THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A COMPARISON OF SOLUTIONS OF A BOUSSINESQ SYSTEM AND THE BENJAMIN-BONA-MAHONY EQUATION

A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
degree of
DOCTOR OF PHILOSOPHY

By
ABDULRAHMAN A. ALAZMAN
Norman, Oklahoma
2000

UMI Number: 9964763

UMI®

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

A COMPARISON OF SOLUTIONS OF A BOUSSINESQ SYSTEM AND THE BENJAMIN-BONA-MAHONY EQUATION

A DISSERTATION

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

Approved By

John P. Albert

[illegible]

Kevin A. Grasse

Marilyn Breen

Md. Musharrafuzzaman

ACKNOWLEDGEMENTS

This dissertation came after so much encouragement and constant support from so many people to whom I owe so much respect and thanks. Before I start this thesis, I would like to mention some of them.

As a matter of fact, I want to start by deeply thanking Allah (my Lord) so much for enabling me to make this small contribution to the science of mathematics. Truly, I want to express my appreciation to my family, my academic advisor at OU, my advisor at the Saudi Arabian Educational Mission, my Advisory Committee Members, and the people at King Saud University in Riyadh, Saudi Arabia.

My father Abdullateef, who passed away four years ago, was a true leader and encouraged me to study hard and get as high an education as I can. My mother is a strong base for me and she is making supplication to Allah that I become successful in my life. I would like to thank my wife and my three wonderful children Arwa, Afnan, and Norah for their continuous support to me before, during, and after this dissertation. My family was behind me from the get-go to the finish line, day and night.

I really want to send special thanks to my academic advisor who is also the chairman of my Advisory Committee, professor John Albert; without him and his family this research would not be possible. He was really there when I needed him. Also, I want to say a word of appreciation to Dr. Ali Babikir, with whose support and advice I was able to get to this point. In addition, I would like to take this opportunity to extend several thanks to my Advisory Committee Members, Professor Marilyn Breen, Professor Kevin Grasse, Professor Walter Wei, and

Professor Musharraf Zaman, for their encouragement and fruitful comments. Last but not least, I want to mention King Saud University, and to thank this great university for giving me this chance to learn, explore, and study. Finally, besides the people I have mentioned, I have left out even more to whom I also owe so much.

TABLE OF CONTENTS

Chapter I

INTRODUCTION

A variety of model equations have been derived for the propagation of long, bidirectional, weakly nonlinear surface water waves (cf. [7]). Many of them go by the name of Boussinesq equations, in honor of the researcher who in 1871 derived the system

$$\begin{aligned} N_t + V_x + \epsilon(NV)_x &= 0 \\ V_t + N_x + \epsilon VV_x - \frac{1}{3}\epsilon V_{xxt} &= 0 \end{aligned} \tag{1.1}$$

to describe two-dimensional water waves in a channel with a flat bottom. In (1.1), $N = N(x,t)$ represents the elevation of the free surface above the channel bottom and $V = V(x,t)$ represents the horizontal velocity in the layer of fluid whose height above the bottom is equal to $h/\sqrt{3}$, where h is the depth of the undisturbed fluid. (See Figure 1.) Here the dependent and independent variables have been nondimensionalized and scaled so that, if the motion under consideration is a wave with small amplitude a and long wavelength l relative to h, then N, V, and their derivatives will have size on the order of 1, while the small dimensionless parameter ϵ is equal to both a/h and h^2/l^2. In the formal derivation of this system, terms containing quadratic or higher powers of ϵ have been ignored.

A situation in which waves governed by (1.1) might be observed is the following. Consider a channel with rectangular cross-section, and let the x-axis be oriented in the direction along the length of the channel. Suppose a piston at one

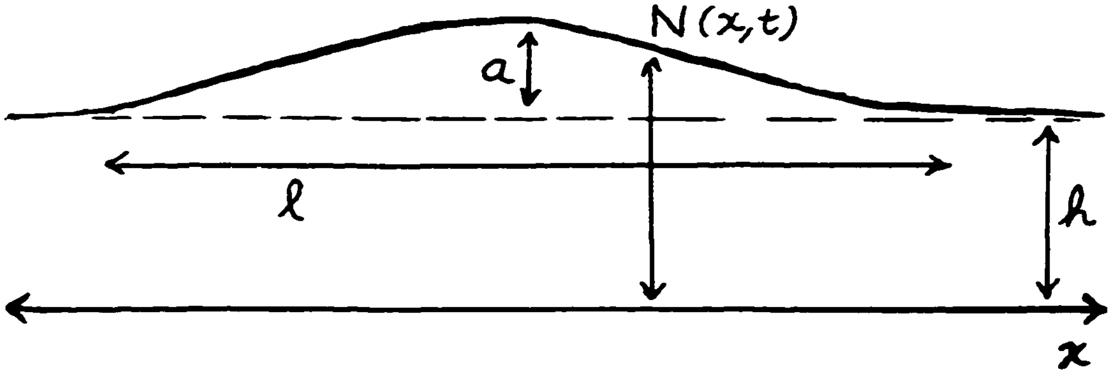

Figure 1. Physical meaning of the variables N, a, h, and l.

end of the channel moves back and forth horizontally, generating waves in such a way that the functions describing the height of the free surface and the velocity field within the water are at all times independent of the crosswise coordinate. If the flow is irrotational, and viscosity and surface tension are negligible, then in the regime of small amplitude and large wavelength described above, the waves created will be governed approximately by (1.1).

In this work, we will consider the pure initial-value problem for (1.1), in which the values of N and V are specified at time $t = 0$ for all $x \in (-\infty, \infty)$, and the values $N(x,t)$ and $V(x,t)$ are sought at all later times $t > 0$ and all $x \in (-\infty, \infty)$. Our assumptions on the initial data will imply in particular that $N(x,0)$ and $V(x,0)$ tend to zero as $x \to -\infty$ and $x \to \infty$. These assumptions would be reasonable, for instance, if one were modeling the behavior of a localized wave which propagates from one end of a channel into a quiet region located far from the ends of the channel.

Besides (1.1), there are many other formally equivalent model systems which can be obtained either by choosing different physical variables as the dependent variables (for example, one could use the velocity at the free surface rather than at the height $h/\sqrt{3}$), or by introducing terms which formally have size ϵ^2 or smaller but which change the character of the equations. One such equivalent model system is that treated in [7]. Another whole family of equivalent systems, treated in [5], are of the form

$$(1.2)\qquad \begin{aligned} &N_t + V_x + \epsilon(VN)_x + \epsilon\left(\frac{\theta^2}{2} - \frac{1}{6}\right)(\lambda V_{xxx} - (1-\lambda)N_{xxt}) = 0 \\ &V_t + N_x + \epsilon VV_x + \epsilon\left(-\frac{1}{2} + \frac{\theta^2}{2}\right)(\mu N_{xxx} - (1-\mu)V_{xxt}) = 0. \end{aligned}$$

Here N and ϵ are the same as in (1.1), while V now represents the horizontal velocity in the layer of fluid whose height above the bottom is θh, where $0 \le \theta \le 1$. The values of λ and μ can be chosen arbitrarily.

In this work, we consider the special case of (1.2) given by the choices $\theta^2 = 2/3$, $\lambda = 0$, and $\mu = 0$. Then (1.2) takes the form

$$(1.3)\qquad \begin{aligned} N_t + V_x + \epsilon(VN)_x - \frac{1}{6}\epsilon N_{xxt} = 0 \\ V_t + N_x + \epsilon VV_x - \frac{1}{6}\epsilon V_{xxt} = 0. \end{aligned}$$

Following the terminology of [5], we will refer to this system as the regularized Boussinesq system.

One reason for studying this particular form of the Boussinesq system is that it is easier to establish existence results for the initial-value problem for system (1.3) than for other systems of type (1.2). In Chapter 2 below, we prove the local existence (in time) of solutions of (1.3) for arbitrary initial data in the Sobolev space $H^1 \times H^1$. The proof requires only a standard application of the Banach Contraction-Mapping Theorem (cf. [3], where a similar result is proved for periodic initial data, and [14]). For other systems of type (1.2), it is considerably more difficult to prove local existence of solutions, and in fact many systems of type (1.2) are probably ill-posed in the sense that solutions exist only for special initial data [5].

For unidirectional waves, that is, waves which propagate either mostly to the right or mostly to the left, the system (1.3) can be formally reduced to a single equation

$$N_t + N_x + \frac{3}{2}\epsilon N N_x + \frac{1}{6}\epsilon N_{xxx} = 0, \tag{1.4}$$

called the Korteweg-de Vries equation [11]. If N satisfies (1.4) and V is given by

$$V = N - \frac{1}{4}\epsilon N^2,$$

then N and V will satisfy a system which agrees with (1.3) except for terms of quadratic and higher order in ϵ. Alternatively, since from (1.4) we see that N_t and $-N_x$ are the same up to terms of order ϵ, then by ignoring terms of order ϵ^2, we can replace (1.4) by the equation

$$N_t + N_x + \frac{3}{2}\epsilon N N_x - \frac{1}{6}\epsilon N_{xxt} = 0. \tag{1.5}$$

Equation (1.5) is called the Benjamin-Bona-Mahony equation [2]. Its relation to the Korteweg-de Vries equation is similar to that between the regularized Boussinesq system (1.3) and other systems of type (1.2). See [Albert/Bona] for a discussion of the advantages and disadvantages of (1.5) versus (1.4) as model equations.

Another advantage of system (1.3) over other systems of type (1.2) is that (1.3) is easier to solve numerically. In fact, J. Bona and M. Chen have performed careful numerical studies of system (1.3) which have revealed various interesting properties of its solutions [3]. In particular, these studies have shown that (1.3) has solutions which behave in many ways like solutions of (1.4) and (1.5). This suggests that the formal derivation of (1.4) or (1.5) from (1.3) is based on correct principles, and that solutions of (1.4) or (1.5) are close to unidirectional solutions of (1.3) when ϵ is small. In Chapter 4 below, we prove that this is indeed the case. Since the mathematical theory of equations (1.4) and (1.5) is much further advanced than that of system (1.3), such a result is of interest: it allows us to use our better knowledge of (1.4) and (1.5) to obtain quantitative information about solutions of (1.3). Thus our comparison result can be thought of as a theoretical explanation and confimation of the numerical work of Bona and Chen. It should be noted, however, that the agreement between solutions of (1.3) and (1.5) which has been seen by Bona and Chen is actually much better than we can currently explain by our theory. This point is discussed further in Chapter 5.

Our comparison result has precedents in work done by a number of other authors on comparisons between (1.4), (1.5) and the Euler equations for water waves (the Euler equations are the fundamental equations for water waves, expressing the laws of conservation of mass and momentum). Since (1.3), (1.4) and (1.5) are derived formally from the Euler equations under the assumption that ϵ is small (cf. chapter 13 of [15]), one would expect that when ϵ is small, solutions of all these equations are close (in an appropriate normed space) to solutions of the Euler equations, and hence in particular are close to each other. Moreover, one would expect that if $[0, T(\epsilon)]$ is the time interval on which solutions remain close, then $T(\epsilon) \to \infty$ as $\epsilon \to 0$.

In fact, Craig [9] has shown that solutions of (1.4) are close to solutions of the Euler equations on time scales of length $1/\epsilon$. More precisely, he proves that under certain smallness conditions on the initial data, a solution of the Euler equations will exist for $0 \leq t \leq 1/\epsilon$; and that if one further assumes (i) a boundedness condition (independent of ϵ) on higher-order Sobolev norms of this solution and (ii) a condition on the solution which implies its unidirectionality, then as $\epsilon \to 0$ the solution tends, in an appropriate Sobolev norm, to a solution of (1.4) for each $t \in [0, 1/\epsilon]$. Moreover, since Bona, Pritchard and Scott have shown in Theorem 1 of [4] that solutions of (1.4) are close to those of (1.5) on time scales of length $1/\epsilon$, it follows that solutions of (1.5) are also close to those of the Euler equations on these time scales.

It remains an open question, however, whether the solutions of (1.3), or of any other system of type (1.2), resemble solutions of the Euler equations. In this work we address the somewhat easier question of whether solutions of (1.3) resemble those of (1.5) or (1.4) on long time scales. Consider solutions of (1.3) with initial data

$$\begin{aligned} N(x,0) &= g(x) \\ V(x,0) &= h(x), \end{aligned} \tag{1.6}$$

and a solution $U(x,t)$ of (1.5) with initial data $U(x,0) = g(x)$. To guarantee unidirectionality of the solution (N, V), we assume that

$$h = g - \frac{1}{4}\epsilon g^2. \tag{1.7}$$

Then the following theorem (restated below as Theorem 4.1.1 and proved in Chapter 4) establishes a comparison between these solutions on a long time scale. (See the next section for an explanation of the notation used for the function spaces in this theorem.)

Theorem. Suppose g is in H^m ($m \geq 7$), and h is given by (1.7). Then there exists a constant $C_1 > 0$, depending only on $\|g\|_7$, such that for all $\epsilon > 0$, system (1.3) with initial conditions (1.6) has a solution pair (N, V) in X_T^m, where

$$T = C_1 \epsilon^{-1/2}. \tag{1.8}$$

Moreover, if U solves (1.5) with $U(x, 0) = g(x)$, and W is defined by

$$W = U - \frac{1}{4}\epsilon U^2$$

for all $t \geq 0$, then for $0 \leq j \leq m - 7$, we have

$$\|(N, V) - (U, W)\|_{\mathcal{H}_j} \leq C\epsilon^2 t \quad \text{for} \quad 0 \leq t \leq T, \tag{1.9}$$

where C depends only on $\|g\|_{j+7}$.

Note that, as a consequence of (1.9) and the Sobolev embedding theorem, it follows easily that

$$\sup_{x \in \mathbf{R}} (|N(x,t) - U(x,t)| + |V(x,t) - W(x,t))|) \leq C\epsilon^2 t$$

for $0 \leq t \leq T$, so that (1.9) implies in particular that the solutions (N, V) and (U, W) are close in a pointwise sense.

We remark also that the estimate in (1.9) is what one would expect to hold for the difference between solutions of two systems which differ by terms of order ϵ^2. For example, consider the simple equations $u_t + u_x = 0$ and $\tilde{u}_t + \tilde{u}_x = \epsilon^2 \tilde{u}$, posed with the same initial data $u(x,0) = \tilde{u}(x,0) = g(x)$. The explicit solutions of these equations are $u(x,t) = g(x-t)$ and $\tilde{u}(x,t) = g(x-t)e^{\epsilon^2 t}$, so for any j one has

$$\|u - \tilde{u}\|_{\mathcal{H}_j} \leq C\epsilon^2 t$$

for all $t \in [0, 1/\epsilon^2]$.

Notice that in the preceding example, the estimate between the difference of the two solutions held on a time interval $1/\epsilon^2$, which is much longer than the time interval $1/\sqrt{\epsilon}$ mentioned in (1.8). We suspect that the result of the preceding Theorem is also true on a longer time scale than $1/\sqrt{\epsilon}$. (Recall that the comparison result between (1.4) and (1.5) in [4] holds on a time interval of length $1/\epsilon$.) However, we have not been able to prove this.

The method we use to prove Theorem 4.1.1 is based on that used by Bona, Pritchard and Scott [4] to compare (1.4) and (1.5). It consists of studying the equation satisfied by the difference between the solutions of the two equations being compared, and using Gronwall's Lemma to obtain estimates on the size of this difference. One technical feature of our analysis which differs from that of [4] is that we have to use a form of Gronwall's Lemma which is valid only on finite time intervals; in this our analysis resembles that of [1]. We also have been influenced by the paper [6], which helps to make clear the role played in comparison results by suitable a priori estimates for the individual equations being compared. In order to obtain our comparison result, it was necessary to first obtain such a priori estimates for system (1.3). This is done in Chapter 3 below, where we also include a review of some a priori estimates we need for equation (1.5).

The exposition below is organized as follows. Chapter 2 describes our notation and contains some well-known results from functional analysis which we include here for the reader's convenience. Chapter 3 concerns a priori estimates on (1.3) and (1.5). In Section 3.1, we prove a local well-posedness result for (1.3) by applying the contraction-mapping principle to an equivalent system of integral equations. In section 3.2, we review certain known results for (1.5), with the intent of establishing uniform bounds (independent of ϵ) on the growth of higher Sobolev norms of its solutions over long time intervals. Chapter 4 contains the statement and proof of our main result, the comparison result between (1.3) and (1.5). Concluding remarks are in Chapter 5.

Chapter II

NOTATION AND PRELIMINARIES

2.1. Notation

In this work the following notation will be used.

For $1 \le p \le \infty$, we define the space L^p to be the set of all functions f on $\mathbf{R}$ satisfying $|f|_p < \infty$, where the norm $|f|_p$ is defined by

$$|f|_p = \left[\int_{-\infty}^{\infty} |f(x)|^p \, dx\right]^{1/p}$$

if $1 \le p < \infty$, and

$$|f|_\infty = \text{ess sup}_{x\in\mathbf{R}}|f(x)|.$$

For any real number s, we define the L^2-based Sobolev space H^s to be the set of all functions f on $\mathbf{R}$ satisfying $\|f\|_s < \infty$, where the norm $\|f\|_s$ is defined by

$$\|f\|_s = \left[\int_{-\infty}^{\infty} (1+k^2)^s |\widehat{f}(k)|^2 \, dk\right]^{1/2} .$$

Here hats are used to denote Fourier transforms, which are defined by the formula

$$\widehat{f}(k) = \int_{-\infty}^{\infty} e^{-ikx} f(x) \, dx.$$

If X and Y are two Banach spaces, then $X \times Y$ denotes the Banach space consisting of ordered pairs (x, y) with $x \in X$ and $y \in Y$, where $\|(x,y)\|_{X\times Y}$ is defined as $\max(\|x\|_X, \|y\|_Y)$.

In particular, we will often work in the space $H^m \times H^m$, which will henceforth be denoted as $\mathcal{H}^m$.

If X is a Banach space and T is a positive real number, we define $C(0,T;X)$ to be the Banach space of all continuous X-valued maps h from $[0,T]$ to X, with norm given by $\|h\|_{C(0,T;X)} = \sup_{0\leq t\leq T} \|h(t)\|_X$. We define $C(0,\infty;X)$ to be the union of all sets $C(0,T;X)$ as T ranges over the set of all positive numbers. For $0 < T \leq \infty$, we will use X_T^m as an abbreviation for $C(0,T;\mathcal{H}^m)$.

If $K(x)$ and $g(x)$ are functions defined for $x \in \mathbf{R}$, we define the convolution of K and g to be the function f given by

$$f(x) = \int_{-\infty}^{\infty} K(x-y)g(y)\, dy.$$

We use the notation

$$f = K * g$$

for convolutions. The operation of convolution is related to that of Fourier transformation via the formula

$$\widehat{K * g}(k) = \widehat{K}(k)\widehat{g}(k).$$

2.2. Review of some results from functional analysis

Here we recall for the reader's convenience some standard results which will be used below.

Proposition 2.2.1 (Banach Contraction-Mapping Theorem). Suppose B is a closed ball in a Banach space X, and $A : X \mapsto X$ is such that

(i) $A(B) \subset B$, and

(ii) there exists $\theta < 1$ such that

$$\|Av\|_X \leq \theta \|v\|_X \quad \text{for all v in} \quad B.$$

Then there exists some $v \in B$ such that $Av = v$.

Proposition 2.2.2. If X is a Banach space, and $\alpha[0,1] \mapsto X$ is a continous map, then $\int_0^1 \alpha(s)\, ds$ is well-defined as an element of X and

$$\| \int_0^1 \alpha(s)\, ds \|_X \leq \int_0^1 \|\alpha(s)\|_X ds.$$

Proposition 2.2.3. Suppose $m \geq 1$ is an integer. Then there exists a constant C_m such that for all f and g in H^m, the product function fg is also in H^m and satisfies

$$\|fg\|_m \leq C_m \|f\|_m \|g\|_m.$$

The proofs of the preceding three results are elementary, and are here omitted. For Proposition 2.2.1, the reader may refer to Theorem V.18 of [12]. Proposition

2.2.2, which contains Minkowski's integral integrality as a special case, is a basic consequence of the definition of the integral of a Banach-space-valued function; cf. Theorem 3.29 of [13]. Proposition 2.2.3 is easily proved by appealing to the alternative definition of H^m as the set of all functions whose derivatives up to order m exist and are in L^2, and using the Liebnitz rule for derivatives of products.

We conclude this section with a version of Gronwall's Lemma.

Lemma 2.2.4 (Gronwall's Lemma).

(i) Let $\alpha, \beta > 0$ be given. if $T > 0$ and $A(t)$ is any non-negative continuous function defined on $[0, T]$ satisfying $A(0) = 0$ and

$$A^2(t) \le \int_0^t \left(\alpha A(s) + \beta A^2(s)\right)\, ds \quad \text{for} \quad 0 \le t \le T,$$

then

$$A(t) \le \frac{\alpha}{\beta}(e^{\frac{1}{2}\beta t} - 1) \quad \text{for} \quad 0 \le t \le T.$$

(ii) Let $\alpha, \beta > 0$ and $\rho > 1$ be given. Define

$$T = \beta^{-1/\rho} \alpha^{(1-\rho)/\rho} \int_0^\infty (1 + z^\rho)^{-1}\, dz.$$

Then there exists a constant $C = C(\rho) > 0$, which is independent of α and β, such that if $0 \le T_1 \le T$ and $A(t)$ is any non-negative continuous function defined on $[0, T_1]$ satisfying $A(0) = 0$ and

$$A^2(t) \le \int_0^t \left(\alpha A(s) + \beta A^{\rho+1}(s)\right)\, ds \quad \text{for} \quad 0 \le t \le T_1,$$

then

$$A(t) \leq C\alpha t \quad \text{for} \quad 0 \leq t \leq T_1.$$

Proof. Define $B(t)$ on $[0, T]$ by

$$B^2(t) = \int_0^t \left(\alpha A(s) + \beta A^2(s)\right)\, ds.$$

Then

$$\frac{d}{dt}(B^2(t)) = \alpha A(t) + \beta A^2(t).$$

Now, by assumption $A(t) \leq B(t)$, so

$$\frac{d}{dt}(B^2(t)) \leq \alpha B(t) + \beta B^2(t)$$

and hence

$$\frac{dB}{dt} \leq \frac{1}{2}(\alpha + \beta B(t)).$$

Now dividing both sides of the last inequality by $\alpha + \beta B(t)$, and integrating both sides of the resulting inequality from 0 to t, we obtain

$$\frac{1}{\beta}\log(1 + \beta B(t)/\alpha) \leq \frac{t}{2}.$$

Therefore

$$B(t) \leq \frac{\alpha}{\beta}(e^{\frac{1}{2}\beta t} - 1)$$

holds for all $t \in [0, T]$. Since $A(t) \leq B(t)$, this proves part (i).

To prove part (ii), let $B(t)$ be defined on $[0, T_1]$ by

$$B^2(t) = \int_0^t \left(\alpha A(s) + \beta A^{\rho+1}(s)\right)\, ds.$$

Then

$$\frac{d}{dt}(B^2(t)) = \alpha A(t) + \beta A^{\rho+1}(t).$$

Again we have $A(t) \leq B(t)$, so

$$\frac{d}{dt}(B^2(t)) \leq \alpha B(t) + \beta B^{\rho+1}(t).$$

But, as shown in Lemma 2 of [1] , if $T_1 \leq T$ then the preceding inequality can be integrated to yield $B(t) \leq C\alpha t$ for $t \in [0, T_1]$, where C depends only on ρ. Since $A(t) \leq B(t)$, the desired result follows.

Chapter III

WELL-POSEDNESS AND REGULARITY RESULTS

3.1. The regularized Boussinesq system

Consider the following regularized Boussinesq system:

$$\begin{aligned} N_t + V_x + \epsilon(NV)_x - \frac{1}{6}\epsilon N_{xxt} &= 0 \\ V_t + N_x + \epsilon V V_x - \frac{1}{6}\epsilon V_{xxt} &= 0. \end{aligned} \tag{3.1.1}$$

In [3], well-posedness results are proved for solutions of (3.1.1) on a finite interval $0 \le x \le L$, subject to boundary conditions at $x = 0$ and $x = L$. Here, we are concerned rather with the pure initial-value problem. Thus we look for solutions of (3.1.1) defined for $x \in \mathbf{R}$, which take prescribed initial data

$$\begin{aligned} N(x,0) &= \phi(x) \\ V(x,0) &= \psi(x) \end{aligned} \tag{3.1.2}$$

defined for $x \in \mathbf{R}$.

In this section, the dependence of (3.1.2) on the parameter ϵ will not play a role in our analysis. In fact, for the purpose of proving the well-posedness theorem which appears below (Theorem 3.1.2), we could just as well have considered the system

$$\begin{aligned} \eta_t + u_x + (\eta u)_x - \frac{1}{6}\eta_{xxt} &= 0 \\ u_t + \eta_x + u u_x - \frac{1}{6}u_{xxt} &= 0, \end{aligned}$$

which is obtained from (3.1.1) by means of the scaling

$$\eta(x,t) = \epsilon N(\sqrt{\epsilon}x, \sqrt{\epsilon}t)$$

$$u(x,t) = \epsilon V(\sqrt{\epsilon}x, \sqrt{\epsilon}t).$$

(Indeed, the latter form of the system is the one considered by Bona and Chen.) However, in Chapter IV of this work, the role of ϵ will become very important, and so we have chosen to keep it visible in the analysis which follows.

The approach we take to studying the well-posedness of (3.1.1), (3.1.2) is the same as that used for the periodic problem in [3], which in turn is the same approach used for the scalar Benjamin-Bona-Mahony equation in [2,3]. Namely, instead of working directly with (3.1.1) and (3.1.2), we first reformulate the problem as a system of integral equations, which will then be solved using the Banach Contraction-mapping Theorem, Theorem 2.2.1.

The first step in reformulating the initial-value problem is to solve equations (3.1.1) for N_t and V_t by making use of the integral form of the operator $-\partial_x(1 - \frac{1}{6}\epsilon\partial_x^2)^{-1}$. This operator is in fact a well-behaved convolution operator, whose properties are studied in the following Lemma.

Lemma 3.1.1. Suppose $f \in H^m$, where m is any non-negative integer. Then there exists a unique $g \in H^{m+1}$ which satisfies the equation

$$(1 - \frac{1}{6}\epsilon\partial_x^2)g = -\partial_x f. \tag{3.1.3}$$

This function g is given by the formula

$$g = K_\epsilon * f,$$

where

$$K_\epsilon(x) = \frac{3}{\epsilon}(\operatorname{sgn} x)e^{-\sqrt{6x^2/\epsilon}}.$$

Further, one has the estimate

$$\|g\|_m = \|K_\epsilon * f\|_m \leq \sqrt{\frac{3}{2\epsilon}}\,\|f\|_m.$$

Proof. By taking the Fourier transform of both sides of (3.1.3), one arrives at the equation

$$\widehat{g}(k) = \left(\frac{-ik}{1+\epsilon k^2/6}\right)\widehat{f}(k).$$

We therefore define $K_\epsilon(x)$ to be the function whose Fourier transform is given by

$$\widehat{K_\epsilon}(k) = \frac{-ik}{1+\epsilon k^2/6}.$$

An elementary computation shows that K_ϵ is then given by the formula displayed in the statement of the Lemma.

Now notice that

$$\begin{aligned}
\|K_\epsilon * f\|_{m+1} &= \left(\int_{-\infty}^{\infty}(1+k^2)^{m+1}|\widehat{K_\epsilon * f}(k)|^2\,dk\right)^{\frac{1}{2}} \\
&\leq \left(\int_{-\infty}^{\infty}(1+k^2)|\widehat{K_\epsilon}(k)|^2(1+k^2)^m|\widehat{f}(k)|^2\,dk\right)^{\frac{1}{2}} \\
&\leq \left(\sup_{k\in\mathbf{R}}(1+k^2)^{\frac{1}{2}}|\widehat{K_\epsilon}(k)|\right)\|f\|_m \leq \frac{7}{\epsilon}\|f\|_m.
\end{aligned}$$

This shows that g is in H^{m+1}.

The stated estimate on $\|g\|_m$ is obtained by a computation similar to that done just above:

$$\begin{aligned}\|K_\epsilon * f\|_m &= \left(\int_{-\infty}^{\infty} (1+k^2)^m |\widehat{K_\epsilon * f}(k)|^2 \, dk\right)^{\frac{1}{2}} \\ &\le \left(\int_{-\infty}^{\infty} |\widehat{K_\epsilon}(k)|^2 (1+k^2)^m |\widehat{f}(k)|^2 \, dk\right)^{\frac{1}{2}} \\ &\le \left(\sup_{k\in\mathbf{R}} |\widehat{K_\epsilon}(k)|\right) \|f\|_m = \sqrt{\frac{3}{2\epsilon}}\, \|f\|_m.\end{aligned}$$

This completes the proof of the lemma.

Now rewrite (3.3.1) in the form

$$\begin{aligned}(1-\frac{1}{6}\epsilon\partial_x^2)N_t &= -\partial_x(V+\epsilon NV) \\ (1-\frac{1}{6}\epsilon\partial_x^2)V_t &= -\partial_x(N+\frac{1}{2}\epsilon V^2),\end{aligned}$$

and apply Lemma 3.1.1. After performing an integration with respect to t and using (3.1.2), we arrive at the following useful reformulation of the initial-value problem:

$$\begin{aligned}N &= \phi + \int_0^t K_\epsilon * (V+\epsilon NV)\, ds \\ V &= \psi + \int_0^t K_\epsilon * (N + \frac{1}{2}\epsilon V^2)\, ds.\end{aligned} \tag{3.1.4}$$

Suppose $(\phi,\psi) \in \mathcal{H}^m$ for some $m \ge 1$, and suppose $T > 0$. If $(N,V) \in X_T^m$ is such that (3.1.4) holds (in the sense of equality in $\mathcal{H}^m$ for each $t \in [0,T]$, then (N,V) also satisfies (3.1.1) in the sense of distributions. Conversely, any distribution solution of (3.1.1) in X_T^m will satisfy (3.1.4) for some ϕ and ψ. Therefore, to prove well-posedness of (3.1.1) in X_T^m, it will suffice to consider (3.1.4).

We can now state and prove the main result of this chapter.

Theorem 3.1.2. Suppose (ϕ, ψ) is in $\mathcal{H}^1$.

(a) There exists $T > 0$ such that (3.1.4) has a solution in X_T^1. In fact, for a given value of ϵ, T can be chosen to depend only on the norm of (ϕ, ψ) in $\mathcal{H}^1$.

(b) For any $T > 0$, there is at most one solution of (3.1.4) in X_T^1.

(c) Let T_0 be the supremum of the set of all $T > 0$ such that (3.1.4) has a solution in X_T^1. Then either $T_0 = +\infty$ or

$$\lim_{t \uparrow T_0} \|(N(\cdot, t), V(\cdot, t))\|_{\mathcal{H}^1} = +\infty.$$

(d) Suppose (ϕ, ψ) is in $\mathcal{H}^m$ for some integer $m \geq 1$. Let T_0 be as defined in (c), and suppose $T \in [0, T_0)$. Then $(N, V) \in X_T^m$, and $(\partial_t^k N, \partial_t^k V) \in X_T^{m+1}$ for every integer $k > 0$.

(e) The correspondence $(\phi, \psi) \mapsto (N, V)$ defined by (d) is a continuous mapping of $\mathcal{H}^m$ into X_T^m for every $T < T_0$; and, if $k > 0$, the correspondence $(\phi, \psi) \mapsto (\partial_t^k N, \partial_t^k V)$ is a continuous mapping of $\mathcal{H}^m$ into X_T^{m+1} for every $T < T_0$. More precisely, the first of these assertions means that for every $T < T_0$ and $\delta > 0$, there exists an open ball B in $\mathcal{H}^m$ containing (ϕ, ψ) such that for each initial datum $(\tilde{\phi}, \tilde{\psi})$ in B, there is a solution $(\tilde{N}, \tilde{V}) \in X_T^m$ of (3.1.4) (with (ϕ, ψ) replaced by $(\tilde{\phi}, \tilde{\psi})$), satisfying

$$\|(N, V) - (\tilde{N}, \tilde{V})\|_{X_T^m} \leq \delta.$$

The second assertion is to be interpreted similarly.

Proof. Let T be a positive number, as yet unspecified. If $h = (N, V)$ is an element of X_T^1, denote by Ah the pair of functions on the right-hand side of equations (3.1.4), i.e.,

$$Ah_1 = \phi + \int_0^t K_\epsilon * (V + \epsilon NV)\, ds$$
$$Ah_2 = \psi + \int_0^t K_\epsilon * (N + \frac{1}{2}\epsilon V^2)\, ds.$$

It is easy to see that this defines a map A from X_T^1 into X_T^1. We now show that if T is sufficiently small, then A is a contraction map on a certain ball in X_T^1.

Let $b = \|(\phi, \psi)\|_{\mathcal{H}^1}$, and set $R = 2b$. Define B_R to be the ball with radius R centered at the origin $\theta = (0, 0)$ in X_T^1.

Suppose $h_1 = (N_1, V_1)$ and $h_2 = (N_2, V_2)$ are two elements of B_R. Then using Proposition 2.2.2, Proposition 2.2.3, and Lemma 3.1.1, we find that

$$\|Ah_1 - Ah_2\|_{X_T^1} \leq \sqrt{\frac{3}{2\epsilon}}\, T(1 + 2\epsilon R)\|h_1 - h_2\|_{X_T^1}.$$

Also,

$$\begin{aligned}\|Ah_1\|_{X_T^1} &= \|Ah_1 - A\theta + (\phi, \psi)\|_{X_T^1} \\ &\leq \sqrt{\frac{3}{2\epsilon}}\, T(1 + 2\epsilon R)\|h_1\|_{X_T^1} + b \\ &\leq \sqrt{\frac{3}{2\epsilon}}\, T(1 + 2\epsilon R)R + b.\end{aligned}$$

Now choose

$$T = \frac{\sqrt{\epsilon}}{\sqrt{6}(1 + 2\epsilon R)},$$

so that

$$\sqrt{\frac{3}{2\epsilon}}\, T(1 + 2\epsilon R) = \frac{1}{2}.$$

(Since $R = 2b = 2\|(\phi, \psi)\|_{\mathcal{H}^1}$, it is indeed the case that the value of T depends only on the norm of (ϕ, ψ) in $\mathcal{H}^1$.) Then, for all h_1 and h_2 in B_R,

$$\|Ah_1 - Ah_2\|_{X_T^1} \leq \frac{1}{2}\|h_1 - h_2\|_{X_{T_0}^1}$$

and

$$\|Ah_1\|_{X_T^1} \leq \frac{1}{2}R + b = R.$$

Hence, for this choice of T, A maps B_R into B_R and is a contraction mapping on B_R. An application of Proposition 2.2.1 then yields the existence of a fixed point of A in B_R. This completes the proof of part (a) of the Theorem.

To prove part (b), suppose that h_1 and h_2 are two solutions of (3.1.4) in X_T^1. Then for any $t \in [0, T]$, one has

$$\begin{aligned} \|h_1(t) - h_2(t)\|_{\mathcal{H}^1} &= \|Ah_1(t) - Ah_2(t)\|_{\mathcal{H}^1} \\ &\leq \sqrt{\frac{3}{2\epsilon}}\,(1 + \|h_1\|_{X_T^1} + \|h_2\|_{X_T^1}) \int_0^t \|h_1(\tau) - h_2(\tau)\|_{\mathcal{H}^1}\, d\tau. \end{aligned}$$

It then follows from Lemma 2.2.4(i), with $\alpha = 0$ and $\beta = \sqrt{\frac{3}{2\epsilon}}\,(1 + \|h_1\|_{X_T^1} + \|h_2\|_{X_T^1})$, that $\|h_1(t) - h_2(t)\|_{\mathcal{H}^1} = 0$ for $0 \leq t \leq T$. Hence $h_1 \equiv h_2$, and (b) is proved.

To prove (c), suppose on the contrary that $T_0 < \infty$ and that there exists a number $M > 0$ and a sequence t_n approaching T_0 from below such that

$$\|(N(x, t_n), V(x, t_n))\|_{\mathcal{H}^1} \leq M \quad \text{for } n = 1, 2, 3, \ldots$$

By part (a), there exists some $T = T(M) > 0$ such that if $(\bar{\phi}, \bar{\psi}) \in \mathcal{H}^1$ and $\|(\bar{\phi}, \bar{\psi})\|_{\mathcal{H}^1} \leq M$, then (3.1.4) has a solution $(\bar{N}(x,t), \bar{V}(x,t))$ in X_T^1. Choose t_n

such that $T_0 - t_n < T$, and choose $(\tilde{\phi}, \tilde{\psi}) = (N(x,t_n), V(x,t_n))$. Then, extending the solution (N,V) to the interval $[0, t_n + T]$ by defining $N(x,t) = \tilde{N}(x, t - t_n)$ and $V(x,t) = \tilde{V}(x, t - t_n)$ for $t_n \le t \le t_n + T$, we obtain a solution of (3.1.4) in $X^1_{(t_n+T)}$. But since $t_n + T > T_0$, this contradicts the maximality of T_0.

To prove (d), we will first show that (N,V) is in $X^j_{T_0}$ for $1 \le j \le m$ by using induction on j. The case $j = 1$ has already been done above in (a). Suppose now that it has been shown that $(N,V) \in X^{j-1}_{T_0}$, where $2 \le j \le m$; we wish to show that $(N,V) \in X^j_{T_0}$. From Lemma 3.1.1 and Proposition 2.2.3, it follows that $K_\epsilon * (V + \epsilon NV)$ and $K_\epsilon * (N + \frac{1}{2}\epsilon V^2)$ are functions in H^j for each $t \in [0, T_0]$. Hence, since ϕ and ψ are in $H^m \subset H^j$, (3.1.4) implies that N and V are in H^j for each $t \in [0, T_0]$. Moreover, if $0 \le t \le t' \le T_0$, then from (3.1.4) and Lemma 3.1.1 we obtain the estimate

$$
\begin{aligned}
\|N(x,t) - N(x,t')\|_j &\le \int_t^{t'} \|K_\epsilon * (V + \epsilon NV)\|_j \, ds \\
&\le \sqrt{\frac{3}{2\epsilon}} \, \|V + \epsilon NV\|_{X^j_{T_0}} |t' - t| \\
&\le \sqrt{\frac{3}{2\epsilon}} \, \|V\|_{X^j_{T_0}} \left(1 + \|\epsilon N\|_{X^j_{T_0}}\right) |t' - t|,
\end{aligned}
$$

and a similar estimate for $\|V(x,t) - V(x,t')\|_j$. It follows from these estimates that (N,V) is a continuous map from $[0,T_0]$ to $\mathcal{H}^j$, so $(N,V) \in X^j_{T_0}$.

The assertions of (d) concerning $(\partial_t^k N, \partial_t^k V)$ will be proved using induction on k. First, take the distributional derivative of (3.1.4) with respect to t to obtain the equations

$$
\begin{aligned}
N_t &= K_\epsilon * (V + \epsilon NV) \\
V_t &= K_\epsilon * (N + \epsilon \frac{1}{2} V^2).
\end{aligned}
$$

Since N and V are in H^m, then $V+\epsilon NV$ and $N+\frac{1}{2}\epsilon V^2$ are also in H^m, so by Lemma 3.1.1 it follows that N_t and V_t are in H^{m+1}. Next, suppose it has been shown that $\partial_t^k V$ is in H^{m+1}, where k is any positive integer. Taking the k-fold derivative of (3.1.4) with respect to t, we get the equations

$$\partial_t^{k+1} N = K_\epsilon * \left(\partial_t^k V + \partial_t^k(\epsilon NV)\right)$$

$$\partial_t^{k+1} V = K_\epsilon * \left(\partial_t^k N + \frac{1}{2}\epsilon \partial_t^k (V^2)\right),$$

and from Lemma 3.1.1 it follows that $\partial_t^{k+1}N$ and $\partial_t^{k+1}V$ are in H^{m+1}. This completes the proof of (d).

We now prove (e). Let $T < T_0$ and $\delta > 0$ be given. From parts (a) through (d), we know that there exists an open ball B' containing (ϕ,ψ) such that if $(\tilde{\phi},\tilde{\psi}) \in B'$ then equations (3.1.4) (with (ϕ,ψ) replaced by $(\tilde{\phi},\tilde{\psi})$) has a solution $(\tilde{N},\tilde{V}) \in X_T^m$. Let p and q be functions of x and t defined by $(p,q) = (\tilde{N},\tilde{V}) - (N,V)$. Then p and q satisfy the equations

$$\begin{aligned} p(x,t) &= p(x,0) + \int_0^t K_\epsilon * (q + \epsilon p\tilde{V} + \epsilon qN)\, ds \\ q(x,t) &= q(x,0) + \int_0^t K_\epsilon * (p + \frac{1}{2}\epsilon q(\tilde{V}+V))\, ds. \end{aligned} \tag{3.1.5}$$

It follows that

$$\|p(t)\|_m \le \|p(0)\|_m + C\int_0^t (\|p(s)\|_m + \|q(s)\|_m)\, dx$$

$$\|q(t)\|_m \le \|q(0)\|_m + C\int_0^t (\|p(s)\|_m + \|q(s)\|_m)\, dx,$$

where p and q are now viewed as H^m-valued functions of t, and the constants C depend on ϵ and the norms in X_T^m of V, $\tilde{V}$, and $\tilde{N}$. Therefore, if we define $W(t)$ by

$$W(t) = \|p(t)\|_m + \|q(t)\|_m,$$

then $W(t)$ satisfies the inequality

$$W(t) \le W(0) + C\int_0^t W(s)\, ds$$

for $0 \le t \le T$. It then follows from Lemma 2.2.4(i) that

$$W(t) \le W(0)e^{Ct},$$

or, in other words,

$$\|\tilde{N} - N\|_m + \|\tilde{V} - V\|_m \le \left(\|\tilde{\phi} - \phi\|_m + \|\tilde{\psi} - \psi\|_m\right) e^{Ct}.$$

Since the right-hand side can be made less than δ for all $t \in [0, T]$ by taking $\|\tilde{\phi} - \phi\|_m + \|\tilde{\psi} - \psi\|_m$ sufficiently small, this completes the proof of the first assertion of (e).

Finally, to prove the second assertion of (e), we differentiate both sides of (3.1.5) repeatedly with respect to t to obtain

$$\partial_t^k p = K_\epsilon * \partial_t^{k-1}(q + \epsilon p\tilde{V} + \epsilon qN)$$

$$\partial_t^k q = K_\epsilon * \partial_t^{k-1}(p + \frac{1}{2}\epsilon q(\tilde{V} + V)).$$

For $k = 1$, this leads to the estimates

$$\|p_t\|_{m+1} \le C(\|q\|_m + \|p\|_m)$$

$$\|q_t\|_{m+1} \le C(\|p\|_m + \|q\|_m).$$

But since we have already shown that $\|p\|_m$ and $\|q\|_m$ can be made arbitrarily small for $t \in [0, T]$ by choosing $\|\tilde{\phi} - \phi\|_m + \|\tilde{\psi} - \psi\|_m$ sufficiently small, this then implies that $\|p_t\|_{m+1}$ and $\|q(t)\|_{m+1}$ can be made arbitrarily small. This proves the second assertion of (e) for the case $k = 1$. The proof for general k then proceeds by an obvious induction argument, whose details we omit.

3.2. The Benjamin-Bona-Mahony equation

In this section we consider the Benjamin-Bona-Mahony equation

$$U_t + U_x + \frac{3}{2}\epsilon U U_x - \frac{1}{6}\epsilon U_{xxt} = 0. \tag{3.2.1}$$

with initial condition

$$U(x,0) = g(x). \tag{3.2.2}$$

The following well-posedness result for (3.2.1) is an immediate consequence of Proposition 2 of [4].

Theorem 3.2.1. Let $g \in H^m$ where $m \geq 1$. Then for every $T > 0$, (3.2.1) has a unique solution U in $C(0,T;H^1)$ such that $U(x,0) = g(x)$ on $\mathbb{R}$. Moreover, $U \in C(0,T;H^m)$, and for every integer $k > 0$, $\partial_t^k U \in C(0,T;H^{m+1})$. For each $T > 0$, the correspondence $g \mapsto U$ is a continuous mapping of H^m to $C(0,T;H^m)$ while, if $k > 0$, the correspondence $g \mapsto \partial_t^k U$ is a continuous mapping of H^m into $C(0,T;H^{m+1})$.

We remark that Proposition 2 of [4] actually refers to the equation

$$\eta_t + \eta_x + \eta\eta_x - \eta_{xxt} = 0,$$

which can be obtained from (3.2.1) by making the transformation

$$\eta(x,t) = \frac{3}{2}\epsilon U(\sqrt{\epsilon/6}\, x, \sqrt{\epsilon/6}\, t).$$

Obviously, however, the result is preserved if the variables are rescaled in this way.

In Chapter 4 below, we will also need an ϵ-independent bound on the growth of Sobolev norms of the solution of (3.2.1) on long time intervals. This bound is provided by Theorem 3.2.3 below. Before stating and proving Theorem 3.2.3, we need the following lemma concerning the behavior of the operator $-(1-\frac{1}{2}\epsilon\partial_x^2)^{-1}$. (Compare with Lemma 3.1 above.)

Lemma 3.2.2. Suppose $f \in H^m$, where m is any non-negative integer. Then there exists a unique $g \in H^{m+1}$ which satisfies the equation

$$(1-\frac{1}{6}\epsilon\partial_x^2)g = -f. \tag{3.2.3}$$

This function g is given by the formula

$$g = J_\epsilon * f,$$

where

$$J_\epsilon(x) = -\sqrt{\frac{3}{2\epsilon}}\, e^{-\sqrt{6x^2/\epsilon}}.$$

Further, one has, for any $\epsilon > 0$,

$$\|g\|_m = \|J_\epsilon * f\|_m \leq \|f\|_m.$$

Proof. Taking the Fourier transform of both sides of (3.2.3) yields the equation

$$\widehat{g}(k) = \left(\frac{-1}{1+\epsilon k^2/6}\right)\widehat{f}(k).$$

Define $J_\epsilon(x)$ to be the function whose Fourier transform is given by

$$\widehat{J}_\epsilon(k) = \frac{-1}{1+\epsilon k^2/6}.$$

The function J_ϵ is obviously related to the function K_ϵ defined in Lemma 3.1 by the relation $\partial_x J_\epsilon = K_\epsilon$. Therefore we can obtain the formula for J_ϵ by taking the antiderivative of K_ϵ, with constant of integration chosen so that J_ϵ is an integrable function. This leads to the formula displayed above for J_ϵ. The final inequality in the lemma is now proved by an argument similar to that used in the proof of Lemma 3.1:

$$\begin{aligned}\|J_\epsilon * f\|_m &= \left(\int_{-\infty}^{\infty}(1+k^2)^m|\widehat{J_\epsilon * f}(k)|^2\,dk\right)^{\frac{1}{2}}\\ &\le \left(\int_{-\infty}^{\infty}|\widehat{J}_\epsilon(k)|^2(1+k^2)^m|\widehat{f}(k)|^2\,dk\right)^{\frac{1}{2}}\\ &\le \left(\sup_{k\in\mathbf{R}}|\widehat{J}_\epsilon(k)|\right)\|f\|_m \le \|f\|_m.\end{aligned}$$

Theorem 3.2.3. Let m be any integer such that $m \ge 1$. Then there exists an increasing function $M : \mathbf{R} \to \mathbf{R}$ with the following property: for every $g \in H^{m+5}$, the solution U of (3.2.1) given by the preceding theorem satisfies

$$\|\partial_t^l U(\cdot,t)\|_{m-l} \le M(\|g\|_{m+5})$$

for every $t \in [0, 1/\epsilon]$ and every integer l satisfying $0 \le l \le m$.

Proof. Suppose $g \in H^m$, and consider first the solution u of the initial-value problem for the Korteweg-de Vries equation

$$\begin{aligned} u_t + uu_x + u_{xxx} &= 0 \\ u(x,0) &= u_0(x). \end{aligned} \tag{3.2.4}$$

Theorem 1 of [8] states that for every $T > 0$, the problem (3.2.4) has a unique solution in $C(0, T; H^m)$, provided $m \geq 3$. (Actually, the analysis in [8] treats the equation $u_t + uu_x + u_{xxx} = 0$, but for our purposes this equation is equivalent to equation (3.2.4), since one can be transformed into the other by means of the change of variables $x' = x - t$, $t' = t$.) The extension of Theorem 1 of [8] to all $m \geq 1$ was accomplished in [10].

As explained in the discussion on p. 577 of [8] (q.v. for further references), there exist a countable number of explicitly-defined functionals $I_0, I_1, I_2, \ldots$ which, at least formally, are conserved under the flow defined by the Korteweg-de Vries equation (2.4). These functionals are of the form

$$I_k = \int_{-\infty}^{\infty} P_k(u)\, dx,$$

where P_k is a polynomial function of u and its partial derivatives with respect to x. In fact, P_k consists of a linear combination of monomials

$$(\partial_x^0 u)^{a_0} (\partial_x^1 u)^{a_1} \ldots (\partial_x^p u)^{a_p},$$

in which the "rank" of each monomial, defined as $\sum_{i=0}^{p}(1 + \frac{1}{2}i)a_i$, is equal to $k+2$. In Prop. 6, p. 578 of [8] it is shown that if $m \geq 3$, then the solution of (3.2.4) satisfies $I_k(u(\cdot, t)) = I_k(u_0)$ for all $t \geq 0$. Again, it follows from the work of [10] that this result holds for all $m \geq 1$.

Now defining a new dependent variable $\tilde{u}$ by

$$u(x,t) = \frac{3}{2}\epsilon\tilde{u}(\sqrt{\epsilon/6}\ (x+t), \sqrt{\epsilon/6}\ t),$$

one sees that solutions u of (3.2.4) correspond to solutions $\tilde{u}$ of the initial-value problem

$$\begin{aligned} \tilde{u}_t + \tilde{u}_x + \frac{3}{2}\epsilon\tilde{u}\tilde{u}_x + \frac{1}{6}\epsilon\tilde{u}_{xxx} &= 0 \\ \tilde{u}(x,0) &= \tilde{u}_0(x), \end{aligned} \tag{3.2.5}$$

where $u_0(x) = \epsilon\tilde{u}_0(\sqrt{\epsilon}x)$. Moreover, because P_k is a polynomial whose monomials all have rank $k+2$, one has that

$$P_k(u(x,t)) = \epsilon^{k+2}P_k(\tilde{u}(\sqrt{\epsilon}\,(x+t), \sqrt{\epsilon}\,t))$$

for all x and t. It follows easily that I_k is also a conserved functional for (3.2.5). Hence, by the same argument as used on p. 577 of [8], we can conclude that there exists an increasing function $M^{(1)} : \mathbf{R} \to \mathbf{R}$ such that the solution of (3.2.5) exists and satisfies

$$\|\partial_x^k\tilde{u}\|_{L^2} \leq M^{(1)}(\|\tilde{u}_0\|_k) \tag{3.2.6}$$

for all $0 \leq k \leq m$.

Now take $\tilde{u}_0 = g$, and consider the relation between $\tilde{u}$ and the solution U of the initial-value problem (3.2.1), (3.2.2). When rescaled into the variables used here, Theorem 1 of [4] states that, if $k \geq 0$ and $g \in H^{k+5}$, then there exists an increasing function $M_k^{(2)} : \mathbf{R} \to \mathbf{R}$ such that for every $\epsilon \in (0,1]$ and every $t \in [0, 1/\epsilon]$,

$$\|\partial_x^k U - \partial_x^k\tilde{u}\|_{L^2} \leq M_k^{(2)}(\|g\|_{k+5})\epsilon^2 t.$$

In particular, it follows that

$$\|\partial_x^k U - \partial_x^k \tilde{u}\|_{L^2} \leq M_k^{(2)}(\|g\|_{k+5})\epsilon \leq M_k^{(2)}(\|g\|_{k+5})$$

for every $t \in [0, 1/\epsilon]$. Combining this result with (3.2.6), we obtain that

$$\|\partial_x^k U\|_{L^2} \leq \tilde{M}^{(1)}(\|g\|_k) + M_k^{(2)}(\|g\|_{k+5})$$

for all $0 \leq k \leq m$. It follows that there exists a function $M : \mathbf{R} \to \mathbf{R}$ such that

$$\|U\|_m \leq M_m(\|g\|_{m+5}).$$

This proves that there exists a function M such that the conclusion of the Theorem holds in the case $l = 0$. We now use induction to show that a function M exists for which the conclusion of the theorem holds for all $0 \leq l \leq m$. In the remainder of this proof, we use the letter M to stand for any increasing real-valued function; the function possibly being different from one line to the next.

To complete the proof by induction, we assume that the desired statement has been proved for $l \leq l_0$, where $0 \leq l_0 \leq m-1$; and we wish to prove it for $l = l_0 + 1$. Differentiating both sides of (3.2.1) l_0 times with respect to t, we obtain the equation

$$\partial_t^{l_0+1} U - \frac{1}{6}\epsilon \partial_x^2 \partial_t^{l_0+1} U = -\partial_t^{l_0}(U_x + \frac{3}{2}\epsilon U U_x).$$

It then follows from Lemma 3.2.2 that

(3.2.7)

$$\|\partial_t^{l_0+1} U\|_{m-l_0-1} = \|J_\epsilon * \partial_t^{l_0}(U_x + \frac{3}{2}\epsilon U U_x)\|_{m-l_0-1} \leq \|\partial_t^{l_0}(U_x + \frac{3}{2}\epsilon U U_x)\|_{m-l_0-1}.$$

Now, by the inductive hypothesis, we have that $\partial_t^{l_0} U \in H^{m-l_0}$, with

$$\|\partial_t^{l_0} U\|_{m-l_0} \leq M(\|g\|_{m+5}).$$

Hence $\partial_t^{l_0} U_x \in H^{m-l_0-1}$, with

$$\|\partial_t^{l_0} U_x\|_{m-l_0-1} \leq \|\partial_t^{l_0} U\|_{m-l_0} \leq M(\|g\|_{m+5}).$$

Moreover, expanding $\partial_t^{l_0}(UU_x)$ using the Leibniz product rule for derivatives, and applying the inductive hypothesis and Proposition 2.2.3, we have also that $\partial_t^{l_0}(UU_x) \in H^{m-l_0-1}$, with

$$\|\partial_t^{l_0}(UU_x)\|_{m-l_0-1} \leq M(\|g\|_{m+5}).$$

Combining these last two estimates with (3.2.7), we obtain

$$\|\partial_t^{l_0+1} U\|_{m-l_0-1} \leq M(\|g\|_{m+5}),$$

as desired.

Chapter IV

THE COMPARISON RESULT

4.1. Statement of the result

Consider again the regularized Boussinesq system

$$\begin{aligned} N_t + V_x + \epsilon(VN)_x - \frac{1}{6}\epsilon N_{xxt} &= 0 \\ V_t + N_x + \epsilon VV_x - \frac{1}{6}\epsilon V_{xxt} &= 0. \end{aligned} \tag{4.1.1}$$

The comparison result proved in this chapter will show that solutions of (4.1.1) with appropriate initial data will resemble solutions of the Benjamin-Bona-Mahony equation over long time intervals. The initial conditions we consider for (4.1.1) are

$$\begin{aligned} N(x,0) &= g(x) \\ V(x,0) &= h(x), \end{aligned} \tag{4.1.2}$$

where h is related to g by

$$h = g - \frac{1}{4}\epsilon g^2. \tag{4.1.3}$$

The meaning of this restriction on the initial data will become clear below. Essentially, it guarantees that the corresponding solution of (4.1.1) is a right-moving wave, rather than a bidirectional wave.

The form of the Benjamin-Bona-Mahony equation we consider is

$$U_t + U_x + \frac{3}{2}\epsilon UU_x - \frac{1}{6}\epsilon U_{xxt} = 0 \tag{4.1.4}$$

with initial condition

$$U(x,0) = g(x). \tag{4.1.5}$$

It will turn out that N resembles U, while V will resemble the function $W(x,t)$ defined for $x \in \mathbf{R}$ and $t \geq 0$ by

$$W = U - \frac{1}{4}\epsilon U^2. \tag{4.1.6}$$

We remark first that for $(g,h) \in \mathcal{H}^m$ $(m \geq 1)$, the well-posedness theory of the preceding chapter only establishes the existence of a solution to (4.1.1), (4.1.2) on a short time interval. In fact, an examination of the formula obtained for T in the proof of part (a) of Theorem 3.1.2 shows that the length of the time interval on which we have proved existence is no longer than a constant times $\sqrt{\epsilon}$. In particular, the time interval of existence decreases to zero as $\epsilon \to 0$, which is the opposite of what we expect from a satisfactory existence theory. Fortunately, the comparison result we now establish actually also asserts that, under the restrictions we have imposed on the initial data, a solution to (4.1.1) and (4.1.2) will actually exist on a time interval of length proportional to $1/\sqrt{\epsilon}$.

Theorem 4.1.1. Suppose g is in H^m $(m \geq 7)$, and h is given by (4.1.3). Then there exists a constant $C_1 > 0$, depending only on $\|g\|_7$, such that for all $\epsilon \in (0,1]$, system (4.1.1) with initial conditions (4.1.2) has a solution pair (N,V) in X_T^m, where

$$T = C_1\epsilon^{-1/2}. \tag{4.1.7}$$

Moreover, if U is the solution of (4.1.4), (4.1.5), and W is defined by (4.1.6), then for $0 \leq j \leq m-7$,

$$\|(N,V)-(U,W)\|_{\mathcal{H}_j} \leq C\epsilon^2 t \quad \text{for} \quad 0 \leq t \leq T, \tag{4.1.8}$$

where C depends only on $\|g\|_{j+7}$.

The proof of this result will be given in the next two sections.

4.2. Proof in the case $j = 0$.

To prove Theorem 4.1.2, we first transform (4.1.4) and (4.1.6) into a form permitting easier comparison with (4.1.1). Differentiating (4.1.6) with respect to x, and using the resulting equation to replace U_x by W_x in (4.1.4), one arrives at the equation

$$U_t + W_x + \epsilon(UW)_x - \frac{1}{6}\epsilon U_{xxt} + \epsilon^2 G_1 = 0, \tag{4.2.1}$$

where

$$G_1 = \frac{3}{4}U^2 U_x.$$

On the other hand, differentiating (4.1.6) with respect to t, and eliminating the U_t term in the resulting equation by using (4.1.4), one obtains

$$W_t + U_x + \frac{1}{2}\epsilon UU_t + \frac{3}{2}\epsilon UU_x - \frac{1}{6}\epsilon U_{xxt} = 0.$$

Using (4.1.4) to replace U_t by $-U_x - \frac{3}{2}\epsilon UU_x + \frac{1}{6}\epsilon U_{xxt}$ in the third term on the left-hand side, we obtain

$$W_t + U_x + \epsilon UU_x - \frac{1}{6}\epsilon U_{xxt} - \frac{3}{4}\epsilon^2 U^2 U_x + \frac{1}{12}\epsilon^2 UU_{xxt} = 0.$$

Now, use (4.1.6) to replace U by $W + \frac{1}{4}\epsilon U^2$ in the third and fourth terms on the left-hand side of the preceding equation. The result can be written in the form

$$W_t + U_x + \epsilon WW_x - \frac{1}{6}\epsilon W_{xxt} + \epsilon^2 G_2 + \epsilon^3 G_3 = 0, \tag{4.2.2}$$

where

$$G_2 = \frac{1}{2}WUU_x + \frac{1}{4}W_xU^2 - \frac{3}{4}U^2U_x - \frac{1}{6}U_xU_{xt} - \frac{1}{12}U_tU_{xx}$$

and

$$G_3 = \frac{1}{8}U^3U_x.$$

Note that equations (4.2.1) and (4.2.2) have the same form as the equations in system (4.1.1), except for the presence of the higher-order terms on the right-hand sides. Indeed, in showing that two systems are equivalent up to higher-order terms, we have essentially repeated the formal derivation of the Korteweg-de Vries equation from the Boussinesq system which appears on pp. 465-466 of [15].

To compare the solutions of the two systems, we will study the functions $m = N - U$ and $n = V - W$. Notice first that, according to Theorem 3.1.2, the system (4.1.1) has a solution in X_T^m for any $T < T_0$, where $T_0 > 0$ is defined as in Theorem 3.1.2(c). Further, the regularity properties of N and V which are guaranteed by Theorem 3.1.2 are sufficient to justify the manipulations which follow.

Subtracting (4.2.1) and (4.2.2) from (4.1.1), we find that m and n satisfy the equations

$$m_t + n_x + \epsilon(mn)_x + \epsilon(Un)_x + \epsilon(Wm)_x - \frac{1}{6}\epsilon m_{xxt} = \epsilon^2 G_1 \tag{4.2.3}$$

and

$$n_t + m_x + \epsilon(nn_x) + \epsilon(Wn)_x - \frac{1}{6}\epsilon n_{xxt} = \epsilon^2 G_2 + \epsilon^3 G_3. \tag{4.2.4}$$

Now multiply (4.2.3) by m and (4.2.4) by n, and add the results to obtain

$$
\begin{aligned}
&mm_t + nn_t + (mn)_x - \frac{1}{6}\epsilon m m_{xxt} - \frac{1}{6}\epsilon n n_{xxt} \\
&\qquad + \epsilon\left(m(mn)_x + m(Un)_x + m(Wm)_x + n^2 n_x + n(Wn)_x\right) \\
&\qquad = \epsilon^2(mG_1 + nG_2) + \epsilon^3 nG_3.
\end{aligned}
\tag{4.2.5}
$$

Next, integrate (4.2.5) with respect to x over the real line, and then with respect to t from 0 to t. Since mn vanishes at $x = \pm\infty$, the third term on the left-hand side of (4.2.5) integrates to zero. After integrating by parts with respect to x in the fourth and fifth terms, we obtain finally that

(4.2.6)

$$
\begin{aligned}
&\frac{1}{2}\int_{-\infty}^{\infty}\left[m^2 + n^2 + \frac{1}{6}\epsilon(m_x)^2 + \frac{1}{6}\epsilon(n_x)^2\right]\,dx = \\
&\quad -\epsilon\int_0^t\int_{-\infty}^{\infty}\left(m(mn)_x + m(Un)_x + m(Wm)_x + n^2 n_x + n(Wn)_x\right)\,dx\,ds \\
&\quad + \epsilon^2\int_0^t\int_{-\infty}^{\infty}(mG_1 + nG_2)\,dx\,ds + \epsilon^3\int_0^t\int_{-\infty}^{\infty} nG_3\,dx\,ds.
\end{aligned}
$$

Now define the positive quantity $A(t)$ by

$$
A^2(t) = \int_{-\infty}^{\infty}\left[m^2 + n^2 + \frac{1}{6}\epsilon(m_x)^2 + \frac{1}{6}\epsilon(n_x)^2\right]\,dx;
$$

i.e., $A^2(t)$ is (except for the factor of $1/2$) the quantity appearing on the left-hand side of (4.2.6). Note that obviously

$$
|m|_2 \le A,
$$

$$
|n|_2 \le A,
$$

$$
|m_x|_2 \le \sqrt{6}\epsilon^{-1/2}A,
$$

and

$$|n_x|_2 \le \sqrt{6}\epsilon^{-1/2}A.$$

We also have the elementary estimate

$$|m|_\infty \le |m|_2^{1/2}|m_x|_2^{1/2} \le 6^{1/4}\epsilon^{-1/4}A,$$

and a similar estimate for $|n|_\infty$.

Our next goal is to deduce from (4.2.6) the inequality

$$A^2(t) \le C\int_0^t \left(\epsilon^2 A(s) + \epsilon^{1/2}A^2(s) + A^3(s)\right)\, ds, \tag{4.2.7}$$

where C denotes a constant which may depend on norms of U and its derivatives, but is independent of ϵ and t. (In what follows, we will use the same letter C for any such constant; the value of C possibly changing from one occurrence to the next).

To prove (4.2.7), we begin by considering the term multiplied by a factor of ϵ on the right-hand side of (4.2.6). The integral with respect to x contained in this term can be broken into five integrals numbered (I) through (V), corresponding to each of five terms in the integrand; we will estimate each of these five integrals separately.

For integral number (I), we have, using Hölder's inequality,

$$\begin{aligned}\epsilon\left|\int_{-\infty}^{\infty} m(mn)_x\, dx\right| &\le \epsilon\left(|m|_\infty|m|_2|n_x|_2 + |m|_\infty|m_x|_2|n|_2\right)\\ &\le C\epsilon^{1/4}A^3 \le CA^3.\end{aligned}$$

For integral number (II), we have

$$\epsilon\left|\int_{-\infty}^{\infty} m(Un)_x\, dx\right| \le \epsilon\left(|U|_\infty |m|_2 |n_x|_2 + |U_x|_\infty |m|_2 |n|_2\right)$$
$$\le C(\epsilon^{1/2}A^2 + \epsilon A^2)$$
$$\le C\epsilon^{1/2}A^2.$$

Applying integration by parts to integral number (III), we obtain

$$\epsilon\int_{-\infty}^{\infty} m(Wm)_x\, dx = -\epsilon\int_{-\infty}^{\infty} Wmm_x\, dx$$
$$= -\frac{1}{2}\epsilon\int_{-\infty}^{\infty} W(m^2)_x\, dx$$
$$= \frac{1}{2}\epsilon\int_{-\infty}^{\infty} W_x m^2\, dx.$$

Therefore

$$\epsilon\left|\int_{-\infty}^{\infty} m(Wm)_x\, dx\right| \le \frac{1}{2}\epsilon |W_x|_\infty |m|_2^2 \le C\epsilon A^2 \le C\epsilon^{1/2}A^2.$$

(The use of the notation C for $\frac{1}{2}|W_x|_\infty$ is justified by (4.1.6), which shows that norms of W may be estimated in terms of norms of U.)

Integral (V) is handled in the same way as (III), and so is also bounded by $C\epsilon^{1/2}A^2$. Finally, integral (IV) equals zero, since $n^2 n_x$ is the derivative of $\frac{1}{3}n^3$, which vanishes at $x = \pm\infty$.

Combining the above estimates, we see that the term on the right-hand side of (4.2.6) containing a factor of ϵ can be bounded by an expression of the form appearing on the right-hand side of (4.2.7). To complete the proof of (4.2.7), then, it remains only to show that the remaining terms on the right-hand side of (4.2.6) can be bounded by such an expression. But from the definitions of G_1, G_2, and

G_3, it is easy to see that $|G_1|_2$, $|G_2|_2$, and $|G_3|_2$ are each bounded by a constant times $\|U\|_2 + \|U_t\|_1$. Hence, by Hölder's inequality,

$$\epsilon^2 \left| \int_{-\infty}^{\infty} (mG_1 + nG_2)\ dx \right| + \epsilon^3 \left| \int_{-\infty}^{\infty} nG_3\ dx \right| \leq C\epsilon^2(|m|_2 + |n|_2) + C\epsilon^3 |n|_2 \leq C\epsilon^2 A,$$

and the proof of (4.2.7) is complete.

We have now verified that (4.2.7) holds for all $t \in [0, T_0)$. Next, we apply Gronwall's Lemma to (4.2.7). First, note that by Young's inequality,

$$\epsilon^{1/2} A^2 = (\epsilon A^{1/2})(\epsilon^{-1/2} A^{3/2}) \leq \epsilon^2 A + \epsilon^{-1} A^3,$$

so (4.2.7) implies that

$$A^2(t) \leq C \int_0^t \left(\epsilon^2 A(s) + \epsilon^{-1} A^3(s)\right)\ ds.$$

Therefore we may apply Lemma 2.2.4(ii) with $\alpha = C\epsilon^2$, $\beta = C\epsilon^{-1}$, $\rho = 2$, and $T_1 = \min(T_0, T)$, where

$$T = \alpha^{-1/2} \beta^{-1/2} \int_0^{\infty} \frac{1}{1 + z^2}\ dz.$$

We then conclude that the inequality

$$A(t) \leq C\epsilon^2 t \tag{4.2.8}$$

holds for all $t \in [0, T_1]$.

It follows now that in fact $T_1 = T$. For if not, then we would have $T_1 = T_0$;

and so for all $t \leq T_0$, using Theorem 3.2.3, we have

$$\begin{aligned}
\|(N,V)\|_{\mathcal{H}^1} &\leq \|(m,n)\|_{\mathcal{H}^1} + \|(U,W)\|_{\mathcal{H}^1} \\
&\leq |m|_2 + |n|_2 + |m_x|_2 + |n_x|_2 + M(\|g\|_6) \\
&\leq C\epsilon^{-1/2}A(t) + M(\|g\|_6) \\
&\leq C\epsilon^{3/2}T_0 + M(\|g\|_6) < \infty,
\end{aligned}$$

which contradicts Theorem 3.1.2(c).

We have now proved that (4.2.8) holds for all $t \in [0,T]$. From the definition of T, we see that $T = C_1\epsilon^{-1/2}$, where C_1 depends only on the constants C appearing in α and β. But an examination of the estimates used to obtain (4.2.7) shows that these constants C actually depend only on the norms $\|U\|_2$ and $\|U_t\|_1$. Hence C_1 also depends only on these quantities. Moreover, from (4.1.4), we have

$$U_t = F * (U_x + \frac{3}{2}\epsilon U U_x),$$

where F is the function whose Fourier transform is

$$\widehat{F}(k) = 1/(1 + \frac{1}{6}\epsilon k^2).$$

Since $|\widehat{F}(k)| \leq 1$ for all k, it follows that

$$\|U_t\|_1 \leq \|U_x + \frac{3}{2}\epsilon U U_x\|_1 \leq C(\|U\|_2 + \|U\|_2^2),$$

where C is independent of ϵ and U. Therefore C_1 actually depends only on $\|U\|_2$. Finally, by applying Theorem 3.2.3 (taking C_1 smaller, if necessary, so that $T = C_1\epsilon^{-1/2}$ is less than $1/\epsilon$), we can bound $\|U\|_2$ for all $t \leq T$ by a quantity $M(\|g\|_7)$

which does not depend on ϵ. Therefore the constant C_1 in (4.1.7) depends only on $\|g\|_7$.

Another consequence of the argument just given is that the constant C in (4.2.8) depends only on $\|g\|_7$. Since

$$\|(N, V) - (U, W)\|_{\mathcal{H}_0} \leq A(t),$$

this proves (4.1.8) in the case $j = 0$.

4.3. Proof in the case $j \geq 1$.

To prove Theorem 4.1.1 for $j \geq 1$, we use induction. Define the quantity $A_j(t)$ by

$$A_j^2(t) = \int_{-\infty}^{\infty} \left[\sum_{k=0}^{j} (m_{(k)}^2 + n_{(k)}^2) + \frac{1}{6}\epsilon(m_{(j+1)}^2 + n_{(j+1)}^2) \right] dx.$$

(Here subscripts in parentheses are used to denote partial derivatives with respect to x; so that $m_(2)$, for example, denotes $\frac{\partial^2 m}{\partial x^2}$.) We wish to prove that $A_j(t) \leq C\epsilon^2 t$ for $0 \leq t \leq T$, where T is defined in (4.1.7). In Section 4.2 this was proved to be true for $j = 0$; we now assume it that has been proved for $j - 1$, and will show that it holds for j.

First, differentiate equations (4.2.3) and (4.2.4) j times with respect to x to obtain

$$\begin{aligned} &\partial_t m_{(j)} + n_{(j+1)} + \epsilon(mn)_{(j+1)} \\ &\qquad + \epsilon(Un)_{(j+1)} + \epsilon(Wm)_{(j+1)} - \frac{1}{6}\epsilon\partial_t m_{(j+2)} = \epsilon^2 (G_1)_{(j)} \\ &\partial_t n_{(j)} + m_{(j+1)} + \epsilon(nn_x)_{(j)} \\ &\qquad + \epsilon(Wn)_{(j+1)} - \frac{1}{6}\epsilon\partial_t n_{(j+2)} = \epsilon^2 (G_2)_{(j)} + \epsilon^3 (G_3)_{(j)} \end{aligned} \tag{4.3.1}$$

Now multiply the first equation in (4.3.1) by $m_{(j)}$ and the second by $n_{(j)}$ to obtain

$$(4.3.2)\qquad \begin{aligned} &\partial_t m_{(j)} m_{(j)} + n_{(j+1)} m_{(j)} + \epsilon (mn)_{(j+1)} m_{(j)} \\ &\qquad + \epsilon (Un)_{(j+1)} m_{(j)} + \epsilon (Wm)_{(j+1)} m_{(j)} - \frac{1}{6}\epsilon \partial_t m_{(j+2)} m_{(j)} \\ &\qquad = \epsilon^2 (G_1)_{(j)} m_{(j)} \\ &\partial_t n_{(j)} n_{(j)} + m_{(j+1)} n_{(j)} + \epsilon (nn_x)_{(j)} n_{(j)} \\ &\qquad + \epsilon (Wn)_{(j+1)} n_{(j)} - \frac{1}{6}\epsilon \partial_t n_{(j+2)} n_{(j)} \\ &\qquad = \epsilon^2 (G_2)_{(j)} n_{(j)} + \epsilon^3 (G_3)_{(j)} n_{(j)} \end{aligned}$$

Next, add the two equations in (4.3.2), and then integrate over $(-\infty, \infty) \times [0, t]$. Since the sum $n_{(j+1)} m_{(j)} + m_{(j+1)} n_{(j)}$ integrates to zero, the result can be written as

$$(4.3.3)\qquad \begin{aligned} &\frac{1}{2}\int_{-\infty}^{\infty} \left[m_{(j)}^2 + n_{(j)}^2 + \frac{1}{6}\epsilon (m_{(j+2)})^2 + \frac{1}{6}\epsilon (n_{(j+2)})^2 \right] \, dx \\ &\qquad = (I) + (II) + (III) + (IV), \end{aligned}$$

where

$$\begin{aligned} (I) &= -\epsilon \int_0^t \int_{-\infty}^{\infty} \left((mn)_{(j+1)} m_{(j)} + (nn_x)_{(j)} n_{(j)} \right) \, dx \, ds \\ (II) &= -\epsilon \int_0^t \int_{-\infty}^{\infty} \left((Un)_{(j+1)} m_{(j)} + (Wm)_{(j+1)} m_{(j)} + (Wn)_{(j+1)} n_{(j)} \right) \, dx \, ds \\ (III) &= \epsilon^2 \int_0^t \int_{-\infty}^{\infty} \left((G_1)_{(j)} m_{(j)} + (G_2)_{(j)} n_{(j)} \right) \, dx \, ds \\ (IV) &= \epsilon^3 \int_0^t \int_{-\infty}^{\infty} (G_3)_{(j)} n_{(j)} \, dx \, ds. \end{aligned}$$

As in Section 4.2, we want to deduce from (4.3.3) the inequality

$$(4.3.4)\qquad A_j^2(t) \le C \int_0^t \left(\epsilon^2 A_j(s) + \epsilon^{1/2} A_j^2(s) + A_j^3(s) \right) \, ds,$$

where C depends on norms of U and its derivatives, but is independent of ϵ and t. We have at our disposal the estimates

$$|m_{(k)}|_2 + |n_{(k)}|_2 \le A_j, \tag{4.3.5}$$

valid for all $k \le j$, and

$$|m_{(j+2)}|_2 + |n_{(j+2)}|_2 \le \sqrt{6}\epsilon^{-\frac{1}{2}} A_j. \tag{4.3.6}$$

Let us now consider the integrals (I), (II), (III), (IV) on the right side of (4.3.3). Using Liebniz' rule for the derivative of a product, we see that the integrand of (I) can be written as the sum of terms of the form $m_{(k)}m_{(i)}n_{(l)}$ or $n_{(k)}n_{(i)}n_{(l)}$, which each of k, i, and l is less than or equal to $j+1$, and no two can both equal $j+1$ in any term. Therefore, Hölder's inequality together with (4.3.5) and (4.3.6) allows us to conclude that

$$|(I)| \le C\epsilon^{1/2} A_{(j)}^3 \le C A_{(j)}^3.$$

Similar arguments applied to the integral (II) in (4.3.3) show that

$$|(II)| \le C\epsilon^{1/2} A_{(j)}^2.$$

Finally, combining Hölder's inequality and (4.3.5), we obtain

$$|(III)| \le C\epsilon^2 A_{(j)}$$

and

$$|(IV)| \le C\epsilon^3 A_{(j)} \le C\epsilon^2 A_{(j)}.$$

Combining these estimates, we obtain (4.3.4), which is valid for all $t \in [0, T)$. An examination of the estimates shows that the constant C appearing in (4.3.4) can be taken to depend only on $\|U\|_{j+2} + \|U_t\|_{j+1}$; and as shown in the preceding Section, the latter quantity can in turn be bounded by a constant times $\|U\|_{j+2}$. Therefore, by Theorem 3.2.3, C can be chosen to depend only on $\|g\|_{j+7}$.

Finally, applying Gronwall's Lemma 2.2.4(ii) to (4.3.4) as before, we conclude that

$$A_j(t) \leq C\epsilon^2 t$$

holds for all $t \in [0, T]$, where C depends only on $\|g\|_{j+7}$. This then completes the proof of Theorem 4.1.1.

Chapter V

CONCLUDING REMARKS

The comparison result proved in Chapter 4 establishes that solutions of the Benjamin-Bona-Mahony equation (1.5) are close to solutions of the regularized Boussinesq system (1.3) on a time interval of length $1/\sqrt{\epsilon}$. This result is a theoretical confirmation of the recent numerical work of Bona and Chen [3]. However, the results of [3] suggest that the comparison result should hold on much longer time intervals, perhaps as long as $1/\epsilon^2$. In fact, for many initial data, the solutions to (1.3) and (1.5) found by Bona and Chen are virtually indistinguishable to the eye, even when graphed over a time interval which is long enough for many significant wave interactions to occur. (In general, different solutions of (1.3), with small amplitudes and long wavelengths measured by the parameter ϵ are expected to require a time interval of at least $1/\epsilon$ on which to interact significantly.) It is not yet clear how to extend Theorem 4.1.1 to these longer time intervals. (A similar question raised by Bona, Pritchard, Scott [4] for equations (1.4) and (1.5), namely the question of whether solutions of (1.4) and (1.5) remain close on a time interval of length $1/\epsilon^2$, has remained open to this day.) One way to obtain an improved result would be to use the same argument as above, but with a more favorable estimate than (4.2.7) for the quantity $A(t)$. To get such an improved estimate on A, one would need to establish better a priori estimates on the solutions of (1.3).

Another, easier way to extend Theorem 4.1.1 would be to prove it for other systems of the form (1.2). As explained in [5], there are other systems of type (1.2) besides (1.3) for which a suitable existence theory can be established. One of them, the Bona-Smith system, is even known to have solutions which exist globally in time. We chose here to study (1.3) rather than the Bona-Smith system because (1.3), with its advantageous numerical properties, was the system studied numerically by Bona and Chen. However, a result like Theorem 4.1.1, or an even better one, could probably be proved in the same way for the Bona-Smith system. Similarly, Theorem 4.1.1 could also be improved, in the sense that less regularity would be required for the initial data g, if we had compared (1.3) to (1.4) instead of comparing (1.3) to (1.5). The reason for this is that for the Korteweg-de Vries equation (1.4), a version of Theorem 3.2.3 holds which requires less regularity on g. In fact, for a solution U of (1.4) one can assert that $\|U\|_m$ is bounded by a constant M_m which depends only on $\|g\|_m$, and that the bound holds for all $t \geq 0$.

REFERENCES

[1] J. Albert and J. Bona, *Comparisons between model equations for long waves*, J. Nonlinear Sci. **1** (1991), 345–374.

[2] T. Benjamin, J. Bona, J. Mahony, *Model equations for long waves in nonlinear dispersive systems*, Phil. Trans. Roy. Soc. London Ser. A **272** (1972), 47–78.

[3] J. Bona and M. Chen, *A Boussinesq system for two-way propagation of non-linear dispersive waves*, Phys. D **116** (1998), 191–224.

[4] J. Bona, W. Pritchard, and L. Scott, *A comparison of solutions of two model equations for long waves*, Fluid dynamics in astrophysics and geophysics (Chicago, Ill., 1981), Lectures in Applied Mathematics vol. 20, American Mathematical Society, Providence, RI, 1983, 235–267.

[5] J. Bona, J.-C. Saut, and J. Toland, *Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media*, preprint (2000).

[6] J. Bona and M. Scialom, *The effect of change in the nonlinearity and the dispersion relation of model equations for long waves*, Canad. Appl. Math. Quart. **3** (1995), 1–41.

[7] J. Bona and R. Smith, *A model for the two-way propagation of water waves in a channel*, Math. Proc. Camb. Phil. Soc. **79** (1976), 167–182.

[8] J. Bona and R. Smith, *The initial-value problem for the Korteweg-de Vries equation*, Philos. Trans. Roy. Soc. London Ser. A **278** (1975), 555–601.

[9] W. Craig, *An existence theory for water waves and the Boussinesq and Kor-*

teweg-de Vries scaling limits, Comm. Partial Differential Equations **10** (1985), 787–1003.

[10] C. Kenig, G. Ponce, and L. Vega, *Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle*, Comm. Pure Appl. Math. **46** (1993), 527–620.

[11] D. Korteweg and G. de Vries, *On the change of form of long waves advancing in a rectangular channel, and on a new type of long stationary waves*, Philos. Mag. **39** (1895), 422-443.

[12] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Volume 1: Functional Analysis*, Academic Press, New York, 1972.

[13] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1973.

[14] R. Showalter, *Well-posed problems for some nonlinear dispersive waves*, J. Math. Pures et Appl. **56** (1977), 123–135.

[15] G. Whitham, *Linear and Nonlinear Waves*, Wiley-Interscience, New York, 1974.