(f(X)\) is a given slowly varying function. Initially,

\[
A(z, 0) = \sqrt{2}r \text{sech} rz e^{-i \frac{U}{2}z} \tag{2.4}
\]

is a solitary wave with the speed \(U\) and amplitude \(\sqrt{2}r\). Due to the slowly varying potential field, this wave will undergo slow changes as it travels through. A multiple scale perturbation analysis is used to determine this slow evolution.

The appropriate form of solution \(A\) is

\[
A = e^{i(r^2 - \frac{U^2}{4} - f(X))(\theta - \theta_0) - i(\sigma - \sigma_0)} q(\theta, X, \varepsilon), \tag{2.5}
\]
Figure 12: The interaction of two initially well separated solitary waves for $\beta = \frac{2}{3}$. The initial conditions are given by (2.1) and (2.2) with $r_{A_0} = 1, U_{A_0} = 1, x_{A_0} = -8; r_{B_0} = 0.8, U_{B_0} = -1, x_{B_0} = 8$. $0 \leq t \leq 16$. The left figure is the $|A|$ plot, and the right figure is the $|B|$ plot. Note that the two waves pass through each other after the interaction and small daughter waves are generated.

Figure 13: The interaction of two initially well separated solitary waves for $\beta = \frac{2}{3}$. The initial conditions are given by (2.1) and (2.2) with $r_{A_0} = 1, U_{A_0} = 0.5, x_{A_0} = -8; r_{B_0} = 0.8, U_{B_0} = -0.5, x_{B_0} = 8$. $0 \leq t \leq 32$. The left figure is the $|A|$ plot, and the right figure is the $|B|$ plot. Note that the daughter waves become larger when the approach velocity of the two waves gets smaller.
Figure 14: The interaction of two initially well separated solitary waves for $\beta = \frac{1}{3}$. The initial conditions are given by (2.1) and (2.2) with $r_{A0} = 1, U_{A0} = 1, x_{A0} = -10; r_{B0} = 0.6, U_{B0} = -1, x_{B0} = 10$. $0 \leq t \leq 20$. The left figure is the $|A|$ plot, and the right figure is the $|B|$ plot. Note that when the initial $|B|$ wave gets flatter, the $|A|$ wave sheds less radiation and contains a smaller daughter wave.
Figure 15: The interaction of two initially well separated solitary waves for $\beta = -\frac{2}{3}$. The initial conditions are the same as in Figure 12. $0 \leq t \leq 16$. The left figure is the $|A|$ plot, and the right figure is the $|B|$ plot. Note that the two solitary waves are reflected off of each other after the interaction.

Figure 16: The interaction of two initially well separated solitary waves for $\beta = -\frac{2}{3}$. The initial conditions are given by (2.1) and (2.2) with $r_{A0} = 1, U_{A0} = 1, x_{A0} = -10; r_{B0} = 0.6, U_{B0} = -1, x_{B0} = 10$. $0 \leq t \leq 20$. The left figure is the $|A|$ plot, and the right figure is the $|B|$ plot.
Figure 17: The interaction of two initially overlapped solitary waves for $\beta = 0.2$. The initial conditions are given by (2.1) and (2.2) with $r_{A0} = 1, U_{A0} = 0.6, x_{A0} = 0; r_{B0} = 1, U_{B0} = -0.6, x_{B0} = 0$. $0 \leq t \leq 16$. The left figure is the $|A|$ plot, and the right figure is the $|B|$ plot. These two solitary waves are trapped by each other and undergo an oscillating motion.

Figure 18: The interaction of two initially overlapped solitary waves for $\beta = -0.2$. The initial conditions are the same as in Figure 17. $0 \leq t \leq 16$. The left figure is the $|A|$ plot, and the right figure is the $|B|$ plot. These two solitary waves escape from each other.
where

\[ \frac{\partial \theta}{\partial x} = \frac{1}{U}, \quad \frac{\partial \theta}{\partial t} = -1, \tag{2.6} \]

\[ \frac{\partial \sigma}{\partial x} = \frac{1}{U}\{r^2 + \frac{U^2}{4} - f(X)\}, \quad \frac{\partial \sigma}{\partial t} = 0. \tag{2.7} \]

Here, \( U, r, \theta_0 \) and \( \sigma_0 \) are functions of the long space variable \( X \).

Substituting (2.5) into (2.3), we have

\[ \frac{q^{\theta\theta}}{U^2} - r^2 q + q^2 q^* = \epsilon F(q), \tag{2.8} \]

with

\[ F(q) = -\left\{U[r^2 - \frac{U^2}{4} - f(X)]_x(\theta - \theta_0)q + \frac{2q_{\theta x}}{U} + \left(\frac{1}{U}\right)_x q_{\theta} \right. \]
\[ \left. -U[r^2 - \frac{U^2}{4} - f(X)]_{\theta \theta x}q + U_{\theta \theta x}q \right\} \]
\[ -i\left\{\frac{1}{U}[r^2 - \frac{U^2}{4} - f(X)]_x[q + 2(\theta - \theta_0)q_{\theta}] - \frac{Ux}{2} q - U_{q x} \right. \]
\[ \left. -\frac{2}{U}[r^2 - \frac{U^2}{4} - f(X)]_{\theta \theta x}q_{\theta} + \frac{2}{U} \sigma_0 q_{\theta \theta} \right\}. \tag{2.9} \]

\( q(\theta, X, \epsilon) \) can be expanded in the form

\[ q(\theta, X, \epsilon) = q_0(\theta, X) + \epsilon q_1(\theta, X) + \ldots, \tag{2.10} \]

where

\[ q_0 = \sqrt{2}\r sech rU(\theta - \theta_0) \tag{2.11} \]

is the leading order solution of (2.8).

At order \( \epsilon \), from (2.8) and (2.9) we have

\[ \frac{1}{U^2} q_{1\theta\theta} - r^2 q_1 + 2q_0^2 q_1 + q_0^2 q_1^* = F_1, \tag{2.12} \]

where \( F_1 = F(q_0) \). Setting \( q_1 = \phi_1 + i\psi_1 \), where \( \phi_1 \) and \( \psi_1 \) are real valued functions, we obtain

\[ L\phi_1 \equiv \frac{1}{U^2} \phi_{1\theta\theta} - r^2 \phi_1 + 3q_0^2 \phi_1 = \Re F_1, \tag{2.13} \]

\[ M\psi_1 \equiv \frac{1}{U^2} \psi_{1\theta\theta} - r^2 \psi_1 + q_0^2 \psi_1 = \Im F_1. \tag{2.14} \]

Notice that the operators \( L \) and \( M \) are self-adjoint and \( Lq_0 = 0, \ Mq_0 = 0 \). For Eqs. (2.13) and (2.14) to have localized solutions \( \phi_1 \) and \( \psi_1 \) around the solitary wave, the solvability conditions

\[ \int_{-\infty}^{\infty} q_{0\theta} \Re F_1 d\theta = 0, \tag{2.15} \]

\[ \int_{-\infty}^{\infty} q_0 \Im F_1 d\theta = 0 \tag{2.16} \]
have to be satisfied. From (2.15) and (2.16) we obtain the evolution equations for \( r \) and \( U \) as

\[
\frac{dr}{dX} = 0, \quad \frac{d(f(X) + \frac{U^2}{4})}{dX} = 0. \tag{2.17}
\]

Therefore, to the leading order approximation

\[
r = \text{constant}, \tag{2.18}
\]

\[
f(X) + \frac{U^2}{4} = \text{constant}. \tag{2.19}
\]

The relations (2.18) and (2.19) are asymptotically accurate for \( \epsilon \ll 1 \). Numerical results show that they are quite good even for \( \epsilon = O(1) \).

With (2.18) and (2.19), the solutions \( \phi_1 \) and \( \psi_1 \) can be easily determined from (2.13) and (2.14). The slow evolution equations for \( \theta_0 \) and \( \sigma_0 \) are obtained by imposing the orthogonality conditions

\[
\int_{-\infty}^{\infty} q_0 \phi_1 d\theta = 0, \quad \int_{-\infty}^{\infty} q_0 \phi_1 d\theta = 0, \tag{2.20}
\]

but the details will not be pursued here.

### III.2.2 The relations (2.18), (2.19) and the conservation laws

An alternative derivation of the relations (2.18) and (2.19) can be given by using the conservation laws. From Eq. (2.3), we have the following conservation relations:

1. Mass conservation,

\[
\int_{-\infty}^{\infty} |A|^2 dz = \text{constant}. \tag{2.21}
\]

2. Momentum conservation,

\[
\frac{d}{dt} i \int_{-\infty}^{\infty} (AA_x^* - A_x^* A) dz = 2\epsilon \int_{-\infty}^{\infty} |A|^2 f_x dz. \tag{2.22}
\]

3. Energy conservation,

\[
\int_{-\infty}^{\infty} (|A_x|^2 + \frac{1}{2} |A|^4 - f(X)|A|^2) dz = \text{constant}. \tag{2.23}
\]

Locally, a solitary wave with the speed \( U \) and amplitude \( \sqrt{2r} \) is of the form

\[
A = \sqrt{2r} \text{sech} (z - x_0 - Ut)e^{-i \frac{U}{2}(z - Ut) - i(r^2 + \frac{U^2}{4})t + if(X)t - i\theta_0}. \tag{2.24}
\]

Since the potential field varies slowly, it is reasonable to assume that this form almost remains the same. Making use of Eq. (2.24), it is found that

\[
\int_{-\infty}^{\infty} |A|^2 dz = 4r, \tag{2.25}
\]
\[ \int_{-\infty}^{\infty} (|A_x|^2 + \frac{1}{2} |A|^4 - f(X)|A|^2) dx \approx 2\pi (\frac{U_0^2}{4} + f(X)), \tag{2.26} \]

where \( X \) in the right hand side of Eq. (2.26) is implied to be the center position of the solitary wave. The mass and energy conservation (2.21) and (2.23) readily recover the relations (2.18) and (2.19).

It needs to be noted that the momentum conservation relation (2.22) is automatically satisfied by (2.18) and (2.19).

### III.2.3 Discussion

The relations (2.18) and (2.19) we derived for Eq. (2.3) offer much insight into the motion of a solitary wave in a slowly varying potential field. The relation (2.18) indicates that the amplitude of the solitary wave does not change as it travels through the potential field. The relation (2.19) indicates that the speed of the wave will change as \( f(X) \) varies. When \( f(X) \) decreases, the wave will accelerate; when \( f(X) \) increases, the wave will slow down. In the latter case, if \( f(X) \) gets large enough, the solitary wave will lose all its speed and come to a stop. The analogy to the motion of a particle in a potential field suggests that this wave can not stay there since such a state is unstable (the proof involves some algebra but the idea is simple). What the wave does is that it will turn direction and go the other way, i.e. it is reflected by the potential field. The condition for reflection is that

\[ \frac{U_0^2}{4} + f(X_0) < \max(f(X)), \tag{2.27} \]

where \( U_0 \) is the initial wave velocity, and \( X_0 \) is the initial center position of the wave.

The above general results can be further confirmed in the special case with \( f = \alpha_0 X \equiv \alpha z \), where Eq. (2.3) is exactly solvable (Chen & Liu 1976). So in this case, \( \alpha \) need not be small. With the change of variables

\[ \xi = z + \alpha t^2, \quad A = \phi(\xi, t)e^{i\alpha z + \frac{1}{2} i\alpha^2 t^2}, \tag{2.28} \]

Eq. (2.3) is reduced to the nonlinear Schrödinger equation

\[ i\phi_t = \phi_{\xi\xi} + \phi^2 \phi^*. \tag{2.29} \]

If initially, \( A \) is a solitary wave with the speed \( U_0 \) and amplitude \( \sqrt{2} r_0 \), i.e.

\[ A(z, 0) = \sqrt{2} r_0 \text{sech } r_0 z e^{-i \frac{U_0}{2} z}, \tag{2.30} \]

the exact solution at later time is

\[ A(z, t) = \sqrt{2} r_0 \text{sech } r_0 (z - U_0 t + \alpha t^2) e^{-i \frac{U_0}{2} z + i\alpha z t - i \left( \frac{\alpha^2}{4} t^2 - \frac{U_0^2}{4} t + \frac{1}{2} i \alpha U_0 t^2 + \frac{1}{4} \alpha^2 t^4 \right)}. \tag{2.31} \]

Suppose \( U_0 > 0 \). When \( \alpha < 0 \), the solitary wave accelerates along the \( z \) direction and goes straight to infinity. When \( \alpha > 0 \), this wave first slows down and comes to a full stop at \( t = \frac{U_0}{2\alpha} \). Then it turns around and moves in the opposite direction. So, it is reflected as expected. Notice the rather surprising result that there has been no amplitude change as the solitary wave travels through this varying potential field. When \( \alpha \) is small, to the leading order, this exact solution is consistent with the relations (2.18) and (2.19) we just derived.
Another interesting special case is that \( f = \alpha_1 X^2 \). In this case, Eq. (2.3) is not exactly solvable, so our perturbation results are particularly useful. When the relation (2.19) is differentiated twice with respect to \( T \), we get

\[
U_{TT} + 4\alpha_1 U = 0. \tag{2.32}
\]

So when \( \alpha_1 < 0 \), the velocity \( U \) increases exponentially, and the solitary wave accelerates to infinity. If \( \alpha_1 > 0 \), this wave is trapped and oscillates about the position \( X = 0 \) with the period \( T_0 = \pi / \sqrt{\alpha_1} \epsilon \).

The knowledge gained from Eq. (2.3) can be readily used to qualitatively and quantitatively explain the numerical interaction behaviors presented earlier.

Suppose the two solitary waves, initially given by (2.1) and (2.2) and well separated, come together. If we choose a new coordinate system moving with velocity \( U_{B0} \), the \( B \) wave is fixed relative to this new system, and the \( A \) wave can be considered as moving into the varying “potential field” with \( f = -\beta |B|^2 \) and at initial speed \( U_{A0} - U_{B0} \). Although this “potential field” is no longer slowly varying, nor is it steady, the qualitative interaction behaviors can still be understood by the results we gained for Eq. (2.3). For example, when \( \beta > 0 \), the \( A \) wave passes through as in Fig. (12). Its reshaping and the generation of daughter waves are due to the non-slowly varying “potential field”. When \( \beta < 0 \), the \( A \) wave is reflected back if \( U_{A0} - U_{B0} \) is not large, as in Fig. (15). In this case, the relation (2.19) for Eq. (2.3) gives the reflection condition

\[
\frac{(U_{A0} - U_{B0})^2}{4} < -2\beta r_{B0}, \tag{2.33}
\]

which is a pretty good estimate when compared with the numerical data. The motion of the \( B \) wave can be similarly analyzed and will not be repeated here.

Suppose, initially, the two solitary waves given by (2.1) and (2.2) are overlapped, i.e. \( x_{A0} = x_{B0} \). Analysis based on (2.18) and (2.19) shows that, when \( \beta < 0 \), each wave gains speed as it leaves the other. Therefore, they both escape as in Fig. 18. When \( \beta > 0 \), each wave loses speed as it tries to leave the other. If \( U_{A0} - U_{B0} \) is small, they trap each other and form a bound, oscillating state, as in Fig. 17. The trapping condition is approximately given by

\[
\frac{(U_{A0} - U_{B0})^2}{4} < 2\beta \min(r_{A0}, r_{B0}). \tag{2.34}
\]

Since these two waves stay together and interact for a long time, they are able to retain their initial forms without much reshaping and radiation only when \( \beta > 0 \) is small.

### III.2.4 The determination of daughter wave profiles

The generation of daughter waves, as shown in Fig. 12, is a distinctive feature when \( \beta > 0 \). In this section, we set out to determine the profiles of these small daughter waves.

Suppose, initially, we have two well separated, decoupled solitary waves of the form (2.1) and (2.2) approaching each other. After interaction the two waves will pass through each other with some reshaping and radiation shedding, and small daughter waves are created.

One observation is that if \( \beta = 0 \) or 1, the interaction between these two solitary waves is elastic. The two waves will recover themselves after the interaction without reshaping and radiation shedding,
and no daughter waves are generated. For other positive values of $\beta$, provided that it is not too large, it is reasonable to assume that the reshaping is slight, the radiation is negligible, and the daughter waves are small.

Under these assumptions, we can analytically determine the profiles of these daughter waves. Without loss of generality, consider the right going wave initially with only the $A$ component (2.1).

After the interaction, the $A$ component still dominates the wave, and its "sech" profile is slightly modified. A small daughter wave in the $B$ component is generated and it travels together with the $A$ component. At the leading order, the $A$ component is approximated by the original "sech" profile, and this small daughter wave in the $B$ component is governed by the linear Schrödinger equation

$$iB_t = B_{xx} + 2\beta r_A^2 \text{sech}^2 r_A(x - U_A t)B.$$  \hspace{1cm} (2.35)

With the change of variables

$$\tilde{z} = x - U_A t, \quad B = e^{-iz_A^2 z + i\frac{v_A^2}{4} t} B,$$  \hspace{1cm} (2.36)

and the bars dropped, Eq. (2.35) reduces to

$$iB_t = B_{xx} + 2\beta r_A^2 \text{sech}^2 r_A z B.$$  \hspace{1cm} (2.37)

The linear Schrödinger equation (2.37) possesses a few discrete eigensolutions which decay to zero as $z$ goes to $\pm \infty$. It also allows eigensolutions with continuous eigenvalues, which remain finite and periodic at $z = \pm \infty$. For our purpose, since the daughter waves are always localized, we only consider those discrete eigensolutions.

Suppose $B = e^{i\omega t} b(z)$, then

$$b_{xx} + (2\beta r_A^2 \text{sech}^2 r_A z + \omega) b = 0.$$  \hspace{1cm} (2.38)

Eq. (2.38) is exactly solvable (Landau & Lifshitz). With the change of variables

$$u = \frac{b}{\text{sech}^2 r_A z}, \quad s = \frac{1}{2} \left( \sqrt{1 + 8\beta} - 1 \right),$$  \hspace{1cm} (2.39)

and

$$\xi = \sinh^2 r_A z,$$  \hspace{1cm} (2.40)

Eq. (2.38) reduces to the hypergeometric equation

$$\xi(1 + \xi) u'' + [(1 - s)\xi + \frac{1}{2}i(1 - s)] u' + \frac{1}{4} (s^2 - \epsilon^2) u = 0,$$  \hspace{1cm} (2.41)

with $\epsilon = \sqrt{-\omega/r_A}$.

The even and odd (in $z$) solutions of Eq. (2.41) are

$$u_1 = F(-\frac{1}{2}s + \frac{1}{2}i\epsilon, -\frac{1}{2}s - \frac{1}{2}i\epsilon, \frac{1}{2}, -\xi),$$  \hspace{1cm} (2.42)

$$u_2 = \sqrt{\xi} F(-\frac{1}{2}s + \frac{1}{2}i\epsilon + \frac{1}{2}, -\frac{1}{2}s - \frac{1}{2}i\epsilon + \frac{1}{2}, -\frac{3}{2}, -\xi).$$  \hspace{1cm} (2.43)

In order that $b = (1 + \xi)^{-\frac{1}{2}} u_1$ should reduce to zero as $\xi \to \infty$, the parameter $-\frac{1}{2}s + \frac{1}{2}i\epsilon$ must be a negative integer or zero; then $F$ is a polynomial of degree $\frac{1}{2}s - \frac{1}{2}i\epsilon$ and $b$ tends to zero as $\xi^{-\frac{1}{2}}\epsilon$.
as $\xi \to \infty$. Similarly, for $b = (1 + \xi)^{-\frac{1}{2}}u_2$, this condition is satisfied if $-\frac{1}{2}s + \frac{1}{2} \varepsilon + \frac{1}{2}$ is a negative integer.

Thus the eigenvalues $\omega$ are determined by $s - \varepsilon = n$, or

$$\omega = -\frac{r_A^2}{4}\left[\sqrt{1 + 8\beta} - (1 + 2n)\right]^2,$$

(2.44)

where $n$ takes positive values starting from zero. There is a finite number of eigenvalues, determined by the conditions $\varepsilon > 0$, i.e. $n < \frac{\sqrt{1 + 8\beta} - 1}{2}$.

1. When $\beta < 0$, no such $n$ exists and the daughter waves of this kind do not exist.

2. When $0 < \beta \leq 1$, $n = 0$ and $\omega = -\frac{r_A^2}{4}[\sqrt{1 + 8\beta} - 1]^2$. So the daughter waves only take the profile

$$b(x) = \text{sech}^s r_A x, \quad s = \frac{1}{2}(\sqrt{1 + 8\beta} - 1),$$

(2.45)

which is an even function in $x$.

3. When $1 < \beta \leq 3$, $n = 0$ or 1. In this case, the daughter waves may take one of the two profiles

$$b_1(x) = \text{sech}^s r_A x, \quad \omega_1 = -\frac{r_A^2}{4}[\sqrt{1 + 8\beta} - 1]^2$$

(2.46)

and

$$b_2(x) = \text{sech}^s r_A x \sinh r_A x, \quad \omega_2 = -\frac{r_A^2}{4}[\sqrt{1 + 8\beta} - 3]^2$$

(2.47)

with $s = \frac{1}{2}(\sqrt{1 + 8\beta} - 1)$. Note that $b_1(x)$ is an even function in $x$ and $b_2(x)$ is an odd function.

The daughter waves in Fig. 12 are of the shape given by Eq. (2.45) since $\beta = \frac{2}{3}$ there.
III.3 Stability of equally mixed solitary wave mode

The equally mixed solitary wave modes of Eqs. (1.30) and (1.31) are given by

\[ A = B = \sqrt{\frac{2}{1 + \beta}} \text{sech} \, r(x - Ut) \, e^{-i \frac{U}{2} x - i \left( r^2 - \frac{U^2}{4} \right)t}. \]  

(3.1)

The stability of this state is of interest and is studied here.

Before the analysis is carried out, some simplifications are made. With the change of variables

\[ \bar{x} = r(x - Ut), \quad \bar{t} = r^2 t, \]  

(3.2)

\[ A = r \, e^{-i \frac{U}{2} x + i \frac{U^2}{4} t} \bar{A}, \quad B = r \, e^{-i \frac{U}{2} x + i \frac{U^2}{4} t} \bar{B}, \]  

(3.3)

and the bars dropped, Eqs. (1.30) and (1.31) remain the same. But the mixed mode (3.1) is reduced to

\[ A = B = e^{-it} \sqrt{\frac{2}{1 + \beta}} \text{sech} \, z. \]  

(3.4)

The linear stability analysis of the simplified mixed mode (3.4) is now carried out. Suppose the solution (3.4) is slightly perturbed with

\[ A = e^{-it} \left\{ \sqrt{\frac{2}{1 + \beta}} \text{sech} \, z + \tilde{A} \right\}, \]  

(3.5)

\[ B = e^{-it} \left\{ \sqrt{\frac{2}{1 + \beta}} \text{sech} \, z + \tilde{B} \right\}. \]  

(3.6)

It is readily found that \( \tilde{A} \) and \( \tilde{B} \) are governed by the following linear equations

\[ i\tilde{A}_t = \tilde{A}_{zz} - \tilde{A} + \frac{2(2 + \beta)}{1 + \beta} \text{sech}^2 z \tilde{A} + \frac{2}{1 + \beta} \text{sech}^2 z \tilde{A}^* + \frac{2\beta}{1 + \beta} \text{sech}^2 z (\tilde{B} + \tilde{B}^*), \]  

(3.7)

\[ i\tilde{B}_t = \tilde{B}_{zz} - \tilde{B} + \frac{2(2 + \beta)}{1 + \beta} \text{sech}^2 z \tilde{B} + \frac{2}{1 + \beta} \text{sech}^2 z \tilde{B}^* + \frac{2\beta}{1 + \beta} \text{sech}^2 z (\tilde{A} + \tilde{A}^*). \]  

(3.8)

The general solution of the above two equations is difficult to obtain. But due to the symmetric nature of these equations, special solutions of the form

\[ \tilde{A} = -\tilde{B} \]  

(3.9)

are allowed. These special solutions satisfy the equation

\[ i\tilde{A}_t = \tilde{A}_{zz} - \tilde{A} + \frac{4}{1 + \beta} \text{sech}^2 z \tilde{A} + \frac{2(1 - \beta)}{1 + \beta} \text{sech}^2 z \tilde{A}^*. \]  

(3.10)

When \( \beta = 0 \), Eq. (3.10) becomes

\[ i\tilde{A}_t = \tilde{A}_{zz} - \tilde{A} + 4 \text{sech}^2 z \tilde{A} + 2 \text{sech}^2 z \tilde{A}^*. \]  

(3.11)
One observation is the fact that
\[ \tilde{A} = \text{sech} \, z \tanh z \] (3.12)
is a special eigensolution of Eq. (3.11) with the eigenvalue 0.

When \( \beta \) is perturbed away from zero, this eigensolution (3.12) changes accordingly. When \( \beta \) is small, its dependence on \( \beta \) can be determined by the following procedure.

To order \( \beta \), Eq. (3.10) can be rewritten as
\[ i\tilde{A}_t = \tilde{A}_{xx} - \tilde{A} + 4 \text{sech}^2 x \tilde{A} + 2 \text{sech}^2 x \tilde{A}^* - 4\beta \text{sech}^2 x (\tilde{A} + \tilde{A}^*). \] (3.13)

We choose to work with the variables \( \tilde{A} \) and \( \tilde{A}^* \) instead of \( \text{Re}(A) \) and \( \text{Im}(A) \). With the notation \( \tilde{A}^* \equiv \tilde{\psi} \), the equations governing \( \tilde{A} \) and \( \tilde{\psi} \) are
\[ i\tilde{A}_t = \tilde{A}_{xx} - \tilde{A} + 4 \text{sech}^2 x \tilde{A} + 2 \text{sech}^2 x \tilde{\psi} - 4\beta \text{sech}^2 x (\tilde{A} + \tilde{\psi}), \] (3.14)
\[ -i\tilde{\psi}_t = \tilde{\psi}_{xx} - \tilde{\psi} + 4 \text{sech}^2 x \tilde{\psi} + 2 \text{sech}^2 x \tilde{A} - 4\beta \text{sech}^2 x (\tilde{A} + \tilde{\psi}). \] (3.15)

The eigensolutions of the above two linear equations are of the form
\[ \begin{pmatrix} \tilde{A} \\ \tilde{\psi} \end{pmatrix} = e^{i\lambda t} \begin{pmatrix} \phi(x) \\ \psi(x) \end{pmatrix} \] (3.16)
with \( \phi \) and \( \psi \) satisfying the equations
\[ \phi_{xx} - \phi + 4 \text{sech}^2 x \phi + 2 \text{sech}^2 x \psi - 4\beta \text{sech}^2 x (\phi + \psi) + \lambda \phi = 0, \] (3.17)
\[ \psi_{xx} - \psi + 4 \text{sech}^2 x \psi + 2 \text{sech}^2 x \phi - 4\beta \text{sech}^2 x (\phi + \psi) - \lambda \psi = 0. \] (3.18)

The appropriate perturbation expansions for \( \lambda, \phi, \) and \( \psi \) are
\[ \phi = \phi_0 + \sqrt{\beta} \phi_1 + \beta \phi_2 + \ldots, \] (3.19)
\[ \psi = \psi_0 + \sqrt{\beta} \psi_1 + \beta \psi_2 + \ldots, \] (3.20)
\[ \lambda = \sqrt{\beta} \lambda_1 + \beta \lambda_2 + \ldots, \] (3.21)
where
\[ \phi_0 = \psi_0 = \text{sech} \, z \tanh z. \] (3.22)

At order \( \sqrt{\beta} \),
\[ \phi_{1xx} - \phi_1 + 4 \text{sech}^2 x \phi_1 + 2 \text{sech}^2 x \psi_1 = -\lambda_1 \phi_0, \] (3.23)
\[ \psi_{1xx} - \psi_1 + 4 \text{sech}^2 x \psi_1 + 2 \text{sech}^2 x \phi_1 = \lambda_1 \psi_0. \] (3.24)

The homogeneous equations have two localized solutions
\[ \begin{pmatrix} \phi_0 \\ \psi_0 \end{pmatrix} \text{ and } \begin{pmatrix} \text{i sech } z \\ -\text{i sech } z \end{pmatrix}. \] (3.25)

So, the solvability conditions of Eqs. (3.23) and (3.24) are
\[ \int_{-\infty}^{\infty} (-\lambda_1 \phi_0^2 + \lambda_1 \psi_0^2) dx = 0, \] (3.26)
\[ \int_{-\infty}^{\infty} (-i\lambda_1 \phi_0 \sech z - i\lambda_1 \psi_0 \sech z) \, dz = 0, \] (3.27)

which are satisfied automatically in view of Eq. (3.22).

The functions \( \phi_1 \) and \( \psi_1 \) are determined from the equations (3.23) and (3.24). To uniquely determine them, normalization conditions are needed. We require that they are orthogonal to the homogeneous solutions (3.25), i.e.

\[ \int_{-\infty}^{\infty} (\phi_0 \phi_1 + \psi_0 \psi_1) \, dz = 0, \] (3.28)

\[ \int_{-\infty}^{\infty} i \sech z (\phi_1 - \psi_1) \, dz = 0. \] (3.29)

Under such conditions, \( \phi_1 \) and \( \psi_1 \) are found to be

\[ \phi_1 = -\psi_1 = \frac{1}{2} \lambda_1 z \sech z. \] (3.30)

At order \( \beta \),

\[ \phi_{2zz} - \phi_2 + 4 \sech^2 z \phi_2 + 2 \sech^2 z \psi_2 = 4 \sech^2 z (\phi_0 + \psi_0) - \lambda_2 \phi_0 - \lambda_1 \phi_1, \] (3.31)

\[ \psi_{2zz} - \psi_2 + 4 \sech^2 z \psi_2 + 2 \sech^2 z \phi_2 = 4 \sech^2 z (\phi_0 + \psi_0) + \lambda_2 \psi_0 + \lambda_1 \psi_1. \] (3.32)

Similarly, the solvability conditions for these nonhomogeneous equations are

\[ \int_{-\infty}^{\infty} \{ [4 \sech^2 z (\phi_0 + \psi_0) - \lambda_2 \phi_0 - \lambda_1 \phi_1] \phi_0 + [4 \sech^2 z (\phi_0 + \psi_0) + \lambda_2 \psi_0 + \lambda_1 \psi_1] \psi_0 \} \, dz = 0, \] (3.33)

\[ \int_{-\infty}^{\infty} i \{ [4 \sech^2 z (\phi_0 + \psi_0) - \lambda_2 \phi_0 - \lambda_1 \phi_1] \sech z - [4 \sech^2 z (\phi_0 + \psi_0) + \lambda_2 \psi_0 + \lambda_1 \psi_1] \sech z \} \, dz = 0. \] (3.34)

Eq. (3.34) is satisfied automatically. Eq. (3.33) is readily simplified to be

\[ \lambda_1^2 = \frac{8 \int_{-\infty}^{\infty} \sech^4 z \tanh^2 z \, dz}{\int_{-\infty}^{\infty} z \sech^2 z \tanh z \, dz} = \frac{64}{15}. \] (3.35)

The above results can be summarized as follows. When \( \beta \) is small, two special solutions of the stability equations (3.7) and (3.8) are given by the following perturbative series:

\[ \tilde{A} = -\tilde{B} = e^{i\lambda z} \phi(z), \] (3.36)

\[ \lambda = \lambda_1 \sqrt{\beta} + O(\beta), \] (3.37)

\[ \phi(z) = \sech z \tanh z + \frac{1}{2} \lambda_1 \sqrt{\beta} z \sech z + O(\beta), \] (3.38)

\[ \lambda_1 = \pm \sqrt{\frac{64}{15}}. \] (3.39)
Notice that if $\beta > 0$, they are both stable. But if $\beta < 0$, one of them becomes unstable. The unstable eigenfunction is

$$\phi(x) = \operatorname{sech} x \tanh x - \frac{1}{2} \sqrt{\frac{64}{15}} \sqrt{-\beta} x \operatorname{sech} x + O(\beta).$$  \hspace{1cm} (3.40)

Its growth rate

$$\rho = \sqrt{\frac{64}{15}} \sqrt{-\beta} + O(\beta) = 2.07 \sqrt{-\beta} + O(\beta).$$  \hspace{1cm} (3.41)

The above analysis readily shows that the equally mixed solitary wave mode (3.1) is unstable when $\beta$ is small negative. When $\beta < 0$ but $|\beta|$ is not small, numerical results show that it is still unstable, which is expected. When $\beta > 0$, numerical results show that it is always stable, which is consistent with our analysis for small positive $\beta$.

It is interesting to note here that the special unstable disturbance (3.36) with $\phi(x)$ and $i\lambda$ given by (3.40) and (3.41) is actually the most unstable disturbance of the stability equations (3.7) and (3.8) for small negative $\beta$. This fact is very clear from the numerical data.

### III.4 Summary

Solitary wave interactions governed by the symmetric coupled nonlinear Schrödinger equations have been studied. Various interaction behaviors have been qualitatively and quantitatively explained by numerical and analytical methods. The stability of the equally mixed solitary wave modes has also been determined.
References


