CANADIAN APPLIED
MATHEMATICS QUARTERLY
Volume 3, Number 1, Winter 1995

THE EFFECT OF CHANGE IN THE NONLINEARITY AND THE DISPERSION RELATION OF MODEL EQUATIONS FOR LONG WAVES

J.L. BONA AND M. SCIALOM

ABSTRACT. The purpose of this paper is to understand the dependence of solutions of nonlinear, dispersive wave equations on the nonlinearity and the dispersion relation. This program of study is carried out here in the relatively specific, but practically important context of Korteweg-de Vries-type equations. In the last part of the paper, it is shown how the results for the Korteweg-de Vries equation and its relatives may be adapted to other classes of model equations such as nonlinear Schrödinger-type equations and regularized long-wave equations. The general thrust of our results is that small perturbations of a given dispersion relation or nonlinearity make only a small difference in the solution over a relatively long time scale. While not unexpected, this kind of theorem is useful as a guide to model builders in showing what sort of approximations can be countenanced without affecting the resulting solutions in an intolerable way.

1. Introduction. Considered herein are wave equations featuring nonlinearity and dispersion. The results of the investigation to be reported presently apply to several classes of model equations. To fix ideas, interest will first be focused on generalized Korteweg-de Vries equations having the form

$$u_t + u_x + f(u)_x - Mu_x = 0, \tag{1.1}$$

where $f : \mathbf{R} \to \mathbf{R}$ and M is a Fourier multiplier operator defined via its Fourier transform as

$$\widehat{Mh}(k) = m(k)\widehat{h}(k). \tag{1.2}$$

At a later stage, the discussion will be broadened to include other types of equations including the regularized long-wave equations, certain

Received by the editors on November 1, 1994.
Work partially supported by the Keck Foundation and the National Science Foundation, USA and by CNPq and FAPESP, Brasil.
The paper was completed while the first author was a CNRS visitor at the Centre de Mathématiques et de Leurs Applications, Ecole Normale Superieure de Cachan.

types of nonlinear Schrödinger equations, and higher-order water-wave models.

Equations of the form (1.1) arise in a wide range of physical contexts as models for the propagation of waves (cf. [**2, 5, 31, 46**]). Because these models apply to problems in the theory of wave propagation, we shall refer to the independent variables x and t as the spatial and temporal variable, respectively. Typically both nonlinear effects, modeled by f and dispersive and sometimes dissipative effects, modeled by M, are only approximations to a more complete accounting of these aspects of wave propagation. (Even though dissipative effects modeled by a non-zero imaginary part of m are included in principle in our development, most of the examples in view have m real-valued. In such cases, the operator M is purely dispersive, and we will often refer to it as the dispersion operator. When m has a non-trivial imaginary part, the operator associated to the imaginary part will be called the dissipation operator.)

In consequence of this state of affairs, it becomes interesting to understand to what extent the detailed structure of the nonlinearity or the dispersion and dissipation is reflected in solutions of the equations. Attention will be given to this issue in the context of the pure initial-value problem for (1.1), in which the dependent variable u is specified for all x at some fixed time t, say $t = 0$, so that

$$u(x,0) = \varphi(x) \tag{1.3}$$

for all $x \in \mathbf{R}$. The specification (1.3) corresponds to determining the wave profile everywhere at some given instant of time. While not always the most practical specification, it tends to be the easiest to understand theoretically and consequently attracts a lot of attention. In the present work we shall consider variations of both the nonlinearity and the dispersion relation. The question that will be posed is, for a fixed initial datum φ, if the dispersion relation m is perturbed, or if the nonlinearity f is changed, what can be said about the resulting variation of the solution u?

A more complete view is now offered of the issues that come to the fore in the remainder of this paper. First, recall that when the model equation (1.1) is written in the form presented, where f and m are order-one quantities, then they are usually formally valid only for small-amplitude, long-wavelength waves in which nonlinear and dispersive effects are relatively small, and approximately balanced in strength. This means that for small values of the dependent variable u and for small wavenumbers k, the effects of f and M should be relatively small, but, should have the same order of magnitude. Suppose the nonlinearity f and the dispersion relation m are homogeneous so that $f(u) = u^{p+1}$ and $m(k) = |k|^\alpha$. Let ε be a representative value of the amplitude of the motions in question and λ a typical value of the wavelength, where it is presumed that both these quantities have been non-dimensionalized with respect to an underlying length scale present in the problem. In these circumstances the initial wave profile φ is naturally scaled as

$$\varphi(x) = \varepsilon\psi(\lambda^{-1}x), \tag{1.4}$$

where ψ and its derivatives are of order one. Then the conditions that nonlinear and dispersive effects are small and balanced are the requirements that $\varepsilon^p\lambda^\alpha$ is of order one, while ε and λ^{-1} are both small. The quantity $S = \varepsilon^p\lambda^\alpha$ is a natural generalization of the classical Stokes or Ursell number of shallow-water theory (cf. [**48, 39, 50, 51, 11**]). If the small parameter δ is defined to be ε^p, then λ has order $\delta^{-1/\alpha}$ and the relation (1.4) can be expressed in terms of the single parameter δ as

$$\varphi(x) = \delta^{1/p}\psi(\delta^{1/\alpha}x). \tag{1.5}$$

If the dependent variable u is rescaled in the form

$$u(x,t) = \delta^{1/p}v(\delta^{1/\alpha}x, \delta^{1/\alpha}t), \tag{1.6}$$

then v satisfies the initial-value problem

$$v_t + v_x + \delta v^p v_x - \delta M_\delta v_x = 0, \qquad v(x,0) = \psi(x), \tag{1.7}$$

where ψ is as above. Written in this way, it is apparent that the nonlinear term $v^p v_x$ and the dispersive term $-M_\delta v_x$ represent balanced, small corrections to the basic uni-directional wave equation $v_t + v_x = 0$. Typically, when such wave equations are derived as models, there are higher-order effects of nonlinearity, dispersion and perhaps dissipation that have been ignored. In the variables used in (1.7), the terms corresponding to these effects come with a δ^2 attached. Thus in these variables, one expects that nonlinear and dispersive effects will accumulate and may have an order-one effect on the wave profile on a time scale of order $1/\delta$, while the ignored terms could produce order-one effects in a time scale of order $1/\delta^2$ (see the discussion of Albert and Bona [3]). Thus interesting nonlinear and dispersive effects will occur on time scales of order $1/\delta$ whilst neglected effects may render the model invalid as a description of reality on a time scale of order $1/\delta^2$. Translating these observations into the corresponding aspects relating to the variables appearing in (1.1), it is seen that

(i) nonlinear and dispersive effects may accumulate to make an order-one relative difference to the wave profile on a time scale of order $1/\delta^{1+1/\alpha}$, and

(ii) neglected effects may render the model invalid on a time scale of order $1/\delta^{2+1/\alpha}$.

These time scales will appear later in the more technical portion of our discussion.

Even in cases where f or M is not homogeneous, the above considerations may still apply if f and m have the form of a homogeneous part plus a remainder which is higher order in the respective dependent variable. Examples of this situation appear in Sections 4 and 5.

The central question that will attract attention here is the following. With initial data as in (1.5), suppose two different dispersion relations m_1 and m_2 or two different nonlinearities f_1 and f_2 to be given, and let u_1 and u_2 be the corresponding solutions of (1.1) emanating from a given initial value φ. For relatively small values of δ, it is expected that both u_1 and u_2 will be small, but, depending on the difference $m_1 - m_2$ and $f_1 - f_2$, it may be that $u_1 - u_2$ is smaller still, at least over certain time intervals. A result of this sort may be interpreted as saying that the difference between using m_2 and f_2 rather than m_1 and f_1 is relatively negligible, at least over certain time intervals. As will appear below, this time interval is often large, proportional to an inverse power of δ, and, under reasonable hypotheses, coincides with the time scale mentioned above over which interesting nonlinear and dispersive effects appear at the leading order.

The plan of the paper is as follows. Section 2 contains a brief explanation of our notational conventions together with theorems attesting to the well-posedness of the initial-value problems for (1.1). In Section 3, a general theorem of comparison for equations of the form (1.1) is formulated and proved. This relatively straightforward result is the mathematical heart of our theory. Detailed commentary on particular comparisons made with the use of this result, together with interpretation in terms of the physical problems being modeled comprise Section 4. Section 5 contains similar results to those obtained in Sections 2, 3, and 4 for other classes of equations. The paper closes with some commentary and suggestions for further research in Section 6.

2. Well posedness of initial-value problems.

This section contains results preliminary to the main theory for Korteweg-de Vries-type equations enunciated in Section 3 and used in Section 4. We begin with a few remarks about notation.

Notation. The notation employed throughout will be that which is currently standard in the theory of partial differential equations. Thus $L_p = L_p(\mathbf{R})$, for $1 \le p \le \infty$ is the usual Banach space of pth-power integrable functions (essentially bounded functions if $p = \infty$) whose norm is denoted by $|\cdot|_p$. As in (1.3) a circumflex adorning a function connotes that function's Fourier transform. The solutions of the initial-value problem (1.1)–(1.3) which will be discussed are, for each instant of time, members of Sobolev spaces H^s for various $s \ge 0$. If $f \in H^s$, then its norm is

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

Notice that the L_2-norm has two, different notations, between which systematic preference will be given to $|\cdot|_2$. If X is any Banach space, the space $C(a,b;X)$ is the collection of continuous maps $u : [a,b] \to X$ with the norm

$$\|u\|_{C(a,b;X)} = \sup_{a \le t \le b} \|u(t)\|_X,$$

where $\|\ \|_X$ connotes the norm on X.

Turning now to the well-posedness of the initial-value problem for equations of type (1.1), the following general result applies.

Theorem 2.1. *Let $s \geq 2$ be fixed and suppose $f : \mathbf{R} \to \mathbf{R}$ is a C^∞-function and the symbol $m = m_1 + i m_2$ of the operator M to have the properties that m_1, m_2 are real-valued, m_1 is an even function, and m_2 is an odd function with $m_2(\xi) \geq 0$ for $\xi \geq 0$. Suppose also that m satisfies the growth condition*

$$|m(k)| \leq C(1 + |k|^r) \tag{2.1}$$

for some constant C and some positive number r. Let initial data u_0 be given in H^s. Then there is a $T = T(\|u_0\|_s) > 0$ and a unique solution

$$u \in C(0,T;H^s) \cap C^1(0,T;H^{s-r'}) \tag{2.2}$$

of the equation (1.1) such that $u(\cdot,t) \to u_0$ in H^s as $t \to 0$, where $r' = \max\{r,1\}$. Moreover, the correspondence that associates to initial data u_0 the solution u is continuous from H^s to the function class in (2.2).

Remark 2.2. This theorem is a straightforward consequence of Kato's general theory for quasi-linear evolution equations [**34–36**]. In Kato's theory, attention is given to abstract evolutionequations of the form

$$\frac{du}{dt} + A(u,t)u + F(u) = 0, \quad \text{for } 0 \leq t \leq T,$$
$$u(0) = u_0,$$

where $u(t)$ takes values in some reflexive Banach space X and A is a mapping from $[0,T] \times W$ into $G(X,1,\beta)$, the generators of C_0-semigroups on X such that $\|e^{-sA(t,y)}\|_{\mathcal{B}(X)} \leq e^{\beta s}$, for some real number β, and where W is an open set in a smaller reflexive Banach space Y. The symbol $\mathcal{B}(X)$ connotes the Banach algebra of bounded operators on X. If $s \geq r'$, then the theory developed in Kato's first paper [**34**] suffices, (with $A(y,t) = M\partial_x + f'(y)\partial_x$, $F \equiv 0$, $Y = H^s$ and $X = H^{s-r'}$) while if $s < r'$, then the technique introduced in Kato [**35**] of conjugating the solution with the associated linear semigroup allows the earlier theory [**34**] to be applied. (Note that the operator M maps real-valued functions to real-valued functions since its symbol m has an even real part and an odd imaginary part.)

It is worth mention that for the Korteweg-de Vries equation (KdV-equation, commonly)

$$u_t + uu_x + u_{xxx} = 0,$$

and for certain other equations having the form (1.1), the presently available theory for the pure initial-wave problem is considerably more refined than is suggested by Theorem 2.1. In particular, the KdV-equation possesses an unexpected smoothing effect which was noticed in the works of Cohen Murray [**22**] and Sachs [**43**], and which received general attention in Kato [**36**]. The most recent theory is due to Bourgain [**17,18**], Constantin and Saut [**23**], Ginibre and Tsutsumi [**26**], Ginibre and Velo [**27**] and Kenig, Ponce and Vega [**37, 38**], and we may safely refer the interested reader to these papers for further references and detailed statements of results.

While Kato's theory is very powerful in terms of the range of its applicability, the dependence of the existence time T on the size of the initial data u_0 is rather complex and usually far from optimal. It will be useful later when questions of time scales arise to have an indication of the size of T as it relates to $\|u_0\|_s$, say. Because of the relatively simple form of (1.1), such an estimate is not hard to supply.

It is a consequence of Kato's theory that if we have in hand an *a priori* deduced finite bound on $\|u(\cdot,t)\|_s$ for suitable values of s and any t in some interval $[0,T_0]$, then the maximum existence time T for the solution of (1.1) starting at u_0 is at least T_0. Therefore, if we can provide such a bound on a time interval $[0,T_0]$ it will follow that the solution of (1.1) emanating from u_0 exists at least for $0 \leq t \leq T_0$ and that it respects the bound in question. These remarks set the stage for the following corollary to Theorem 2.1 in which the nonlinearity is specialized to the homogeneous case to obtain a simpler statement.

Corollary 2.3. *Let s, m and u_0 be as in Theorem 2.1 and suppose $f(z) = z^{p+1}$ for a positive integer p. Then the solution u of (1.1) corresponding to these specifications exists at least on the time interval $[0,T_0]$, and on this interval*

$$|u(\cdot,t)|_2 \leq |u_0|_2 \tag{2.3}$$

and

$$|u_{xx}(\cdot,t)|_2 \leq C_p|u_0''|_2$$

where

$$T_0 = \frac{d_p}{|u_0|_2^{(3p-2)/4}|u_0''|_2^{(p+2)/4}} \tag{2.4}$$

and C_p and d_p are constants depending only on p. Moreover, there are constants D_j, depending only on p and $\|u_0\|_2$, $2 \le j \le s$, such that

$$|\partial_x^j u(\cdot,t)|_2 \le D_j |\partial_x^j u_0|_2$$

for $2 \le j \le s$ and $0 \le t \le T_0$.

Proof. As discussed above, it suffices to deduce bounds on the H^s-norms of solutions that are valid on the advertised time intervals. The energy-type estimates we shall use in establishing the desired results may be derived formally as if the solutions in question were C^∞-functions all of whose partial derivatives lie in L_2. The resulting inequalities are then justified for solutions of finite regularity on the basis of the continuous dependence result in Theorem 2.1. In the calculations below, the dispersive and sometimes dissipative term $-Mu_x$ plays essentially no role.

The general relation that is effective here is to take the pth-derivative with respect to the spatial variable x, multiply the result by $\partial_x^s u$ and then integrate with respect to x over the entire line. Since the operator m_1 corresponding to the real part M_1 of the symbol m is self adjoint, this series of operations leads to the differential relation

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|\partial_x^s u(\cdot,t)|_2^2 = &-\int_{-\infty}^{\infty} \partial_x^s u(x,t)\partial_x^{s+1}u^{p+1}(x,t)\,dx \\ &-\int_{-\infty}^{\infty}(\xi^2)^s\xi m_2(\xi)|\widehat{u}(\xi,t)|^2\,d\xi.\end{aligned} \tag{2.5}$$

For $s=0$ in (2.5) we obtain that the L_2-norm is a decreasing function of t so that

$$|u(\cdot,t)|_2 \le |u_0|_2 \tag{2.6}$$

for all t for which the solution exists. Of course, if $m_2 \equiv 0$, the L_2-norm is time-independent and equality holds in (2.6). Using (2.6) and (2.5), the case $s=1$ does not appear to provide information in the absence of a more specific assumption on m_2, but one may obtain bounds on higher-order, L_2-based seminorms. For example, if $s=2$, integration by parts in (2.5) leads to the relation

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|u_{xx}|_2^2 = &-\frac{5(p+1)p}{2}\int_{-\infty}^{\infty} u^{p-1}u_x u_{xx}^2\,dx \\ &+\frac{(p+1)p(p-1)(p-2)}{4}\int_{-\infty}^{\infty} u^{p-3}u_x^5\,dx \\ &-\int_{-\infty}^{\infty}\xi^5 m_2(\xi)|\widehat{u}(\xi,t)|^2\,d\xi.\end{aligned} \tag{2.7}$$

Dropping the last term in (2.7) and using straightforward estimates and interpolation of function classes between L_2 and H^2 produces the differential inequality

$$\frac{1}{2}\frac{d}{dt}|u_{xx}|_2^2 \le c_p|u|_2^{(3p-2)/4}|u_{xx}|_2^{(p+10)/4},$$

where c_p is a constant depending only on p. Using (2.6) leads to

$$\frac{1}{2}\frac{d}{dt}|u_{xx}|_2^2 \le c_p|u_0|_2^{(3p-2)/4}|u_{xx}|_2^{(p+10)/4}.$$

The latter differential inequality gives directly the upper bound

$$|u_{xx}(\cdot,t)|_2^{(p+2)/4} \le \frac{|u_0''|_2^{(p+2)/4}}{1-((p+2)/4)c_p|u_0|_2^{(3p-2)/4}|u_0''|_2^{(p+2)/4}t} \tag{2.8}$$

for $0 \le t < d_p/|u_0|_2^{(3p-2)/4}|u_0''|_2^{(p+2)/4}$, where $d_p = 4/(p+2)c_p$ is a constant depending only on p.

In consequence, if $t \le T_0 = 2/(p+2)c_p|u_0|_2^{(3p-2)/4}|u_0''|_2^{(p+2)/4}$, then it follows that

$$|u_{xx}(\cdot,t)|_2 \le 2^{4/(p+2)}|u_0''|_2. \tag{2.9}$$

Once this H^2-bound is in hand, it is straightforward to derive bounds on H^s-seminorms, $s>2$, which depend only on the H^2-bound already in hand, and which therefore apply exactly on the time interval $[0,T_0]$.

We pass over the details, which parallel those presented in the case $s = 2$.

This concludes the proof of the corollary. □

Remark 2.4. If in the last result, the initial data has small amplitude and long wavelength, then it can be written in the form $u_0(x) = \varepsilon\varphi(\lambda^{-1}x)$ where φ is an order-one function. In this case, a calculation shows that $T_0 = C\lambda\varepsilon^{-p}$, where C depends on p and on norms of φ, and so is an order-one quantity that is independent of ε and λ. Note in particular that if the amplitude and wavelength are balanced in the way explained in Section 1, then $\varepsilon = \delta^{1/p}$ and $\lambda = \delta^{-1/\alpha}$ so that T_0 has order $\delta^{-(1+\alpha)/\alpha}$, which is exactly the time scale over which formal considerations indicate that the weak effect of nonlinearity and dispersion can have an order-one effect on the wave profile.

Remark 2.5. In many cases, including most of the examples that appear in Section 4, the imaginary part m_2 of the symbol m vanishes. In these cases, one can sometimes derive bounds that do not become infinite in finite time by making use of the dispersion operator M. In fact, for sufficiently smooth solutions u of (1.1), the quantity

$$\int_{-\infty}^{\infty}\left[\frac{1}{2}u(x,t)Mu(x,t) - F(u(x,t))\right]dx \tag{2.10}$$

is time-independent, where $F' = f$ and $F(0) = 0$. In conjunction with (2.5), and for suitable symbols m and nonlinearities f, (2.10) yields time-independent bounds on the semi-norm $(\int uMu\,dx)^{1/2}$ which depend only on $|u_0|_2$ and the value of this semi-norm at u_0. If the symbol $m(\xi)$ grows fast enough as $\xi \to \pm\infty$, these bounds translate into global solutions of (1.1) (see [**36**, **1**]).

3. The comparison theorem.

The principal technical result for equations of type (1.1) is formulated and proved below as Theorem 3.2.

Suppose there are given two symbols m and n corresponding to dispersion operators M and N, respectively. Consider the initial-value problems

$$u_t + u^p u_x - Mu_x = 0, \tag{3.1}$$

$$v_t + v^p v_x + bv^q v_x - Nv_x = 0, \tag{3.2}$$

where p and q are integers and $q > p$, with initial data

$$u(x,0) = v(x,0) = \varepsilon\varphi(\varepsilon^\gamma x), \tag{3.3}$$

where γ is a fixed, positive number. Rescale u and v by the relations

$$\begin{aligned} u(x,t) &= \varepsilon U(\varepsilon^\gamma x, \varepsilon^\beta t),\\ v(x,t) &= \varepsilon V(\varepsilon^\gamma x, \varepsilon^\beta t), \end{aligned} \tag{3.4}$$

where $\beta = p + \gamma$. Then U and V satisfy the initial-value problems

$$\begin{aligned} U_t + U^p U_x - M_\varepsilon U_x &= 0,\\ V_t + V^p V_x + b\varepsilon^{q-p} V^q V_x - N_\varepsilon V_x &= 0, \end{aligned} \tag{3.5}$$

with

$$U(x,0) = V(x,0) = \varphi(x).$$

Here, if F is a function of the spatial variable x, then $M_\varepsilon F$ is defined by its Fourier transform as

$$\widehat{M_\varepsilon F}(\xi) = \varepsilon^{-p} m(\varepsilon^\gamma \xi)\widehat{F}(\xi) = m_\varepsilon(\xi)\widehat{F}(\xi), \tag{3.6}$$

and similarly for N_ε. It is to this pair of rescaled initial-value problems that the next result speaks.

Lemma 3.1. *Suppose γ and p are such that for all ξ and sufficiently small ε the symbols m_ε and n_ε defined in (3.6) satisfy*

$$|m_\varepsilon(\xi) - n_\varepsilon(\xi)| \le \varepsilon |P_{r-1}(\xi)|, \tag{3.7}$$

for all $\xi \in \mathbf{R}$, where $r \ge 2$ is an integer, and P_{r-1} is a polynomial of degree $r-1$. Let $\varphi \in H^{k+r}$ where $k \ge 0$. Suppose that the initial-value problems (3.5) are well posed in the sense expressed in Theorem 2.1, in $C(0,T;H^{k+r})$ for some $T > 0$ and that the H^{k+r}-norms of U and V are bounded on $[0,T]$ with a bound that depends only on the norm of φ in H^{k+r}, and not on ε at least for ε small. Then there exists an $\varepsilon_0 > 0$ and constants B_j, such that for $0 \le t \le \min\{1,T\}$ and $0 < \varepsilon \le \varepsilon_0$,

$$\left|\partial_x^j\left(U(\cdot,t) - V(\cdot,t)\right)\right|_2 \le \varepsilon t B_j, \tag{3.8}$$

for $0 \le j \le k$. The constants B_j depend only on T and the norms of the solutions U and V in $C(0,T;H^{k+r})$.

Proof. The method employed here to establish this technical fact is to define w as the difference $U-V$ and then apply energy-type arguments for its estimation. Toward this end, note first that w satisfies the initial-value problem

(3.9)
$$\begin{aligned} w_t &+ \sum_{i=0}^{p} \frac{1}{p+1-i}\binom{p}{i} V^i (w^{p+1-i})_x \\ &+ \sum_{i=0}^{p-1} \frac{1}{i+1}\binom{p}{i} (V^{i+1})_x w^{p-i} - M_\varepsilon w_x \\ &= (M_\varepsilon - N_\varepsilon)V_x - b\varepsilon^{q-p} V^q V_x, \qquad w(x,0) \equiv 0, \end{aligned}$$

at least in the sense of tempered distributions. The inequalities in (3.8) will be established by induction on j. In what follows, calculations will be made as if the solutions U and V are C^∞-functions, all of whose derivatives lie in L_2. The formulas that result therefrom will only involve spatial derivatives of order less than or equal to $k+r$. As in Theorem 2.1, these formulas may then be justified for initial data φ in H^{k+r} by taking recourse to the continuous-dependence result in Theorem 2.1.

To begin, we write a master, energy-type relation. Differentiate the equation in (3.9) j times with respect to x and multiply the result by $\partial_x^j w = w_{(j)}$, where a new notation has been introduced for partial derivatives with respect to the spatial variable x. Upon integrating the equation that arises from the just-described operations with respect to x over the entire real line, with respect to t over the interval $[0,t]$ where $t \le T$, and after suitable integrations by parts, using the fact that $w(\cdot,0) \equiv 0$, there appear the relationships

(3.10)
$$\begin{aligned} |w_{(j)}(\cdot,t)|_2^2 = &-\sum_{i=0}^{p} 2\alpha_i \int_0^t ([V^i(w^{p+1-i})_x]_{(j)}, w_{(j)})\,ds \\ &- \sum_{i=0}^{p-1} 2\beta_i \int_0^t ([(V^{i+1})_x w^{p-i}]_{(j)}, w_{(j)})\,ds \\ &+ 2\int_0^t ((M_\varepsilon - N_\varepsilon)V_{(j+1)}, w_{(j)})\,ds \\ &- 2b\varepsilon^{q-p}\int_0^t ((V^q V_x)_{(j)}, w_{(j)})\,ds \end{aligned}$$

for $0 \le j \le k$, where

$$\alpha_i = \frac{\binom{p}{i}}{p+1-i} \quad \text{and} \quad \beta_i = \frac{\binom{p}{i}}{i+1}.$$

As mentioned above, the calculations leading to (3.10) are straightforwardly justified.

Formula (3.10) will be used inductively to derive the bounds advertised in (3.8). Consider first the case $j=0$ for which (3.10) may be written in the form

(3.11)
$$\begin{aligned} |w(\cdot,t)|_2^2 = &-\sum_{i=0}^{p} 2\alpha_i \int_0^t \int_{-\infty}^{\infty} V^i (w^{p+1-i})_x w\,dx\,ds \\ &- \sum_{i=0}^{p-1} 2\beta_i \int_0^t \int_{-\infty}^{\infty} (V^{i+1})_x w^{p-i+1}\,dx\,ds \\ &+ 2\int_0^t \int_{-\infty}^{\infty} (M_\varepsilon - N_\varepsilon)V_x w\,dx\,ds \\ &- 2b\varepsilon^{q-p}\int_0^t \int_{-\infty}^{\infty} V^q V_x w\,dx\,ds. \end{aligned}$$

Note that in the first sum, the term corresponding to $i=0$ is zero. Estimating the first and second terms on the right-hand side of (3.11) in a standard way and applying Plancherel's theorem, the Cauchy-Schwarz inequality and (3.7) to the third term on the right leads to the inequality

(3.12)
$$\begin{aligned} |w(\cdot,t)|_2^2 \le\ & 2\sum_{i=1}^{p} \frac{p-i+1}{p-i+2}\alpha_i \int_0^t |(V^i)_x|_\infty |w^{p-i}|_\infty |w|_2^2\,ds \\ &+ 2\sum_{i=0}^{p-1} \beta_i \int_0^t |(V^{i+1})_x|_\infty |w^{p-i-1}|_\infty |w|_2^2\,ds \end{aligned}$$

$$+ \varepsilon c_0 \int_0^t \|V\|_r |w|_2 \, ds$$
$$+ \frac{2b}{q+1} \varepsilon^{q-p} \int_0^t \|V^{q+1}\|_1 |w|_2 \, ds,$$

where

$$c_0 = \sup_{\xi \in \mathbf{R}} \frac{2|\xi P_{r-1}(\xi)|}{(1+\xi^2)^{r/2}}. \tag{3.13}$$

If we define

$$A_0 = \max_{0 \le t \le T} \Big(2\sum_{i=0}^{p} \frac{p-i+1}{p-i+2} \alpha_i |(V^i)_x(\cdot,t)|_\infty |w^{p-i}(\cdot,t)|_\infty$$
$$+ 2\sum_{i=0}^{p-1} \beta_i |V^{i+1}(\cdot,t)|_\infty |w^{p-i-1}(\cdot,t)|_\infty \Big)$$
$$\le 2\sum_{i=0}^{p} \alpha_i i \|V\|^{i-1}_{C(0,T;H^1)} \|V\|_{C(0,T;H^2)} \|w\|^{p-i}_{C(0,T;H^1)}$$
$$+ 2\sum_{i=0}^{p-1} \beta_i \|V\|^{i+1}_{C(0,T;H^1)} \|w\|^{p-i-1}_{C(0,T;H^1)}$$

and

$$C_0 = \max_{0 \le t \le T} \Big(c_0 \|V(\cdot,t)\|_r + \frac{2b}{q+1} \varepsilon^{q-p-1} \|V^{q+1}(\cdot,t)\|_1 \Big),$$

where c_0 is as in (3.13), then Gronwall's lemma implies that

$$|w(\cdot,t)|_2 \le \varepsilon C_0 \frac{e^{A_0 t/2} - 1}{A_0} \le \varepsilon B_0 t \tag{3.14}$$

provided that $0 \le t \le T$. Notice that B_0 depends on T and on the norm of the solutions U and V. Also note that (3.14) may be used to obtain a better estimate of A_0, namely that

$$A_0 \le 2p\|V\|^{p-1}_{C(0,T;H^1)} \|V\|_{C(0,T;H^2)} + 2\|V\|^p_{C(0,T;H^1)} + 0(\varepsilon)$$

as $\varepsilon \downarrow 0$. However, this refinement has no bearing on the issues at the fore here. Consider now the case $j = 1$ for which the master relation (3.10) may be put as

$$|w_x(\cdot,t)|_2^2 = -\sum_{i=1}^{p} (p-i+1)\alpha_i \int_0^t \int_{-\infty}^{\infty} (V^i)_x w^{p-i} w_x^2 \, dx\, ds$$
$$- \sum_{i=0}^{p-1} (p-i+1)(p-i)\alpha_i \int_0^t \int_{-\infty}^{\infty} V^i w^{p-i-1} w_x^3 \, dx\, ds$$
$$- \sum_{i=0}^{p-1} 2\beta_i \int_0^t \int_{-\infty}^{\infty} (V^{i+1})_{xx} w^{p-i} w_x \, dx\, ds$$
$$- \sum_{i=0}^{p-1} 2\beta_i (p-i) \int_0^t \int_{-\infty}^{\infty} (V^{i+1})_x w^{p-i-1} w_x^2 \, dx\, ds$$
$$+ 2\int_0^t \int_{-\infty}^{\infty} (M_\varepsilon - N_\varepsilon) V_{xx} w_x \, dx\, ds$$
$$+ \frac{2b}{q+1} \varepsilon^{q-p} \int_0^t \int_{-\infty}^{\infty} (V^{q+1})_{xx} w_x \, dx\, ds.$$

The second term on the right-hand side of the last relation is estimated using the embedding of $H^{1/6}$ into L_3 and interpolation. The other terms are estimated in obvious ways, the upshot being the inequality

$$|w_x(\cdot,t)|_2^2 \le \sum_{i=1}^{p} (p-i+1)\alpha_i \int_0^t |(V^i)_x|_\infty |w|_\infty^{p-i} |w_x|_2^2 \, ds$$
$$+ k_0 \sum_{i=0}^{p-1} (p-i+1)(p-i)\alpha_i \int_0^t |V|_\infty^i |w|_\infty^{p-i-1} \|w_x\|_{1/2} |w_x|_2^2 \, ds$$
$$+ \sum_{i=0}^{p-1} 2\beta_i \int_0^t |(V^{i+1})_{xx}|_\infty |w|_\infty^{p-i-1} |w|_2 |w_x|_2 \, ds \tag{3.15}$$
$$+ \sum_{i=0}^{p-1} 2(p-i)\beta_i \int_0^t |(V^{i+1})_x|_\infty |w|_\infty^{p-i-1} |w_x|_2^2 \, ds$$
$$+ \varepsilon c_0 \int_0^t \|V_x\|_r |w_x|_2 \, ds$$

$$+\frac{2b}{q+1}\varepsilon^{q-p}\int_0^t |(V^{q+1})_{xx}|_2|w_x|_2\,ds,$$

where k_0 is an embedding constant and c_0 is defined in (3.13). The third term on the right-hand side of (3.15) is further bounded above using (3.14) as follows:

$$(3.16)\quad \int_0^t |(V^{i+1})_{xx}|_\infty |w|_\infty^{p-i-1}|w|_2|w_x|_2\,ds \leq \varepsilon B_0\int_0^t |(V^{i+1})_{xx}|_\infty|w|_\infty^{p-i-1}|w_x|_2\,ds.$$

Much as before, if we let

$$A_1=\max_{0\leq t\leq T}\Big(\sum_{i=1}^{p}(p-i+1)\alpha_i|(V^i)_x|_\infty|w|_\infty^{p-i} + k_0\sum_{i=0}^{p-1}(p-i+1)(p-i)\alpha_i|V|_\infty^i|w|_\infty^{p-i-1}\|w_x\|_{1/2} + \sum_{i=0}^{p-1}2(p-i)\beta_i|(V^{i+1})_x|_\infty|w|_\infty^{p-i-1}\Big)$$

and

$$C_1=\max_{0\leq t\leq T}\Big(B_0\sum_{i=0}^{p-1}|(V^{i+1})_{xx}|_\infty|w^{p-i+1}|_\infty + c_0\|V_x\|_r + \frac{2b}{q+1}\varepsilon^{q-p-1}\|V^{q+1}\|_2\Big),$$

then Gronwall's inequality applied to (3.15) and (3.16) yields

$$(3.17)\qquad |w_x(\cdot,t)|_2\leq \varepsilon C_1\frac{e^{A_1t/2}-1}{A_1}\leq \varepsilon B_1 t$$

provided that $0\leq t\leq T$. Again, B_1 depends on T and on norms of the solutions U and V.

The proof is finished by an inductive step wherein the desired result (3.8) is assumed to be valid for $j<m$ where $m<k$, and then on that

basis the result is established for $j=m$. To this end, attention is given to the master relation (3.10) with $j=m$. First note that

$$(3.18)\quad \int_{-\infty}^{\infty}(V^i(w^{p-i+1})_x)_{(m)}w_{(m)}\,dx = \sum_{j=0}^{m}\binom{m}{j}\int_{-\infty}^{\infty}(V^i)_{(j)}(w^{p-i+1})_{(m-j+1)}w_{(m)}\,dx$$
$$= \int_{-\infty}^{\infty}V^i(w^{p-i+1})_{(m+1)}w_{(m)}\,dx + m\int_{-\infty}^{\infty}(V^i)_x(w^{p-i+1})_{(m)}w_{(m)}\,dx + \sum_{j=2}^{m}\binom{m}{j}\int_{-\infty}^{\infty}(V^i)_{(j)}(w^{p-i+1})_{(m-j+1)}w_{(m)}\,dx.$$

Expanding the derivatives of powers of w and integrating by parts the one term where $w_{(m+1)}$ appears, we come to an expression of the form

$$\int_{-\infty}^{\infty}F\,w_{(m)}^2\,dx+\int_{-\infty}^{\infty}G\,w_{(m)}\,dx,$$

where F is a polynomial in V, V_x, w and w_x and G is a polynomial in $V, V_x,\dots,V_{(m)}, w, w_x,\cdots,w_{(m-1)}$. Estimating the L_∞-norm of F and the L_2-norm of G, and using the induction hypothesis leads to an inequality of the form

$$\Big|\int_{-\infty}^{\infty}(V^i(w^{p-i+1})_x)_{(m)}(\cdot,t)w_{(m)}(\cdot,t)\,dx\Big| \leq a_i|w_{(m)}(\cdot,t)|_2^2+\varepsilon b_i|w_{(m)}(\cdot,t)|_2.$$

Making similar estimates of terms in the other sums on the right-hand side of (3.10), and combining these with the bounds

$$\Big|\int_{-\infty}^{\infty}(M_\varepsilon-N_\varepsilon)V_{(m+1)}w_{(m)}\,dx\Big|\leq \varepsilon c_0\|V\|_{r+m}|w_{(m)}|_2,$$

and

$$\Big|b\varepsilon^{q-p}\int_{-\infty}^{\infty}(V^{q+1})_{(m+1)}w_{(m)}\,dx\Big|\leq b\varepsilon^{q-p}\|V^{q+1}\|_{(m+1)}|w_{(m)}|_2,$$

there appears the inequality

$$|w_{(m)}(\cdot,t)|_2^2 \le A_m \int_0^t |w_{(m)}(\cdot,s)|_2^2\,ds + \varepsilon C_m \int_0^t |w_{(m)}(\cdot,s)|_2\,ds \tag{3.19}$$

from which it follows that

$$|w_{(m)}(\cdot,t)|_2 \le \varepsilon C_m \frac{e^{A_m t/2}-1}{A_m} \le \varepsilon B_m t \tag{3.20}$$

for $0 \le t \le T$. The constant B_m depends as before on norms of U and V, on T, and on the previous constants $B_0, \dots, B_{m-1}$. The inductive step being established, it is concluded that (3.8) holds for all $j \le k$.

The proof of the lemma is thereby finished. □

As an immediate corollary of this lemma and the transformations (3.4), the principal technical result to be used in the later sections is obtained.

Theorem 3.2. *Suppose that condition* (3.7) *is valid for the scaled operators M_ε and N_ε for some value of r and fixed values of p and γ, and let $\beta = p+\gamma$. Suppose the associated equations* (3.5) *scaled via* (3.4) *with these values of γ and β are both well posed on a time interval $[0,T]$ and that T and the $C(0,T;H^{k+r})$-norms of the solutions depend only on the H^{k+r}-norm of the initial data φ for some k with $k+r \ge 2$, and is independent of ε sufficiently small. Let $\varphi \in H^{k+r}$ be given and let u_ε and v_ε be the solutions of* (3.1) *and* (3.2)*, respectively, with initial data as in* (3.3)*. Then there are constants B_j, $0 \le j \le k$ which depend only on the norm of φ in H^{k+r} such that for $0 \le j \le k$,*

$$|\partial_x^j(u_\varepsilon - v_\varepsilon)|_2 \le \varepsilon^{2+\gamma(j-1/2)} B_j \varepsilon^\beta t \tag{3.21}$$

provided $0 \le t \le T\varepsilon^{-\beta}$. By interpolation, therefore, it follows that

$$|\partial_x^j(u_\varepsilon - v_\varepsilon)|_\infty \le \varepsilon^{2+\gamma j} C_j \varepsilon^\beta t \tag{3.22}$$

for $0 \le j < k$ and $0 \le t \le T\varepsilon^{-\beta}$, where $C_j = (B_j B_{j+1})^{1/2}$.

Remark 3.3. If we take $\lambda = \varepsilon^{-\gamma}$ in Remark 2.4 following the proof of Corollary 2.3, then a lower bound on the time scale over which (3.1) or (3.2) is well posed is $\varepsilon^{-p-\gamma} = \varepsilon^{-\beta}$, thus showing that the relevant initial-value problems are well posed over the advertised time scales.

Remark 3.4. It is worth specializing Theorem 3.2 to the case where the symbol m of the dispersion operator M is homogeneous and the symbol n of N is a higher-order perturbation of m. Thus consider the case $m(\xi) = |\xi|^\alpha$ and $n(\xi) = |\xi|^\alpha + a|\xi|^\nu$ where $\gamma\nu \ge p+1$. In this case, it is straightforward to see that

$$|m_\varepsilon(\xi) - n_\varepsilon(\xi)| \le a\varepsilon^{\gamma\nu - p}|\xi|^\nu \le a\varepsilon|\xi|^\nu,$$

as $\varepsilon \downarrow 0$, where m_ε and n_ε are the symbols of M_ε and N_ε, respectively (see (3.6)). In consequence, the condition (3.7) holds and the conclusions of Theorem 3.2 hold. It would thus appear that larger values of γ lead to better results, and that the power α appearing in the symbol m is irrelevant to the considerations. However, for an arbitrary value of γ, the symbol $m_\varepsilon(\xi)$ has the form $\varepsilon^{\gamma\alpha - p}|\xi|^\alpha$, whereas the nonlinearity has no ε-dependence at all. If $\gamma\alpha$ is less than p, the dispersive term features an inverse power of ε. In this case the detailed structure of the solution is expected to be dominated by dispersive effects for small values of ε, despite the gross, ε-independent bounds available from Theorem 3.2. Alternatively, if $\gamma\alpha$ is greater than p, the dispersive term has a positive power of ε attached, and it will follow from the arguments presented in the next section that dispersive effects do not make an order-one relative contribution to the wave profile on the time scale during which nonlinear effects accumulate to make an order one relative difference. Thus on one hand when $\gamma\alpha - p < 0$, we find a singular situation in which dispersive effects become increasingly dominant as ε becomes smaller, and on the other hand when $\gamma\alpha - p > 0$, dispersive effects become negligible for small ε. The choice $\gamma = p/\alpha$ is the only one that avoids both these situations and formally renders nonlinear and dispersive effects at the same order of approximation. Notice that with this choice, the power ν of the perturbation is no longer unrestrained. Instead, one has $\nu \ge (1+1/p)\alpha > \alpha$, which accords naturally with the idea that the perturbation should be of higher order.

4. Applications to equations of KdV type.

In this section a number of interesting examples will be set forth which show the efficacy of the theory developed in Section 3 to equations of KdV-type, as in (1.1) or more particularly, of the type displayed in (3.1). As already noted in Remark 3.3, the initial-value problem for such equations automatically satisfies the hypotheses of Theorem 3.2 concerning existence of solutions and of ε-independent bounds over an appropriate time scale. In consequence, the conclusions about comparison of solutions of two such equations enunciated in Theorem 3.2 will be available as soon as condition (3.7) is verified for the two dispersion relations associated to the relevant equations.

The examples listed below all arise naturally in practically important situations that require modeling of long waves.

A. *Higher-Order corrections of the dispersion relation.* The situation envisioned here is perhaps the most straightforward application of Theorem 3.2. The idea is that the two symbols m and n of the operators M and N are the same except that one has a higher-order correction to the modelling of dispersion. A paradigm for this situation is provided by the surface water-wave problem in which the full, linearized dispersion relation is

$$c(k) = \left(\frac{\tanh(k)}{k}\right)^{1/2} \tag{4.1}$$

in suitably normalized variables. The Korteweg-de Vries equation is obtained when c is replaced by the first two terms in its Taylor expansion about the origin, namely

$$c_{KdV}(k) = 1 - k^2/6, \tag{4.2}$$

which will be a good approximation provided only small values of k (long waves) are in question. It may happen that one needs to model dispersive effects more accurately, however, while still staying in the realm of long-wave models, and in this situation it is natural to take an additional term in the Taylor expansion, namely

$$\tilde{c}(k) = 1 - \frac{1}{6}k^2 + \frac{19}{360}k^4 \tag{4.3}$$

(cf. Abdelouhab et al. [1]). The model equations corresponding to (4.2) and (4.3) are

$$u_t + u_x + \frac{3}{2}uu_x + \frac{1}{6}u_{xxx} = 0$$

and

$$v_t + v_x + \frac{3}{2}vv_x + \frac{1}{6}v_{xxx} + \frac{19}{360}v_{xxxxx} = 0,$$

respectively. (The factor 3/2 in the nonlinear term comes out naturally in the usual non-dimensionalization of variables, as in Benjamin et al. [6]). The natural scaling for small-amplitude long waves on the surface of shallow water is that of (3.4) with $\gamma = 1/2$ and $\beta = 3/2$ (cf. Bona and Smith [16]). By moving to a travelling frame of reference, the linear translational term u_x can be eliminated, and one is then left with the two equations

$$u_t + \frac{3}{2}uu_x + \frac{1}{6}u_{xxx} = 0, \tag{4.4a}$$

$$v_t + \frac{3}{2}vv_x + \frac{1}{6}v_{xxx} + \frac{19}{360}v_{xxxxx} = 0 \tag{4.4b.}$$

Letting M connote the operator $-\partial_x^2/6$ and N the operator $-\partial_x^2/6 - 19\partial_x^4/360$, we find ourselves in the situation envisioned in (3.1)–(3.2) with $p = 1$. As mentioned already, the initial-value problems for the equations in (4.4) satisfy the hypotheses about existence of solutions. (Indeed, in this particular case, the relevant initial-value problems are globally well posed in H^k for $k \geq 0$ and possess time-independent bounds in H^j for $j = 0, 1, 2 \cdots, k$ for (4.4) and for $k = 0, 1, 2$ for (4.4) (see [**16, 34, 14, 18, 36, 37**]). Hence to apply Remark 3.4 in this situation it is only necessary to check condition (3.7) of Lemma 3.1. A straightforward calculation reveals that the symbols m_ε and n_ε of the associated operators M_ε and N_ε satisfy the relation

$$|m_\varepsilon(\xi) - n_\varepsilon(\xi)| = \frac{19}{360}\varepsilon\xi^4,$$

for all $\xi \in \mathbf{R}$. Thus taking $r = 5$ and supposing the initial data φ lies in H^{k+5} for some non-negative value of k, it is deduced at once that

$$|\partial_x^j(u - v)|_\infty \leq C_j\varepsilon^{2+j/2+3/2}t \tag{4.5}$$

for $0 \leq j < k$, provided $0 \leq t \leq \varepsilon^{-3/2}$, with similar estimates for the L_2-norm of the difference. One concludes that the inclusion of higher-order dispersive effects is without formal consequence at the level of modelling inherent in either equation in (4.4). In the Korteweg-de Vries equation written as in (4.4a) and with small-amplitude, long wavelength initial data $\varepsilon g(\varepsilon^{1/2}x)$, we know that nonlinear and dispersive effects accumulate to make an order-one relative contribution to the wave profile at time t of order $\varepsilon^{-3/2}$ (see Bona et al. [12]). Equally, at time t of order $\varepsilon^{-5/2}$, the error terms inherent in the Korteweg-de Vries model could in principle, and do in fact make an order-one relative contribution to the wave profile (see again Bona et al. [12]), thus rendering the Korteweg-de Vries approximation invalid. The same remarks apply to the extended model because higher-order nonlinear effects have not been included. Now, referring to (4.5), we see that while u and v are both of order ε, their difference is of order ε^2 at $t = \varepsilon^{-3/2}$. As ε^2 is the order that would be contributed by the neglected terms in either model, it is inferred that the effect of the higher-order dispersion relation in the extended model is of no consequence on time scales over which the neglected effects remain relatively insignificant.

B. *Dissipative effects.* It is often useful in both theoretical and practical investigations to include dissipative effects in a model that accounts for nonlinearity and dispersion. Such equations may take the form

$$v_t + v^p v_x - M v_x + L v = 0. \tag{4.6}$$

In the notation of Theorem 2.1, $l(\xi) = \xi m_2(\xi)$ where m_2 is the imaginary part of $m = m_1 + i m_2$. Because $m_2(\xi) \geq 0$ for $\xi \geq 0$ and m_2 is assumed to be an odd function, it follows that l is even and non-negative valued (cf. [**7, 8, 9, 10, 11, 25**]). Frequently encountered examples include the KdV-Burgers equation

$$v_t + v v_x + v_{xxx} - \nu v_{xx} = 0, \tag{4.7}$$

where $\nu > 0$ and the parabolic regularization

$$v_t + v v_x + v_{xxx} + \nu v_{xxxx} = 0 \tag{4.8}$$

of the KdV equation (cf. [**4, 28, 32, 33**] and the references contained therein for (4.7) and Abdelouhab et al. [**1**], Iório [**30**], Saut [**44**] and Temam [**49**] for (4.8)).

Here we consider the effect of perturbing equation (3.1) by a homogeneous dissipative operator L whose symbol $l(\xi) = b|\xi|^\mu$ for $b > 0$ and some $\mu \geq 0$. Suppose the dispersion relation $m(\xi) = |\xi|^\alpha$ is also homogeneous. As explained in the Introduction, the scaling appropriate to the unperturbed initial-value problem

$$u_t + u^p u_x - M u_x = 0, \qquad u(x,0) = \varepsilon\psi(\lambda^{-1}x),$$

is that in which $\varepsilon^p \lambda^\alpha$ is order one, or what is the same, the scaling for which λ has the same order as $\varepsilon^{-p/\alpha}$. Taking $\gamma = p/\alpha$, $\beta = p + \gamma$ and $N = M + L\partial_x^{-1}$, we find that

$$m_\varepsilon(\xi) = |\xi|^\alpha \quad \text{and} \quad n_\varepsilon(\xi) = |\xi|^\alpha + ib\varepsilon^{p((\mu-\alpha-1)/\alpha)}|\xi|^{\mu-1}\operatorname{sgn}(\xi).$$

Thus the relative effect of the dissipation operator L depends upon the exponent $\theta = p((\mu - \alpha - 1)/\alpha)$. In case $\mu > \alpha + 1$ so that θ is positive, then

$$|m_\varepsilon(\xi) - n_\varepsilon(\xi)| \leq b\varepsilon^\theta |\xi|^{\mu-1},$$

and Theorem 3.2 may be applied with ε^θ in place of ε, to deduce that

$$|\partial_x^j(u - v)|_\infty \leq \varepsilon^{1+\theta+(p/\alpha)j} C_j \varepsilon^\beta t \tag{4.9}$$

for $0 \leq t \leq \varepsilon^{-\beta}$.

For example, suppose $\mu = 4$, $\alpha = 2$ and $p = 1$ so that (4.8) is being viewed as a perturbation of the KdV equation. Then $\theta = 1/2$ and $|u - v|_\infty \leq C\varepsilon^3 t$. Thus while u and v are both order ε in the L_∞-norm, their difference is of order $\varepsilon^{3/2}$ at $t = \varepsilon^{-3/2}$. On the other hand, the Burgers-type dissipation in (4.7) has $\theta = -1/2$, and consequently this term does not constitute a small perturbation of the KdV operator.

C. *Higher-order correction of the nonlinearity.* In this subsection, consideration is given to the effect of including a higher-order nonlinear term in the model equation. A case that arises often in practice is the inclusion of a cubic nonlinearity in the Korteweg-de Vries equation. The appearance of $u^2 u_x$ as the next term in the approximation of nonlinear effects is explained in Benjamin et al. [**6**], where it is argued heuristically in effect that this term will generically account for nonlinear effects at the second order.

Attention is thus given to the two equations

$$u_t + u u_x + u_{xxx} = 0, \quad \text{and} \quad v_t + v v_x + c v^2 v_x + v_{xxx} = 0 \tag{4.10}$$

where c is a nonzero real number. Taking the scaling appropriate to the Korteweg-de Vries equation, namely $\gamma = 1/2$ and $\beta = 3/2$ in (3.4) and applying Theorem 3.2 leads directly to the conclusion that for $k \geq 2$, $g \in H^k$ and $0 \leq j < k$,

$$|\partial_x^j(u-v)|_\infty \leq t\, C_j \varepsilon^{2+j/2+3/2}, \tag{4.11}$$

at least for $0 \leq t \leq \varepsilon^{-3/2}$, just as in (4.5). Referring to the discussion in Section 4A, it is concluded that during the time period over which significant alteration of the initial profile takes place due to nonlinear and dispersive effects, the cubic nonlinearity remains relatively negligible for data that satisfies the basic Korteweg-de Vries-type scaling.

D. *Comparison between the Korteweg-de Vries equation and Smith's equation.* This comparison is a little more subtle than the simple perturbations featured in Sections 4A–C. The evolution equation

$$u_t + uu_x - Mu_x = 0, \tag{4.12}$$

where the symbol m of M is $m(\xi) = \sqrt{1+\xi^2} - 1$, was derived by Smith [47] as a model for continental shelf waves. (The form of the symbol in (4.12) corrects a minor oversight in Smith's paper.) Because m is smooth and has the approximate form $\xi^2/2$ near $\xi = 0$, it is natural to ask whether or not an appropriate version of the Korteweg-de Vries equation might be just as good as a model for the phenomena in question. This depends upon the scaling assumption that applies to the initial data. If the waves to which the model is to be applied are adequately represented by the scaling $\varepsilon g(\varepsilon^{1/2}x)$, as is implicitly assumed by Smith, then we will show now that one might as well use the Korteweg-de Vries equation as a model.

Turning to a detailed analysis of the last assertion, we attempt to apply Theorem 3.2 to equation (4.12) and the Korteweg-de Vries equation in the form

$$v_t + vv_x + v_{xxx}/2 = 0. \tag{4.13}$$

Again we choose $\gamma = 1/2$ and $\beta = 3/2$ which is consistent with the use of the Korteweg-de Vries equation. The crux of the matter is to establish condition (3.7) in Lemma 3.1. As the symbol m is not homogeneous, this is slightly more complicated than for the perturbation featured in Section 4A. The operator M_ε has a symbol m_ε given by

$$m_\varepsilon(\xi) = \frac{\sqrt{1+\varepsilon\xi^2} - 1}{\varepsilon},$$

whereas if $N = -\partial_x^2/2$, its symbol n_ε is $\xi^2/2$, and consequently

$$|m_\varepsilon(\xi) - n_\varepsilon(\xi)| \leq \varepsilon\xi^4/4.$$

Thus, supposing that $g \in H^{k+5}$ for some $k > 0$, it may be inferred from Theorem 3.2 that

$$|\partial_x^j(u-v)|_\infty \leq C_j \varepsilon^{2+j/2+3/2} t \tag{4.14}$$

for $0 \leq t \leq \varepsilon^{-3/2}$, for $0 \leq j \leq k-1$.

The conclusion one deduces from (4.14) is that for data that has an amplitude to wavelength relationship well approximated by the form $\varepsilon g(\varepsilon^{1/2}x)$, it does not matter whether (4.12) or (4.13) is used to model the wave evolution. Both give the same answer to within the inherent order of accuracy of either equation on the long time scale over which nonlinear and dispersive effects make an order-one relative contribution to the wave profile.

E. *Comparison between the Korteweg-de Vries equation and the intermediate long-wave equation.* This is an interesting application of the general theory to a comparison between the Korteweg-de Vries equation and the intermediate long-wave equation. The physical setting in which the latter equation arises is in a stratified fluid bounded above and below by planar, rigid boundaries. The fluid is assumed to consist of a homogeneous top layer of thickness H_1 and density ρ_1 lying over a homogeneous bottom layer of thickness H_2 and density $\rho_2 > \rho_1$. Provision can be made for a thin, transition layer separating the two homogeneous layers. In the special cases where either H_1 and H_2 are nearly equal, or in the case where one of H_1 or H_2 is much smaller than the other, and assuming also that small-amplitude long waves propagate in a single direction without variation in the direction transverse to that of the primary motion, one derives the model equation

$$u_t + uu_x - L_H u_x = 0, \tag{4.15}$$

where, for $H > 0$, the operator L_H is a Fourier multiplier operator as in (1.2) with symbol

$$l_H(\xi) = \frac{1}{H}\Big(\xi\coth(H\xi) - \frac{1}{H}\Big). \tag{4.16}$$

The parameter H is taken to be $H_1 \cong H_2$ in case H_1 and H_2 are nearly equal, whereas if $H_1 \ll H_2$, then $H = H_2$ or if $H_2 \ll H_1$, then $H = H_1$. For a discussion of the derivation of this model, see Kubota et al. [40] and the more recent commentary of Bona and Rose [13]. It is known that if a sufficiently smooth initial datum is specified, then in the scaling represented in (4.16) the solution $u = u_H$ tends to the solution of the Korteweg-de Vries equation with the same initial datum uniformly on bounded time intervals as H tends to zero (Abdelouhab et al. [1]). The result below is a much more precise rendition of the last-quoted theorem in that it provides detailed estimates of the difference between solutions of the intermediate long-wave equation and the Korteweg-de Vries equation over long time scales.

To this last-mentioned end, the theory developed in Section 3 is again brought to bear. We attempt to apply Theorem 3.2 with $M = L_H$ as above and $N = -\partial_x^2/3$ corresponding to the Korteweg-de Vries equation, and with $\gamma = 1/2$ and $\beta = 3/2$. A straightforward computation reveals that the symbols m_ε and n_ε of the operators M_ε and N_ε that appear after the change of variables (3.4) with the above values of γ and β, are

$$\frac{1}{\varepsilon H}\Big(\varepsilon^{1/2}\xi\coth(\varepsilon^{1/2}H\xi) - \frac{1}{H}\Big) \quad \text{and} \quad \frac{1}{3}\xi^2, \tag{4.17}$$

respectively. Using the elementary fact that

$$\coth(x) = 1/x + x/3 + x^3 h(x)$$

where h is a smooth function which is bounded on the real axis, it is easy to establish that

$$|m_\varepsilon(\xi) - n_\varepsilon(\xi)| \le \varepsilon c_H \xi^4, \tag{4.18}$$

where $c_H = H^2 h(\varepsilon^{1/2}H\xi)$. Hence, if an initial datum φ is presented in H^{k+5}, then there are constants C_j depending only on the norm of φ in this space such that for $0 \le t \le \varepsilon^{-3/2}$ and $0 \le j \le k-1$,

$$|\partial_x^j(u - v)|_\infty \le C_j H^2 \varepsilon^{2+j/2+3/2} t, \tag{4.19}$$

where u and v are the solutions of the intermediate long-wave equation and the Korteweg-de Vries equation, with initial data $\varepsilon\varphi(\varepsilon^{1/2}x)$, respectively. Notice that in this case, the result is uniform on any bounded interval of values of H, and that it improves when truly small values of H are considered, a fact consistent with the aforementioned result of Abdelouhab et al. [1].

F. *Comparison between the intermediate long-wave equation and the Benjamin-Ono equation.* In this comparison, interest is turned to large rather than small values of the parameter H appearing in the intermediate long-wave equation. According to the results of Abdelouhab et al. [1], if the intermediate long-wave equation is scaled appropriately, then solutions converge to the solution of the Benjamin-Ono equation in the limit as $H \to \infty$, uniformly on bounded time intervals. As remarked in the just-cited reference, the scaling that leads to the Benjamin-Ono limit is not the same as that which leads to the KdV equation. This will be reflected in the H-dependence of the forthcoming estimates, as it was in (4.18) when comparing the intermediate long-wave equation to the Korteweg-de Vries equation.

The intermediate long-wave equation is taken in the form (4.15) where now the symbol of the operator L_H is

$$l_H(\xi) = \xi\ \coth\ (\xi H) - \frac{1}{H}. \tag{4.20}$$

This evolution equation is to be compared with the Benjamin-Ono equation

$$v_t + vv_x - Nv_x = 0, \tag{4.21}$$

where the symbol n of the operator N is $n(\xi) = |\xi|$. Theorem 3.2 is applied with $\gamma = 1$ and $\beta = 2$. This is the natural scaling for the Benjamin-Ono equation when it is written in the form (4.21). Indeed, the evolution equation (4.21) is invariant under this scaling.

To apply our theory, it remains to estimate appropriately the left-hand side of (3.7) in the present circumstances. Reference to Lemma 4.1 in Abdelouhab et al. [1] leads to the conclusion

$$|\xi\coth(\mu\xi) - |\xi|| \le \frac{1}{\mu}, \tag{4.22}$$

uniformly for $\xi \in \mathbf{R}$. (Note that by evaluating the left-hand side of (4.22) at $\xi = 1/\mu$, one sees that this inequality is sharp in its dependence upon μ). Combining the last three relations with the change of variables (3.4) leads to the estimate

$$|l_{H,\varepsilon}(\xi) - n_\varepsilon(\xi)| \le C\frac{1}{\varepsilon H}. \tag{4.23}$$

An interesting point comes to the fore now, namely the relationship between the small amplitude parameter ε and the large depth parameter H. This issue has been discussed in some detail by Bona and Rose [13], the conclusion being that in order to achieve the Benjamin-Ono limit, it is natural to take $H = \varepsilon^{-2}$. With this proviso, (4.23) implies (3.7) with $r = 1$.

5. Other types of model equations. The foregoing theory is easily adapted to other types of nonlinear, dispersive wave equations. Included in the list of obvious candidates are regularized long-wave equations, certain versions of nonlinear Schrödinger equations, Boussinesq equations and Boussinesq systems. It is our purpose here to indicate how the conclusions presented earlier pertaining to Korteweg-de Vries-type equations carry over to some of these other models.

A. *Regularized long-wave equations.* As explained in Benjamin et al. [6] and Albert and Bona [3], the regularized long-wave equations

$$u_t + u_x + f(u)_x + Mu_t = 0 \tag{5.1}$$

are often useful as models in contexts where Korteweg-de Vries-type equations arise. Here, f and M are the operators defined in (1.1)–(1.2). Indeed, in the last-quoted reference, a theoretical discussion was undertaken of the relationship between the initial-value problems for (1.1) and (5.1) for small amplitude, long wavelength data.

A theory entirely analogous to that presented for models of the Korteweg-de Vries type may be developed for regularized long-wave equations. Such a theory may be constructed by following the arguments of Section 3 more or less line for line. The one difference that manifests itself is that (5.1) may be written in the pseudo-parabolic form

$$(I + M)u_t + u_x + f(u)_x = 0,$$

or equivalently

$$u_t + (I + M)^{-1}\partial_x(u + f(u)) = 0. \tag{5.2}$$

Because $(I + M)^{-1}$ is a smoothing operator, the theory of the initial-value problem for regularized long-wave equations is more straightforward than for the associated Korteweg-de Vries-type equation. A local existence theory along the lines of our Theorem 2.1 is easily concluded (see Albert and Bona [**3**, Lemma 3]). Moreover, if the symbol m of the dispersion operator M satisfies $m(\xi) \ge a|\xi|^\alpha$ for some positive constant a, at least for large values of $|\xi|$, then the initial-value problem is globally well posed in any L_2-based Sobolev space stronger than the space H_m provided that (i) $\alpha > 1$, or (ii) $\alpha = 1$ and f' grows at most cubically (see Albert and Bona [**3**, Theorem 2]). The space H_m is composed of those L_2-functions h such that

$$\|h\|^2_{H_m} = \int_{-\infty}^{\infty} (1 + m(\xi))|\hat{h}(\xi)|^2 d\xi$$

is finite.

Another way to make comparisons between solutions to the initial-value problems for the equation

$$u_t + u_x + u^p u_x + Mu_t = 0 \tag{5.3}$$

and, say,

$$v_t + v_x + v^p v_x + v^q v_x + Nv_t = 0, \tag{5.4}$$

is to use the theory of Albert and Bona [3] to compare (5.3) with (3.1) and (5.4) with (3.2), then use the theory of Section 3 to compare (3.1) and (3.2). The outcome of this somewhat round-about argument is the same as that gleaned by the more direct method of simply following the calculations in Section 3, but applying them to equations of the form displayed in (5.1). Here is a formal statement of the result in view. The context to which this result applies is that surrounding Lemma 3.1 and Theorem 3.2. That is, relative to (5.3) and (5.4), it is assumed that the initial data satisfies

$$u(x,0) = v(x,0) = \varepsilon\varphi(\varepsilon^\gamma x), \tag{5.5}$$

as in (3.3). Letting $\beta = p + \gamma$ as before, we further assume condition (3.7) to hold for the rescaled operators M_ε and N_ε as defined in (3.6). In this situation, and assuming (5.3) and (5.4) are well posed on the appropriate time interval, the following analog of Theorem 3.2 obtains.

Theorem 5.1. *In the situation delineated above, let u_ε and v_ε be the solutions of* (5.3) *and* (5.4), *respectively, that correspond to initial data as in* (5.5) *where $\varphi \in H^k$ for some $k \geq 1$. Then there are constants B_j, $0 \leq j \leq k$ which depend only on the norm of φ in H^k such that*

$$|\partial_x^j(u_\varepsilon - v_\varepsilon)|_2 \leq B_j \varepsilon^{2+\gamma(j-1/2)} \varepsilon^\beta t \tag{5.6}$$

for $0 \leq j \leq k$, provided $0 \leq t \leq \varepsilon^{-\beta}$. By interpolation, it therefore follows that

$$\left|\partial_x^j(u_\varepsilon - v_\varepsilon)\right|_\infty \leq C_j \varepsilon^{2+\gamma j} \varepsilon^\beta t \tag{5.7}$$

for $0 \leq j \leq k$ and $0 \leq t \leq \varepsilon^{-\beta}$, where $C_j^2 = B_j B_{j+1}$ for all the relevant values of j.

Remark 5.2. The examples provided in Section 4 for Korteweg-de Vries-type equations all have immediate analogues for the associated regularized long-wave equations.

B. *Nonlinear Schrödinger equations.* Attention is now turned to the initial-value problem for a class of nonlinear Schrödinger equations, namely

$$iu_t - Mu + g(u) = 0 \quad \text{and} \quad u(\cdot, 0) = \varphi(\cdot), \tag{5.8}$$

where $g : \mathbf{C} \to \mathbf{C}$, M is a Fourier multiplier operator as in (1.2), and $u = u(x,t)$ is a complex-valued function of the variables $x \in \mathbf{R}^n$ and $t \in \mathbf{R}$. Equations of the type exhibited in (5.8) are natural generalizations of the cubic Schrödinger equation

$$iu_t - \Delta u + |u|^2 u = 0 \tag{5.9}$$

that arises as an approximate model in several physical contexts. The initial-value problem for equations of the type

$$iu_t - \Delta u + f(x,u) = 0 \tag{5.10}$$

has been the object of considerable effort for the last several years, and the theory is well developed in certain respects, though much remains to be done. A recent and comprehensive review of the theory for (5.10) together with an extensive bibliography may be found in Cazenave [20]. A much studied generalization of the cubic Schrödinger equation has the nonlinearity $|u|^\sigma u$, where $\sigma > 0$. If σ is not an even integer, then such a nonlinearity has limited regularity at the origin, and this in turn will usually limit the regularity of the solution, regardless of how smooth the initial data proves to be. The interplay between the value of σ and how smooth solutions can be has been carefully exposed by Cazenave and Weissler [21].

We first consider the small-amplitude, long wavelength scaling associated with (5.8). Assume as before that the initial data φ has the form $\varphi(x) = \varepsilon\psi(\lambda^{-1}x)$. Suppose both the nonlinearity and the symbol m of the dispersion operator M are homogeneous; say $m(\xi) = |\xi|^\alpha$ and $g(z) = |z|^{2\nu} z$, with $\nu \geq 1$. The condition that nonlinear and dispersive effects be small and balanced is that ε and λ^{-1} are both small, but at the same time $S = \varepsilon^{2\nu}\lambda^\alpha$ is order one. This sort of condition is still valid as the requirement that weak nonlinear and dispersive effects be balanced, even for nonhomogeneous nonlinearities and dispersions provided they are dominated by homogeneous parts for small values of the independent variables. Guided by the relationship that S be order one, it is natural to make the change of variables

$$W(x,t) = \varepsilon u(\varepsilon^{2\nu/\alpha} x, \varepsilon^{2\nu-1} t),$$

corresponding to the choice $\lambda = \varepsilon^{-2\nu/\alpha}$ and an associated time scale. The new dependent variable is a solution of the equation

$$iW_t - \varepsilon MW + \varepsilon |W|^{2\nu} W = 0 \tag{5.11}$$

with order-one initial data $W(x,0) = \psi(x)$. As before, one infers formally from (5.11) that on a time scale of order $1/\varepsilon$, the nonlinear and dispersive effects can have an order-one effect on the wave profile. Assuming that neglected effects come in at order ε^2, we infer in the same formal way that these may have an order-one effect on the time scale $1/\varepsilon^2$. Thus, in these variables, one sees the model to present interesting nonlinear and dispersive effects on a time scale of order $1/\varepsilon$, but to be formally invalid on a time scale of order $1/\varepsilon^2$. In the original variables in which the initial data has small amplitude and long wavelength, this means that nonlinear and dispersive effects make an order-one relative difference to the wave profile on a time scale of order $\varepsilon^{-2\nu}$, while neglected effects may make substantial contributions on a time scale of order $\varepsilon^{-2\nu-1}$.

A local existence theory for (5.8) can be provided under various hypotheses on the symbol m and the nonlinearity g. For the comparison results that are the principal goal here, relatively smooth solutions are needed. In consequence, we assume here that g is a C^∞-mapping of $\mathbf{R}^2$ into $\mathbf{R}^2$. While this severely restricts the type of nonlinearity encompassed by our theory, it nevertheless applies to an interesting example to be presented shortly. In the case where g is smooth, a local existence theory is easily provided via semigroup theory (cf. [**34, 36**]). In applying this theory, it is perhaps easiest to write u in terms of its real and imaginary parts $v + iw$. Then v and w satisfy the system

$$w_t + Mv - g_1(v,w) = 0, \qquad v_t - Mw + g_2(v,w) = 0, \tag{5.12}$$

where $g = g_1 + i\,g_2$. If one writes U for the column vector (v,w), then (5.12) has the form

$$U_t + A(U) = F(U) \tag{5.13a}$$

where

$$A = \begin{pmatrix} 0 & -M \\ M & 0 \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} g_2 \\ -g_1 \end{pmatrix}. \tag{5.13b}$$

(The mappings A and F correspond to those with the same appellation in Kato [**34**].) The linear portion of (5.13) corresponds to the evolution equation

$$z_{tt} + M^2 z = 0, \tag{5.14}$$

which is satisfied by both dependent variables. This is a generalized version of the 'good' Boussinesq equation. It is purely dispersive in character, and the evolution it generates defines an isometry in any of the L_2-based spaces H^s. It therefore certainly satisfies the criteria imposed on the operator A in Kato [**34**]. Because g is taken to be smooth, g_1 and g_2 define locally Lipschitz mappings of H^s provided that $s > n/2$ where n is the dimension of the underlying spatial domain $\mathbf{R}^n$. These simple facts combined with Kato's theory yield the following local existence theory, which suffices for the present purposes.

Proposition 5.1. *Let $s > n/2$, let $g : \mathbf{C} \to \mathbf{C}$ be smooth, and suppose the real-valued, even symbol m of the dispersion operator M to satisfy the polynomial growth condition* (2.1). *Then corresponding to initial data $\varphi \in H^s$, there exists a $T = T(\|\varphi\|_s) > 0$ and a unique solution u in $C(0,T;H^s)$ of the initial-value problem* (5.1). *Moreover, T and the mapping that associates to φ the solution u of* (5.1) *with initial data φ are continuous from H^s into $\mathbf{R}^+$ and $C(0,T;H^s)$, respectively.*

As in Section 2, one can provide an explicit lower bound on the time interval T of existence by differentiating equation (5.8) with respect to the spatial variable s times, multiplying by $\partial_x^s\bar{u}$, integrating over $\mathbf{R}^n$ and considering the imaginary part of the result. The resulting differential inequality is of such a form that, for suitable values of s, it yields a finite upper bound on $|\partial_x^s u(\cdot,t)|_2$ on a time interval inversely proportional to the H^s-norm of the initial data φ. In the case of a homogeneous nonlinearity $g(u) = \lambda|u|^{2k}u$, the form of the upper bound can be presented explicitly.

Attention is now given to comparisons between two different Schrödinger equations. We content ourselves here with comparing solutions of the following, relatively concrete Schrödinger equations, namely

$$iu_t + a|u|^{2k}u - Mu = 0 \tag{5.15}$$

and

$$iv_t + a|v|^{2k}v + b|v|^{2q}v - Nv = 0, \tag{5.16}$$

where k and q are integers with $q > k$, and M and N are Fourier multiplier operators with symbols m and n, respectively, as in (1.2).

Imagine both equations to be presented with the same initial data $\varphi(x)$, and guided by the earlier remarks about the balance between nonlinear and dispersive effects, suppose $\varphi(x) = \varepsilon\psi(\varepsilon^{2k/\alpha}x)$ for $x \in \mathbf{R}^n$, where $\alpha > 0$ reflects the lowest order portion of the symbol m of

M. Make the changes of variables

(5.17) $$u(x,t) = \varepsilon U(\varepsilon^{2k/\alpha}x, \varepsilon^{2k}t), \qquad v(x,t) = \varepsilon V(\varepsilon^{2k/\alpha}x, \varepsilon^{2k}t).$$

The new dependent variables U and V are found to satisfy the initial-value problems

(5.18) $$\begin{gathered} iU_t + a|U|^{2k}U - M_\varepsilon U = 0, \\ iV_t + a|V|^{2k}V + b\varepsilon^{2(q-k)}|V|^{2q}V - N_\varepsilon V = 0, \\ U(x,0) = V(x,0) = \psi(x), \end{gathered}$$

for $x \in \mathbf{R}^n$, where a and b are constants and M_ε and N_ε are defined as before (see (3.6)). In this circumstance, the following result applies.

Theorem 5.2. *Suppose that for all ξ and sufficiently small ε the dispersion operators M_ε and N_ε in* (5.18) *have symbols m_ε and n_ε which satisfy the inequality*

(5.19) $$|m_\varepsilon(\xi) - n_\varepsilon(\xi)| \le \varepsilon^\mu |P_r(\xi)|,$$

where r is a nonnegative integer, P_r is a polynomial of degree r, and $\mu \ge 1$. Suppose $\psi \in H^{s+r}$ where $s > 0$ and $s + r > n/2$ and that both initial-value problems in (5.18) *are well posed in the sense of Proposition* 5.1 *in $C(0,T;H^{s+r})$ for some $T > 0$. Then there exists $\varepsilon_0 > 0$ and constants B_j depending only on norms of ψ, $0 \le j \le s$, such that for $0 \le t \le \min\{1,T\}$ and $0 < \varepsilon \le \varepsilon_0$,*

(5.20) $$|\partial_x^j(U(\cdot,t) - V(\cdot,t))|_2 \le \varepsilon^\nu t B_j,$$

where $\nu = \min\{\mu, 2(q-k)\}$.

As an immediate corollary of Theorem 5.2, we have the following result stated in terms of the original variables u and v.

Corollary 5.3. *Suppose the hypotheses of Theorem* 5.2 *to hold and that $T \ge 1$. Let u and v be the solution of* (5.15) *and* (5.16), *respectively, with initial data $\varphi(x) = \varepsilon\psi(\varepsilon^{2k/\alpha}x)$. Then for $0 \le t \le \varepsilon^{-2k}$, one has*

(5.21) $$|\partial_x^j(u - v)|_2 \le B_j t \varepsilon^{1+\nu+2k+(k/\alpha)(2j-1)}$$

for $0 \le j \le s$. By interpolation, it follows that

(5.22) $$|\partial_x^j(u - v)|_\infty \le C_j t \varepsilon^{1+\nu+2k+2kj/\alpha}$$

for $0 \le t \le \varepsilon^{-2k}$ and $0 \le j \le s-1$.

Remark 5.4. Consider the special case wherein $\alpha = 2$, $m|\xi| = |\xi|^2$, and $k = 1$ that corresponds to taking the cubic Schrödinger equation (5.9) as the initial model. An interesting perturbation of this model is the equation

(5.23) $$iv_t + |v|^2 v + b|v|^4 v + \Delta v - \Delta^2 v = 0$$

corresponding to adding higher order nonlinear and dispersive terms. In this case, $q = 2$, $\mu = 2$ and $r = 4$, and the estimates in (5.22) becomes

(5.24) $$|\partial_x^j(u - v)|_\infty \le C_j t \varepsilon^{5+j}$$

for $0 \le t \le \varepsilon^{-2}$. As $\partial_x^j u$ and $\partial_x^j v$ both have size of order ε^{1+j}, we see that at a time t of order ε^{-2}, the difference $\partial_x^j(u - v)$ has size of order ε^{3+j}, showing very convincingly that the higher order terms in (5.23) are not playing a serious role on this time scale.

C. *Higher-order water-wave equation compared to the Korteweg-de Vries equation.* As explained earlier, the Korteweg-de Vries equation represents only a relatively low-order approximation to the complete description of small-amplitude, long-wavelength surface water waves. Starting with the full Euler equations for two-dimensional waves on the surface of a perfect, irrotational fluid being acted upon by gravity, one may systematically derive approximations to any formal order. Carrying the approximation to the next order after that which yields the Korteweg-de Vries equation leads to the model

(5.25) $$\begin{gathered} v_t + \frac{3}{2}vv_x + \frac{1}{6}v_{xxx} - \frac{3}{8}v^2 v_x + \frac{23}{24}v_x v_{xx} + \frac{5}{12}vv_{xxx} - \frac{19}{360}v_{xxxxx} = 0, \\ v(x,0) = \varepsilon\psi(\lambda x), \end{gathered}$$

(cf. Olver [**41**]). Here $v(x,t)$ is proportional to the local amplitude of the wave expressed as the deviation of the free surface from its equilibrium position at the spatial point x along the channel at time t. The parameters ε and λ are measures of the amplitude and wavelength, respectively, and they satisfy the conditions noted in the introduction that ε and λ^{-1} are small and the Stokes number $S = \varepsilon\lambda^2$ is of order one. As before, ψ and its derivatives are all of order one and surface-tension effects have been neglected, as is often appropriate for long waves.

According to the formal ideas outlined in part 4A, one expects the last four terms in equation (5.25) to comprise higher-order corrections to the basic Kortweg-de Vries equation. Consequently, it is to be expected that these terms will make no substantial contribution to the evolution of small-amplitude, long wavelength initial data over the time scale corresponding to the validity of the Korteweg-de Vries equation. (In this regard, it is worth noting that Craig **[24]** has shown conclusively that an appropriate solution of the Korteweg-de Vries equation provides a good approximation to the solution of the full Euler equations over the formal time scale $0 \le t \le \varepsilon^{-3/2}$.)

Indeed, the content of parts 4A and 4B is that the inclusion of the fourth and seventh terms in equation (5.25) is without consequence in that they only make a small relative contribution to the wave profile. Equation (5.25) and the Korteweg-de Vries equation as written in (4.4a) have the form (3.1) and (3.2) with $p = 1$, $M = -\partial_x^2$, $q = 2$ and $N = -\partial_x^2 - \partial_x^4$ except for the fifth and sixth terms v_xv_{xx} and vv_{xxx} in (5.25). While not covered explicitly by our Theorem 3.2, these terms may also be determined to make a negligible contribution by the same sort of energy estimates used previously. Indeed let v be a sufficiently smooth solution of (5.25) and u the solution of the Korteweg-de Vries equation in (4.4a) with the same initial data. Questions of existence, uniqueness and appropriate bounds on solutions corresponding to regular initial data have already been dealt with by Saut **[44]** and Ponce **[42]**. After performing the change of variables (3.4) with $\gamma = 1/2$ and $\beta = 3/2$ we obtain new dependent variables U and V, and letting $w = U - V$, it is ascertained that w satisfies the forced evolution equation

$$(5.26)\quad \begin{aligned} w_t + \frac{3}{2}(Vw)_x + \frac{3}{2}ww_x + \frac{1}{6}w_{xxx} + \frac{3}{8}\varepsilon V^2V_x - \frac{23}{24}\varepsilon V_xV_{xx} \\ - \frac{5}{12}\varepsilon VV_{xxx} - \frac{19}{360}\varepsilon V_{xxxxx} = 0, \end{aligned}$$

with zero initial data at $t = 0$. This is the analog of equation (3.9) which formed the basis for the proof of Lemma 3.1. In this case we see explicitly that $q = 2$ and that

$$(5.27)\quad |m^\varepsilon(\xi) - n^\varepsilon(\xi)| \le \frac{19}{360}\varepsilon\xi^4.$$

Hence the general energy relation (3.10) will be the same in this case save for the additional terms which are

$$(5.28)\quad \begin{aligned} \frac{23}{24}\varepsilon \int_0^t \int_{-\infty}^{\infty} (V_xV_{xx})_{(j)}w_{(j)}\,dx \\ + \frac{5}{12}\varepsilon \int_0^t \int_{-\infty}^{\infty} (VV_{xxx})_{(j)}w_{(j)}\,dx\,ds. \end{aligned}$$

Thus if one adds to our previously defined constant C_j the quantity

$$\frac{23}{24}\sum_{k=0}^{j}\binom{j}{k}|V_{(k+1)}|_\infty\|V\|_{j-k+2} + \frac{5}{12}\sum_{k=0}^{j}\binom{j}{k}|V_{(k)}|_\infty\|V\|_{j-k+3},$$

then inequality (3.19) will hold for the present version of $w_{(j)}$. In consequence, (3.20) will also hold for the $w_{(j)}$ under consideration here and this translates into the relations

$$|\partial_x^j(u - v)|_\infty \le C_j\varepsilon^{2+j/2+3/2}t$$

for $0 \le t \le \varepsilon^{-3/2}$ provided $0 \le j \le k - 1$, where the initial data ψ lies in H^{k+5} with $k \ge 0$, and the constant C_j still depends only on the norms of u and v in $C(0, T; H^{k+5})$. As mentioned above, appropriate bounds on norms of v have already been provided in the works of Saut **[44]** and Ponce **[42]**. Thus we obtain the same result as that which appeared in (4.5), and the conclusion is that all the higher-order terms are sensibly irrelevant to the waves' evolution over the time interval $0 \le t \le \varepsilon^{-3/2}$ in which the nonlinear and dispersive terms uu_x and u_{xxx} make a relative contribution of order one to the shape and speed of propagation of the initial wave profile.

It deserves remark in the context of (5.25) that this equation contains the complete formal description of the propagation of small-amplitude long waves on the surface of an ideal fluid through order ε^2, where ε measures the wave amplitude. One therefore expects this equation to faithfully track solutions of the Euler equation over the time scale $\varepsilon^{-5/2}$ in the variables pertaining to (5.25), and hence to conclude that it may have interest as a model beyond that of the KdV equation. This puts (5.25) in a substantially different status than the perturbed equations in (4.4b) and (4.6). The recently derived model

$$u_t - 3uu_x + 2u_x u_{xx} + uu_{xxx} - u_{xxt} = 0$$

of Camassa and Holm [**19**], which has some of the features of the second-order model (5.25), suffers in the same way as (4.4b) and (4.6) from being an incomplete description at second order, and hence our theory would show it to be no more effective then the KdV equation as a model for the propagation of small-amplitude long waves on the surface of an ideal fluid under the force of gravity.

5. Conclusion. In the body of this paper, we have sought to understand in a mathematically exact manner the consequences of adding higher-order terms to model equations for long-wave propagation. We find that in many practically interesting situations, the inclusion of such terms is without consequence relative to certain, long time scales.

Such results are potentially valuable in several respects. As exemplified in Section 4D, one may sometimes conclude that a simpler model yields comparable results, so preferring the KdV equation to Smiths equation when $a\lambda^2$ is of order one. Alternatively, one may add a term or otherwise modify the model equation for mathematical or numerical convenience in certain ways without changing the accuracy appreciably. This is seen clearly in the arguments leading to the family (5.2) of regularized long-wave equations (Benjamin et al. [**6**]), and is also present implicitly in commentary pointing to the study of the KdV equation with an extended dispersion relation as in (4.4b).

Similar theory can be worked out for comparing solutions of scalar Boussinesq equations such as

$$u_{tt} = u_{xx} + (u^2 + u_{xx})_{xx},$$

though there are a few interesting points that intrude into the analysis. However, Boussinesq systems that take account of the two-way propagation of waves present more difficulty than one might initially

expect. Such systems present interesting modeling and mathematical issues that seem worth further study. The interested reader may consult the recent script of Bona et al. [**15**] for more commentary on these matters and an introduction to the existing literature.

REFERENCES

1. L. Abdelouhab, J.L. Bona, M. Felland and J.-C. Saut, *Non-local models for nonlinear, dispersive waves*, Physica D **40** (1989), 360–392.

2. M.J. Ablowitz and P.A. Clarkson, *Solitons, nonlinear evolution equations and inverse scattering*, London Math. Soc. Lecture Notes Ser. **199**, Cambridge Univ. Press, Cambridge-New York, 1991.

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

4. C.A. Amick, J.L. Bona and M.E. Schonbek, *Decay of solutions of some nonlinear wave equations*, J. Differential Equations **81** (1989), 1–49.

5. T.B. Benjamin, *Lectures on nonlinear wave motion*, Lectures Appl. Math. **15** (1974), 3–47.

6. T.B. Benjamin, J.L. Bona and J.J. Mahony, *Model equations for long waves in nonlinear dispersive systems*, Phil. Trans. Royal Soc. London Sect. A **272** (1972), 47–78.

7. P. Biler, *Asymptotic behavior in time of some equations generalizing Korteweg-de Vries-Burgers*, Bull. Polish Acad. Sci. **32** (1984), 275–281.

8. ———, *Large-time behavior of periodic solutions of two dispersive equations of Korteweg-de Vries type*, Bull. Polish Acad. Sci. **32** (1984), 401–405.

9. J.L. Bona and L. Luo, *Decay of solutions to nonlinear, dispersive wave equations*, Differential Integral Equations **6** (1993), 961–980.

10. ———, *More results on the decay of solutions to nonlinear, dispersive wave equations*, preprint.

11. J.L. Bona, W.G. Pritchard and L.R. Scott, *An evaluation of a model equation for water waves*, Phil. Trans. Royal Soc. London Sect. A **302** (1981), 457–510.

12. ———, *A comparison of solutions of two model equations for long waves*, Lectures Appl. Math. **20** (1983), 235–267.

13. J.L. Bona and M. Rose, *General models for the propagation of long waves*, preprint.

14. J.L. Bona and J.-C. Saut, *Dispersive blow-up of solutions of generalized Korteweg-de Vries equations*, J. Differential Equations **103** (1993), 3–57.

15. J.L. Bona, J.-C. Saut and J.F. Toland, *Boussinesq equations and other systems for small-amplitude long waves in nonlinear dispersive media*, preprint.

16. J.L. Bona and R. Smith, *The initial-value problem for the Korteweg-de Vries equation*, Phil. Trans. Royal Soc. London Sect. A **278** (1975), 555–604.

17. J. Bourgain, *Fourier transform restriction phenomena for certain lattice bsets and applications to non-linear evolution equations, Part I: Schrödinger uations*, Geom. Funct. Anal. **3** (1993), 107–156.

18. ———, *Fourier transform restriction phenomena for certain lattice subsets id applications to non-linear evolution equations, Part II: The KdV equation*, eom. Funct. Anal. **3** (1993), 209–262.

19. R. Camassa and D.H. Holm, *An integrable shallow water equation with peaked lutions*, Phys. Rev. Letters, to appear.

20. T. Cazenave, *An introduction to nonlinear Schrödinger equations*, Textos de étodos Matemáticos **22**, Instituto de Matemática-Universidade Federal do Rio de ineiro, Rio de Janeiro, 1989.

21. T. Cazenave and F.B. Weissler, *The Cauchy problem for the critical nonlinear chrödinger equation in H^s*, Nonlinear Anal. **14** (1990), 807–836.

22. A. Cohen-Murray, *Solutions of the Korteweg-de Vries equation for irregular ata*, Duke Math. J. **45** (1978), 149–181.

23. P. Constantin and J.-C. Saut, *Local smoothing properties of dispersive juations*, J. American Math. Soc. **1** (1988), 413–446.

24. W. Craig, *An existence theory for water waves and the Boussinesq and orteweg-de Vries scaling limits*, Comm. Partial Differential Equations **10** (1985), 87–1003.

25. D.B. Dix, *The dissipation of nonlinear dispersive waves: The case of symptotically weak nonlinearity*, Comm. Partial Differential Equations **17** (1992), 665–1693.

26. J. Ginibre and Y. Tsutsumi, *Uniqueness for the generalized Korteweg-de 'ries equation*, SIAM J. Math. Anal. **20** (1989), 1388–1425.

27. J. Ginibre and G. Velo, *Smoothing properties and retarded estimates for some ispersive evolution equations*, to appear.

28. H. Grad and P.N. Hu, *Unified shock profile in a plasma*, Phys. Fluids **10** 1967), 2596–2599.

29. R.J. Iório, *On the Cauchy problem for the Benjamin-Ono equation*, Comm. 'artial Differential Equations **11** (1986), 1031–1081.

30. ———, *KdV, BO and friends in weighted Sobolev spaces*, Lecture Notes in Aath. **1450** (1990), 104–121.

31. A. Jeffrey and T. Kakutani, *Weak nonlinear dispersive waves: A discussion entered around the Korteweg-de Vries equation*, SIAM Rev. **14** (1972), 582–643.

32. R.S. Johnson, *A nonlinear equation incorporating damping and dispersion*, I. Fluid Mech. **42** (1970), 49–60.

33. ———, *Shallow water waves on a viscous fluid—the undular bore*, Phys. 'luids **15** (1972), 1693–1699.

34. T. Kato, *Quasi-linear equations of evolution with applications to partial lifferential equations*, Lecture Notes in Math. **448** (1975), 25–70.

35. ———, *On the Korteweg-de Vries equation*, Manuscripta Math. **29** (1979), 39–99.

36. ———, *On the Cauchy problem for the (generalized) Korteweg-de Vries equation*, Stud. Appl. Math., Adv. Math. Suppl. Studies **8** (1983), 93–128.

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

38. ———, *The Cauchy problem for the Korteweg-de Vries equation in Sobolev spaces of negative indices*, Duke Math. J. **71** (1993), 1–21.

39. D.J. 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 wave*, Phil. Mag. **39** (1895), 422–443.

40. T. Kubota, D.R.S. Ko and L.D. Dobbs, *Weakly nonlinear, long internal gravity waves in stratified fluids of finite depth*, J. Hydronautics **12** (1978), 157–165.

41. P.J. Olver, *Hamiltonian and non-Hamiltonian models for water waves*, in *Trends in applications of pure mathematics to mechanics* (P. G. Ciarlet and M. Reseau, eds.), Palaiseau, France, Lecture Notes in Physics **195** (1983), 273–290.

42. G. Ponce, *Lax pairs and higher order models for water waves*, J. Differential Equations **102** (1991), 360–381.

43. R. Sachs, *Classical solutions of the Korteweg-de Vries equation for non-smooth initial data via inverse scattering*, Comm. Partial Differential Equations **10** (1985), 29–98.

44. J.-C. Saut, *Sur quelques généralisations de l'equation de Korteweg-de Vries*, J. Math. Pures Appl. **58** (1979), 21–61.

45. J.-C. Saut and R. Temam, *Remarks on the Korteweg-de Vries equation*, Israel J. Math. **24** (1976), 78–87.

46. A.C. Scott, F.Y.F. Chu and D.W. McLaughlin, *The soliton: A new concept in applied science*, Proc. IEEE **61** (1973), 1443–1483.

47. R. Smith, *Nonlinear Kelvin and continental-shelf waves*, J. Fluid Mech. **57** (1972), 379–391.

48. G.G. Stokes, *On the theory of oscillatory waves*, Trans. Camb. Phil. Soc. **8** (1849), 441–455.

49. R. Temam, *Sur un problème non linéaire*, J. Math. Pures Appl. **48** (1969), 159–172.

50. F. Ursell, *The long-wave paradox in the theory of gravity waves*, Proc. Camb. Phil. Soc. **49** (1953), 685–694.

51. G.B. Whitham, *Linear and nonlinear waves*, John Wiley & Sons, New York, 1974.

Department of Mathematics, University of Texas at Austin, Austin, TX, 78712, USA

Departamento de Matemática, Universidade Estadual de Campinas, Campinas, 13081, SP, Brasil