Internal Solitons in the Ocean

by


Woods Hole Oceanographic Institution
Woods Hole, MA 02543

January 2006

Technical Report

Funding was provided by the Office of Naval Research under Contracts No. N00014-04-10146 and N00014-04-10720

Approved for public release; distribution unlimited.
WHOI-2006-04

Internal Solitons in the Ocean

by


Woods Hole Oceanographic Institution
Woods Hole, Massachusetts 02543

January 2006

Technical Report

Funding was provided by the Office of Naval Research under Contracts
No. N00014-04-10146 and N00014-04-10720

Reproduction in whole or in part is permitted for any purpose of the United States Government. This report should be cited as Woods Hole Oceanog. Inst. Tech. Rept.,
WHOI-2006-04.

Approved for public release; distribution unlimited.

Approved for Distribution:

James F. Lynch, Chair
Department of Applied Ocean Physics and Engineering
INTERNAL SOLITONS IN THE OCEAN

John R. Apel ¹
Johns Hopkins University Applied Physics Laboratory
Laurel, Maryland, 20723 USA
and
Global Ocean Associates, Rockville, MD 20853 USA

Lev A. Ostrovsky
Zel Technologies/Earth System Research Laboratory
Boulder, Colorado, 80305 USA
and
Institute of Applied Physics, Russian Academy of Sciences, Nizhny Novgorod, Russia

Yury A. Stepanyants
Australian Nuclear Science and Technology Organization, PMB 1
Menai (Sydney), NSW, 2234, Australia
and
Institute of Applied Physics, Russian Academy of Sciences, Nizhny Novgorod, Russia

James F. Lynch
Applied Ocean Physics and Engineering Department
Woods Hole Oceanographic Institution
Woods Hole, MA, 02543 USA

February 13, 2006

¹Deceased
Abstract

Nonlinear internal waves in the ocean are discussed (a) from the standpoint of soliton theory and (b) from the viewpoint of experimental measurements. First, theoretical models for internal solitary waves in the ocean are briefly described. Various nonlinear analytical solutions are treated, commencing with the well-known Boussinesq and Korteweg–de Vries equations. Then certain generalizations are considered, including effects of cubic nonlinearity, Earth’s rotation, cylindrical divergence, dissipation, shear flows, and others. Recent theoretical models for strongly nonlinear internal waves are outlined. Second, examples of experimental evidence for the existence of solitons in the upper ocean are presented; the data include radar and optical images and in situ measurements of waveforms, propagation speeds, and dispersion characteristics. Third, and finally, action of internal solitons on sound wave propagation is discussed.

This review paper is intended for researchers from diverse backgrounds, including acousticians, who may not be familiar in detail with soliton theory. Thus, it includes an outline of the basics of soliton theory. At the same time, recent theoretical and observational results are described which can also make this review useful for mainstream oceanographers and theoreticians.
Contents

1 Preface 4

2 Introduction and Overview 4

3 Theoretical Models 6
   3.1 Basic equations ............................................................... 6
   3.2 Shallow-water models ..................................................... 8
   3.3 Internal Waves in Nonrotating Fluids .................................. 10
      3.3.1 The Korteweg-de Vries (KdV) Equation ......................... 11
      3.3.2 The Extended and Modified Korteweg-de Vries Equations .... 15
      3.3.3 The Benjamin-Ono Equation ........................................ 21
      3.3.4 The Joseph–Kubota–Ko–Dobbs (JKKD) Equation ............... 22
   3.4 Soliton Propagation Under Perturbations. ............................ 24
   3.5 Refraction and Diffraction of Solitons ................................ 26
   3.6 Internal Waves on Shear Flows ......................................... 28
   3.7 Nonlinear Waves in Rotating Ocean .................................... 38
   3.8 Strongly Nonlinear Waves ............................................... 43
      3.8.1 Non-dispersive Waves and Evolution Equations ................ 50
      3.8.2 Simplified Evolution Equation (β-model) ....................... 51
      3.8.3 Deep Lower Layer .................................................... 53

4 Experimental observations in the oceans 53
   4.1 Internal Solitons Near the Continents ................................ 54
   4.2 Internal Waves in the Deep Ocean ...................................... 68
   4.3 Surface Signatures of Internal Waves .................................. 75

5 Effects of Non-Linear Internal Waves on Sound Waves in the Ocean 78

6 Concluding Remarks 85

7 References 86
1 Preface

The first incarnation of this review paper appeared in 1995 as a Report of the Applied Physics Laboratory of John Hopkins University (APEL ET AL., 1995). It was planned then to continue working on the material and publish it in a refereed journal. However, these plans were frozen when one of the authors and the actual initiator of this project, Prof. John Apel, passed away.

Recently, we received a suggestion to publish this material both as a technical report and as a review article in the Journal of the Acoustical Society of America (JASA), with the motivation being that the acoustical monitoring of internal solitary waves had become one of the leading topics in acoustical oceanography. We agreed, realizing that both the theory and observations of internal solitons have progressed enormously since 1995. Thus, along with preserving most of the previous material of the paper, we tried to update it in order to reflect, at least briefly, the main new results in the area. This took another few years, and while doing that, we had to restrict ourselves in adding too many new parts; otherwise the text threatened to grow out of our control. As a result, the basic material and older results are still represented more comprehensively than the results of the last 8–10 years. Still we hope that, first, we managed to concisely present or at least mention most of the important new achievements and, second, that such an imbalance is not important to acousticians and other professionals who are not directly involved in ocean hydrodynamics. On the other hand, for those who are involved in physical oceanography, the paper can give some useful information regarding the present status of the problem and also the corresponding references. All this seems worth the effort due to the richness of the topic. Indeed, ocean internal solitary waves are arguably the most ubiquitously observed type of solitons in geophysics, and they affect many important oceanic processes, especially in the coastal zones. As a result, their studies by various means, including acoustic ones, is an exciting enterprize. Note also that a review of laboratory experiments with internal solitary waves has recently been published by two of the authors (OSTROVSKY & STEPANYANTS, 2005) [see also in (GRUE, 2005)], so that we omit this important issue here.

2 Introduction and Overview

It has been known for over a century that in the island archipelagos of the Far East, there are occasionally seen on the surface of the sea long, isolated stripes of highly agitated features that are defined by audibly breaking waves and white water (WALLACE, 1869). These features propagate past vessels at speeds that are at times in excess of two knots; they are not usually associated with any nearby bottom feature to which one might attribute their origin, but are indeed often seen in quite deep water. In the nautical literature and charts, they are sometimes identified as “tide rips”. In Arctic and sub-Arctic regions, especially near the mouths of fjords or river flows into the sea, analogous phenomena of lower intensity are known, dating back perhaps even to the Roman reports of “sticky water,” but certainly a recognized phenomenon since Viking times (EKMAN, 1904).
It is now understood that many of these features are surface manifestations of internal gravity waves, sometimes only weakly nonlinear but quite often highly nonlinear excitations in the form of "solitary waves" or "solitons." Their soliton-like nature (steady propagation, preserving shape) has only relatively recently been established, with two of us rhetorically questioning in 1989 whether internal solitons actually exist in the ocean (OSTROVSKY & STEPANYANTS, 1989). Now it is widely accepted view that they (or at least structures close to solitary waves) exist as ubiquitous features in the upper ocean, and that they may be seen at scores of locations around the globe with a wide variety of in situ and remote sensors.

This paper sets forth (a) the basic theoretical formulations and characteristics of solitons in a stratified, sheared, rotating fluid and (b) some of the observational and experimental evidence for their existence.

Isolated nonlinear surface waves of great durability were first reported propagating in a shallow, unstratified Scottish canal by John Scott Russell in 1838 and 1844, but their correct theoretical description was offered much later, in 1870s by BOUSSINESQ and RAYLEIGH and in 1895 by KÖRTEWEG and DE VRIES [see, e.g., MILES, 1980]. More recent reviews have set forth many of the interesting characteristics of solitons in general, such as their ability to preserve shapes and amplitudes upon interaction, as elastic particles do (SCOTT ET AL., 1973; ABLowitz & SEGUR, 1981; DODD ET AL., 1982).

Recognition of the nonlinear and, more specifically, the solitary character of oceanic internal waves on continental shelf waters appears to have first been made in the 1960s and early 1970s (LEE, 1961; ZIEGENBEIN, 1969, 1970; HALPERN, 1971; LEE & BEARDSLEY, 1974; APEL ET AL., 1975A), and extensive investigations into the phenomenon have since been made by many groups of workers. The bibliography includes references to these works that will be cited later in their proper contexts. A number of experimental data concerning internal wave (IW) solitons in the ocean may be found in, e.g., (OSTROVSKY & STEPANYANTS, 1989; APEL, 1995), and later in (DUDA & FARMER, 1999; SABININ & SEREBRYANY); see also the Internet Atlas of internal solitons (JACKSON & APEL, 2004).

The creation of solitons relies on the existence of both intrinsic dispersion and nonlinearity in the medium. If, through nonlinear effects, the speed of the wave increases depending on the local displacement, the long wave (simple wave) steepens toward a shock-like condition. In a dispersive system, however, unlike in non-dispersive acoustics, this shock formation is resisted by dispersion, i.e. the difference between phase velocities of the various Fourier components making up the wave, which tends to broaden the steepening fronts. A soliton then represents a balance between these two factors, with a wave of permanent shape resulting that propagates at a speed dependent on its amplitude, the layer depths, and the density contrast, among other factors. In many cases, a soliton train (a "solibore") is formed rather than a single soliton.

This simple picture, although providing a conceptual framework for discussing solitons,

---

2Notwithstanding the formal definition of solitons as nonlinear pulses which remain unchanged upon interaction (SCOTT ET AL., 1973), we shall use the name soliton for any stable, non-dissipative (or weakly dissipative) solitary formations, not only for brevity but also because we (and a number of others) believe that even if solitary waves interact by emitting some radiation (as is typically the case in non-integrable mathematical models), they still reveal the properties of a particle, which is the reason for the term "soliton".
must be enriched by a more thorough theoretical treatment of the many facets of solitary waves.

In the recent years, the “family” of observed internal solitary waves has been significantly extended, and to address this and other issues, a special Workshop on internal solitary waves was held in 1998 (Duda & Farmer, 1999). New observations have confirmed that internal solitary waves in coastal zones are often strongly nonlinear, so that the most usable weakly-nonlinear theoretical models fail to describe them adequately.

The atmosphere also supports nonlinear internal waves, most notably the lee-wave/ lenticular cloud phenomenon found downwind of sharp gradients in mountain ranges (see Smith, 1988; Christie et al., 1978; Christie, 1989; Rottman & Grimshaw, 2002 and references therein); we do not discuss atmospheric internal waves here, however.

The practical importance of internal waves (IWs) is evident, as strong IWs can provide an intensive mixing in both the upper ocean and in shallow areas, can affect biological processes, as well as radar signals, play a role in underwater acoustics and underwater navigation, etc. Military aspects of the problem are of interest as well; apart from the seemingly anecdotal information circulated in 1970s on the IW role in submarine catastrophes, it should be noted that some recent publications have been supported under Naval auspices, such as the aforementioned Workshop (Duda & Farmer, 1999).

We shall concentrate on internal solitons in the sea, with Section 3 developing the theoretical aspects, Section 4 giving a summary of observational data (in situ and remote), and their discussion. Finally, Section 5 briefly outlines the impact of internal solitons on acoustic waves.

3 Theoretical Models

3.1 Basic equations

The description of internal gravity waves in water is, in general, based on the equations of hydrodynamics for an incompressible, stratified fluid in a gravity field:

\[
\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U} \cdot \nabla) \mathbf{U} + \frac{\nabla p}{\rho} + (\mathbf{f} \times \mathbf{U}) = -\mathbf{g},
\]

(1)

\[
\frac{\partial \rho}{\partial t} + (\mathbf{U} \cdot \nabla) \rho = 0,
\]

(2)

\[
\nabla \cdot \mathbf{U} = 0
\]

(3)

Here the basic variables are: \( \mathbf{U} = (u, v, w) \) is the fluid velocity vector (\( w \) is its vertical component), \( p \) is the fluid pressure, \( \rho \) is its density, \( \mathbf{g} \) is the gravitational acceleration, and \( \mathbf{f} \) is the Earth’s angular frequency vector.

In the ocean, the static density variations are very small, typically less than one percent. This enables one to somewhat simplify the problem by using the Boussinesq approximation.
Let us represent the density field $\rho = \rho_0(z) + \rho'$ as the sum of a large, equilibrium, depth-dependent part $\rho_0(z)$, and a small variable part $\rho'(\mathbf{r}, t)$, where $\mathbf{r} = (x, y, z)$ is the position coordinate, with $x$ and $y$ lying in the horizontal plane, and $z$ directed upward. According to Boussinesq approximation, vertical variations of the static density, $\rho_0(z)$, are neglected in all terms except the buoyancy term proportional to $d\rho_0/dz$ which is, in fact, responsible for the existence of internal waves. Boundary conditions of zero vertical displacement are applied at the bottom, $z = -H$, and at the horizontal surface $z = 0$ that corresponds to the unperturbed water surface, (the “rigid lid” approximation, an analog of Boussinesq approximation for the boundary condition).

The hydrodynamical equations written in the Boussinesq approximation and its ancillary relationships then have the forms
\[ \nabla \cdot \mathbf{u} + \frac{\partial w}{\partial z} = 0, \quad (4) \]

\[ \rho_0 \frac{\partial \mathbf{u}}{\partial t} + \nabla p' + \rho_0 \mathbf{f} \times \mathbf{u} = - \left[ \rho_0 w \frac{\partial \mathbf{u}}{\partial z} + \rho_0 (\mathbf{u} \cdot \nabla) \mathbf{u} \right] \equiv s_1, \quad (5) \]

\[ \frac{\partial \rho}{\partial t} + w \frac{d\rho_0}{dz} = - \left[ w \frac{\partial \rho'}{\partial z} + (\mathbf{u} \cdot \nabla) \rho' \right] \equiv s_2, \quad (6) \]

\[ \frac{\partial \rho'}{\partial z} + g \rho' = - \rho_0 \left[ w \frac{\partial w}{\partial z} + (\mathbf{u} \cdot \nabla) w \right] - \rho_0 \frac{\partial w}{\partial t} \equiv s_3, \quad (7) \]

Here the variables are: \( \mathbf{u} = (u, v) \) is the horizontal fluid velocity vector; \( w \) is its vertical component; \( p' \) is the fluid pressure perturbation; \( f = 2\Omega \sin \varphi \) is the so-called Coriolis parameter or radian frequency; \( (\varphi \) is the geographic latitude and \( \Omega \) is the angular velocity of the Earth’s rotation\(^3\), and \( \nabla \) is the two-dimensional (2D) gradient operator acting on the horizontal plane \((x, y)\). For the derivation of these relationships see, e.g. PHILLIPS (1977), LeBLOND & MYSAK (1978), MIROPOLOSKY (1981) or APEL (1987).

### 3.2 Shallow-water models

Most of the studies devoted to internal solitons deal with moderate-amplitude waves for which the velocity variations in the wave are small compared with the wave phase velocity; this permits us to take into account only linear and quadratic terms in the theory. It is also typically supposed that the characteristic horizontal scale of the wave is large compared with either the depth of the basin or the thickness of the layers where the perturbation mode is localized. In other words, dispersion and nonlinearity are relatively small and comparable in magnitude. These restrictions mean that the right-hand parts of the previous equations specified as \( s_{1,2,3} \) are small, which permits one to use perturbation theory\(^4\). We begin from this approximation, keeping in mind that strongly nonlinear processes also exist in the oceans, and they will be addressed further in this paper.

\[ w = \sum_{m=1}^{\infty} W_m(z)w_m(x, y, t), \quad \mathbf{u} = \sum_{m=1}^{\infty} C_m \frac{dW_m}{dz} \mathbf{U}_m(x, y, t), \quad (8) \]

and similarly for other variables. Vertical displacement of the isopycnal surfaces (those of equal density) is given by \( \xi(x, y, z, t) = \sum_{m=1}^{\infty} \eta_m(x, y, t)W_m(z) \). Here \( C_m \) are constants. The orthogonal eigenfunctions \( W_m \) satisfy the boundary-value problem in the linear, nondispersive approximation:

\(^3\)Here the so-called traditional approximation is used, where only the vertical component of \( \mathbf{f} \) is taken into account, which is valid for long waves (see the references cited in this paragraph).

\(^4\)The “nonhydrostatic” linear term \( \rho_0 \partial w / \partial t \) in \( s_3 \) is small provided the wavelength is large in comparison with the vertical scale.
\[
\frac{d^2 W}{dz^2} + \frac{N^2(z)}{c^2} W = 0, \\
\tag{9}
\]

with boundary conditions \( W(0) = W(−H) = 0 \). From this, the eigenvalues \( c = c_m \) and the eigenfunctions \( W_m \) can be found; note that \( c_m \) has the meaning of a long-wave velocity for each internal mode. The important quantity

\[
N(z) = \sqrt{-\frac{g}{\rho_0} \frac{dp_0}{dz}} \\
\tag{10}
\]

is the Brunt–Väisälä or buoyancy frequency, the rate at which a stably stratified column of water oscillates under the combined influence of gravity and buoyancy forces.

Two simple cases are often considered for the modal problem. The first is the case \( N = \) constant, which occurs when the function \( \rho_0(z) \) is an exponential. For small density variations, this exponential function can be considered as a linear one. In this case \( W(z) \) is a harmonic function, and \( c = c_m \approx NH/m\pi \), where \( m = 1, 2, \ldots \). From here it follows that the first mode is the fastest.

Another very useful model, which will be often considered below, is a fluid consisting of two layers, with upper layer having thickness \( h_1 \) and density \( \rho_1 \), and the lower one, of thickness \( h_2 = H - h_1 \) and density \( \rho_2 > \rho_1 \). This models a sharp jump of the density, a pycnocline, typical of many areas of the ocean. Again, the density difference, \( \delta \rho = \rho_2 - \rho_1 \), is supposed small, \( \delta \rho \ll \rho_1,2 \). In this case, only one internal mode exists and has the following long-wave speed

\[
c = \sqrt{\frac{g \delta \rho}{\rho_m} \frac{h_1 h_2}{h_1 + h_2}}, \\
\tag{11}
\]

where \( \rho_m = \frac{1}{2}(\rho_1 + \rho_2) \) is the mean density of the fluid.

In the general case, after solving (9), approximate equations describing the dependence of physical values on \( x, y, \) and \( t \) in long waves can be derived with the use of different perturbation schemes. Here we briefly describe a rather general model suggested by OSTROVSKY (1978), that reduces the problem to the solution of a system of coupled evolution equations in a form analogous to the Boussinesq equations (which should not be confused with the Boussinesq approximation) for long, weakly nonlinear surface waves. A variable \( \eta \) is used that characterizes the vertical displacement of an isopycnic surface from their equilibrium levels. Along this undulating surface,

\[
w = \frac{\partial \eta}{\partial t} + (U \cdot \nabla) \eta, \\
\tag{12}
\]

so that at \( z = \) const we have

\[
w \simeq \frac{\partial \eta}{\partial t} + (U \cdot \nabla) \eta + \frac{\partial^2 \eta}{\partial z \partial t} \eta \\
\tag{13}
\]

if we neglect nonlinear terms of the third and higher orders. Substituting this into the basic set of equations, orthogonalizing them [i.e. multiplying each equation by \( W \) or \( dW/dz \) and
integrating over $z$ at the interval $[-H, 0]$, and then invoking some elementary transformations, we finally obtain a system of 2D coupled equations. In the absence of any resonance interactions, each mode can be considered as independent, which yields the following system:

$$\frac{\partial \eta}{\partial t} + H (\nabla U) + \frac{\sigma}{2} (\nabla \cdot \eta U) = 0,$$  \hspace{1cm} (14)

$$\frac{\partial U}{\partial t} + \frac{c^2}{H} \nabla \eta + (f \times U) \left( 1 - \frac{\sigma \eta}{2H} \right) + \sigma \left[ (U \nabla) U - \frac{1}{2H} \frac{\partial \eta U}{\partial t} \right] + DH \nabla \frac{\partial^2 \eta}{\partial t^2} = 0. \hspace{1cm} (15)$$

Here $\sigma$ and $D$ are non-dimensional parameters describing nonlinearity and high-frequency dispersion, respectively. For each mode, they are determined by

$$\sigma = \frac{H}{Q} \int_{-H}^{0} \left( \frac{dW}{dz} \right)^3 d z, \hspace{1cm} D = \frac{1}{H^2 Q} \int_{-H}^{0} W^2 d z, \hspace{1cm} Q = \int_{-H}^{0} \left( \frac{dW}{dz} \right)^2 d z. \hspace{1cm} (16)$$

Equations (14) and (15) are the extensions of the Boussinesq equations, well known for surface waves, to the internal modes. A known peculiarity should be noted here: for the case of $N(z) = \text{const}$, the nonlinear parameter $\sigma$ is zero, so that the nonlinearity vanishes in these equations, and reveals itself only in either the next (cubic) approximation or by going beyond the Boussinesq and/or rigid lid approximations.

At small nonlinearity, only a weak mode coupling exists, that usually leads to small corrections to the shape of the soliton and to its velocity, as long as there is no resonant coupling between different modes, such as occurs, for instance, when their phase velocities are close to one another. If the latter is not the case, one may consider each mode separately. However, there are important cases of complex density profiles wherein strong mode coupling may occur. Some effects of neighboring mode coupling on the propagation of the Korteweg-de Vries (KdV) soliton were evaluated in the paper by VLASENKO (1994). There it was shown that the influence of an $n$-th mode on the fixed $m$-th mode decreases in inverse proportion to $|n - m|$.

It is interesting that the system (14), (15), which here describes internal wave modes, is also applicable to long-wavelength Rossby (or planetary/potential vorticity) waves that exist when the Coriolis parameter $f$ depends on the horizontal coordinate $y$ (the latitude) via $f \approx f_0 + \beta y$ (PEDLOSKY, 1987). In this case, $\beta$ describes the variation of Coriolis frequency with latitude ($\beta$-plane approximation).

### 3.3 Internal Waves in Nonrotating Fluids

Let us first examine the well-investigated case of internal waves propagating in an arbitrarily stratified but nonrotating fluid, thus taking $f = 0$. Suppose that the associated linear eigenvalue problem has already been solved and that the modal speeds $c_m$ are known. Let us now take into account small dispersion and small nonlinearity. Then for one-dimensional
waves propagating along the $x$-axis, the Boussinesq set of equations for IW, (14) and (15), reduces to

$$\frac{\partial \eta}{\partial t} + H \frac{\partial U}{\partial x} = -\frac{\sigma}{2} \frac{\partial (\eta U)}{\partial x},$$  \hspace{1cm} (17)

$$\frac{\partial U}{\partial t} + \frac{c^2}{H} \frac{\partial \eta}{\partial x} = -\sigma \left[ U \frac{\partial U}{\partial x} - \frac{1}{2H} \frac{\partial (\eta U)}{\partial t} \right] - DH \frac{\partial^3 \eta}{\partial x \partial t^2}.$$  \hspace{1cm} (18)

If one considers a solution of this set in the form of a stationary solitary wave vanishing at infinity and depending on one variable $\xi = x - Vt$, one obtains a soliton in the implicit form

$$\sqrt{\frac{2}{DH^2}}(\xi - \xi_0) = \pm \sqrt{1 + \zeta_0} \left[ 2 \arctan \sqrt{\frac{1 + \zeta}{\zeta_0 - \zeta} - \frac{1}{\sqrt{\zeta_0}}} \ln \left| \frac{\sqrt{\zeta_0 - \zeta} + \sqrt{\zeta_0(1 + \zeta)}}{\sqrt{\zeta_0 - \zeta} - \sqrt{\zeta_0(1 + \zeta)}} \right] \right].$$  \hspace{1cm} (19)

Here $\zeta = \frac{\eta_0}{2H}$. Formally this solution is valid even for strong nonlinearity which, however, contradicts to the applicability of Eqs. (17) and (18) derived under the condition of small nonlinearity. Still, this solution could be of some interest from the mathematical viewpoint. In the small-amplitude limit, $\zeta \ll 1$, Eq. (19) reduces to the explicit form discussed below. Evidently, the signs $\pm$ in Eq. (19) correspond to the waves propagating in opposite directions along the axis $x$.

3.3.1 The Korteweg–de Vries (KdV) Equation

For a progressive wave propagating in the positive direction of axis $x$, the classical Korteweg–de Vries equation (KORTEWE & DE VRIES, 1895) widely discussed in literature (see, e.g. WHITHAM, 1974; MIROPOLSKY, 1981; ABDLOWITZ & SEGUR, 1981) readily follows from the Boussinesq set of equations:

$$\frac{\partial \eta}{\partial t} + c \frac{\partial \eta}{\partial x} + \alpha \eta \frac{\partial \eta}{\partial x} + \beta \frac{\partial^3 \eta}{\partial x^3} = 0,$$  \hspace{1cm} (20)

where the re-scaled nonlinear and dispersion parameters ($\alpha$ and $\beta$ respectively) are

$$\alpha = \frac{3c \sigma}{2H}, \quad \beta = \frac{cDH^2}{2}$$  \hspace{1cm} (21)

with $\sigma$ and $D$ given by Eq. (16). The important quantities $\alpha$ and $\beta$ are known as environmental parameters and incorporate the effects of buoyancy (density stratification), shear currents in general (see below) and depth via their effects on the eigenfunction profiles, $W(z)$.

The well-known solitary solution to Eq. (20) is

$$\eta(x, t) = \eta_0 \text{sech}^2 \frac{x - Vt}{\Delta},$$  \hspace{1cm} (22)
the nonlinear velocity \( V \) and the characteristic width \( \Delta \) of this soliton being related to the linear speed \( c \) and the amplitude of the displacement \( \eta_0 \) by

\[
V = c + \frac{\alpha \eta_0}{3}, \quad \Delta^2 = \frac{12 \beta}{\alpha \eta_0}.
\] (23)

The dispersion parameter \( \beta \) is always positive for oceanic gravity waves (although for capillary waves on a surface of thin liquid films, this parameter may be negative). The sign of the nonlinear parameter \( \alpha \) may be both positive and negative. The combination of parameters \( \alpha \) and \( \beta \) determines the soliton polarity; namely, the sign of \( \eta_0 \) is such that \( \Delta^2 \) in Eq. (23) is positive. Thus, if \( \alpha \) is negative, so will be \( \eta_0 \), i.e. the soliton is a wave of isopycnal depression. This appears to be the usual case where a shallow pycnocline overlies deeper water. However, in shallow seas with strong mixing, the reverse situation may occur, with the pycnocline being located near the bottom. In this case \( \alpha \) and \( \eta_0 \) are both positive.

Let us consider the aforementioned two-layer model where \( \rho(z) = \rho_1 \) for \( 0 > z > -h_1 \) and \( \rho(z) = \rho_2 > \rho_1 \) for \( -h_1 > z > -H \). In this case we have

\[
c = \left[ \frac{g(\rho_2 - \rho_1)h_1h_2}{\rho_2h_1 + \rho_1h_2} \right]^{1/2} \approx \left[ \frac{g\delta \rho}{\rho_m} \frac{h_1h_2}{h_1 + h_2} \right]^{1/2},
\]

\[
\alpha = \frac{3c}{2h_1h_2} \frac{\rho_2h_1^2 - \rho_1h_2^2}{\rho_2h_1 + \rho_1h_2} \approx \frac{3}{2} \frac{h_1 - h_2}{h_1h_2},
\]

\[
\beta = \frac{c_1 h_2 \rho_1 h_1 + \rho_2 h_2}{6 \rho_{2h_1 + \rho_1h_2}} \approx \frac{c h_1 h_2}{6}.
\] (26)

The relations on the right are valid for the ocean, where \( \delta \rho = \rho_2 - \rho_1 \) is always small. As seen from (25), solitons propagating on a thin upper layer over a deeper lower layer are always negative, i.e. depressions, whereas solitons riding on near-bottom layers are elevations\(^5\).

The one-dimensional KdV equation (20) can be derived directly from the hydrodynamic equations (4)–(7) in their 2D form [see, e.g. (GRIMSHAW ET AL., 2002b) and references therein]. However, the Boussinesq-type equations (14) and (15) have their own value. They are valid for arbitrary stratification and also allow various generalizations of the KdV equation, such as the Kadomtsev–Petviashvili equation shown below. The soliton solution to the Boussinesq equations, Eq. (19), can be presented in the form of Eq. (22) for small-amplitude perturbations but relationships between the amplitude, \( \eta_0 \), velocity, \( V \), and half-width, \( \Delta \), of a soliton, are slightly different. In particular, for the two-layer model the solitary wave solution was firstly derived by Keulegan, (1953) from the corresponding two-layer Boussinesq equations. Instead of (23), he obtained

\[
V = c \sqrt{1 + \frac{h_1 - h_2}{h_1h_2} \eta_0}, \quad \Delta^2 = \frac{4}{3} \frac{h_1^2h_2^2}{(h_1 + h_2)\eta_0}.
\] (27)

\(^5\)Note that surface wave solitons in natural basins or estuaries would always be humps and never depressions. However, capillary solitons on a surface of thin films may be of negative polarity, i.e. they may represent surface depressions. Such solitons were recently observed in laboratory experiments (FALCON ET AL., 2002).
In the limit of $\eta_0 \to 0$ these formulae reduce to the corresponding expressions (23) for the KdV soliton.

The KdV equation belongs to the class of completely integrable systems. It was a subject of intense study during the past five or so decades. Currently it is one of the most thoroughly studied of nonlinear equations, and we shall not go into details which can be easily found in numerous books and reviews [see, e.g., (SCOTT ET AL., 1973; WHITHAM, 1974; MILES, 1980; ABLÒWITZ & SEGUR, 1981; DODD ET AL., 1982)]. Rather, we will just list a few salient points of interest. Note first that it belongs to the class of exactly integrable equations for which an infinite set of integrals of motion exists. A remarkable process worth noting is the interaction of KdV solitons, from which they escape unchanged, similar to two colliding rigid particles, only acquiring an additional delay (phase shift) at a given distance (hence the name of soliton). Another important feature of the KdV equation is that solitons can arise from arbitrary localized perturbations having the same polarity as a soliton. Moreover, if the total “mass” of an initial perturbation,

$$M = \int_{-\infty}^{\infty} \eta(x, t) \, dx,$$

is nonzero and its sign coincides with the soliton polarity, at least one soliton will emerge, even for a small-amplitude and small-width perturbation (KARPMAN, 1973). In particular, an initial delta-impulse, $\eta(x, 0) = \eta_0 \delta(x)$, where $\delta(x)$ is Dirac delta-function, always evolves into one soliton followed by a dispersive “tail” (ABLOWITZ & SEGUR, 1981). Perturbations with the opposite sign of mass never generate solitons but rather disperse into a long oscillatory wavetrain, whose amplitude eventually tends to zero. The number and parameters of solitons produced by an initial pulse can be calculated exactly by the inverse scattering method or evaluated approximately by means of perturbation techniques [see, e.g., (KARPMAN, 1973; ABLÒWITZ & SEGUR, 1981)]. The result depends on the value of the Ursell parameter, $U_r = \alpha A_0 L_0^2 / \beta$, where $A_0$ and $L_0$ are the amplitude and characteristic width of the initial perturbation, respectively. Some examples of experimental observations of these processes will be illustrated in the forthcoming Sections. Here we will only mention that for an actual KdV soliton (22) whose characteristic width, $\Delta$, is related to the amplitude, $\eta_0$, according to Eq. (23), the Ursell parameter is equal to 12, independent of the soliton amplitude.

Transient processes. The single pulse solution to the KdV equation in the form of Eq. (22) is very simple, and readily provides physical insight when examined. However, a common observation in the ocean is of wave-trains consisting of several oscillations with wavelengths, crest lengths, and amplitudes varying from the front to the rear of the wavetrain (as schematically shown in Fig. 1 plotted from a simple theoretical model of evolution of an initial step-function within the framework of the KdV equation). As these oscillations, especially the few frontal ones, are very close to being a series of solitons (indeed, each oscillation develops

---

6 This parameter is known in nonlinear wave theory as the similarity parameter of the KdV equation (KARPMAN, 1973). In this paper we use the term “Ursell parameter” from the surface-wave terminology and apply it to the general water waves.
into independent soliton at infinity), and the entire perturbation represents an undular bore, it is sometimes called solibore (APEL, 2003).

Figure 1: Disintegration of a stepwise perturbation into a train of solitons within the framework of the KdV equation (a simplistic scheme of solibore formation). Axes are in arbitrary units.

KORTEWEG AND DE VRIES (1895) had already found periodic solutions to their equation in the form of the so-called “cnoidal” waves, which involve the Jacobi elliptic function \( cn_s(x) \). This function has a nonlinear parameter, \( s \), that characterizes the degree of non-linearity, with \( 0 \leq s \leq 1 \). For the KdV equation, the cnoidal solution is given by [see, e.g. (KARPMAN, 1973)]:

\[
\eta(x, t) = \eta_m + \eta_0 cn_s^2[k_0(x - Vt)]
\]  

(29)

In the above solution \( \eta_0 \) is the wave magnitude, \( \eta_m \) is the constant background and \( k_0 \) is the wavenumber; the soliton velocity \( V \) can be expressed in terms of these parameters. This solution reduces to a harmonic wave when \( s \rightarrow 0 \) and \( cn_s x \rightarrow \cos x \), and to a solitary wave when \( s \rightarrow 1 \) and \( cn_s x \rightarrow \text{sech}^2 x \). Thus, the soliton can be considered as a limit of a periodic wave train at \( s = 1 \).

However, there is still some ocean phenomenology missing from the cnoidal solution. Specifically, it does not describe transient processes such as the onset and the long-term trailing edge displacement of the isopycnal surfaces behind the wave group. Using an approximate approach suggested by WHITHAM (1974), GUREVICH & PITAЕVKII (1973) have constructed a self-similar solution for the evolution of an initially stepwise perturbation into a train of oscillations with a slow variation of the nonlinear parameter \( s \) within the train.
In the process of evolution, these oscillations become deep at the front of the perturbation forming a set of separated impulses, each close to a soliton, and eventually decrease to a constant trailing edge. A recent review on further development of the Gurevich & Pitaevskii approach can be found in EL ET AL., 2005. An exact analytical solution describing the disintegration of a stepwise initial perturbation into solitons was obtained by KRUSLOV (1975, 1976) using an inverse scattering method. Figure 1 illustrates such a process.

To describe an oceanic nonlinear wave train with oscillatory behavior at the leading edge and a constant depression at the trailing edge, APEL (2003) has applied the Gurevich & Pitaevskii approach to modeling the internal solibores. For a stepwise initial impulse exerted on a fluid at \( t = 0 \), the solution of the KdV equation is given by:

\[
\eta(x, t) = \eta_m + \eta_0 \left\{ \text{dn}^2_s[k_0(x - Vt)] - (1 - s^2) \right\},
\]

where \( \text{dn}_s(x) \) is another periodic elliptic function (Jacobi delta-amplitude) which tends to unity at \( s \to 0 \) and to \( \text{sech}^2(x) \) when \( s \to 1 \); and \( s \) is a slowly varying function of \( x \) and \( t \).

This solution, which Apel has named the “dnoidal” wave, can be suitable for describing weakly nonlinear internal tides. For this case, initial tidal perturbations have a finite duration and are relatively smooth. Thus, the process of solilobore formation includes a stage of wave steepening and the subsequent formation of oscillations. The first stage may be described by the equation of a “simple wave” which is in fact a KdV equation with \( \beta = 0 \), i.e. the dispersionless KdV equation. Each point of such a wave propagates at its own velocity, \( c + \alpha \eta \), until the wave front becomes steep [see the details in (APEL, 2003; EL ET AL., 2005)]. At that point, the dispersion effects must be taken into account, which leads to the formation of solitons at the frontal zone of each tidal period, which smoothly transfer to a dnoidal wave with variable parameters at the rear. To accommodate this relaxation back to the equilibrium state, Apel introduced an “internal tide recovery function” \( I(x, t) \), which multiplies the dnoidal solution. This function takes the dnoidal solution back to equilibrium, using just one adjustable parameter which is the time required for the relaxation to occur.

### 3.3.2 The Extended and Modified Korteweg–de Vries Equations

It follows from the above that for \( \rho_2 h_1^2 < \rho_1 h_2^2 \) (that is practically \( h_1 < h_2 \) for the ocean), solitons cause the interface (the pycnocline) to descend, and vice versa if the inequality is reversed. Of some interest is the special case when \( h_2^2/h_1^2 \approx \rho_2/\rho_1 \approx 1 \), i.e. the interface is close to the middle of the water layer. In this case the nonlinear coefficient \( \alpha \) is small or even equal to zero. As mentioned before, in this case one must either abandon the Boussinesq and rigid lid approximations or take into account higher-order nonlinear terms in the evolution equations. In the latter case, the extended Korteweg–de Vries (eKdV) equation (also called the combined KdV and Gardner equation), having both quadratic and cubic nonlinearities, results (LEE & BEARDSLEY, 1974; DJORDJEVIC & REDEKOPP, 1978; KAKUTANI & YAMASAKI, 1978; MILES, 1979; GEAR & GRIMSHAW, 1983; SMYTH & HOLLOWAY, 1988; LAMB & YAN, 1966):
\[
\frac{\partial \eta}{\partial t} + \left( c + \alpha \eta + \alpha_1 \eta^2 \right) \frac{\partial \eta}{\partial x} + \beta \frac{\partial^3 \eta}{\partial x^3} = 0,
\]

where for the case of two-layer fluid the second nonlinear coefficient is

\[
\alpha_1 = \frac{3c}{h_1^2 h_2} \left[ \frac{7}{8} \left( \frac{\rho_2 h_1^2 - \rho_1 h_2^2}{\rho_2 h_1 + \rho_1 h_2} \right)^2 - \frac{\rho_2 h_1^3 + \rho_1 h_2^3}{\rho_2 h_1 + \rho_1 h_2} \right] \approx -\frac{3}{8} c \frac{(h_1 + h_2)^2 + 4h_1 h_2}{h_1^2 h_2}.\]

The last expression is again valid for the case of close densities which we shall consider below.

This equation, as well as its generalization containing a combination of higher-order nonlinear and dispersive terms, was derived by many authors starting from the paper by Lee & Beardsley (1974). A contemporary derivation, convenient for applications, can be found, e.g. in (DJORDJEVIC & REDEKOPP, 1978; GRIMSHAW ET AL., 2002b).

As follows from Eq. (32), within the framework of two-layer model, \( \alpha_1 \) is always negative. However, in the general case the coefficient \( \alpha_1 \) may be either negative or positive (GRIMSHAW ET AL., 1997; TALIPova ET AL., 1999).

If the pycnocline is located just at the critical level so that the parameter \( \alpha \) is exactly zero, Eq. (31) reduces to the well-known modified Korteweg–de Vries (mKdV) equation. In the geophysically most interesting case, when \( \alpha_1 < 0 \) and \( \beta > 0 \), this equation has no stationary solitary wave solutions asymptotically vanishing at \( x \to \pm \infty \). However, it has a particular solution that is a type of stepwise transition, which can be considered to be a soliton in a more general sense. Such a solution is usually called a kink and has the form of a bore moving into a depression area (PERELMAN ET AL., 1974; ONO, 1976; ROMANOVA, 1979; MILES, 1981; GROSSE, 1984; FUNAKOSHI & OIKAWA, 1986):

\[
\eta = \pm \eta_0 \tanh \left( \frac{x - vt}{\Delta} \right),
\]

where now

\[
V = c + \frac{\alpha_1 \eta_0^2}{3}, \quad \text{and} \quad \Delta^2 = -\frac{6 \beta}{\alpha_1 \eta_0^2}.\]

Note that the vertical velocity component \( w \approx \partial \eta / \partial t \) has the form of a localized pulse, so that it may properly be treated as a soliton. The specific feature of such a kink is that its velocity, \( V \), is always less then the linear velocity, \( c \), due to \( \alpha_1 < 0 \).

Internal bores in a two-layer fluid were recently considered in (DIAS & VANDEN-BROECK, 2003). They studied numerically the structure of steady-state bores of arbitrary amplitude within the framework of the primitive set of hydrodynamic equations for inviscid fluid. Two families of kink-type solutions were found in the form of the elevation and depression fronts. These solutions included bores of limiting configuration, when the interface reaches either the upper or lower boundary (i.e. free surface or bottom). Limiting depression bores always have monotonic fronts, whereas their counterparts in elevation may have non-monotonic profiles.
The hydrodynamic stability of internal solitons and bores has been only scarcely studied. However, waves in a two-layer fluid always create a discontinuity of the tangential velocity at a pycnocline. In particular, for a bore such a velocity jump extends far behind its front. In general, in such a shear flow, the Kelvin–Helmholtz instability does exist (e.g. STEPANYANTS & FABRIKANT, 1996). This issue is interesting because some observations of internal bores in lakes, seas, and oceans have already been reported (see, e.g. THORPE, 1971; WINANT, 1974; IVANOV & KONYAEV, 1976).

The mKdV equation also has soliton-type solutions, but only those propagating on a constant nonzero pedestal (PERELMAN ET AL., 1974; ONO, 1976; ROMANOVA, 1979; FUKANOSHI & OIKAWA, 1986). These solutions are interesting not so much by themselves as they are within the framework of the eKdV Eq. (31) with \( \alpha \neq 0 \) (note that under the transformation \( \eta = u - \alpha / 2 \alpha_1 \), Eq. (31) can be reduced to the mKdV form). In this case the solitary solution of Eq. (31) can be written in the form of a kink-antikink pair of a stationary shape:

\[
\eta(x, t) = -\frac{\alpha \nu}{\alpha_1 2} \left[ \tanh \left( \frac{x - V t}{\Delta} + \phi \right) - \tanh \left( \frac{x - V t}{\Delta} - \phi \right) \right],
\]

or, equivalently,

\[
\eta(x, t) = -\frac{\alpha \nu}{\alpha_1 2} \frac{\sinh(2\phi)}{\cosh^2[(x - V t)/\Delta] + \sinh^2\phi},
\]

where \( \nu \) is a free dimensionless parameter with the range \( 0 < \nu < 1 \), and the remaining parameters are

\[
\phi(\nu) = \frac{1}{4} \ln \left( \frac{1 + \nu}{1 - \nu} \right), \quad \Delta = \sqrt{\frac{-24\alpha_1 \beta}{\alpha^2 \nu^2}}, \quad V = \frac{-\alpha^2 \nu^2}{6\alpha_1}.
\]

In contrast with the kink described by (33) and (34), the velocity of this soliton is always greater than the linear velocity \( c \). This family of solutions has rather interesting properties. The amplitude of the soliton, \( \eta_0 = - \left( \alpha / \alpha_1 \right) \nu \tanh \phi \), varies from zero up to a maximum of \( |\alpha / \alpha_1| \), in contrast to the amplitude of the KdV soliton, which in principle can range from zero to infinity. When soliton amplitude approaches its maximum value, its width increases so that the soliton profile changes from the bell shape to the rectangular shape. In the limit \( \nu \rightarrow 1 \), the eKdV soliton tends to two infinitely separated kinks.

In the near-critical situation when \( h_1 \approx h_2 = h \) and the eKdV equation is indeed applicable, the amplitude does not exceed \( |h_2 - h_1|/2 \), and the velocity cannot exceed the value of

\[
V_{\text{max}} = c - \frac{\alpha^2}{6\alpha_1} \simeq c \left[ 1 + \frac{(h_2 - h_1)^2}{8h^2} \right].
\]

Note that if the ratio \( h_1 / h_2 \) is as close to unity as \( \rho_2 / \rho_1 \), the latter ratio must be taken into account in these formulae. For example, the maximal soliton amplitude is \( |\rho_2 h_1 - \rho_1 h_2|/2\rho_2 \).
Furthermore, the solution (35) has in fact two spatial scales: that of the hyperbolic tangent profile, as characterized by the parameter $\Delta$, and a distance between these profiles characterized by the parameter $S = \phi(\nu)\Delta$. In general, the actual width of the soliton is determined by these parameters, which in turn depend on the hydrodynamic environment and the amplitude through the free parameter $\nu$. Figure 2 shows normalized shapes of solitons for three values of the modified free parameter $\epsilon \equiv 1 - \nu$. The evolution from a classical KdV soliton when $\nu$ is small and the characteristic total width $D \approx 2\Delta$ to the flat-top kink–antikink construction at $\nu \to 1$, in which case $D \approx 2S$, is clear.

![Normalized wave shapes in the eKdV equation (35) for three values of the parameter $\epsilon = 1 - \nu$](image)

Figure 2: Normalized wave shapes in the eKdV equation (35) for three values of the parameter $\epsilon = 1 - \nu$: $1 - \epsilon = 10^{-1}$ (close to the KdV case); $2 - \epsilon = 10^{-4}$; $3 - \epsilon = 10^{-7}$.

The width of the soliton increases in both limits: $\nu \to 0$ and $\nu \to 1$. Hence, for some $\nu = \nu_m$ there exists a minimum value of $D$. Figure 3 depicts $D_{0.5}$, the full width of the soliton at the half its maximum amplitude, as a function of the amplitude, $\eta_0$. The minimum of $D_{0.5}$ occurs at $\nu \approx 0.9$, when the amplitude is about 0.56 of the maximum. A more detailed discussion of the dependency between $D$ and $\eta_0$ both for weakly nonlinear perturbations, described by the eKdV equation, and for more intensive perturbations, described by the primitive Eulerian equations, can be found in the paper by Funakoshi & Oikawa (1986).
Figure 3: Dependency of the characteristic width, \( \widetilde{D}_{0.5} \), of eKdV solitons, Eq. (35), on amplitude \( \tilde{\eta}_0 \) in dimensionless variables: \( \widetilde{D}_{0.5} = D_{0.5}|\alpha|/(24\alpha_1\beta)^{1/2} \); \( \tilde{\eta}_0 = \eta_0\alpha_1/\alpha \).
From Eq. (37) it follows that

$$\eta_0 = -\frac{\alpha \nu \sqrt{1+\nu} - \sqrt{1-\nu}}{\alpha_1 \sqrt{1+\nu} + \sqrt{1+\nu}}$$

(39)

with $\nu$ related to $\phi$ and $\Delta$ by Eqs. (37).

An interesting observation was made in (GRIMSHAW ET AL., 1997) and (TALIPOVA ET AL., 1999). It has been shown that given a certain hydrology (density stratification), the cubic nonlinear coefficient $\alpha_1$ may in fact become positive. In particular, this is the case of a three-layer fluid with two density jumps located symmetrically with respect to the middle of the fluid layer. The coefficients of the eKdV equation (31) for one of two internal modes in the Boussinesq approximation ($\rho_1 \approx \rho_2$) are

$$c = \sqrt{\frac{\Delta \rho}{\rho} gh}, \quad \alpha = 0, \quad \alpha_1 = -\frac{3c}{4\rho^2} \left(13 - \frac{9H}{2h}\right), \quad \beta = \frac{ch}{4} \left(H - \frac{4h}{3}\right),$$

(40)

where $H$ is the total fluid depth and $h < H/2$ is the thickness of the upper and the lower layers. At the critical thickness, $h_{cr} = 9H/26$, both quadratic and cubic nonlinear coefficients vanish (however for another mode, the coefficient $\alpha$ remains nonzero). Apparently, the next order nonlinearity must be taken into account in this case.

Higher-order KdV equations containing corrections both to the nonlinear and dispersive terms have been derived for internal waves in a stratified shear flows in many papers beginning from the aforementioned pioneering paper by LEE & BEARDSLEY (1974). For a rather general and convenient form of this derivation, see, e.g. (GRIMSHAW ET AL., 2002b) and (POLOUKHINA ET AL, 2002). In the case of positive cubic nonlinearity (e.g., when $h < 9H/26$, where $H$ is the total fluid depth), solitons of both positive and negative polarities may exist on a zero background. In addition to that, nonstationary solitons, called breathers, are also possible. The evolution of initial pulse-type perturbation may be fairly complex (GRIMSHAW ET AL., 1997).

The mathematical theory of the eKdV equation has been developed in many papers for different combinations of signs of nonlinear and dispersion terms (MILES, 1981; MARCHANT & SMITH, 1996; SLYUNYAEV & PELINOVSKI, 1999; SLYUNYAEV, 2001; GRIMSHAW ET AL., 2002A). It was shown that the eKdV equation and its reduced version, the mKdV equation, are also completely integrable equations as is the usual KdV equation. In particular, SLYUNYAEV & PELINOVSKI (1999) have studied in detail the evolution of an initial pulsed disturbance and analyzed an exact two-soliton interaction in the case when the dispersion and cubic nonlinear coefficients of Eq. (31) are of opposite signs (i.e. $\beta > 0$ and $\alpha_1 < 0$). GORSHKOV & SOUSTOVA (2001) suggested an approximate description of the multi-soliton interaction based on the perturbation theory for solitons and kinks. This theory was applied to an experimental situation in the ocean (GORSHKOV ET AL., 2004).

The evolution of initial perturbations in the case when the dispersion and cubic nonlinear coefficients of Eq. (31) are of the same signs ($\beta > 0$ and $\alpha_1 > 0$) was studied in SLUNYAEV, 2001. And GRIMSHAW ET AL. (2005) have shown that this situation is typical for the shelf zones of the World Ocean.
Although the eKdV equation is valid for small nonlinearity and a specific stratification, sometimes it can be successfully applied to the description of strongly nonlinear internal solitons as a phenomenological model, whereas the usual KdV equation fails to approximate observational and laboratory data (Stepanyants, 1990; Stanton & Ostrovsky, 1998; Michallet & Barthélemy, 1998). The reason for this is a qualitative (but in general not quantitative!) correspondence of the eKdV solitons to those of strongly nonlinear solitary waves in a two-layer fluid. The correspondence relates, in particular, to the non-monotonic dependence of their height on the amplitude and to the existence of a limiting amplitude.

3.3.3 The Benjamin–Ono Equation

An important modification is needed if the wavelength is large compared with one (say, upper) layer but small compared with the other (lower) layer of the ocean, so that one can let $h_2 \to \infty$. These waves can be described by another completely integrable model, namely by the differential-integral Benjamin–Ono (BO) equation [see, e.g., (Ablowitz & Segur, 1981)]:

$$\frac{\partial \eta}{\partial t} + c \frac{\partial \eta}{\partial x} + \alpha \eta \frac{\partial \eta}{\partial x} + \beta \frac{\partial^2}{\partial x^2} \int_{-\infty}^{\infty} \frac{\eta(x', t)}{x - x'} \, dx' = 0,$$

where the symbol $\varphi$ indicates that the principal value of the integral should be taken, and the coefficients are

$$c = \sqrt{(\rho_2 - \rho_1)gh_1 \rho_1}, \quad \alpha = -\frac{3}{2} \frac{c}{h_1}, \quad \beta = \frac{c \rho_2}{2 \rho_1}. \quad (42)$$

Solitons described by this equation are also well known:

$$\eta(x, t) = \frac{\eta_0}{1 + (x - \eta t)^2/\Delta^2}. \quad (43)$$

Their amplitudes $\eta_0$, velocities $V$, and half-widths $\Delta$ are related by

$$V = c + \frac{\alpha \eta_0}{4}, \quad \text{and} \quad \Delta = \frac{4 \beta}{\alpha \eta_0}. \quad (44)$$

The displacement of these solitons is a downgoing motion of the interface when the upper layer is thin, and conversely for the case when the thin layer lies near the bottom (there is a general thumb rule: pycnocline displacement induced by a soliton is directed to the deepest layer).

As in most integrable cases, the BO solitons restore their parameters after a collision. However, unlike the KdV case, the displacement in the BO soliton decreases algebraically, as $x^2$, rather than exponentially (therefore they are often called “algebraic solitons”). Another difference is that BO solitons do not acquire a phase shift after a collision. As was shown by D. Pelinovsky & C. Sulem in 1998, they are stable with respect to small perturbations and can emerge from arbitrary pulse-type initial perturbations of appropriate polarity, i.e., the polarity required for the existence of a BO soliton.
3.3.4 The Joseph–Kubota–Ko–Dobbs (JKKD) Equation

Apparently, Whitham (1974) was the first to point out explicitly that a linear evolution equation can be constructed by applying the inverse Fourier transform to a dispersion relation known for a harmonic wave. Such an equation can be differential or, more generally, integro-differential. For finite-amplitude perturbations it can be supplemented by a nonlinear term derived in the non-dispersive, long-wave approximation, and which accounts for a hydrodynamic nonlinearity of the type $\eta \frac{\partial \eta}{\partial x}$. This approach, albeit not quite consistent, leads to useful model equations in cases when the regular derivation is cumbersome or even impossible.

Following these ideas, Joseph (1977) and Kubota et al. (1978) introduced a model evolution equation, which is termed here the JKKD equation\(^7\). It describes weakly nonlinear perturbations propagating within a thin oceanic layer surrounded from above and below by thick homogeneous layers of fluid of arbitrary depth. A similar equation was later derived rigorously by Segur & Hammack (1982) for a slightly different configuration in which a layer of lighter fluid overlies a layer of heavier fluid. The thickness of one of the layers, say the upper one, $h_1$, is assumed to be small in comparison with the thickness of a lower layer, $h_2$, i.e. $h_1/h_2 \ll 1$. At the same time the perturbation wavelength, $\lambda \gg h_1$, may have an arbitrary relationship with $h_2$, i.e. the total water layer can be either shallow or deep. The resulting evolution equation can be presented in a variety of equivalent forms; one of the simplest is

$$\frac{\partial \eta}{\partial t} + c \frac{\partial \eta}{\partial x} + \alpha \eta \frac{\partial \eta}{\partial x} - \beta \frac{\partial^2 \eta}{\partial x^2} \int_{-\infty}^{\infty} \eta(\xi'/h_2, t) \frac{\tanh \left( \frac{x - x'}{2h_2} \right)}{\tanh \left( \frac{\xi - \xi'}{2h_2} \right)} \, dx' = 0, \quad (45)$$

where, for the two-layer model with a sharp density interface, the parameters $c$ and $\alpha$ are the same as in Eq. (41) and $\beta = c h_1/(4h_2)$.

The dispersion relation corresponding to this equation and relating the wavenumber $k$ with a frequency $\omega$ of the linear perturbations, $\eta \sim \exp (kx - \omega t)$, is

$$\omega = ck \left( 1 - \frac{k h_1}{2 \tanh k h_2} \right). \quad (46)$$

In the shallow-water ($kh_2 \to 0$) and deep-water ($kh_2 \to \infty$) limits this reduces to the KdV and BO dispersion relations, respectively.

The JKKD equation has a solitary solution which has been obtained by many authors and presented in different forms (Joseph, 1977; Chen & Lee, 1979; Segur & Hammack, 1982). One of the forms convenient for practical applications is

$$\eta(x, t) = \frac{\eta_0}{1 + \frac{2}{1 + \cos \frac{\partial x}{\Delta}} \sinh^2 \frac{x - Vt}{\Delta}}, \quad (47)$$

where

---

\(^7\)It is also called the intermediate long wave (ILW) or finite-depth (FD) equation.
\[
\eta_0 = \frac{4 h_1^2}{3 \Delta} \frac{\sin \frac{2k_2}{\Delta}}{1 + \cos \frac{2k_2}{\Delta}}, \quad V = c \left( 1 - \frac{h_1}{\Delta \tan \frac{2k_2}{\Delta}} \right).
\] (48)

Here \(\Delta\) is a free parameter characterizing the soliton width.

Equations (47) and (48) reduce to a KdV soliton, Eqs. (22) and (23), in the limit of \(h_2/\Delta \to 0\), for a fixed \(h_1/h_2\). Meanwhile, as was pointed out by CHEN & LEE (1979), there is no smooth transition from the JKKD soliton to the BO soliton as \(h_2 \to \infty\), in contrast with Joseph’s claim (JOSEPH, 1977). This issue still remains unclear because Eq. (45) tends to either the KdV or the BO limit as \(h_2 \to 0\) or \(h_2 \to \infty\), respectively. CHEN & LEE (1979) also found a periodic solution for the JKKD equation, which reduces to the algebraic BO soliton of the form of Eq. (43) but only for one fixed set of parameters \(\eta_0\), \(V\), and \(\Delta\). SEGUR & HAMMACK (1982) have derived the next-order correction to the wave amplitude for the JKKD equation, and have obtained the corresponding corrections to the JKKD soliton.

The main features of the initial perturbation dynamics within the framework of the JKKD equation are very similar to those described by the KdV model. It is worth noting here that all the models considered above, beginning from the KdV, are analytically integrable, possessing an infinite set of conservation laws and multi-soliton solutions [see, e.g. (ABLOWITZ & SEGUR, 1981)]. Their properties have been thoroughly studied by mathematicians.

However, seemingly crude it may be, the two-layer model often gives a good approximation to situations in which the density and velocity vary continuously but sharply with depth; this occurs especially for the first vertical mode in near-shore regions where moderate depths are the rule. This may be seen, for example, in a comparison between the two-layer model and a more exact model based on a smooth density profile (IVANOV ET AL., 1992) which was measured in the Levant Sea (the eastern Mediterranean). The validity of the two-layer models and their comparison with results from laboratory experiments are discussed in the afore-mentioned report by OSTROVSKY & STEPANYANTS (2005). An optimal adjustment of the two-layer model parameters which gives the best approximation for wave velocities and other observable wave characteristics in the real ocean is discussed in (NAGOVITSYN ET AL., 1990) and (GERKEMA, 1994). However, in other cases the parameters of the corresponding equations must be calculated from the expressions (21), (16) corresponding to a general case of a continuously stratified fluid (see, e.g. (OSTROVSKY & STEPANYANTS, 1989), and references therein).

From an observational viewpoint, it is important to remember that a single measurement of a solitary-like formation does not guarantee that the entity is indeed a soliton. An initial impulse may quickly disintegrate afterwards into something other than a solitary wave. In principle, it is necessary to follow such a wave out to a distance much greater than its spatial width to ensure that its shape remains stationary, which is not a simple task in real experiments. Another criterion for identification of a soliton is based on knowledge of the background density and horizontal velocity profiles. After a theoretical calculation of soliton parameters, one can compare these with the observational data. For example, the product of the characteristic wave width \(\Delta\) and the square root of its height, \(\sqrt{\eta_0}\), must not depend on that height [cf. Eq. (23)], provided the KdV equation is applicable to the
situation considered.

3.4 Soliton Propagation Under Perturbations.

For the conditions generally existing in nature, the simple models considered above are rarely applicable without taking into account a number of perturbing factors, such as dissipation of various origins, wave front curvature, horizontal inhomogeneities, depth variation, and the like. However, these factors are often locally weak enough so that they strongly affect the wave only at distances large compared to a wavelength. Under these conditions, perturbation theory is generally applicable. Such an approach results in the appearance in the model equation of small additive terms, each responsible for a specific perturbing factor. As an example, for waves much longer than the total depth of the controlling layer, the three factors mentioned above may be taken into account within the framework of a generalized “time-like” KdV (TKdV) equation\(^8\) (in the context of internal waves see, for example, Liu et al., 1985):

\[
\frac{\partial \eta}{\partial t} + \frac{1}{c} \frac{\partial \eta}{\partial t} - \frac{\alpha \eta \partial \eta}{c^2} - \frac{\beta \partial \eta}{c \partial t^2} \partial^2 \eta = -\frac{\eta}{2r} - \frac{\eta}{2c} \frac{dc}{dr} + R(\eta). \tag{49}
\]

The terms in the right-hand side of Eq. (49) describe, respectively, the effects of cylindrical divergence (the distance \(r\) from the source is supposed to be much greater than the wavelength), slow variation of the long-wave speed \(c\) along the ray \(r\) due to spatial inhomogeneity (e.g. due to variation of the pycnocline depth), and dissipation.

The latter term depends on a specific mechanism of losses. In particular, a horizontal eddy (turbulent) viscosity \(A_t h\) and a molecular viscosity \(\nu_m\) result in dissipation that is described by a Reynolds-type term, \(R = \frac{\delta \partial^2 \eta}{c \partial t^2}\) (\(\delta\) is proportional to the sum of \(A_t h\) and \(\nu_m\), and usually \(A_t h \gg \nu_m\)). Semi-empirical models accounting for bottom friction are also used, resulting in the terms \(R = \gamma_{Ra} \eta\) (Rayleigh dissipation) and \(R = \gamma_{Ch} |\eta| \eta\) (Chezy dissipation) with \(\gamma_{Ra}, \gamma_{Ch}\) taken as empirical coefficients (Holloway et al., 1997, 1999, 2002; Grimshaw, 2002). A rigorous consideration of viscous effects in the laminar bottom boundary layer leads to inclusion into Eq. (49) of a more complex integral term (Grimshaw, 1981, 2002; Das & Chakrabarti, 1986).

From Eq. (49), an ordinary differential equation for the slow variation of the soliton amplitude \(\eta_0\) over large distances can be obtained. As follows from perturbation theory [see, e.g. (Ostroffsky, 1974)], the first-order solution of Eq. (49) for the soliton amplitude may be obtained by multiplying it by \(\eta\), substituting the soliton (22), (23) with locally constant parameters, and then integrating over infinite limits in time, \(-\infty < t < \infty\). One obtains

\[
\frac{d \eta_0}{dr} = -\frac{2\eta_0}{3r} - \frac{2\eta_0 \frac{dc}{dr}}{3c} - \frac{4\alpha \delta}{45\beta} \eta_0^2, \tag{50}
\]

\(^8\)This term has been introduced by Osborne (1995) for the version of KdV equation with transposed temporal and spatial variables \(t\) and \(x\), whereas that form of the equation has been used by many authors long before. Such equation is relevant to the analysis of time series measured by point sensors in fixed spatial places.
which describes slow variations of the soliton amplitude under the effect of small cylindrical divergence, horizontal inhomogeneity, and eddy and molecular viscosity. The variations of length and width of the soliton are then defined via the local relation, Eq. (23), as before.

As particular cases, we readily obtain the laws of soliton variability due to
(a) cylindrical divergence ($\delta = 0$, $c = \text{const}$):

$$\eta_0 \sim r^{-2/3}, \quad \Delta \sim r^{1/3};$$

(b) a smooth horizontal variation of $c$ in a plane wave:

$$\eta_0 \sim c^{-2/3}.$$  \hspace{1cm} (52)

According to Eq. (23), variations of $\Delta$ are not directly defined by $c$ (or $\eta_0$) but rather by a combination of the parameters $\beta$ and $\alpha$, which may change together with $c$.

(c) the separate effect of Reynolds losses results in the following damping law:

$$\eta_0(r) = \frac{\eta_0(0)}{1 + \eta_0(0)qr},$$

where $\eta_0(0)$ is the initial soliton amplitude, and $q = 4\alpha\delta/\beta$. From Eq. (53) it is seen that soliton damping is non-exponential because of the nonlinearity. Moreover, at large distances, $r \gg [\eta_0(0)q]^{-1}$, the soliton amplitude ceases to depend on its initial value at all\footnote{A similar behavior is known in nonlinear acoustics for weak shock waves having a sawtooth form.}, $\eta_0(r) \sim (qr)^{-1}$. Chezy friction leads to the same law of soliton attenuation (GRIMSHAW, 2002; CAPUTO & STEPANYANTS, 2003), whereas Rayleigh dissipation yields an exponential damping with an exponent different from that for linear waves (OSTROVSKY, 1983; CAPUTO & STEPANYANTS, 2003).

Soliton decay due to energy dissipation in the laminar boundary layer is described by an integral dissipative term [see, e.g., GRIMSHAW, 2002; OSTROVSKY \& STEPANYANTS, 2005 and references therein]:

$$R(\eta) = -\delta_1 \int_{-\infty}^{\infty} \frac{1 - \text{sgn}(t - t') \partial \eta(t', x) / \partial t'}{\sqrt{|t - t'|}} dt'.$$  \hspace{1cm} (54)

The dissipation coefficient $\delta_1$ depends, in general, on many parameters such as the depth, density, and viscosity of the fluid layers (LEONE ET AL., 1982). However, in the Boussinesq approximation with the additional assumption that kinematic viscosities of layers are also equal, $\nu_1 = \nu_2 = \nu_m$, this coefficient may be presented in a relatively simple form (HELFICH, 1992)\footnote{A misprint in the numerator of formula (A.6) of HELFICH’s paper of 1992 should be mentioned, it must be a product of depths rather than their difference.}:

$$\delta_1 = \frac{1}{4c h_1 + h_2} \left[ b + \frac{(1 + b)^2}{2b} + 2 \frac{h_2}{W} (1 + b) \right],$$

\hspace{1cm} (55)
where \( b = h_1/h_2 \) and \( W \) is the width of the tank. The applicability of this dissipation model requires the boundary-layer thickness to be much less than the total water depth.

For such dissipation, the following damping law for soliton amplitude follows from the adiabatic theory:

\[
\eta_0(r) = \frac{\eta_0(0)}{(1 + r/r_{ch})^4},
\]

where \( r_{ch} \) is the characteristic spatial scale of soliton decay (see details in the references cited above). For \( r \gg r_{ch} \) this formula gives \( \eta_0(r) \sim r^{-4} \), and the soliton amplitude also ceases to depend on its initial value [because \( r_{ch} \sim \eta_0^{-1/4}(0) \)].

The perturbed KdV and eKdV equations similar to Eq. (49) were used in numerical modeling of the internal tide transformation on the Australian north-west shelf, the Malin shelf edge (western coast of Sootland), and the Arctic shelf (Laptev Sea) (HOLLOWAY ET AL., 1997; 1999; 2002; GRIMSHAW ET AL., 2005). The spatial variability of the coefficients, the Earth’ rotation and bottom friction effects were taken into account in those papers.

We should remark also that in case (a), Eq. (49) represents the cylindrical KdV equation, which is completely integrable. This equation possesses exact solitary solutions (CALOGERO & DEGASPERIS, 1978; NAKAMURA & CHEN, 1981), whose parameters vary in space due to cylindrical divergence, in agreement with equation (51) as obtained by the asymptotic method. Note finally that to describe soliton transformation, different models may have to be employed sequentially. For example, a cylindrically diverging soliton in the BO model decreases as \( 1/r \); after its broadening due to amplitude decrease, it can eventually be transformed into a cylindrical KdV soliton behaving according to Eq. (51) (STEPANYANTS, 1981).

An interesting kind of dissipation of interfacial internal waves can occur in the deep ocean when a sharp pycnocline is adjoined by a smoothly stratified layer of “infinite” depth (depth much greater than the thickness of the other layer and the characteristic wavelength). In this case, the governing model equation for the interface perturbation is similar to the JKKD equation but contains a more complex integral kernel, which describes both the dispersion and the dissipation due to the radiation of downward propagating bulk internal waves (MASLOWE & REDEKOPP, 1980; GRIMSHAW, 1981) [see also (GRIMSHAW, 2002)]. According to estimates from these papers made for typical oceanic conditions, such solitary waves may decry in a time comparable to their intrinsic time scale, \( T_{inl} \sim \Delta/V \).

### 3.5 Refraction and Diffraction of Solitons

Various generalizations of the KdV equation have been suggested for nonlinear waves having smoothly curved phase fronts. One of the most popular is the Kadomtsev–Petviashvili (KP) equation, which is applicable to a weakly diffracted wave beam, and is based again on adding a small term to the KdV equation describing transverse variations:

\[
\frac{\partial}{\partial x} \left( \frac{\partial \eta}{\partial \xi} + c \frac{\partial \eta}{\partial x} + \alpha \eta \frac{\partial \eta}{\partial x} + \beta \frac{\partial^3 \eta}{\partial x^3} \right) = -\frac{c}{2} \frac{\partial^2 \eta}{\partial y^2},
\]

26
where \( y \) is the coordinate transverse to the propagation direction \( x \). This equation is also known to be completely integrable. Its exact solutions have been studied in numerous papers and books (see, e.g. Ablowitz & Segur, 1981; Infeld & Rowlands, 1990). The main properties of solitary solutions to this equation as applied to oceanic waves (when the dispersion parameter is always positive, \( \beta > 0 \)) are as follows. A plane soliton is stable with respect to transverse perturbations of its front. Multiple soliton interactions can occur when solitons propagate in different directions at small angles to each other. The zone of nonlinear interaction of two solitons can be fairly long in space; the perturbation in this zone looks like a KdV soliton and propagates steadily. The interaction between two solitons has been studied in detail by Newell & Redekopp (1977) and Miles (1977a, b). They found a specific case of resonant interaction of two obliquely propagating solitons where the the zone of nonlinear interaction is infinite and forms another stationary soliton.

For waves propagating at small angles to an arbitrarily chosen direction in the horizontal plane, the 2D analogs of the BO and JKKD equations may also be derived (Ablowitz & Segur, 1981). These equations, however, are apparently not integrable. The transverse stability of a BO soliton was studied and confirmed by Ablowitz & Segur (1981). Two-soliton interaction was studied numerically by Tsuji & Oikawa (2001). They found that the phenomenon similar to the resonant interaction of KP solitons does occur for the BO equation, too. However, the concept of resonant interaction is not so effective for that equation, because the newly generated wave in the zone of nonlinear interaction of two solitons is far from the BO soliton.

It is interesting to note that strong mathematical ties have been found between the KdV and KP equations and their cylindrical analogs, in the sense that their solutions can be expressed via each other [see (Stepanyants, 1989) and references therein].

Another approach to studying two-dimensional effects is the “nonlinear geometrical optics” of solitons, which describes refraction of their fronts. The theory is based on the method elaborated earlier by Whitham (1974) for shock waves and then extended to solitons by Ostrovsky & Shrir (1976). It considers the motion of soliton fronts in orthogonal coordinates \((\Phi, \Psi)\) corresponding to lines of constant phase and normals to them, or “rays,” at succeeding moments of time. This leads to a pair of kinematic equations:

\[
V \frac{\partial \theta}{\partial \Psi} = \frac{\partial R}{\partial \Phi},
\]

\[
R \frac{\partial \theta}{\partial \Phi} = -\frac{\partial V}{\partial \Psi},
\]

where \( R \) is the dimensionless width of the ray tubes, and \( \theta \) is the angle between the rays and some reference direction. In the case considered, this system must be closed by using a dynamic relation reflecting the conservation of soliton energy; in particular, for the KdV equation it is

\[
R(V - c)^2 \partial (V) = \text{const},
\]
where $\Delta$ is the soliton width. From this, the system of Eqs. (58) may be shown to possess two characteristics that define the propagation speed during perturbations of the soliton amplitude and the shape of the wave front. These perturbations, being strong, propagate as simple (Riemann) waves with the possible formation of “shocks,” i.e., sharp jumps in amplitude and frontal direction. As also occurs in the dynamics of compressible gases, one needs a form of dissipation or a “viscosity” to smooth out this shock. As shown by Shrima (1980), weak radiation of small amplitude waves from the soliton front in the process of soliton adjustment to the local hydrological conditions may play the role of such a viscosity in this case. Phenomena suggesting that similar effects may exist for internal solitons have been observed in the Sulu Sea in the Philippines with sonars and radars (Apel et al., 1985).

### 3.6 Internal Waves on Shear Flows

The velocities of shear flows in the ocean are often of the same order of magnitude as the velocities of the IWs, so that energy exchange between these two types of motions may be very effective, and may even result in instabilities. As it is well known (Miles, 1961; Howard, 1961), the main parameter defining the criterion for stability of a stratified fluid is the Richardson number, $\text{Ri}(z) = \left| N/U_0^2(z) \right|^2$, where $U_0(z)$ is the horizontal mean flow speed [see, e.g., (Turner, 1973; Phillips, 1977; LeBlond & Mysak, 1978; MIropol’sky, 1981)].

In the Boussinesq approximation, the modal structure of a linear perturbation is defined by the Taylor-Goldstein equation:

$$\frac{d}{dz} \left[ (U_0 - c)^2 \frac{dW}{dz} \right] + \left[ N^2 - k^2(U_0 - c)^2 \right] W = 0,$$

(60)

where the notation is the same as in Eq. (9). In the long-wave approximation it is a generalization of Eq. (9):

$$\frac{d}{dz} \left[ (U_0 - c)^2 \frac{dW}{dz} \right] + N^2 W = 0.$$

(61)

**Stable flows.** According to the well-known Miles-Howard theorem, a flow with $\text{Ri} > 1/4$ is always stable in the linear approximation. For long IWs in stable shear flows, both KdV and BO equation were obtained by several authors, among them Benney (1966), Lee & Beardsley (1974), Maslowe & Reekopp (1980), Grimshaw (1981), Suvorov (1981), Tung et al. (1981), Grimshaw et al. (2002b), and others. In particular, Maslowe & Reekopp (1980) have considered weakly nonlinear internal waves in continuously stratified shear flows and showed that the generalized KdV and BO equations can be derived for long waves in shallow and deep configurations, respectively, both with and without critical layers (layers where the wave phase velocity coincides with the local velocity of the mean current).
In the Boussinesq approximation, the expressions for the environmental coefficients of the KdV equation (20) can be presented in the form [cf. (21) via (16)]:

$$\alpha = \frac{3}{2Q} \int_{-H}^{0} (U_0 - c)^2 \left( \frac{dW}{dz} \right)^3 dz, \quad \beta = \frac{1}{2Q} \int_{-H}^{0} (U_0 - c)^2 W^2 dz, \quad Q = \int_{-H}^{0} (U_0 - c)^2 \left( \frac{dW}{dz} \right)^2 dz.$$  

(62)

The coefficients are calculated for a fixed internal mode with a given modal number $n$. Note that the above-mentioned situation in which $\alpha = 0$ is also possible here; the corresponding modes satisfy the mKdV equation rather than KdV. Of course, a eKdV equation can also be applicable in this environment. Some interesting models of stratified shear flows that allow exact solutions for the Taylor–Goldstein equation (60) were considered by Weidman & Velarde (1992). Other rare examples of analytically solvable models of shear flows are mentioned in the book by Turner (1973).

Thus, the existence of stable shear flows in stratified fluid leads to the essential complication of the problem because the more complex Taylor–Goldstein equation must be solved for eigenvalues and eigenfunctions and more complex integrals, Eqs. (62), must be calculated for the coefficients of the governing equations. At the same time, some stable shear flows may support new types of waves, the “vorticity waves”, which are very similar to internal gravity waves and also occur inside the fluid. They may exist even in a homogeneous fluid and may play an important role in the dynamics of the upper ocean. Below we present a brief overview of these kind of waves as a specific example of waves on stable shear flows.

Consider a shear flow with a smooth monotonic profile without inflection points as depicted in Fig. 4 (shear flow of a “boundary layer type”). According to the inviscid linear theory of hydrodynamic stability, such a profile is stable with respect to small perturbations. To fix the idea, we at first approximated it by a piecewise-linear function as shown in Fig. 4 by the dashed line. Such a profile may occur, for example, in the upper ocean layer under the action of a wind.

There are several peculiarities of the linear eigenmodes for such piecewise-linear velocity profiles. For a homogeneous fluid the eigenvalue problem (60) reduces to the well-known Rayleigh equation [see, e.g., (Turner, 1973; LeBlond & Mysak, 1978)], which is a particular case of the Taylor–Goldstein equation (60) with $N = 0$. It was established by Rayleigh that for smooth profiles without inflection points (like the one shown by solid line in Fig. 4) the equation does not possess a discrete spectrum of eigenvalues and eigenfunctions, although there exists a continuum spectrum of eigenvalues and eigenfunctions which are known as the Case–Dickey waves\textsuperscript{11}. On the other hand, for a piecewise-linear profile like that shown in Fig. 4 by the dashed line, a discrete eigenvalue does exist. The corresponding eigenmode is related to the vorticity jump at depth $h$ and may be called the vorticity wave (we recall that the vorticity for a parallel shear flow is determined by the derivative of the velocity profile $U'(z)$).

\textsuperscript{11}Such waves are also known in plasma physics as the Van Kampen–Case waves [see, e.g., (Shrira et al., 2001) and references therein.]
Figure 4: Typical smooth monotonic shear flow profile (in some normalized variables) without inflection points (solid line) and its piecewise-linear approximation (dashed line). Dots are experimental data obtained under laboratory conditions. From (SHRIRA ET AL., 2005).

When a piecewise-linear profile is replaced with an even slightly smoothed profile, the discrete eigenvalue of the Rayleigh equation corresponding to the vorticity jump disappears. However, the continuous spectrum of eigenvalues and corresponding eigenmodes still exists (they are a generic feature of any smooth shear flow profile). When the initial perturbation is decomposed onto the set of eigenmodes of the continuous spectrum, their collective behavior for a relatively long time looks like the evolution of the single eigenmode of the discrete spectrum of the piecewise-linear velocity profile (SHRIRA, 1989; SHRIRA & SAZONOV, 2001; SAZONOV & SHRIRA, 2003). As was shown in the cited papers, the nonlinearity further consolidates the eigenmodes of the continuous spectrum and makes their behavior still more similar to the discrete eigenmode. Weakly nonlinear three-dimensional perturbations propagating in the horizontal plane on the background of a “boundary layer type” shear flow can be described by the following Shrira’s equation (SHRIRA, 1989; VORONOVIICH ET AL., 1998B):

$$\frac{\partial A}{\partial t} + c \frac{\partial A}{\partial x} - \alpha A \frac{\partial A}{\partial x} - \beta \frac{\partial}{\partial x} \tilde{G}(A) = 0. \quad (63)$$

where $A(x, y, t)$ is a function which determines the dependence of all variables on the horizontal coordinates (the details can be found in the above cited papers); in particular, the longitudinal horizontal and vertical velocity components are, correspondingly,
\[ u(x, y, z, t) = -A(x, y, t)U'(z) \quad \text{and} \quad w(x, y, z, t) = A(x, y, t)(U(z) - U_0). \]

Other parameters of the equation are \( c = U_0, \alpha = U'(z)|_{z=0}, \beta = (U^2/U')|_{z=0}, \text{and} \ \hat{G}(\ldots) \text{ is in general a 2D integral operator defined by the formula} \]

\[
\hat{G}[\varphi(r)] = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} Q(k)\varphi(r')e^{ik(r-r')} \, dk \, dr', \tag{64}
\]

where \( r = (x, y), \ k = (k_x, k_y) \) and the kernel \( Q(k) = |k|/\tanh(|k|H). \) In the derivation of Eq. (63) the total fluid depth \( H \) can be arbitrary but the thickness of the boundary layer \( h \) is assumed to be small compared to the characteristic wavelength.

For plane perturbations in deep water, the integral kernel simplifies to \( Q(k) = |k| \) and Shrira’s equation (63) reduces to the BO equation. However, the plane waves are highly anisotropic in the horizontal plane because their properties depend on the orientation with respect to the direction of the basic shear flow. For linear harmonic waves the dispersion relation has the form

\[
\omega = (c - \beta|k|)k_x \tag{65}
\]

The characteristic property of Eq. (63) is that in the general case it possesses 2D solitary wave solutions completely localized in space. They were first constructed numerically by Abramyan et al. (1992) [see also (Voronovich et al., 1998b)]. Such solutions, in dimensionless variables (see details in the papers cited) have a circular symmetry. An analysis of the asymptotics of the solitary wave field far from its peak shows that, in the deep water limit, it decays as \( r^{-2} \) (Pelinovsky & Shrira, 1995), similar to the BO-solitons. The characteristic spatial scale of such solitary wave is \( \Delta \sim \beta/(V-c) \), where \( V \) is the solitary wave velocity.

Further investigations (D'yachenko & Kuznetsov, 1994; Pelinovsky & Stepanyants, 1994; Pelinovsky & Shrira, 1995; Gaidashev & Zhdanov, 2004) showed that within the framework of this model the solitary waves are unstable, i.e., upon being slightly disturbed, they gradually build up. This instability has an “explosive” character, i.e., the soliton amplitude turns to infinity in a finite time whereas the wavelength simultaneously tends to zero (such an instability in 2D- and 3D-cases is also called wave collapse). However, for the slightly perturbed 2D solitary wave described above, the parameters change rather gradually, for example, amplitude increases as \( \eta_0 \sim (T_c - t)^{-1/2} \), where \( T_c \) is a characteristic collapse time. In the process of evolution the soliton preserves its shape and adiabatically grows to infinity. Numerical simulations (Shrira et al., 2005) showed that a growing peak shaped like a circular 2D soliton, followed by a small-amplitude wave tail, eventually emerges out of a very broad class of localized initial perturbations (Fig. 5\textsuperscript{12}). At some stage

\footnote{This figure gives rather a qualitative illustration of the process described because a bit more general equation containing smooth longitudinal inhomogeneity was actually simulated.}
the weakly nonlinear evolution equation (63) ceases to be applicable and the further fate of
the perturbations must be, apparently, studied within the framework of the set of primitive
hydrodynamic equations.

Figure 5: Evolution and collapse of a 2D Gaussian initial perturbation as described by
Eq. (63) in deep fluid. From (Shirira et al., 2005).

As the collapse instability is rather weak for solitons, they can exist for a relatively
long time and, perhaps, can be detected in the ocean (similar phenomena can also occur in
the atmospheric boundary layers). However, there are no observations of this in nature so
far, either in the atmosphere or ocean. Perhaps, an implicit indication of the existence of
vorticity waves in the oceanic mixed upper layers is the numerous observations of slicks on
the ocean surface when no internal waves were detected (surface signatures of internal waves
are discussed in detail in subsection 4.3 below). At the same time, wave processes similar to
those described above were observed in laboratory experiments with an air boundary layer
above a flat plate in a wind tunnel (Kachanov et al., 1993) and in a wind-wave water
tank (Shirira et al., 2005).

For oceanic conditions, estimates show that for a typical shear flow profile as presented
in Fig. 4 with \( h = 1 - 10 \text{ m} \), \( U_0 = 10 - 30 \text{ cm/s} \), the parameters of a solitary wave are as follows: the maximum longitudinal fluid velocity perturbation is \( u_{\text{max}} = 1 - 3 \text{ cm/s} \); the wave velocity is \( V = 15 - 60 \text{ cm/s} \); the characteristic wavelength is \( \Lambda \approx 10 \text{ m} \) for \( h = 1 \text{ m} \) and \( \Lambda \approx 100 \text{ m} \) for \( h = 10 \text{ m} \). It is quite realistic to detect such perturbations by, e.g., acoustical methods. It would be of interest to study contribution of such localized formations to the turbulence of the upper ocean.

Another simplification of Eq. (63) can be realized for a shallow ocean when the characteristic wavelength is much greater than the total fluid depth, \( \Lambda \gg H \). Then, the integral kernel \( Q(k) \approx (1 + |k|^2 H^2/3)/H \), and integral operator \( \hat{G} \) reduces to the 2D Laplacian operator. In this case Shpira's equation (63) transforms to one of the generalizations of the KdV equation, namely to the 2D version of the Zakharov–Kuznetsov equation (ZAKHAROV-KUZNETSOV, 1974). Two-dimensional stationary solutions of that equation were studied in many papers with application to different types of waves [see, e.g., (VORONOVICH ET AL., 1998b) and references therein].

Solitary solutions in the general case of an arbitrary depth fluid were constructed numerically in (VORONOVICH ET AL., 1998b). They are qualitatively analogous to the solution for deep water; they decay monotonically and possess circular symmetry in the dimensionless variables.

In the same paper a more general case was considered when in addition to the shear flow there is also a density stratification in the ocean. In this case multiple resonances may occur between the vorticity and internal waves.

Further theoretical development to the study of properties of nonlinear vorticity waves in the context of their strong interaction with internal gravity waves was made in (VORONOVICH ET AL., 1998a; SHIRYA ET AL., 2000; 2004; VORONOVICH ET AL., 2006) where the resonance phenomena were studied in detail and, in particular, wave breaking of the coupled vorticity and internal waves was considered over a sloping bottom. Such situation can occur when, e.g., the wind induced shear current is oriented onshore.

**Unstable flows.** The situation becomes radically different when the basic flow is unstable, a condition that is possible if \( \text{Ri} < 1/4 \). Observational data testify that this does take place in the ocean, although it is difficult to tell how frequently; see, for example, (DES Aublies & SMITH, 1982; Padman & Jones, 1985; Sandström et al., 1989; Bogucki & Garrett, 1993; Boyd et al., 1993). Typically, there are often near-critical conditions in which the Richardson number is close to 1/4 because the surface wind stress drives the near-surface layers and causes significant vertical shear in the mean flow. Physical models of shear flows having smoothly varying density and velocity profiles are usually very difficult to handle in analytical form, even in the linear case. That is why simplified approximations to the stratification and current, such as models having tangential discontinuities and piecewise-linear profiles, are often used in theory (Turner, 1973; LeBlond & Mysak, 1978; Craik, 1985, Redekopp, 2002).

An interesting finding was made by Thorpe (1969) who showed that fluid stratification can destabilize some shear flows known to be stable in a uniform fluid (see also Turner,
A similar phenomenon is known in fluid mechanics since the 1920s, when W. Heisenberg discovered the destabilizing effect of viscosity in boundary layers [see about this, e.g., in (Betchov & Criminale, 1967)]. A physical interpretation of this effect can be done on the basis of a concept of negative energy waves (NEW) (Ostrovsky et al., 1986; Stepanyants & Fabrikant, 1996). Some physical insight of this concept can be obtained with the help of a simple model illustrating onset of a radiative instability of NEW due to radiation of bulk internal waves from the pycnocline towards the deep ocean. This model relevant to some real oceanographic situations is described below (Ostrovsky & Tsimring, 1981; Ostrovsky et al., 1984a).

Let us assume that the upper layer of the fluid is homogeneous over depth $h$, and moves along the $x$-axis over a stratified lower layer of infinite depth, having $N = \text{const}$ (Fig. 6). The upper layer density is $\rho_1$, and its velocity is $U$. At $z < -h$ the stratification is $\rho_2 = \rho_* \exp[-b(z + h)]$, so that the Brunt–Väisälä frequency is $N = \sqrt{g b} = \text{const}$, and the density has a discontinuity, $\Delta \rho = \rho_* - \rho_1$, at $z = -h$.

Figure 6: Sketch of near-surface shear flow with tangential discontinuity of current velocity in the model of a stratified ocean of infinite depth.

In this case, a simple way to derive the evolution equation for soliton dynamics is to start from the dispersion equation for linear, sinusoidal wave trains, which at $z < -h$ have the form $\exp[i(kx + mz - \omega t)]$, where $m$ is the vertical component of the wave vector. Using the linearized hydrodynamic equations, one can obtain a dispersion equation in the form

$$a(\omega - kU)^2 \coth kh + \omega \sqrt{\omega^2 - N^2} - (1 - a)gk = 0,$$

where $a = \rho_1/\rho_*$, and $m^2 = k^2(N^2/\omega^2 - 1)$. The sign of the real part of $m$ should be chosen so as to assure that the wave radiation propagates downward from the level of the discontinuity.
Such a dispersion relation is shown in Fig. 7. One can see that for short waves (large $k$), $\omega$ is complex, a condition that yields an exponential growth of small perturbations. This is a variant of the well-known Kelvin–Helmholtz instability. But the instability also exists for longer waves, as it is also seen from Fig. 7. An important peculiarity of such waves, the negative energy waves, is that their excitation actually decreases the total energy of the system so that adding any real losses to the system leads to an instability (OSTROVSKY & TSIKRING, 1981; OSTROVSKY ET AL., 1986). In the case considered, NEW correspond to the lower branch of the dispersion curve (Fig. 7), and losses are associated with the radiation of bulk internal waves from the interface boundary downward into the lower layer, which again results in a complex frequency, $\omega$.

If $U > U_{cr} = \sqrt{(1 - a)gh/a}$, the instability associated with negative energy waves takes place starting at $k = 0$ (Fig. 7). For the lower branch of the dispersion relation shown schematically in Fig. 7, the dependency $\omega(k)$ can be readily derived from Eq. (66) in the limit $k \to 0$:

$$\omega(k) \approx ck + \beta k^3,$$

where parameters $c$ and $\beta$ are complex in general. It is important to note that the other mode, corresponding to the upper branch of the dispersion curve shown Fig. 7 and having positive energy, is not localized at all: its grows up to infinity as $z \to -\infty$. Actually, one has to solve a nonstationary initial-value problem in order to describe the evolution of these types of waves.

Note that in the absence of the current, the dynamics of long, weakly nonlinear interfacial waves in the above model is described by the BO equation complemented by an integral term responsible for the radiation losses (MASLOWE & REDERKOPP, 1980; GRIMSHAW, 1981, 2002). Naturally, in this model all wave modes are of positive energy. Soliton damping due to radiation has a complex character and, as calculated in the papers cited above, the damping is rather strong for typical oceanic conditions so that the damping time may be comparable with the intrinsic soliton time scale. This is one of possible reasons of relative scarcity of internal solitary waves observed in the deep ocean.

Evidently, there are many other possible mechanisms of wave dissipation, especially in the upper layer. As already mentioned, these losses can be due to viscosity, turbulence, and diffusion. A study of these effects carried out by OSTROVSKY & SOUSTOVA (1979) showed that for long waves, the imaginary part of the frequency is given by

$$\text{Im}(\omega) \approx -\epsilon - \delta_{t}k^2, \quad \epsilon = \frac{\Pi}{glz(1 - a)}.$$  \hspace{1cm} (68)

Here $\Pi$ is the vertical buoyancy flux, $\delta_{t}$ is the turbulent viscosity coefficient, and $l_z$ is the vertical scale of turbulence. Considering the dispersion relation, Eq. (67), together with the dissipative contribution, Eq. (68), one can reconstruct an evolution equation for the long progressive waves in the form of a generalized KdV equation augmented by small dissipative terms. Each term, responsible for nonlinearity, dispersion, and dissipation, is additive.
Figure 7: Real ($\overline{\sigma}_r$) and imaginary ($\overline{\sigma}_i$) parts of the complex dispersion relation, Eq. (66), in dimensionless variables: $\overline{\sigma}_{r,i} = \omega_{r,i}\sqrt{\frac{a}{h}}(1 - a)g$, and $\kappa = kh$, for two values of dimensionless velocity, $\overline{U} = U/\sqrt{(1 - a)gh}$. (1) $\overline{U} < \overline{U}_{cr}$; (2) $\overline{U} > \overline{U}_{cr}$.

and can be derived separately. Such method of derivation of evolution equations from linear dispersion relations is widely used; see, for example, (Whitham, 1974; Korpe1 & Banerjee, 1984). The following model equation was derived in this fashion (Ostrovsky et al., 1984b):
\[
\frac{\partial \eta}{\partial t} + c_r \frac{\partial \eta}{\partial x} + \alpha \eta \frac{\partial \eta}{\partial x} - \beta_r \frac{\partial^3 \eta}{\partial x^3} = -\epsilon \eta + \delta \frac{\partial^2 \eta}{\partial x^2} + \frac{c_i}{\pi} \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \frac{\eta(x', t) \, dx'}{x - x'}.
\]  

(69)

Here \( \eta \) is again the displacement of the interface, \( c_r = \text{Re}(c) \), \( c_i = \text{Im}(c) \), \( \beta_r = \text{Re}(\beta) \), and \( \alpha = 3U_{cr}/2h \); the principal value of the integral is to be taken. It should be noted that the integral term has the same structure as that corresponding to Landau damping of ion acoustic waves in a plasma (Karpman, 1973), but has the opposite sign, i.e. it causes amplification.

For the case when the right-hand side of this equation is small, it is again possible to construct solutions for solitons with slowly varying parameters by multiplying Eq. (69) by \( \eta \) and integrating over \( x \), as was done above. The resulting equation for the soliton amplitude \( \eta_0 \) has the form (Ostrovsky et al., 1984b)

\[
\frac{d\eta_0}{dt} = -\frac{4}{3} \eta_0 + 2.92 \frac{c_i}{\pi} \sqrt{\frac{\alpha \eta_0^3}{12 \beta}} - \frac{4 \alpha \delta \eta_0^2}{45 \beta}.
\]  

(70)

This equation has three equilibrium states determined by zeros of its right-hand side. Two of them, \( \eta_0 = 0 \) and \( \eta_0 = 0 \), are stable, whereas the intermediate state, \( \eta_0 = \eta_0 < \eta_0 = \eta_0 \), is unstable. Thus, there exists a “hard” regime of amplification: Small-magnitude solitons with \( \eta_0 < \eta_0 \) are damped, whereas larger ones asymptotically reach the limiting magnitude \( \eta_0 \) (Fig. 8).

---

**Figure 8:** Soliton amplitude versus time in dimensionless variables as given by Eq. (70) for three different initial conditions (1 - \( u_0 = 0.5 \); 2 - \( u_0 = 0.57 \) and 3 - \( u_0 = 3.0 \)). Normalized variables are: \( u = \eta_0 \frac{\alpha \delta}{13 \pi}, \quad \tau = 4cT/3, \quad \gamma = 2.1. \)

The exact solution to this equation can be presented in the implicit form; in dimensionless variables it reads:
\[
\tau = \ln \frac{u_0}{u} - \frac{\gamma + \sqrt{\gamma^2 - 4}}{\sqrt{\gamma^2 - 4}} \ln \frac{2\sqrt{u_0} - \gamma + \sqrt{\gamma^2 - 4}}{2u - \gamma + \sqrt{\gamma^2 - 4}} + \frac{\gamma - \sqrt{\gamma^2 - 4}}{\gamma^2 - 4} \ln \frac{2\sqrt{u_0} + \gamma + \sqrt{\gamma^2 - 4}}{2u + \gamma + \sqrt{\gamma^2 - 4}},
\]

where the normalized variables are:

\[
\tau = \frac{4}{3} \epsilon \xi, \quad u = \eta_0 \frac{\alpha \delta_t}{15 \beta \epsilon}, \quad \gamma = \frac{3}{8} \frac{2.92 \sqrt{5}}{\pi} \frac{c_i}{\sqrt{\epsilon \delta_t}} \approx 0.78 \frac{c_i}{\sqrt{\epsilon \delta_t}},
\]

and \(u_0 \sim \eta_0(0)\) is the initial value of soliton amplitude. Note also that at \(\epsilon = 0\), we obtain the “soft” regime (\(\eta_{02} = 0\)). Small solitons with \(\eta_{02} < \eta_0 \ll \eta_{03}\) initially grow according the “explosive” law [cf. Eqs. (53) and (56)]

\[
\eta_0 = \frac{\eta_0(0)}{(1 - t/T_c)^2}, \quad T_c = \frac{2\pi}{1.46 c_i \sqrt{3 \beta}} \left( \frac{\alpha \eta_0(0)}{1 - t/T_c} \right)^{1/3},
\]

where \(\eta_0(0)\) is the initial soliton amplitude, and \(T_c\) is the “explosion time” during which \(\eta_0 \to \infty\). However, toward the end of this stage the soliton amplitude growth is slowed down and the amplitude reaches its stationary value, \(\eta_{03}\), as may be seen in Fig. 8.

Estimates of the magnitude of such an effect for the ocean seem reasonable. If we accept the parameters of the model to be \(\Delta \rho/\rho = 10^{-3}\), \(N = 5 \cdot 10^{-3} \text{ s}^{-1}\), and \(U = 1.1U_\sigma\), then for \(h \simeq 50 \text{ m}\), the radiative instability is most pronounced for wavelengths of order 100–1000 m, which correspond (for the wave on the flow) to time scales of 20–200 min. The characteristic time of the development of the instability is of the order of 20 h in this case. This may be considered as one of the possible mechanisms for the generation of internal wave trains in the ocean.

This theory has serious limitations for real cases because of the fact that the maximal growth rate is realized for shorter waves (see Fig. 7), which may result in the generation of billow turbulence and mixing in the region of the density interface. Nevertheless, the larger-scale instability described above may still exist on such a short-scale billow background.

### 3.7 Nonlinear Waves in Rotating Ocean

For the description of mesoscale processes having spatial scales of a few kilometers or more and time durations of an hour or more, the effects of the Earth's rotation become significant. There arise some radically new elements in the behavior of nonlinear waves in this case. The important new variable is the Coriolis parameter \(f\) which defines the lowest possible frequency of surface and internal gravity waves; these are also sometimes called gyroscopic or inertial gravity waves. For frequencies close to \(f\), long-wave dispersion plays a major role\(^\text{13}\). A sketch of the dispersion relation for such linear harmonic waves is presented on Fig. 9. It

\(^{13}\text{An analogous situation occurs for electromagnetic and acoustic waves in waveguides.}\)
should be emphasized that we are restricting our attention to waves with frequencies above \( f \), where \( f \) is assumed to be constant (\( f \)-plane approximation). As was already mentioned, Rossby waves for which the meridional dependence of the Coriolis force must be taken into account (\( \beta \)-plane approximation) also exist in a rotating fluid, but their frequencies lie below \( f \).

Figure 9: Dispersion curve for waves in a rotating fluid. The graph is plotted in the log-log scale in dimensionless variables: \( \overline{\sigma} = \omega/f \), \( \kappa = ck/f \) with \( \beta f^2/c^3 = 10^{-4} \). Line 1 - dispersion curve as described by Eq. (73) below, line 2 - the dispersionless approximation, \( \omega \approx ck \).

Equations (14) and (15) may be used for the description of gyroscopic waves. Moreover, for waves with their frequency spectrum lying in the interval between \( f \) and the maximum Brunt-Väisälä frequency, \( N_{\text{max}} \), but not too close to either of these parameters, both the low- and high-frequency dispersion effects are small. An adequate partial differential equation may again be obtained from the dispersion relation for linear waves in the limit of weak dispersion, namely

\[
\omega \approx ck - \beta k^3 + \frac{f^2}{2ck^3} \tag{73}
\]

from which the evolution equation follows in the form

\[
\frac{\partial}{\partial x} \left( \frac{\partial \eta}{\partial t} + c \frac{\partial \eta}{\partial x} + \alpha \frac{\partial \eta}{\partial x} + \beta \frac{\partial^3 \eta}{\partial x^3} \right) = \frac{f^2}{2c} \eta. \tag{74}
\]

This equation was firstly derived by OSTRÖVSKY (1978), and was then reproduced and analyzed in many subsequent papers [see, e.g. (LEONOV, 1981; REDEKOPP, 1983; GRIMSHAW, 1985; KATSI & AKYLAS, 1987; GERMAIN & RENOUARD, 1991; etc.]). In the absence of rotation (\( f = 0 \)), it reduces to the KdV equation (20), so that Eq. (74)
may be called rotationally modified KdV or briefly rKdV equation. An analogous equation with $\beta = 0$ was obtained for different types of waves, including inertial-gravity waves in the ocean (MUZYLEV, 1982), any kind of waves in random media (BENILOV & PELINOVSKY, 1988), waves in relaxing media (VAKHNENKO, 1999).

Unlike the above model equations, Eq. (74) is apparently not completely integrable. Still, for periodic and localized solutions, it possesses a series of useful integrals such as an energy integral. It also has a “zero-mass” integral, $M = 0$, where for a localized solution the wave “mass” is defined in Eq. (28); for periodic solutions this condition is also true provided the integration is taken over the wave period. Note that for previously considered equations such as the KdV, mKdV, eKdV, BO and JKKD equations, this “mass” integral is an arbitrary constant but not necessarily zero.

Some other definite integrals for this equation, which are actually the constraints, are presented in (GRIMSHAW ET AL., 1998A) including those found by BENILOV (1992) and having the form of momenta. The simplest among them is

$$\int_{-\infty}^{\infty} x\eta(x, t) \, dx = 0.$$  

It is also easy to generalize the rKdV equation for a 2D case to obtain a rotationally modified KP (RKP) equation:

$$\frac{\partial}{\partial x} \left( \frac{\partial \eta}{\partial t} + c \frac{\partial \eta}{\partial x} + \alpha \eta \frac{\partial \eta}{\partial x} + \beta \frac{\partial^3 \eta}{\partial x^3} \right) = \frac{f^2}{2c} \eta - \frac{c}{2} \frac{\partial^2 \eta}{\partial \xi^2}. \quad (75)$$

This equation was also derived by OSTROVSKY (1978) and then reproduced by GRIMSHAW (1985) and other authors.

Exact analytical solutions for even the stationary version of Eq. (74) are unknown (except for the steady wave of the parabolic profile, see below). However, many of the solutions have been investigated numerically by now.

A relatively simple analysis can be performed if the high-frequency dispersion (the term with $\beta$) is neglected, which is possible for sufficiently long waves. Then, a second-order equation results, which for stationary solutions depending on $\xi = x - Vt$ yields

$$\frac{d}{d\xi} \left[ (c - V + \alpha \eta) \frac{d\eta}{d\xi} \right] = \frac{f^2}{2c} \eta. \quad (76)$$

Similar equation arises from the particular case of nonlinear Klein-Gordon equation which describes elastic waves in bending rods propagating in opposite directions [“two-directional” wave equation (RYBAK & SKRJNNIKOV, 1990)].

Stationary solutions of Eq. (76) can be analyzed on the phase plane of the variables $\eta$ and $d\eta/d\xi$ [see details in (OSTROVSKY, 1978; GRIMSHAW ET AL., 1998A; STEPANYANTS, 1999; BOYD & CHEN, 2002; FRAUNIE & STEPANYANTS, 2002; GRIMSHAW, 2002; LIU & VARLAMOV, 2004)].

---

14It is also referred to in many papers as the Ostrovsky equation [see, e.g. (GILMAN ET AL., 1995; NEW & ESTEBAN, 1999; BOYD & CHEN, 2002; FRAUNIE & STEPANYANTS, 2002; GRIMSHAW, 2002; LIU & VARLAMOV, 2004)].
There exists a family of periodic solutions to this reduced equation whose shape varies from sinusoidal to parabolic. The wave of limiting amplitude has sharp crests and is represented by a periodical sequence of parabolic arcs. Note that each arc itself is also a solution of the full Eq. (74) with the high-frequency dispersion:

\[
\eta = \frac{f^2}{12\alpha c} \left[ (\xi - \xi_0)^2 - \frac{\lambda^2}{12} \right], \quad -\frac{\lambda}{2} \leq \xi - \xi_0 \leq \frac{\lambda}{2},
\] (77)

In 1996 Gerkema (GerKema, 1996) derived for small-amplitude long waves in two-layer fluid a “two-directional” generalization of equation Eq. (74), the rotation modified Boussinesq equation which is capable to describe waves in opposite directions. Later, on the basis of Lee & Beardsley (1974) approach, the same equation was derived by New & Esteban (1999) for an arbitrarily stratified fluid. In terms of the lateral component, \( A(x, t) \), of the stream function, \( \Psi(x, z, t) = A(x, t)\Phi(z) \), the equation has the form

\[
\frac{\partial^2 A}{\partial t^2} - c^2 \frac{\partial^2 A}{\partial x^2} - s \frac{\partial^4 A}{\partial t^2 \partial x^2} + f^2 A = r \frac{\partial^2 A^2}{\partial x^2},
\] (78)

where the expressions for coefficients \( r \) and \( s \) can be found in (New & Esteban, 1999). For waves propagating only in one direction this equation readily reduces to Eq. (74) with \( \alpha = r/c \) and \( \beta = cs/2 \) assuming additionally that the rotation effect is of the same order of smallness as the nonlinear and finite-depth dispersion effects. Note, that for long plane waves of small but finite amplitude Eq. (78) is physically equivalent to the Boussinesq set of equations, Eqs. (4)–(7), i.e. it is obtained in the same approximations on the smallness of nonlinearity and finite-depth dispersion. The rotation effect which is supposed to be small enough to neglect by centrifugal effect, still could be essentially greater than the nonlinear and finite-depth dispersion effects. In other words, the Boussinesq set of equations as well as Eq. (78) are applicable not only for dispersionless waves or waves slightly affected by finite-depth dispersion as shown in Fig. (9), but also for infinitely long waves which are strongly affected by Coriolis dispersion.

Internal waves in a two-layer rotating fluid were also studied in (Plougonven & Zeitlin, 2003; Zeitlin et al., 2003). Following Shrira’s approach (Shirira, 1981; 1986) developed for strongly nonlinear surface waves in a rotating fluid, they considered stationary periodic solutions for interfacial waves without high-frequency dispersion and then numerically constructed wave shapes. The analysis of the stability of such nonlinear periodic waves was examined by means of numerical simulation (Bouchut et al., 2004). It has been demonstrated that the nonlinear waves are stable at least with respect to a certain class of initial perturbations and, moreover, large-amplitude waves can evolve from the initial perturbations.

An important peculiarity of equations Eq. (74) and Eq. (78) is that for the high-frequency dispersion characteristic of oceanic waves (when the coefficient \( \beta \) in Eq. (74) is positive), solitary waves in the form of stationary localized pulses cannot exist at all (Leonov, 1981; Galkin & Stepanyants, 1991; Liu & Varlamov, 2004). Physical interpretation of this “antisoliton theorem” is rather simple: due to rotational dispersion, there is always a
resonance (phase synchronism) between a source moving at an arbitrary speed and linear perturbations. This resonance leads to a wave radiation from the soliton so that it can not remain stationary. The same is true for smooth solutions of the reduced form of Eq. (74) at \( \beta = 0 \). However, in the latter case, non-analytic solitary structures with derivative jumps can exist in the form of smooth-head solitons, sharp-crest solitons ("cuspons"), loop solitons and "compactons" (solitons determined on a compact support) (Ostrovsky, 1978; Rybak & Skrynnikov, 1990; Vakhnenko 1992; Stepanyants, 2006; Parkes, 2005). These solitons consist of several matched pieces of separate branches of solutions.

In addition to the stationary solutions mentioned above, some non-stationary solutions for the rKdV equation have also been studied, mostly numerically (Ostrovsky & Stepanyants, 1990; Gilman et al., 1996). It was observed, in particular, that the initial KdV-type solitary perturbation undergoes a "terminal decay", i.e. it completely annihilates (more exactly, transforms into radiation) in a finite time (Grimshaw et al, 1998a; Grimshaw et al., 1998b). The "extinction time" of a KdV soliton was estimated for real oceanic conditions. It amounts about one day for the first internal mode, and is larger for the higher modes. Note, that this is in rough agreement with the observation that solitons can exist in the oceans for a few days (see, e.g., the typical lifetime of a soliton packet, \( \tau_{life} \), in Table 2, Section 4.1).

In the process of soliton decay, at the rear part of the initial impulse, a new soliton-like perturbation is generated from the radiative tail as shown in Fig. 10. Its shape is close to the initial KdV soliton. This impulse shares the fate of the initial soliton, i.e. it terminally decays in a finite time, then generating a new pulse-type perturbation in its rear part, and so on. Such process resembles some sort of the recurrence phenomenon well-known for the KdV model [see, e.g. (Ablowitz & Segur, 1981)]. During this process, the soliton background tends towards a parabolic shape.

Another interesting example of a non-stationary wave propagation within the framework of rKdV equation is shown in Fig. 11 (Gilman et al., 1996). This example shows that despite of antisoliton theorem (Leonov, 1981; Galkin & Stepanyants, 1991) which forbids the existence of stationary solitary waves, the non-stationary large-amplitude solitary waves can co-exist with smooth periodic waves of small amplitude. Solitary wave shape is very close to KdV soliton, and its amplitude and other related parameters vary adiabatically in time and space.

Related equations were also obtained for deep rotating fluids. One of them is a generalization of the BO equation for a very deep ocean having a relatively thin pycnocline (Grimshaw, 1985):

\[
\frac{\partial}{\partial x} \left( \frac{\partial \eta}{\partial t} + c \frac{\partial \eta}{\partial x} + \alpha \eta \frac{\partial \eta}{\partial x} + \frac{\delta}{\pi \partial x^2} \int_{-\infty}^{\infty} \frac{\eta(\zeta, t)}{x-\zeta} d\zeta \right) = \frac{f^2}{2c} \eta.
\]  

(79)

Estimates show that the role of rotation is important for real internal waves in the ocean that have lengths of a few kilometers and more and periods in excess of roughly one hour. Computations of a modification of the Boussinesq-type equations with rotational terms (cf. Eqs.(14) and (15)) have been made by Gerkema (1994) and Gerkema & Zimmerman
Figure 10: Evolution of a KdV soliton on a constant negative background within the framework of the rKdV equation. Numbers near the wave crests indicate the time in dimensionless variables. From Gilman et al., 1996).

(1995). They have shown that the process of nonlinear internal wave generation by a tide can be strongly affected by rotation. In particular, the amount of solitons generated at each tidal period typically decreases due to rotation. GerkeMA (1994) analyzed the role of rotation for some observations and found it to be significant for moderate and high latitudes.

3.8 Strongly Nonlinear Waves

In previous sections, both nonlinearity and dispersion were considered small in the sense that in, e.g. a two-layer fluid, the displacement of the pycnocline is significantly smaller than its equilibrium depth (or, for a pycnocline close to the bottom, than its height over the bottom). Along with a number of observations for which the weakly nonlinear models provide a good approximation, there is also a growing number of data for which they are evidently wrong (see the experimental Section 4 for examples). Note that the transition from the KdV equation to the eKdV-type equation has been suggested by Lee and Beardsley.
(1974) to improve theoretical description of Halpern's observations (HALPERN, 1971). As was mentioned above, the eKdV model sometimes may serve as a phenomenological model for strong solitons because, as in the more consistent theories, it predicts much wider solitons than those which follow from the KdV equation for strong waves.

A more consistent description of strongly nonlinear internal waves can be based on direct numerical simulation for the basic hydrodynamic equations. For 2D steady-state waves in the Boussinesq approximation, the basic Euler equations, Eqs. (4) to (7), can be reduced to a single equation for the stream function, $\Psi$, as shown as early as in 1930s by DUBRIEL-JACOTIN (1932) and later independently by LONG (1953):

$$\frac{\partial^2 \Psi}{\partial \xi^2} + \frac{\partial^2 \Psi}{\partial z^2} + \frac{N^2(z - \Psi/V)}{V^2} \Psi = 0. \tag{80}$$

Here $\xi = x - Vt$, $N(z)$ is the known buoyancy frequency, and the velocity components are $u = \partial \Psi / \partial z$ and $w = -\partial \Psi / \partial x$. It is seen that at $N(z) = \text{const}$, the steady waves (but only
them!) are described by a linear equation, as mentioned above for weakly nonlinear waves.

To describe strongly nonlinear internal waves, direct numerical simulation for the basic hydrodynamic equations has been used. In particular, many numerical works have considered steady waves in a two-layer fluid. For this case, the linear Laplace equation can be used for each layer. The first study using this approach was probably the paper by AMICK & TURNER (1986) [see also (TURNER & VANDEN-BROECK, 1988)]. In addition to a detailed mathematical treatment of the problem, they have shown that there exists a limiting amplitude at which a soliton acquires a flat top and tends to two separated kinks, similar to the case of the eKdV equation but with different parameters. The amplitude and velocity of such a limiting soliton are

\[ \eta_{\text{lim}} = \frac{h_1 - h_2\sqrt{a}}{1 + \sqrt{a}} \approx \frac{h_1 - h_2}{2}, \quad (81) \]

\[ v_{\text{lim}} = \frac{\sqrt{g(1-a)(h_1 + h_2)}}{1 + \sqrt{a}} \approx \frac{\sqrt{g'(h_1 + h_2)}}{2}, \quad (82) \]

where \( a = \rho_1/\rho_2 < 1 \), \( g' = g(1-a) \), and positive displacement is upward. The relations on the right are valid for oceanic conditions where density variation is always small, i.e., \( a \approx 1 \).

Subsequently, direct numerical analysis of the two-layer case as applied to stationary solitary waves was performed by many authors [see, e.g. (EVANS & FORD, 1996; GRUE ET AL., 1999)]. As an example, Fig 12 shows calculations of soliton profiles for the two-layer fluid using the parameters chosen by EVANS & FORD (1996) [see also the paper by Evans in (DUDA & FARMER, 1999)].

More recently, calculations of solitary waves in a sea with smooth stratification have been performed. In particular, VLAŞENKO ET AL. (2000) calculated some practical cases, taking the data from observation. They considered a stratified layer with different density profiles, including those of a pycnocline type and some smoother ones, and calculated stationary soliton structures using the Euler equations for vorticity and density. Qualitatively their results or soliton width and velocity correspond to the two-layer approximation but quantitatively, they differ significantly from both the two-layer, strongly nonlinear model, and from the weakly nonlinear KdV model for the stratifications used in the work.

Also, some non-steady, strongly nonlinear processes have been studied by direct numerical simulation. VLAŞENKO & HUTTER (2002a,b) have modelled the shoaling of long internal waves in a coastal area. These processes include steepening, formation of soliton groups, and breaking, with the associated generation of turbulence (Fig. 13).
Figure 12: Normalized soliton profiles for the two-layer fluid with $h_2/h_1 = 3$, $\rho_1/\rho_2 = 0.997$ (surface is at +1; bottom at −3 on the vertical axis). The profiles shown correspond to different soliton amplitudes, $-\eta_0/h_1 = 0.05$; 0.1; 0.2; 0.3; 0.4; 0.5; 0.6; 0.7; 0.8; 0.9; 0.925; 0.95; 0.975; and 0.99. The dashed horizontal line marks the level of limiting amplitude wave. From (EVANS & FORD, 1996).
Figure 8. Evolution of the density field during internal wave breaking near the lake boundary (along the cross section 3 in Fig. 1). The amplitude of the incoming wave is $a_m = 15$ m. The fluid stratification used is depicted by line 1 in Fig. 3. The density anomaly relative to the free surface is shown. The contour interval equals $0.2 \, \text{kg} \cdot \text{m}^{-3}$, the time scale $T$ equals 83 s. Segment $A - A$ represents a mixing zone with a newly formed fluid stratification after wave breaking.

Figure 9 illustrates the process of overturning of a rear steep front; it is shown at that moment of time when the top of the propagating baroclinic bore outstrips the wave trough. In this situation, the heavier and denser water penetrates into the relatively light water layers and falls down to the wave trough (Fig. 8c). Thus, we may conclude that the reason for wave breaking is the kinematic instability of the wave profile. This conclusion can be seen even more clearly in Fig. 9, where the horizontal and vertical velocities during of an overturning event are presented (the left panels show the fields of density and $u$ and $w$ velocities before overturning and the right panels – after overturning). Before breaking, the maximum of the horizontal velocity (30.1 cm s$^{-1}$) was located at the free surface. During the overturning event, this maximum shifts from the free surface to the deep layers and to the lateral boundary of the steep slope, where the value of the maximum velocity increases right up to wave breaking. At the moment of breaking, the onshore fluid velocities (42.1 cm s$^{-1}$) exceed the wave phase speed (36.4 cm s$^{-1}$).

The overturning event is also clearly seen from the analysis of the evolution of the $w$-field (note the appearance of $\bar{w}$).
LAMB (2002, 2003) has shown that shoaling of a solitary wave can result in the formation of a trapped core that was observed long before in the laboratory experiment by DAVIS & ACRIVOS (1967). STASNA & LAMB (2002) have also numerically investigated soliton propagation on shear currents. For a linear shear current of a constant vorticity, they found asymmetry in the internal solitary wave propagation, specifically the wave propagating in one direction is taller and narrower than the wave propagating in the opposite direction. Their numerical study also showed that the maximum wave amplitude solution approaches the conjugate flow limit or a kink-type solution. At the same time, wave breaking phenomena and shear instability are also possible for some parameters. BREYIANNIS ET AL., 1993 numerically studied internal waves in two-layer fluid with a linear shear flow. Although their method is applicable for the general case, they focussed on surface waves at the air-water interface, where a large density jump occurs, and also added wind of constant velocity in the upper layer.

Direct numerical computations are usually time-costly and, what is perhaps more important, they do not provide physical insight into the problem. At the same time, the analysis of observational data shows that, although wave amplitudes are often so big that there is no small parameter allowing one to construct a weakly nonlinear model, in many cases the wavelength remains larger than the thickness of one of the layers or of the total ocean depth, especially in the coastal zones. For these cases, a long-wave approximation can be developed that uses the corresponding expansion of dispersive terms while keeping a strong nonlinearity. This approach was first suggested for surface waves by WHITHAM (1967) who used a Lagrangian representation of the primitive hydrodynamic equations. He derived a model equation by means of perturbation methods based on the expansion of the Lagrangian (The later work by GREEN & NAGHDI (1976) that included a sloping bottom should also be mentioned). We shall concentrate further in this section on long-wave models in which the explicit dependence on the vertical coordinate is eliminated.

For internal waves, MIYATA (1985, 1988, 2000) was apparently the first who suggested (albeit without a detailed derivation) the long-wave equations for strongly nonlinear, weakly dispersive waves in a two-layer fluid, and analyzed a steady solitary solution of these equations. Miyata’s equations, together with other weakly nonlinear models, KdV, eKdV, BO and JKKD, were examined against numerical calculations in (Miyata, 1988; Michallet & Barthélemy, 1998). It was shown in particular that, as expected, the eKdV works well when the pycnocline is close to the middle of the water layer, whereas Miyata’s model agrees very well with numerical data for fully nonlinear hydrodynamic equations practically in the full range of wave amplitudes (except for those very close to the limiting amplitude where numerical results may themselves become erroneous).

A detailed analysis of the same problem for a two-layer fluid (including the case when the lower layer is infinitely deep) was performed by CHOI & CAMASSA (1996, 1999). For shallow water, these equations (essentially the same as those obtained by Miyata), being reduced to the case of a small density jump, $\Delta \rho \ll \rho_{1,2}$, can be represented in the form

$$
\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x} [(h_1 + \eta) u_1] = 0,
$$

(83)
\[- \frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x} [(h_2 - \eta)u_2] = 0. \]  
(84)

\[\frac{\partial (u_1 - u_2)}{\partial t} + u_1 \frac{\partial u_1}{\partial x} - u_2 \frac{\partial u_2}{\partial x} + g' \frac{\partial \eta}{\partial x} = D, \]  
(85)

where

\[D = \frac{\partial}{\partial x} \left\{ \frac{(h_1 + \eta)^3}{3(h_1 + \eta)} \left[ \frac{\partial^2 u_1}{\partial x^2} + u_1 \frac{\partial^2 u_1}{\partial x^2} - \left( \frac{\partial u_1}{\partial x} \right)^2 \right] \right\} - \frac{\partial}{\partial x} \left\{ \frac{(h_2 - \eta)^3}{3(h_2 - \eta)} \left[ \frac{\partial^2 u_2}{\partial x^2} + u_2 \frac{\partial^2 u_2}{\partial x^2} - \left( \frac{\partial u_2}{\partial x} \right)^2 \right] \right\}. \]

Here the $z$-axis is directed downwards, and $u_{1,2}$ are the horizontal velocities in the layers, each averaged over the layer thickness.

A more consistent derivation of the long-wave, strongly nonlinear model, Eqs. (83)–(85), in the Boussinesq approximation ($\Delta \rho \to 0$) has been provided by (Ostrovsky & Grue, 2003) on the basis of the (aforementioned) Whitham’s Lagrangian approach.

For a stationary soliton in which all dependent variables depend on one “running coordinate” $x - Vt$, this system can be reduced to a second-order ODE, the first integral of which gives

\[\frac{d\eta}{dx} = \mp \eta \sqrt{\frac{3[(h_1 + h_2) - g'(h_1 + \eta)(h_2 - \eta)/V^2]}{h_1^2(h_2 - \eta) + h_2^2(h_1 + \eta)}}, \]  
(86)

where different signs correspond to the frontal and trailing edges of a soliton, respectively. In particular, the soliton velocity is related to its amplitude $\eta_0$ by

\[V^2 = \frac{g'(h_1 + \eta_0)(h_2 - \eta_0)}{h_1 + h_2}. \]  
(87)

Note that this latter expression differs from the linear long-wave velocity, Eq. (11), only in that instead of the non-perturbed depths, $h_{1,2}$; those at the soliton peak, $h_1 + \eta_0$ and $h_2 - \eta_0$, are taken. The maximum possible amplitude of a soliton in this approximation coincides with that presented by Eq. (81). The solitary wave solution of this equation shows very good agreement with laboratory experiments and numerical solutions of the full set of Euler equations (Camassa et al., 2006). Later, Jo & Choi (2002) studied the above system numerically to describe non-stationary processes, such as soliton formation, interaction, and propagation over non-uniform topography.

A similar method was used by Voronovich (2003) to find stationary solitary solutions in a two-layer fluid when each layer is stratified in such a way that the buoyancy frequency, $N$, is constant in each layer (a “2.5-layer model”). For steady waves, each layer is described by a linear equation (80) so that nonlinearity is again due to the interface. Note that in this case the solution may include an internal vortex core.

Yet another model for large-amplitude long interfacial waves in a two-layer fluid of finite depth was derived by Craig et al. (2004). The displacement of the interface between the
layers was assumed to be of small slope, i.e., $\eta_0/\Lambda \ll 1$, where $\eta_0$ is the wave amplitude and $\Lambda$ is the wavelength, but no smallness assumption was made on the wave amplitude. Based on Hamiltonian representations of the primitive set of hydrodynamic equations, a pair of coupled equations for the interface displacement and fluid velocity was derived by means of a perturbation method. The equations contain fairly complex nonlinear-dispersive terms and their solutions have not been analyzed so far. The comparison of this model with the earlier derived ones has also not been analyzed so far.

3.8.1 Non-dispersive Waves and Evolution Equations

The set Eqs. (83) and (85) is two-directional and can be considered as a strongly nonlinear extension of the Boussinesq equations in the two-layer case. In applications, one usually deals with waves propagating in one direction from a source, such as a shelf break, transforming a part of the energy of the barotropic tides into internal waves propagating onshore. Hence, an important problem is to obtain an evolution equation for a wave propagating in one direction, i.e. a strongly nonlinear analog of the KdV or eKdV equations. Besides simplifying analytical considerations and making the result more physically clear, this may also significantly save computer time in practical applications. This seems to be especially attractive for strongly nonlinear waves, the equations of which are typically non-integrable.

This problem was discussed by Ostrovsky (1999) and addressed in detail in Ostrovsky & Grue (2003). The approach starts from the exact non-dispersive limit for long waves when the term $D$ in Eq. (85) is neglected. In this case, a progressive (simple, or Riemann) nonlinear wave exists which propagates with a nonlinear velocity $c(\eta)$, and all variables are functionally related: $u_{1,2} = u_{1,2}(\eta)$, so that

$$\frac{\partial \eta}{\partial t} + C_{\pm}(\eta) \frac{\partial \eta}{\partial x} = 0.$$  

(88)

Here, two possible simple wave velocities exist (Sandström & Quon, 1993):

$$C_{\pm}(\eta) = \frac{1}{h} \left[ \frac{(h_2 - \eta)^2 - (h_1 + \eta)^2}{h_2 - \eta} \right] u_1 \pm \sqrt{g' H (h_1 + \eta) (h_2 - \eta) - \frac{u_1^2 H^2 (h_1 + \eta)}{h_2 - \eta}},$$  

(89)

where $H = h_1 + h_2$, and the variables are related by an equation

$$\frac{du_1}{d\eta} = \frac{C(\eta) - u_1}{h_1 + \eta}.$$  

(90)

As shown by Slunyaev et al. (2003), these relations can be expressed in an explicit form for the simple wave velocity which reads as

$$C(\eta) = \pm c \left\{ 1 + 3 \frac{(h_1 - h_2)(h_1 - h_2 - 2\eta)}{(h_1 + h_2)^2} \left[ \sqrt{\frac{(h_1 - \eta)(h_2 + \eta)}{h_1 h_2}} - \frac{h_2 - h_1 + 2\eta}{h_2 - h_1} \right] \right\},$$  

(91)
where \( c \) is the linear velocity of long waves given by Eq. (24).

As pointed out in the cited paper, the velocity of a simple wave reduces to the linear wave velocity in two cases, when the perturbation is infinitesimal, \( \eta \to 0 \), and when \( \eta = (h_2 - h_1)/2 \); it is worth noting that the latter formula determines the limiting amplitude, \( \eta_{\text{lim}} \), of a soliton.

These results are exact for a non-dispersive wave in a two-layer fluid. Similar to the gas dynamics case (Landau & Lifshitz, 1988), the basic non-dispersive equations can be rewritten in terms of Riemann invariants, \( I_\pm = u_1 + U_\pm(\eta) \), where \( U_\pm \) correspond to the above dependencies between \( u_1 \) and \( \eta \) in a simple wave, and the signs \( \pm \) correspond to the signs at the radical in Eq. (89). As is known from the theory of hyperbolic equations, the progressive wave of Eq. (88) corresponds to the case when one of the invariants, e.g. \( I_- \), turns to a constant, in our case zero. The latter condition defines a relationship between the variables equivalent to Eq. (90), and their use in the equation for \( I_+ \) results in a simple wave Eq. (88).

Returning to the full set of equations (83)–(85) in which the dispersive operator \( D \) is non-zero but small, for the wave propagating in a positive direction, the invariant \( I_- \) is also small, of order \( D \). As a result, Eq. (88) is modified to give (Ostrovsky & Grue, 2003)

\[
\eta_t + C(\eta)\eta_x = \frac{D_1(\eta)}{dI_+/d\eta} = D(\eta) \frac{h_1 + \eta}{C_+ - C_-} \frac{h_2 - \eta}{h_1 + h_2}.
\]

Here \( C_\pm \) correspond to the signs in Eq. (91).

In (Ostrovsky & Grue, 2003) steadily propagating solitary solutions of this equation were found for the specific but practically important case of \( h_1 \ll h_2 \). These solutions differ noticeably from those of the two-directional Choi-Camassa equations considered above.

### 3.8.2 Simplified Evolution Equation (\( \beta \)-model)

It must be emphasized that, unlike in the weakly nonlinear case, the long-wave approximations for surface and internal waves considered above (both two- and one-directional) are of a somewhat contradictory nature. Indeed, these equations combine strong nonlinearity and weak dispersion, whereas a soliton exists as a balance between the nonlinearity and dispersion, so that it is \textit{a priori} unclear whether a strong soliton would be long enough to provide sufficiently small dispersion terms and thus secure the applicability of the shallow-water approximation. Considering this, in (Ostrovsky, 1999) a different approach to obtaining an evolution equation was suggested. Namely, the exact non-dispersive operator is retained unchanged but a dispersive term is represented in a semi-phenomenological form based on slowness of the displacement variation. For strongly nonlinear waves, along with the exact long-wave velocity \( c(\eta) \) defined by (89), we introduce a local dispersion parameter corresponding to that in KdV, but with the instantaneous values of the layer depths:

\[
\beta = \frac{1}{6} C(\eta)(h_1 + \eta)(h_2 - \eta).
\]

As a result, a strongly nonlinear evolution equation (\( \beta \)-model) reads
\[
\frac{\partial \eta}{\partial t} + C(\eta) \frac{\partial \eta}{\partial x} + \frac{\partial}{\partial x} \left[ \beta(\eta) \frac{\partial^2 \eta}{\partial x^2} \right] = 0.
\] (94)

Detailed comparisons between solitary solutions following from different long-wave models, direct fully nonlinear computations, and observational data, were presented in (Ostrovsky & Grue, 2003). An example is shown in Fig. 14. The particle velocity and soliton propagation velocity are approximated very well by the long-wave equations. As regards the soliton profile and width, for a moderate depth ratio, \( h_2/h_1 \), both two-directional Choi–Camassa (CC) model and the evolution \( \beta \)-model are quite satisfactory.

![Graph of \( V/c \) versus \( \eta/h_1 \)](image)

Figure 14: Normalized soliton velocity versus normalized amplitude for \( h_2/h_1 = 20.4 \) (COPE). Solid line – fully nonlinear numerical solution, crosses – \( \beta \)-model (CC-model gives very close results), dotted line – fully nonlinear calculations for \( h_2/h_1 = 500 \) (i.e., practically infinitely deep lower layer). Filled square – observation. From (Ostrovsky & Grue, 2003).

Note that all one-directional models have a common disadvantage: they do not exactly conserve mass and energy at the same time. In particular, Eq. (92) conserves neither, and Eq. (94) conserves only mass (although energy is close to constant, as well). Its modification, the “e-model” also used in (Ostrovsky & Grue, 2003) conserves energy but allows slight variations of mass. Fortunately, solutions of the \( \beta \)- and e-equations are typically close to each other. As a further development, the \( \beta \)-model was expanded to the case of a shelf with a sloping bottom, and was verified by direct numerical simulation in (Vlasenko et al., 2005), where the adiabatic transformation of strong solitons was demonstrated.

For a very large depth ratio their applicability is limited because the basin is not always shallow enough for the solitons. In these cases the following deep-layer model may work better.
### 3.8.3 Deep Lower Layer

Another limiting case when the lower layer is infinitely deep but the wave is still long as compared to the upper layer; for weak nonlinearity, this is the case for which the BO equation, (41), has been derived. In a two-directional, weakly dispersive form the corresponding long-wave equations were obtained by Choi & Camassa (1996). They considered the upper layer nonlinear but non-dispersive and the lower fluid dispersive but linear, and obtained the equations in the form (here we again use the limit of a small density jump across the density interface)

\[
\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x} [(h_1 + \eta) u_1] = 0,
\]

\[
\frac{\partial u_1}{\partial t} + u_1 \frac{\partial u_1}{\partial x} + g \frac{\partial \eta}{\partial x} = \frac{1}{\pi} \varphi \int_{-\infty}^{\infty} \frac{\partial^2 \eta(t, \xi)}{\partial \xi^2} \frac{d\xi}{\xi - x}.
\]

Here, the same restrictions on the applicability of the dispersive term should again be imposed. In (Ostrovsky & Grue, 2003) an evolution equation generalizing the BO equation for strong nonlinearity has been suggested in the form

\[
\frac{\partial \eta}{\partial t} + C(\eta) \frac{\partial \eta}{\partial x} + \frac{1}{\pi} \frac{\partial}{\partial x} \varphi \int_{-\infty}^{\infty} \frac{\partial \eta(t, \xi)}{\partial \xi} \gamma(\eta(t, \xi)) \frac{d\xi}{\xi - x} = 0.
\]

where \( C(\eta) = \sqrt{g' h_1 \left(3\sqrt{1 + \eta/h_1} - 2\right)} \) is the simple wave velocity (89) taken in the limit of \( h_1/h_2 \to 0 \), and \( \gamma(\eta) = \frac{1}{2} C(\eta)(h_1 + \eta) \) generalizes the corresponding parameter of the BO equation in the same manner as it has been done above for the \( \beta \)-model. This equation gives a very good agreement with direct computations, at least for moderate soliton amplitudes, up to \( \eta_0 = 0.8 h_1 \).

In view of the growing number of observations of strong solitons in the ocean (see the experimental part below) and the increasing role of coastal areas, the theory of strongly nonlinear internal waves should be considered as an important branch of theoretical physical oceanography.

### 4 Experimental observations in the oceans

Observations of nonlinear internal waves close to solitons or their groups (solibores) are numerous and these measurements are being actively performed now. It is impossible to even mention all them. Here we discuss some examples, specifically a few of the pioneering ones and also a few recent experiments. More observational data can be found in (Duda & Farmer, 1999; Sabinin & Serébrany, 2005), and in the Internet Atlas (Jackson & Apel, 2004).
4.1 Internal Solitons Near the Continents

With the advent of satellites that carry high-resolution imaging systems (e.g., Landsat, Seasat, and subsequent spacecraft), it has been possible to obtain overviews of certain processes occurring in the sea that have distinguishable surface signatures. Such phenomena include coherent internal waves. The combination of the methods of satellite oceanography, underwater acoustics, and in situ measurements has enabled considerable progress to be made in understanding the kinematics and dynamics of these waves. The Olympian view provided by the satellite remote observations has allowed careful planning of subsequent in situ experiments to be done, with the result that a moderately detailed picture of the birth, evolution, propagation, and decay of the internal waves has been obtained. For those cases in which workers could achieve both (a) the synoptic view provided by satellites and (b) the detailed in-water view given by current meters and acoustic echo-sounders, these coherent waves have most frequently proven to be soliton-like. As a consequence, when soliton characteristics are observed in satellite images for which no concurrent in situ data are available, one can assume with some confidence that a solitary wave is indeed being observed (APEL & GONZALEZ, 1983). In addition, certain features unique to soliton interactions have been seen in satellite imagery, e.g. the spatial phase shifts that occur when two solitons pass through each other [see, e.g., in (APEL ET AL., 1995)]

Such observations are unambiguous proof of the solitary wave character of these oceanic internal waves.

Having established that internal solitons are fairly commonplace in the sea, we can then discuss the observations in that context without continually raising questions about the correctness of the interpretation.

The question as to why an essentially underwater oscillation is so visible on the surface has been extensively studied, and a short treatment of the physical processes resulting in internal wave signatures will be given in Subsection 4.3.

Tidal interaction with bottom features appears to be the dominant mechanism for generation of the coherent oceanic internal waves near the continents; closer to shore, riverine canyons or glacial scours provide secondary generating mechanisms. Also, the boundaries of intense current systems such as the Gulf Stream appear to be sources of coherent wave packets that propagate at large angles to the current direction. Since tides, stratification, and bottom topography are global features, one expects tidally excited internal waves to be ubiquitous wherever stratified waters and shallow bathymetry exist.

At least two mechanisms have been advanced to explain the generation process at the shelf break. The first, formulated by RATTRAY (1960), hypothesizes that the barotropic (i.e. uniform in depth) tide will be scattered into baroclinic (i.e. varying with depth) modes at the shelf break. The second is a kind of lee-wave mechanism, wherein tidal flow directed offshore beyond the shelf break results in an oscillating depression of the pycnocline just offshore of the break (MAXWORTHY, 1979). As the tidal current ellipse is swept out, the reversal of that current releases the lee wave from its down-current, phase-locked position. The wave of depression then propagates opposite to the earlier current direction, i.e. toward

---

15 The results were originally presented at a conference (APEL & LIN, 1991).
shore, where it evolves independently of further tidal action except for advective effects. There is some experimental support for both of these processes at the shelf edge, at sills or similar geometries, and different opinions exist concerning this. Some works suppose that the second mechanism is dominant (Halpern, 1971; Sandström & Elliott, 1984; Apel et al., 1985, Hibya, 1986). The others prefer the first one (Gerkema & Zimmerman, 1995; Vlasenko, 1993). It should be noted that the first scattering mechanism does not have a threshold, whereas the second one does. On the other hand, the second one seems to be better experimentally investigated. For the lee-wave generation mechanism, a necessary condition is that the current velocity $U$ should exceed the local phase speed of the internal wave $c$. That is, the internal Froude number, $Fr$, should exceed unity, where $Fr$ is defined as

$$Fr = \frac{U^2/\Delta l_z}{g\Delta \rho/\rho}.$$  

(98)

Here $\Delta l_z$ is the vertical scale of the density and velocity gradients. Note that this condition correlates with the one needed for shear flow instability (see above) which is based on the Richardson number condition $Ri < 1/4$. The Richardson number can be expressed as $Ri \sim 1/\text{Fr}$. Shear instability can be considered as a third possible mechanism of IW generation (which can be realized independently of the tide-shelf break interaction.); such a process may be responsible for soliton formation in Knight Inlet (Farmer & Armi, 1999). Since internal wave phase speeds are typically of order 0.1 to 1.0 m/s, only moderate tidal currents are required for their formation. Such currents can occur over underwater sills; near continental shelf breaks, islands, straits, and shallow banks; and even at the mid-ocean ridges, where currents such as the Gulf Stream Extension cross the Mid-Atlantic Ridge.

Figure 15 is an image of the ocean made with the 6-cm-wavelength synthetic aperture radar (SAR) on the European Remote Sensing Satellite, ERS-1, taken southeast of New York near the edge of the continental shelf on 18 July 1992. The image is 90 \times 90 km$^2$, and shows quasi-periodic internal wave signatures running diagonally across it, approximately parallel to the edge of the continental shelf. The waves occur in packets separated by some 20 to 25 km and propagate under refractive control of the shallow continental shelf in water depths between roughly 200 to 50 m. Each packet has been generated by the semidiurnal (12.5-h) and diurnal (25-h) tides during the phase when the tidal current is directed offshore. Initially, the offshore flow depresses the pycnocline just seaward of the continental shelf break. As it moves shoreward, the depression begins to steepen and develop undulations, most likely because of the dispersion at the leading edge of the depression which agrees with the description given in the theoretical part here. Within less than one semidiurnal tidal period, the pycnocline depression became fully modulated and has grown into a solitary wave train. It is these wave packets that are visible in Fig. 15, which shows perhaps six groups of waves that represent a history of soliton formation, propagation, and attenuation extending backward in time to at least 75 h. Also visible in the lower right-hand corner is what is believed to be a nascent packet being formed just at the shelf break at the time of the satellite overpass.
Figure 15: Image of New York Bight taken by the European Remote Sensing Satellite ERS-1 on 18 July 1992, showing several packets of internal solitons generated during six previous tidal cycles in the vicinity of the Hudson Canyon. Dimensions 100 × 100 km. Image courtesy of R. D. Chapman and the European Space Agency.

Figure 16 is an acoustic echo-sounder trace of solitary wave packets in the region of Fig. 15, taken with a 20-kHz downward-looking pulsed sonar. The nonlinear character of the waves is clearly visible, with only downgoing pulses appearing and with very little upgoing excursion. The towed echo-sounder made repeated passes across the wave train, thereby allowing questions of stability and wave coherence to be addressed (APEL ET AL., 1975A). Moored and towed current meters and temperature probes give similar detailed pictures of the behavior of the waves with time.

Equations (22) and (23) allow one to test if solitary waves measured in situ are KdV solitons or not. There are different possible indicators showing that the wave is close to a soliton:

i) the conformity of the wave profile with the sech²-shape;
ii) the linear increase of the velocity versus amplitude; and
iii) the inverse proportionality of the solitary wave amplitude to the square of its width,

\[ \Delta^2 \eta_0 = \frac{12 \beta}{\alpha} \simeq \frac{4h_1^2 h_2}{3|h_1 - h_2|}. \]
Figure 16: Acoustic echo-sounder record of solitary wave displacements in the region of Fig. 15. From (Gasparovic et al., 1986).

Usually the KdV criterion works well for moderate-amplitude solitons. Figure 17 shows data taken by Nagovitsyn et al. (1990) in the Sea of Okhotsk (Far East) during the summer. The data have been normalized and are compared with the KdV soliton shape; the agreement is considered to be satisfactory. Analogous comparisons can be found in (Kuznetsov et al., 1984; Sandström & Elliott, 1984).

A schematic of the type of solitary waves appearing in Fig. 15 is given in Fig. 18; an overall vertical profile and a plan view are shown, along with the characteristic length scales. Such packets have strong tendencies for those individual solitons having the largest amplitudes, longest wavelengths, and longest crest lengths to be at the front of the group, with the ones having the smallest attributes appearing at the rear. In theory, the small solitons are followed by linear, dispersive wave trains, but such are difficult to distinguish from the trailing solitons.

Direct observations of the generation of internal solitary waves at the Nova Scotian shelf break have been made by Sandström & Elliott (1984). The wave packets were generated there during the phase of the tide when the current was directed offshore. The evolution of the packet occurred with surprising rapidity; fully developed solitary wave characteristics appear within 2 to 3 h of the packet’s birth.

Internal waves propagating shoreward on the continental shelf soon encounter shoaling bathymetry that affects their propagation speeds and amplitudes. Near the region where
Figure 17: Normalized soliton shapes (different symbols) as measured in the Sea of Okhotsk, in comparison with theoretical profile of KdV soliton (solid line). From (NAGOVITSYN ET AL., 1991).
Figure 18: Schematics of a soliton wave packet on the continental shelf, and some of its dimensions. From (APEL, 1995).
the pycnocline depth is roughly one-half the total depth, they appear to undergo various transformations. These processes have been studied in numerous papers. Depending on local hydrological conditions and the incoming wave, different scenarios of subsequent wave evolution can occur including the creation of sequences of secondary solitons, disperse wave-trains, various types of billows, vorticity formation or just turbulent spots (Pelinovsky & Shavratsky, 1976; 1977; Djordjevic & Redekopp, 1978; Knickerbocker & Newell, 1980; Helfrich et al., 1984; Malomed & Shira, 1991). As mentioned, for the case of a mid-depth pycnocline, the extended KdV equation (31) works well; its application to the interpretation of observation data is described in, e.g. the paper by Holloway et al. (2002). Wave propagation through the point when the pycnocline crosses the middle of the layer due to the depth variation has been observed by, e.g. Orr & Mignerey (2003).

In shallow regions, increased amounts of suspended sediments often exist due to solitons, as has been observed with acoustic echo-sounders (Apel et al., 1975b; Proni & Apel, 1975). Since the bottom currents associated with these waves are large enough to resuspend sediments, the force due to bottom friction is large and the process of breaking on the sloping bottom may be what ultimately destroys the waves. Soliton signatures have generally disappeared from imagery taken in this shallow domain. Soliton processes also apparently inject large amounts of nutrients into the food chain in the shallow region (Sandström & Elliott, 1984), and it is likely that other areas also benefit biologically from the effects of internal waves. Note that in a number of cases, satellite images show more than one family of solitary wave packets undergoing collisions.

Clear evidence of soliton collisions has been obtained using satellite imagery. Figure 19a is a schematic of the SAR image of Fig. 20a that shows two intersecting solitary wave packets that have suffered phase shifts while crossing each other (which agrees with the theoretical results for the interaction of two KdV solitons). Ab initio calculations of the spatial phase shifts using concurrent measurements of density in the theory of Section 2 have yielded the multi-soliton interactions illustrated in Fig. 19b (Apel & Lin, 1991). A comparison between the theoretical and observational phase shifts is shown in Fig. 19c; the agreement between theory and experiment is quite good.

For solitons in midlatitude continental shelf waters during summer conditions, typical scales of the waves (their order of magnitude) in terms of a two-layer approximation are given in Table 1; note that deviations from these quantities can be large. The algebraic quantities appearing in the Table are defined in Fig. 18. The distance $D$ between successive packets is set by the nonlinear phase speed of the leading solitons in the packets. The radius of curvature $R_c$ can be of the order of the distance from the generation region, although in shallower water, the phase fronts become controlled by the bathymetry.
Figure 19: (a) Schematic of phase shifts during encounter of two soliton packets, as derived from SAR image of Fig. 20. (b) Interacting solitons from the analytical KdV model. (c) Calculated and observed phase shifts for several pairs of interacting solitons. From (APEL & LIN, 1991).
Figure 20: (a) Internal solitons in the New York Bight as observed with SAR from Environmental Research Institute of Michigan flown on Canadian CV-580 aircraft. (b) In situ current measurements of internal wave packet in frame (a). (c) Fractional modulations of waves in radar image, comparing observation and theory (the latter is briefly outlined below). From (Gasparovic et al., 1988).
Table 1. Typical Scales for Continental Shelf Solitons

<table>
<thead>
<tr>
<th>L (km)</th>
<th>( \eta_0 ) (m)</th>
<th>( h_1 ) (m)</th>
<th>( h_2 ) (m)</th>
<th>( l_s ) (m)</th>
<th>( \lambda_0 ) (m)</th>
<th>W (km)</th>
<th>D (km)</th>
<th>( R_c ) (km)</th>
<th>( \Delta \rho/\rho )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–5</td>
<td>0–30</td>
<td>5–25</td>
<td>100</td>
<td>100</td>
<td>50–500</td>
<td>0–30</td>
<td>15–25</td>
<td>25–( \infty )</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Scales for dynamical quantities appear in Table 2.

Table 2. Dynamical Quantities

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Brunt–Väisälä frequency, ( N/2\pi )</td>
<td>10 cycles/h</td>
</tr>
<tr>
<td>Radian frequency, ( \omega )</td>
<td>0.001–0.005 rad/s</td>
</tr>
<tr>
<td>Phase speed, ( c )</td>
<td>0.20–1.0 m/s</td>
</tr>
<tr>
<td>Current velocity, ( U )</td>
<td>0.10–1.0 m/s</td>
</tr>
<tr>
<td>Packet lifetime, ( \tau_{life} )</td>
<td>24–48 h</td>
</tr>
<tr>
<td>Interpacket period, ( \tau_{gen} )</td>
<td>12.5–25 h</td>
</tr>
</tbody>
</table>

At the edge of the continental shelf, still other processes are active in generating internal waves. For example, upwelling, especially near regions of strong bathymetric relief such as submarine canyons, appears to be an important source. Figure 21 shows an instance of such a disturbance as observed by the 10-cm-wavelength synthetic aperture radar (SAR) on the Soviet/Russian spacecraft Almaz-1 [see, e.g., (CHELOMÉI ET AL., 1990)]. On the basis
Figure 21: Almaz-1 SAR image upwelling and internal wave generation off the coast of Ireland at the edge of the continental shelf, 5 July. Image courtesy of V. S. Etkin and A. V. Smirnov (1991).
of preliminary calculations, workers analyzing the data believe that the upwelling serves to excite internal disturbances in submarine canyons and similar regions of high relief, which then go on to develop into solitons.

Subsequent observations of large, shoaling deep-water nonlinear internal waves include the Asian Seas International Acoustics Experiment (ASIAEX), a joint acoustics and physical oceanography observational program conducted in the northeastern part of the South China Sea in 2001. In ASIAEX, numerous moored, shipboard, and RADARSAT observations were made of the nonlinear internal wave field, in support of the program’s objective of determining the effect of nonlinear internal waves on acoustic propagation at low frequencies (50 – 1000 Hz) in shallow water (0 – 300 m). The nonlinear waves observed in ASIAEX were among the largest observed in the world’s oceans. Emanating from the Luzon Strait, these waves had crests up to 200 km in lateral extent, and vertical amplitudes ranging from 29 m to over 140 m. The ASIAEX experiment concentrated significant resources in studying the shoaling of these large waves, with heavily instrumented moorings extending from 300 m depth to 70 m depth, as well as high frequency acoustic imaging. The interested reader is referred to articles by Orr & Mignerey (2003), Liu et al. (2004), and Ramp et al. (2004).

Finally, we present an example of a group of strongly nonlinear solitary waves observed in 1995 in the Coastal Ocean Probing Experiment (COPE) off the coast of northern Oregon. Presumably due to the proximity of Columbia River, a sharp and shallow (5 – 7 m deep) pycnocline was formed on which groups of very strong, tide-generated internal waves propagated. A highlight of this experiment was that the in situ observations were performed in two locations along the onshore IW propagation direction separated by 20 km, so that the soliton group evolution could be followed up. At the first site, with a depth of 150 m, measurements were carried out with a CTD probe from the floating platform (FLIP), see Stanton & Ostrovsky (1998). In the second site, with a depth of 60 m, moored thermistor chains were used (Trevorrow, 1998). In both cases, the current velocity of the IW was measured. Remote images were also obtained from coastal X-band and Ka-band Doppler radars with horizontal and vertical polarizations (Kropfli et al., 1998). Details of this experiment can be found in the cited publications.

Figure 22 presents the 14°C isotherm displacement at the FLIP site. The isotherm depressions reach 25 – 30 m from its initial depth of 5 – 7 m. The same group of solitons was registered at the thermistor site after about 6.4 hours of onshore propagation so that mean propagation velocity was about 0.85 m/s. At this site the solitons become somewhat smaller (their amplitudes do not exceed 17 – 18 m that is still a very strong nonlinearity). Peak particle velocity exceeded 0.7 m/s, only slightly smaller than the propagation velocity, which confirms that the nonlinearity was very strong.

Figure 23 shows a radar image of the ocean surface. Amazingly, such images could be obtained even with strong sea surface roughness, at wind speeds of up to 13 m/s. The parameters of observed solitons are far from those predicted by the KdV equations (a somewhat better approximation can sometimes be given by the eKdV, as done in Stanton & Ostrovsky, 1998) but they can be satisfactorily approximated by the strongly nonlinear
models discussed above.

More examples of strongly nonlinear solitons are given in (Duda Farmer, 1999). Note also an earlier observation in Celtic Sea (Pingree & Mardell, 1986) where soliton amplitudes reached 40 m.

![Graph](image)

Figure 22: Temporal record of depth of 14°C isotherm for September 25–26, 1995 (the peaks are actually water depressions). From (Kroffli et al., 1999).
Figure 23: Panoramic radar images of an IW packet generated during strong spring tide on September 25, 1995. Left – Doppler velocity, right – scattering cross-section from vertically polarized radar signals. From (Kropflí et al., 1999).
Note that, in general, localized (pulse-type) perturbations observed in the ocean are not necessarily solitons. They, perhaps, develop into solitons later if the hydrological conditions are appropriate. An example of such an observation is the one made by (INALL ET AL., 2001) in 1995 at the edge of the Malin Shelf (to the west of Scotland) at a site of 145 m total depth. In the time series of recorded pulse-type internal waves, the largest waves had a downgoing vertical displacement of about 25 m on the pycnocline located at the depth of about 30 m, and propagated approximately toward the shelf with speeds of 0.54 – 0.6 m/s. Surprisingly, these large-amplitude waves were only encountered during neap rather than spring tides, so that their relationship with the tide is not clear. According to authors’ estimates, none of the waves observed possessed significant dispersion and hence, could not be treated as solitons. The authors estimated mass transport in the waves and obtained good agreement with the observed transport in the lower layer. In a typical IW packet of solitary waves, a lower layer transport of about 5 m$^2$/s offshore was maintained over a period of about 1.5 h, with a peak of about 20 m$^2$/s. As large-amplitude IWs appeared only sporadically at the halfway point of each tidal cycle, this short lived transport translates into a sustained rate over the whole deployment period of about 0.3 m$^2$/s. This correlates with Huthnance’s idea (HUTHNANCE, 1995) that nonlinear IWs may contribute significantly to cross-shelf exchange processes. (According to Huthnance’s estimate, the typical value of the transport rate is about 1 m$^2$/s for regions with large internal tides).

The last, but not the least interesting, aspect of strong internal solitons in coastal zones is their environmental effects. Examples can be found in (DUDA & FARMER, 1999). In particular, STANTON, (1999) observed a significant increase in turbulent mixing in solitons, whereas LENNERT-CODY & FRANKS (1999) observed strong phytoplankton luminescent activity in solitons.

4.2 Internal Waves in the Deep Ocean

As a broader view of internal soliton activity on the continental shelves was obtained, the interest of researchers expanded to their study in the deep ocean. Various questions were asked: Do solitons exist in the deep sea? What are their typical and extreme parameters? Where are they encountered most frequently? Over what distances can they propagate? What fraction of the total internal wave field energy do they contain? Why are they so coherent as to be recognizable over a large span of space and time scales? What relation do they have to deep-ocean internal waves, whose spectra, as described by GARRETT & MUNK (1975), imply that very little coherence exists, and that energy levels are approximately the same around the globe? We now know that solitary internal waves occur in a wide range of deep-ocean locales, in regions at least as far as 500 km from shore, and perhaps even farther.

Deep-water solitons are typically formed during strong tidal flow over relatively shallow underwater sills that protrude up into the permanent thermocline (e.g. 300 – 500 m); they then radiate away from their sources in narrow straits. The Strait of Gibraltar has been the most thoroughly studied case (FARMER & ARMU, 1988). Figure 24 is a photograph taken from the US Space Shuttle showing three packets of solitons propagating into the
western Mediterranean Sea following their generation at the Gibraltar sill by the combined tidal, surface, and subsurface flows of Mediterranean water out into the Atlantic Ocean (La Violette et al., 1986). The amplitude of these waves is of order 50 m and their inter-soliton distance, or wavelength, is 500 to 2000 m. It should be emphasized that these internal waves may have significantly different characteristics from the continental shelf waves discussed above. The solitons can have larger amplitudes in deep water and much larger scales both across and along their fronts. At the same time, due to a deeper pycnocline position and smoother stratification, the effective nonlinearity in such solitons may remain relatively small, whereas in the shelf zone it can reach significant values as discussed above.

Even larger waves have been detected in the areas of the Guiana Basin in Western Atlantic (Kuznetsov et al., 1984) and Mascaren Ridge in Indian Ocean (Konyaev & Sabinin, 1992). There, soliton-like (but still evolving) depressions of up to 85 – 90 m have been reported to propagate towards the open ocean with velocities of about 1.7–2.5 m/s. A group of solitary waves having an inter-soliton spacing of 20 km and a maximum crest length of over 90 km at a distance of 400 km from their source has been observed with the Landsat Multi-Spectral Scanner in the Sulu Sea in the Philippines, where they radiate from a narrow sill approximately 2 km in width (apel et al., 1985; Liu et al., 1985). In these South East archipelagios, a combination of complex geography, inland seas with differing tidal responses, and a deep pycnocline formed by the stress of the trade winds all work to cause the soliton populations in the region to be extraordinarily dynamic. These large, deep-water solitons have been studied in some detail. Several measurements in the Andaman Sea (Osborne & Burch, 1980) and Sulu Sea (apel et al., 1985; Liu et al., 1985) have shown that solitons exist with amplitudes of up to 70 – 90 m and phase speeds approaching 2.5 m/s, caused by strong tidal flow over underwater sills between islands. Although the unperturbed pycnocline depth is of order 125 to 150 m in that region of the ocean, waves having amplitudes of the order of the pycnocline depth can be generated by the lee-wave process. Such large amplitudes are accompanied by an appreciable nonlinear increase in phase speed, according to Eqs. (23). The observations of solitons with amplitudes up to 60 m were made by Pinkel (1999) during the TOGA-COARE experiment in the Western Pacific. The champion internal solitary wave (ISW) of 120 m amplitude has been apparently observed in the Strait of Gibraltar by H. Bryden [unpublished Cruise Report, April, 1998. See also (Sabinin et al., 2004)].

Detailed observations and analysis of solitary internal wave dynamics in the Sulu Sea were reported by Apel et al. (1985) and Liu et al. (1985). Some 14 days of measurements were made with current meters and thermistors at three locales along the direction of wave propagation, which had been determined earlier using satellite imagery (Fig. 25).

Shipborne instruments, including radar, optical, and acoustic sensors, were also used to follow wave packets from their birth at a sill at the southern entrance to the Sulu Sea to their decay approximately 400 km across the sea. A tidal generation process was established during which (1) the tide initially produces a complex hydrodynamic perturbation at the sill; (2) the disturbance gradually becomes steeper as it propagates into deep water; and (3) the perturbation then forms into a solibore, i.e. an undulatory bore that becomes fully
modulated into a group of several solitons by the time the packet reaches a distance of 200 km from the source region. At that distance, the modal differences in velocity of propagation [see Eq. (9)] have caused a separation of the internal waves by mode to have taken place, with only mode $n = 1$ being observed at larger distances (although higher-order modes are detected closer to the source). Measurements of density, currents, and shear flow profiles
Figure 25: (a) Internal solitons in the Sulu Sea in the Philippines’ area, as observed with the Defense Meteorological Satellite, and seen in reflected sunlight. Borneo is at lower left. Visible are 5 packets of solitons generated by tidal action in the Sibutu Passage. (b) Schematic of phase fronts in frame (a) as observed on two consecutive days in 1979. Current meter moorings were emplaced at SS-1, SS-2, and SS-3 during 1981. From (APEL ET AL., 1985).

allowed the evaluation of the coefficients of the solitary wave equations, Eqs. (22), (23) or the like.

To describe these waves, LIU ET AL. (1985) used a modified form of the JKKD equation that is similar to Eq. (45), but with modifications for cylindrical spreading and Reynolds-type dissipation [as in Eq. (49)]. They calculated the evolution of a wave packet over distances from 90 to 200 km using as the initial data the registered wave perturbation at the point located 90 km away from the source. Their results are shown in Fig. 26 for a time interval of 12.5 h. To fit numerical and observational data for the 200-km site, the authors used an empirical value of the coefficient of horizontal eddy viscosity $A_h = 10 \text{ m}^2/\text{s}$ and obtained fairly good agreement between the theoretical/numerical and experimental results.

The eddy viscosity ansatz parameterizes the interaction of internal waves with turbulence, as well as the radiation of internal wave energy to the surface wave field, which is the source of their surface signatures (see below). Figure 26 clearly shows how the soliton packet evolves out of the initial disturbance into a rank-ordered train of pulses that decay in both amplitude and wavelength toward the rear of the group. Figure 27 shows a comparison of the phase speeds of 18 solitons in the Sulu Sea with the theoretical JKKD model [cf. Eq. (48)]; the
dependency appears to be in accord with the theory. Such comparisons form quite reasonable tests of the soliton character of the waves. Numerical models [see, e.g. (LAMB, 1994)] show that in addition to the tidally generated solitons which are formed near the continental shelves and which then propagate towards shore, a well-defined internal soliton packet that propagates into deep water should also develop. Although such pairs of oppositely directed waves are sometimes seen as the result of tidal flow over sills, no reports of clearly formed solitons propagating into the deep ocean from the continental shelf break have been published to the knowledge of the authors. However, other studies suggest that the shelf-break generation process also launches an offshore-travelling disturbance that propagates down to the bottom, then reflects/refracts upward, and finally reappears near the surface a few hundred kilometers out to sea (NEW, 1988; PINGREE & NEW, 1989; PINGREE ET AL., 1986). Note also an aircraft-based lidar observation of a solibore-like group near Alaska propagating offshore (CHURNSIDE & Ols...
Figure 27: Experimental versus theoretical phase speeds for solitons in the Sulu Sea. From (APEL ET AL., 1985).

TROVSKY, 2005). If such generation actually occurs in nature, then the continental shelves can be considered to be global sources of deep-water internal waves. Indeed, the generation process is likely to be as widespread as is the combination of bathymetry, stratification, and current flow. If this is true, as it appears to be, then the entire rim and island population of an ocean basin can be sources of internal waves.

As known, deep-ocean internal waves are characterized by a more or less global energy spectrum, random phases, and rms (root-mean-square) amplitudes of roughly 5 m (GARRETT & MUNK, 1975). If the process of offshore propagation is also active wherever the generation of an onshore component occurs, then the observed upper-ocean internal wave field could originate in considerable part from the bathymetric features around its margins, and in the upper ocean (seasonal thermocline) they can be nonlinear and soliton-like. A measurement at a given point in the deep sea would show the sum of many waves arriving from a variety of points and directions. Even though the internal waves might individually be composed of solitons, the resultant summed signals would have random amplitudes, phases, and propagation directions, as is observed. Such incoherent fields may not be detectable in an image because they presumably have few patterns recognizable to the eye. However, the available high-resolution SAR images of the open ocean [e.g. APEL (1987), p. 466] suggest that the visible waves are soliton-like and more or less randomly distributed in space.
The energy in such internal waves ultimately derives from the rotational energy of the Earth–Moon–Sun system. Observations of the decay of orbital parameters, as well as measurements of solid-Earth tidal dissipation and bottom friction acting on oceanic tides in shallow seas, suggest that ocean tides by themselves are insufficient to account for the energy losses. Estimates of internal wave dissipation made using Seasat SAR data suggest that perhaps 5% to 10% of the changes in the rotational energy budget could be attributed to internal wave excitation; thus it appears that this process can provide a small but significant fraction of the missing energy (Fu & Holt, 1984). More quantitative estimates of the dissipation rates have been made by Sandström & Oakey (1995), who suggest that shear-flow instabilities may result in dissipation rates near 5.0 \cdot 10^{-4} \text{ W/m}^3. Later estimates vary in their predictions but all of them include internal waves in the balance schemes of tidal energy. It is amusing to consider that space-based observations of events occurring beneath the sea and which have subtle surface signatures should yield information relevant to celestial mechanics.

Figure 28: Map showing sites where internal solitons have been reported (courtesy of Chris Jackson).

A brief compilation of sightings of internal waves thought to be solitons is shown in Fig. 28. The signatures have been recorded by in-situ sensors and, in some cases, observed and photographed simultaneously from aircrafts or satellites. Huthnance (1989) cites a number of cases for which published documentation exists that establishes their internal wave character. However, not all of these sightings are necessarily of solitons but, perhaps, just intense internal waves. The updated information on intense internal waves registered from satellites and the map of their sightings is presented in the Internet Atlas by Jackson & Apel (2004). The global nature of the sightings demonstrates the widespread occurrence
of intense internal waves and, in particular, internal solitons and their trains – solibores. Note that the majority of data reflected on the maps concerns primarily coastal zones and adjoined parts of deep ocean areas. Apparently, this confirms an important role of shelf breaks in the formation of intense nonlinear IWs.

4.3 Surface Signatures of Internal Waves

The visibility of internal waves at the surface is due to the modification of the equilibrium surface wave vector spectrum, $\Psi_{eq}(k)$, by the subsurface currents in the IW, to form a perturbed, nonequilibrium spectrum, $\Psi(k; x, t)$. In general, this spectrum depends on a number of parameters such as currents, wind, the presence of surfactants, long surface waves, etc. As a result, the problem of surface wave modulation is not completely solved yet. Detailed discussion of this problem is beyond the scope of this paper. Below we give only a short outline of mechanisms of forming the surface signatures of internal waves.

In the majority of cases, the basic motion affecting surface waves is the horizontal velocity of the IW current near the surface. This current changes the wave number and, generally, the energy of surface waves. The commonly used description of the modulation processes is grounded in the equation for the conservation of the wave action spectrum (PHILLIPS, 1977; APEL, 1987). The wave action spectral density, $N(k, x, t)$, is wave energy density per unit surface area divided by the Doppler frequency of the wave, $\omega_D = \omega - k \cdot U$, where $U$ is the horizontal current velocity at the surface, namely,

$$N(k, x, t) = \frac{\Psi(k, x, t)}{\omega - k \cdot U},$$

(99)

where $k$ is wave vector, $x = (x, y)$ is surface coordinate, and $\Psi(k, x, t)$ is the nonequilibrium energy spectrum, which varies locally in space and time because of the advective and straining effects of the internal wave currents.

The balance equation for action spectral density can be derived from the Lagrangian description or from the energy density balance equation for surface waves (PHILLIPS, 1977). In the “relaxation time” approximation, it states that the action spectrum changes along characteristics in $(k, x)$ space according to

$$\frac{dN(k, x, t)}{dt} = \frac{\partial N}{\partial t} + \nabla \cdot k N \cdot \frac{dk}{dt} + \nabla_x N \cdot \frac{dx}{dt} = - \frac{N(N - N_{eq})}{\tau},$$

(100)

where $N_{eq}(k)$ is the equilibrium action density, and $\tau$ is the surface wave relaxation time, a phenomenological measure of how long a perturbed wave spectrum takes to relax back to its equilibrium state depending on the wind stress (HUGHES, 1978). This equation may be integrated along characteristics defined by

$$\frac{dx}{dt} = c_g + U,$$

(101)

$$\frac{dk}{dt} = -(k \cdot \nabla_x) U$$

(102)
\( \mathbf{c}_g \) is the wave group velocity vector.

**THOMPSON ET AL. (1988)** give the solution for \( N(\mathbf{k}, \mathbf{x}, t) \) in terms of a perturbation, \( P(\mathbf{k}, \mathbf{x}, t) \), as

\[
P(\mathbf{k}, \mathbf{x}, t) = N_{eq} \int_{-\infty}^{t} \sum_{i,j=1}^{2} k_i \frac{\partial U_{ij}}{\partial \mathbf{x}_i} \frac{\partial}{\partial k_i} \left( \frac{1}{N_{eq}} \right) \exp \left[ - \int_{-\tau}^{t} \frac{dt''}{\tau(k'' \tau)} \right] dt'',
\]

(103)

where the perturbed action density spectrum is

\[
N(\mathbf{k}, \mathbf{x}, t) = \frac{N_{eq}}{1 + P(\mathbf{k}, \mathbf{x}, t)}.
\]

Equation (103) is an one-dimensional temporal integral along the paths defined by Eqs. (101), (102) and defines the distortion of the spectrum as the surface waves are affected by the internal wave field. See (HUGHES, 1978) or (THOMPSON ET AL., 1988) for a more complete treatment of the effects described here.

The signatures of strong internal waves can be considerably enhanced by the effect of “group synchronism”, when the phase velocity of internal wave is close to the group velocity of the surface waves responsible for the formation of radar and optical images (BASOVICH ET AL., 1984, 1986). This is because the phase speeds of the internal wave solitons are typically dozens of cm/s (sometimes over 1 m/s), with the corresponding wavelengths for surface waves lying between several decimeters and few meters. These waves, in turn, can affect the shorter gravity-capillary waves (cascade modulation).

It should be noted that the analysis of the “space-time rays”, Eqs. (101), (102) can often be an informative tool for understanding the character of wave modulation. In an important case of a plane IW when \( U_x = U(x - Vt) \), in the reference frame moving with the wave speed \( V \), the wave packet frequency is \( \Omega = \omega - k_x V \) (\( \omega \) is the wave frequency in the bottom reference frame), and the corresponding dispersion equation reads as

\[
\Omega = k_x(U - V) \pm \sqrt{g(k_x^2 + k_y^2) + \sigma(k_x^2 + k_y^2)^{3/2}},
\]

(105)

where \( \sigma \) is the surface tension coefficient. This relation is essentially the first integral of the system (101),(102). It allows to construct trajectories of wave packets and find variation of their wavelengths. Note that these trajectories can be closed if the packet is trapped at a given IW period due to the group synchronism [see, e.g. (BASOVICH ET AL., 1984, 1986)].

To obtain more specific results, the function \( U(x - Vt) \) must be specified. For moderate-amplitude waves, when the KdV model is applicable, from the displacement Eq. (22) the horizontal velocity on the surface can easily be found in the form

\[
u(x, z = 0, t) = \eta_0 V \frac{dW}{dz} \left|_{z=0} \right. \text{sech}^2\Phi
\]

(106)

where \( \Phi \) is the propagating phase, \( \Phi = (x - Vt)/\Delta \), and \( W(z) \) is the eigenfunction for vertical velocity. For such weakly nonlinear waves \( u/V \ll 1 \), and the solution of the action equation (100) without relaxation \( (\tau \to \infty) \) shows that the wave intensity is distributed according
to the "strain rate", \( \frac{\partial u}{\partial x} \bigg|_{z=0} \) [see, e.g. (THOMPSON ET AL., 1988)]. This strain rate causes the alternating compressive and tensile effects on the surface wave spectrum, that renders the internal waves visible on the ocean surface as roughness changes. In the converging phase of internal current, those surface waves whose group speeds are near the phase speeds of the internal waves are swept together and amplified, whereas that portion of the phase having diverging internal currents exhibits diminution of the overlying surface wave spectral content.

In order to estimate the signal contrast found in an internal wave image, it may be assumed that for small variations of the current strain rate, the relative modulation of the spectrum, \( \delta \Psi / \Psi_{eq} \), is mirrored in the relative modulation of the radar or optical cross-section per unit area of the ocean surface, \( \delta \sigma^0 / \sigma^0 \), viz:

\[
\frac{\delta \sigma^0}{\sigma^0} \approx \frac{\delta N}{N_{eq}} \approx \frac{\delta \Psi}{\Psi_{eq}}.
\]

This means that the formulae presented above can be used to estimate the image contrast, provided the fractional modulation is small. Much of the research work to date involving the analysis of internal wave properties has used these formulae.

As an example, Figure 29 shows internal solitary waves in the New York Bight via (a) their surface signatures in a SAR image, (b) \textit{in situ} current meter measurements made simultaneously with the SAR image, and (c) theoretical cross-sectional modulations computed using the formalism mentioned above (GASPAROVIC ET AL., 1988). The agreement between observation and theory is quite good, as may be seen by comparing the solid and dotted lines in Fig. 29. This figure was also used to evaluate the phase shifts shown in Fig. 19c.

For strongly nonlinear waves, however, Eq. (106) is inapplicable, and instead one can use the calculations outlined above for strong solitons. For such solitons BAKHANOV & OSTROVSKY (2002) used Eq. (106) as a wave shape approximation but with parameters \( \Delta \) and \( V \) characteristic of strong solitons. As a result they have shown that, instead of the strain rate maxima, the decrease of surface wave intensity (sicks) can shift towards the soliton peak (maximal depression) that is in agreement with the COPE observations mentioned above. Indeed, in (KROPFFLI ET AL., 1999), the variations of scattering intensity and isotherm displacements were plotted together to show that for the strong solitons, the scattering minima are from the soliton peaks (Fig. 29). Note that in the same paper it was shown that a similar effect can be produced due to surfactant modulation by the IW current.

This results in alternating regions of enhanced and diminished surface wave spectral density, regions that are rougher and smoother than the average. Electromagnetic radiation incident on the surface is thus scattered differentially by the rough and smooth portions, and an image constructed from such scattered radiation – say a photograph made in visible light or a radar image – will map the roughness variations. This map mainly mirrors the underlying internal currents.

Even more complex situation occurs for short gravity-capillary waves when the wind relaxation cannot be neglected. In this case the IW modulates air flow stress in the lower atmospheric boundary layer which, in turn, affects wave amplification by wind. Whereas the
action of long surface waves on the low atmospheric boundary layer has been a subject of a number of papers, the effect of internal solitons wavetrain on short surface waves via the wind perturbations has been considered in comparatively few publications (Gorshkov et al., 2003).

Finally, internal waves can affect generation of ripples at the crests of longer gravity waves (“parasitic ripples”). This process is important when the gravity wave amplitude is close to its breaking value where the wave crest has a large curvature (Longuet-Higgins, 1995). Due to the sensitivity of ripple generation to the curvature, even for a slight modulation of the “primary” long wave by the IW current, the ripple amplitude can change radically. This version of cascade modulation (IW → longer gravity wave → ripples) was observed in laboratory by Ermakov & Salashin (1994) and theoretically described by Charnotskii et al. (2002).

5 Effects of Non-Linear Internal Waves on Sound Waves in the Ocean

From the point of view of ocean acoustics, nonlinear internal waves are important scatterers of sound. This scattering is highly frequency dependent, unsurprising given the high degree of spatial structure of the solitons. We first define what frequency bands are of interest. Based
on experience, we operationally consider three frequency bands: high frequency \( f \geq 50 \) kHz, mid frequency \( 1 \) kHz \( \leq f \leq 50 \) kHz, and low frequency \( f \leq 1 \) kHz.

In the high frequency band, medium attenuation limits sound propagation to ranges on the order of a few meters to a kilometer. Thus for sonars at high frequencies, the monostatic backscatter geometry is found to be the most useful. Using high frequency sound also allows one to image an ensonified object with good resolution. In the past few decades, high frequency acoustic scattering has been used successfully for imaging the detailed structure of nonlinear internal waves. Such images have provided information on various parameters of the internal waves, including: snapshots of the detailed spatial structure of individual solitons; the space-time evolution of solitons, including their generation, propagation, dispersion, broadening by bathymetry changes, inversion at the so-called “critical depth”; and the dissipation of solitons via turbulent processes and wave breaking.

High frequency acoustic images also give views of how material such sediment and biota are carried by the waves, which is important since these waves can provide non-zero mean transport of a “tracer” due to their non-linearity. In terms of the acoustics problem for high frequency scattering by non-linear internal waves, there are still a number of loose ends. Specifically, there is a well known, yet continually nagging problem of determining which “tracer” of the internal wave has scattered the acoustic signal. Is the scattering due to sound speed or density structure, biota, bubbles, or sediment? These tracers all highlight different parts of the internal wave, and if one is doing imaging, it is obviously necessary to know what one is imaging. Moreover, one is also often interested in the tracers themselves; for example, one might like to know what type and size of biota are being carried along by the wave. The literature on high frequency scattering, while reasonably developed, is not yet overly extensive. We would refer the reader to the work of Orr et al. (2000), Warren et al. (2003), Farmer & Armi (1999), and Moum et al. (2003) as representative samples of this literature, from which one can find references to other work in the field.

Scattering of sound from internal waves at medium frequencies is, at this point in time, an under-developed area of research. The Strait of Gibraltar tomography study by Tiemann et al. (2001a, b) is perhaps the most detailed look at scattering of mid-frequency sound by internal soliton trains reported in the literature. Using frequency modulated sweeps from 1136 to 3409 Hz, the two transmission paths considered, of 14.6 and 20.1 km length, provided a “raypath-averaged” view of the solitons, from which the researchers were able to understand how tidal cycle variability affected both the soliton wave field and also its effect on the acoustics. The effects of solitons on medium frequency acoustics have also been seen in “acoustic navigation nets” working at 8 – 13 kHz, and Headrick & Lynch (2000a, b) report significant travel time fluctuations of short paths (hundreds of meters to a kilometer) due to nonlinear internal waves. Henyey & Ewart and their collaborators [see (Williams et al., 2001)] looked at kilometer scale transmissions in shallow water at mid-frequencies, only using moored towers to constrain observations of the energy to the water column.

Past these studies, the literature is rather sparse. This paucity of results should not continue indefinitely, however, as there are good reasons to look at the scattering of mid-frequency sound by internal solitary waves. For instance, if one is interested in the turbulence
generated both in a nonlinear internal wave and in its wake, which could have scales from centimeters to tens of meters, looking for resonance Bragg scattering from mid-frequency acoustics, which has the same span of wavelengths as the turbulence, might be a viable method. Short range tomography at mid-frequencies, which has been explored by YAMAOKA ET AL. (2002), is also an interesting possibility. Experiments which will look at mid-frequency propagation and scattering in the midst of a strong internal wave field are being planned at present (J. LYNCH AND D. J. TANG, private communication), so that it is probably just a matter of time before the literature in this area will expand.

By far the largest amount of research on the scattering of sound by internal solitary waves has been done in the low frequency regime, and in shallow water, where we operationally define shallow water to be the region from the tidal mixing front (≈ 30 m) to the continental shelf break (≈ 200 m). In shallow water, there is a well known “optimal” frequency of propagation, on the order of a few hundred Hertz, at which one sees a minimum in propagation loss (JENSEN ET AL., 1994). This optimal transmission characteristic makes low frequency an ideal band for shallow water sonar systems. Since the continental shelves are also the home to a plethora of nonlinear internal waves, it becomes inevitable that the interaction between the sound waves and ocean internal waves is strongly observed.

When examining, the interaction of sound with the coastal soliton field, there are a number of different issues to consider. First, we note that we must treat both the amplitude and phase of a scattered signal. This can to some extent be done separately, as the scattering characteristics for these two basic quantities is often independent, at least to first order. Also, in treating these variables, we further note that the most common acoustic measurements are of intensity and pulse travel time fluctuations, two secondary quantities, rather than of the amplitude and phase directly. Another important consideration in looking at the acoustic scattering by internal waves is that the acoustic scattering is very different for source-to-receiver geometries which go across the wavefronts of the internal wave packets as opposed to those geometries which are along the IW wavefronts. We will examine this next, in the context of acoustic normal mode theory, which is a natural and physically insightful descriptor for low frequency, shallow water sound. We will use 2-D range dependent mode theory here for notational simplicity, noting that a fully 3-D treatment of the acoustic field is needed for some of the effects we will discuss.

When dealing with a range dependent ocean acoustic waveguide, the Helmholtz equation is non-separable; however, a variant of the usual separation of variables technique, called “partial separation of variables”, can be employed. Specifically, the range dependent Helmholtz equation is:

\[ \nabla^2 \psi(r, z) + k^2(r, z) \psi(r, z) = 0, \]  

(108)

where \( z \) is the vertical coordinate, \( r \) is the horizontal range, \( \psi(r, z) \) is the range dependent normal mode field, and \( k(r, z) = \omega/c(z, r) \) is the total wavenumber, which carries within it the description of the range-dependent sound speed of the ocean.

For this case, we stipulate the partially separable (modal) solution:
\[ \psi(r, z) = \sum_n R_n(r) \phi_n(r, z). \]  

Inserting this solution into Eq. (108) results in a “local normal mode” equation

\[
\left[ \frac{\partial^2}{\partial z^2} + k^2(r, z) - k_n^2 \right] \phi_n(r, z) = 0
\]

(110)

and a set of coupled equations for the radial part of the solution

\[
R''_m(r) + \frac{1}{r} R'_m(r) + k^2_m(r) R_m(r) = - \sum_n \left[ A_{mn} R_n + B_{mn} \left( \frac{R_n}{r} + 2 R'_n \right) \right],
\]

(111)

where the prime signifies a range derivative and

\[
A_{mn} = \int_0^\infty \rho(z) \phi_m(z, r) \phi'_n(z, r) \, dz \quad \text{and} \quad B_{mn} = \int_0^\infty \rho(z) \varphi_m(z, r) \varphi'_n(z, r) \, dz
\]

(112)

are the mode coupling coefficients.

These “coupled mode equations” are well known in ocean acoustics, and so we will just refer the reader to some of the standard texts if more information is desired about them (JENSEN ET AL., 1994; KATZNELSON & PETNIKOV, 2002). The solutions to these equations are usually generated numerically, via codes like the well-known KRAKEN code (PORTER, 1991).

It is worth noting the details of the weakly range dependent solution to the coupled mode equations, obtained by setting the \( A_{mn} \) and \( B_{mn} \) terms of Eq. (111) equal to zero. In this limit, the so-called “adiabatic mode” solution for pressure is as follows:

\[
p(z, r) \sim \psi(z, r) = C \sum_m \frac{\phi'_m(z_r) \phi_m(z_r)}{R \int_0^R k_m(r) \, dr} \exp \left\{ i \int_0^R k_m(r) \, dr \right\} \exp \left\{ - \int_0^R \alpha_m(r) \, dr \right\}.
\]

(113)

This particular modal solution is germane to propagation along the internal wavefront, where the range dependence of the medium parameters is slow, so that the coupling coefficients remain small. The adiabatic solution clearly shows how the variability of the ocean medium (in our case the internal waves) produces amplitude and phase (and thus travel time) fluctuations. For amplitude, the passage of the internal waves over the source and receiver positions produces a distortion of the normal modes at those positions, which thus changes the received pressure \( p(z, r) \). For phase, \( k_m(r) \) varies over the source to receiver range, which changes the phase, \( \phi_m = \int_0^R k_m(r) \, dr \). Such adiabatic amplitude and phase fluctuation effects are well known, and examined in detail for the case of surface gravity waves.
For the coupled mode (across IW crest) propagation case, we get both the adiabatic effects noted above (since the adiabatic solution is, in essence, the lowest order coupled mode solution) as well as some additional effects. In the case of strong range dependence (i.e. large $\partial c_s(z, r)/\partial r \to \text{large}$, where $c_s$ is the sound speed in the water column and “large” means of the order of 10 m/s per km or more, the same order as the deep ocean vertical sound speed change), the normal modes exchange energy along the source-to-receiver track. This leads to modal arrivals at a distant receiver which have shared characteristics of the modes along the path. As an example, consider the travel time of a mode one reception at a receiver (assuming we can filter the signal so as to identify modes) in a waveguide that only supports two modes. For the case of no coupling, mode one arrives at its expected travel time, i.e. $t_1 = R/v_1^G$. However, with coupling, the mode one arrival can start out in mode two, travel to a point $x$ where it encounters a soliton and couples into mode one, and then continues propagating to the receiver as that mode. This “coupled mode one arrival” has the arrival time $t_{1p}^p = x/v_2^G + (R-x)/v_1^G$, which is intermediate between the uncoupled mode one and mode two arrival times. It is easy to see that if there are “scatterers” densely distributed along the source-to-receiver (S/R) path, one will see a mode one arrival which is spread out in time between the usual mode one and mode two arrival times. This “time spreading” is a well known phenomenon in shallow water pulse propagation, where the coupled mode arrival spread between the arrival times of the fastest and slowest uncoupled modes.

Turning to soliton induced coupled mode effects on amplitude, it is found that the biggest contribution to amplitude fluctuation is caused by the difference in medium attenuation for different modes (or “differential attenuation”). In general (though exceptions can be found), the low order trapped acoustic modes attenuate slowly, whereas the higher order trapped modes (and certainly the continuum modes) attenuate far more quickly, due to enhanced boundary interaction. Thus, if a low mode couples to a high mode, more propagation loss is seen – the opposite is true for a high mode coupling to a low mode.

Let us look a little further at phase and travel time scattering effects. “Pulse wander”, which is the variation in arrival time of a pulse with no change in the shape of the pulse, is mathematically the frequency derivative of the phase (fluctuation) integral shown in Eq. (113). It has been shown by Lynch et al. (1996) that pulse wander effects are significantly larger for along IW wavefront propagation than for across wavefront, an effect which is readily understood by examining the phase integral in Eq. (113). The integrand of that integral oscillates quickly and largely cancels for across IW wavefront propagation, whereas it is relatively constant along an IW wavefront. The wander also shows a distinct mode number dependence, which is associated with where the acoustic mode vertical turning points are located in relation to the maximum amplitude points of the internal wave modes (which are usually dominated by mode one, as previously discussed.) Wander effects for path lengths of 25 – 50 km at frequencies of 100 – 500 Hz tend to be of order 10 msec (or less) along IW wavefronts, and about 1 msec (or less) across the wavefronts.

The travel time spread, which is caused by the mode coupling due to across IW wavefront propagation, shows an interesting effect which has been dubbed “near receiver dominance”. Specifically, when an internal wave or packet of internal waves is between an acoustic source
and receiver, and moreover is close to the receiver, then the time spread seen is a maximum. The explanation for this is seen by taking the limit \( x \to R \) in our previous two mode example. In this limit, the arrival time difference between the “undisturbed mode one” and the “coupled mode one” is a maximum. This spreading effect was clearly seen in the 1995 SWARM experiment (APEL ET AL., 1997) by HEADRICK ET AL. (2000a,b), who showed an M2 tidal signal in the spread of pulsed signals due to the passage of the nonlinear internal tide by the receiver.

We next turn to the acoustic amplitude scattering effects of the internal waves. We will first look at the across IW wavefront propagation geometry, simply because that was the geometry that was first examined experimentally and theoretically, and is better understood at this point in time. Undoubtedly the best known shallow water acoustics experiments on sound scattering by internal waves are the Yellow Sea series of experiments reported by ZHOU ET AL. (1991). In these experiments, ZHOU ET AL. (1991) reported seeing anomalously high propagation losses versus frequency, up to 30 – 40 dB, a huge amount. These anomalous losses were attributed to resonant Bragg scattering from a strong internal wave train with evenly spaced internal wave solitons. (This is not the usual soliton wavetrain one sees, but it is what exists in the Yellow Sea.) Resonant scattering occurs when projection of the IW wavelength along the acoustic path, \( \Lambda_{IW} \), is equal to the acoustic mode interference distance (commonly called the “mode cycle distance”), \( \Delta_{mn} \), where \( \Delta_{mn} = 2\pi/(k_m - k_n) \).

The internal wave train acts as a 2D Bragg crystal lattice, which give both frequency and azimuth dependence to the scattering field. At resonance, the predominantly lower mode energy created at the source was transferred to higher modes, which then attenuated more quickly, thus greatly increasing the total energy losses reported. (We again note that the opposite effect from this can happen, depending upon the depth of the source relative to the stratification of the ocean. If the source is in warmer, higher sound speed water, higher acoustic modes are preferentially excited at the source. IW coupling then transfers much of their energy to the low modes, which decay more slowly, producing less loss.) This resonance mechanism is a robust and easily understandable one. Moreover, it can be generalized to non-regularly spaced internal wave trains, simply by examining the spatial wavenumber spectrum of a soliton wave train, and then matching these wavenumber components to the acoustic mode cycle distances. PREISIG & DUDA (1997) [see also (DUDA & PREISIG, 1999)] have shown, via numerical simulations, the details of a second mechanism for strong scattering, specifically the resonance of pairs of acoustic modes with the widths of the individual solitons. This mechanism is universal and is, in fact, a more detailed version of the Bragg mechanism just discussed. Since strong IW scattering mechanisms exist in nature, substantial across-wavefront, coupled mode acoustic scattering is the rule, rather than the exception.

Next we come to propagation along IW wavefront. This topic has been explored mostly theoretically, though recent experimental data has confirmed some of the theory. An interesting geometry which recently has been considered is the “along IW wavefront” geometry shown in Fig. 30. This creates “horizontally ducted propagation” of acoustic normal modes between internal wave solitons. Two neighboring solitons in a wavetrain create high sound-speed regions, with a comparatively low speed region in between. This creates a horizontal
duct (in the $x$-$y$ plane) for each vertical acoustic normal mode. The detailed scattering by this duct was predicted via the theory by Katznelson & Pereselkov (2000), by computer modelling by Finette et al. and recently has seen a striking confirmation in the SWARM experimental data taken by Badiey et al. (2002, 2005). This ducting effect is a strong one, easily producing 6 – 8 dB level increases in low frequency, broadband transmissions. Moreover, since this is a fully three-dimensional effect, the focusing of energy gives amplitude fluctuations over and above the 5.6 dB that one expects from saturated normal mode multipath interference in two dimensions (in the $x$-$z$ plane). Additionally, the description of such propagation requires fully 3-D theory and numerics, and goes beyond the scope of the simple 2-D equations presented previously.

![Diagram of ducting sound between internal waves](image)

Figure 30: Ducting of sound between internal waves.

There are two additional “along-internal-wavefront” scattering effects that have been theoretically predicted by Pierce & Lynch (2003) and Colosi et al. (2004), though not yet unambiguously observed. They are the so-called “Lloyd’s mirror” (Katznelson & Petnikov, 2002) and “whispering gallery mode” effects. In the first effect, one sees the interference between a direct arrival and a totally internally reflected arrival which has been glancingly reflected off an internal wave or the leading edge of a packet of internal waves. The second effect, the whispering gallery effect, is somewhat more complicated. To begin with, sound propagating along-shelf is being refracted seaward by seaward sloping bathymetry. Then internal waves just seaward of the acoustic paths reflect the acoustic energy back in the shoreward direction, given small grazing angle incidence. This results in the sound being trapped between the slope pushing it seaward and the IW’s reflecting it shoreward. It is the ocean acoustic analog of a whispering gallery except that the curved rays act like the curved walls of a whispering gallery, whereas the IW’s act as the wall reflector. This effect may have in fact been observed (though from a front, not internal waves) via the enhanced noise level at a receiver in the SWARM experiment; however, this evidence is preliminary at best, and detailed experiments are needed to verify this.

To conclude, we note that further oceanographic measurements of nonlinear internal
waves are needed to better understand the scattering of acoustic signals. To begin with, the fully 3D structure and evolution of internal wavetrains in range variable bathymetry and hydrology needs to be understood, as this defines the basic scattering entity. Second, the turbulence generated by the nonlinear internal waves needs to be better measured, as it also promises to be important to mid-frequency acoustic scattering. Finally, the relative strength of the coastal nonlinear versus linear internal wave field needs to be described, as both species of waves can be important to acoustic scattering (SPERRY ET AL., 2003).

6 Concluding Remarks

In this review paper we tried to outline different aspects of ISWs including theoretical models and field data, and the ISW action on acoustic wave propagation. Natural observations confirm that intense ISWs or their trains (solibores) do exist both in shallow and deep ocean areas, and in many cases their parameters are close to those predicted theoretically. There is not only academic interest to ISWs: they are able to provide strong vertical mixing, transport particles, affect turbulence and biological life, and even possibly interfere with underwater navigation. They affect surface waves, thus creating surface “slicks” that are visible by optical devices and radars and sometimes by naked eye. They influence propagation of acoustic signals in water and may form specific conditions for ducting sound waves.

Intensive theoretical and observational studies over the last decade have confirmed that ISWs are a widespread phenomenon throughout the oceans, especially in coastal zones where they are generated by barotropic tides. They probably absorb a noticeable part of the total tidal energy. Internal solitons are, so to say, the “extremes” of the internal wave spectrum, their magnitudes possibly reaching many dozens of meters. They are probably may be at least partially responsible for the fact that the IW spectrum in upper ocean differs from the Garrett–Munk spectrum characteristic of deeper water layers (the main thermocline area).

In spite of an impressive recent progress in studies of ISWs, some important questions remain unanswered or at least not completely clear. One of them is how the ISW are generated in the ocean. The main source of oceanic ISWs is apparently internal tides. However, the relative roles of specific mechanisms in their formation (lee waves, scattering of barotropic tides by bottom features), as well as the roles of shear instability, the Earth’s rotation, etc., that are discussed in literature, are not clear in many cases. Also the energy sinks for solitons are not always known.

Another problem is the ISWs propagation along inhomogeneous paths, e.g. onshore propagation from the shelf break zone. There exist a number of theoretical models for these processes, especially for weakly nonlinear waves described by the KdV or KP equations with slowly varying parameters. Experimentally, however, most in situ observations cover one or, in few cases, two observation points, with little knowledge of what occurs at other points. On the other hand, remote observations from aircraft and satellites can provide a panoramic pattern of the surface “slicks,” but without detailed in situ data. Some exceptions have been described above.

Other questions relate to statistical properties of solitons: where are most active zones
of soliton generation located, what are the typical and extreme parameters of solitons in different areas, and, in cases when solitons are randomly distributed, what are their statistical and spectral characteristics. We already mentioned the Atlas of solitons available on the Internet, but it only describes several typical areas; perhaps a more comprehensive atlas could be created in the form of a map like that shown in Fig. 28 but much more detailed.

Very close to the aforementioned statistical problems is the problem of the sudden appearance of internal waves of giant amplitude (so called “freak” or “rogue” waves). In recent years this problem was intensively attacked by many researchers in reference to surface waves, due to their practical importance to navigation and coastal engineering. Essential progress in this field has been achieved in understanding of the nature of freak waves, their statistical features, possible mechanisms of generation, etc. [see, e.g., (KHARIF & PELINOVSKY, 2003; KURKIN & PELINOVSKY, 2004) and references cited therein]. A similar problem may be topical for internal waves, too, as they can pose a danger for submarines, oil and gas platforms, pipelines and other engineering constructions in the coastal zones. The problem of freak internal wave studies has only recently attracted the attention of researchers, and only the first theoretical steps had been undertaken in this direction (KURKIN & PELINOVSKY, 2004). Based on the working definition of freak waves as waves whose amplitudes exceed the average background by more than 2 – 2.2 times, many observed solitary IWs can be referred as freak waves. We should note that the long-wave models considered in this review seem to be more justified for describing freak internal waves than the corresponding shallow-water models used for modeling freak surface waves as the latter usually appear in the deep ocean rather than in shallow seas.

And certainly the numerous possible mechanisms of ISW action with the electromagnetic and acoustic fields must be better understood: indeed, the interpretation of remote sensing data crucially depends on the corresponding models.

Finally, the action of ISWs on other processes such as, e.g. biological life in upper ocean and shallow seas, has not been sufficiently addressed yet; this is a very promising and practically important area of oceanography.

Nonetheless, from we already know, we can arguably state that ISWs are the most regularly and clearly observed kind of solitons in natural conditions! One can foresee much further progress in this area of theoretical and experimental oceanography.

**Acknowledgment.** The authors are indebted to A. Newhall for his enormous help with the manuscript preparation for publication. The authors are also thankful to E. Pelinovsky and V. Shriira for useful advice and discussions of some aspects of this review.

### 7 References


temporal dependence of sound propagation due to shallow water internal waves, IEEE

Measurement and modeling of 3-D sound intensity variations due to shallow water


18. Basovich, A. Ya., V. V. Bakhanov, and V. I. Talanov, 1984, The influence of intense
internal waves on wind-wave spectra, in Nonlinear and Turbulent Processes in Physics,
Proc. 2nd Int. Workshop on Nonlinear and Turbulent Processes in Physics, Kiev,

19. Basovich, A. Ya., V. V. Bakhanov, and V. I. Talanov, 1987, Transformation of wind-
driven wave spectra by short internal wave trains, Izv. AN SSSR. Fizika Atmosfery i
Okeana, 23, n. 7, 694–706 (in Russian). (Engl. transl.: Izvestiya. Atmospheric and
Oceanic Physics, 23, n. 7, 520–528.)

nonlinear fluctuating media without dispersion, ZhETF, 94, n. 1, 175–185 (in Russian).
(Engl. transl.: Sov. Phys. JETP, 67, n. 1, 98–103.)

21. Benilov, E. S., 1992, On surface waves in a shallow channel with an uneven bottom,


Academic Press.


25. Bouchut, F., J. Le Sommer, and V. Zeitlin, 2004, Frontal geostrophic adjustment and
nonlinear wave phenomena in one-dimensional rotating shallow water. Part 2. High-
resolution numerical simulations, J. Fluid Mech., 514, 35–63.

lating waves of the rotation-modified Korteweg–de Vries (“Ostrovsky”) equation, Wave
Motion, 35, 141–155.

27. Boyd, T. J., D. S. Luther, R. A. Knox, and M. C. Hendershott, 1993, High-frequency
internal waves in the strongly sheared currents of the upper equatorial Pacific: Obser-

88


95


96

137. La Violette, P. E., T. H. Kinder, and D. W. Green, III, 1986, Measurements of Internal Waves in the Strait of Gibraltar Using a Shore-Based Radar, Report 118, Naval Ocean Research and Development Activity, Stennis Space Center, MS.


166. Miyata, M., 2000, A note on broad narrow solitary waves, IPRC Report 00-01, SOEST, University of Hawaii, Honolulu, 00-05, 47 p.


98


212. Porter, M. B., 1991, The KRAKEN normal mode program, Report SM-245, the SACLANT Undersea Research Centre, La Spezia, Italy.


236. Shira, V. I., V. V. Voronovich, and I. A. Sazonov, 2000, Wave breaking due to internal wave–shear flow resonance over a sloping bottom, J. Fluid Mech., 425, 187–211.


Nonlinear internal waves in the ocean are discussed (a) from the standpoint of soliton theory and (b) from the viewpoint of experimental measurements. First, theoretical models for internal solitary waves in the ocean are briefly described. Various nonlinear analytical solutions are treated, commencing with the well-known Boussinesq and Korteweg-deVries equations. Then certain generalizations are considered, including effects of cubic nonlinearity, Earth’s rotation, cylindrical divergence, dissipation, shear flows, and others. Recent theoretical models for strongly nonlinear internal waves are outlined. Second, examples of experimental evidence for the existence of solitons in the upper ocean are presented; the data include radar and optical images and in situ measurements of waveforms, propagation speeds, and dispersion characteristics. Third, and finally, action of internal solitons on sound wave propagation is discussed. This review paper is intended for researchers from diverse backgrounds, including acousticians, who may not be familiar in detail with soliton theory. Thus, it includes an outline of the basics of soliton theory. At the same time, recent theoretical and observational results are described which can also make this review useful for mainstream oceanographers and theoreticians.

Approved for public release; distribution unlimited.