PERGAMON

Nonlinear Analysis 38 (1999) 625–647

Nonlinear Analysis

Comparison of model equations for small-amplitude long waves

Jerry L. Bona[a,b,*], Hongqiu Chen[b,c]

[a] *Texas Institute for Computational and Applied Mathematics, The University of Texas at Austin, Austin, TX 78712, USA*

[b] *Department of Mathematics, The University of Texas at Austin, Austin, TX 78712, USA*

[c] *Doctoral Program in Mathematics, The Graduate School and University Center of CUNY, New York, NY 10036, USA*

Received 19 October 1998; accepted 11 January 1999

Abstract

Consider a body of water of finite depth under the influence of gravity, bounded below by a flat, impermeable surface. If viscous and surface tension effects are ignored, and assuming that the flow is incompressible and irrotational, the fluid motion is governed by the Euler equations together with suitable boundary conditions on the rigid surfaces and on the air-water interface. In special regimes, the Euler equations admit of simpler, approximate models that describe pretty well the fluid response to a disturbance. In situations where the wavelength is long and the amplitude is small relative to the undisturbed depth, and if the Stokes number is of order one, then various model equations have been derived. Two of the most standard are the KdV-equation

$$u_t + u_x + uu_x + u_{xxx} = 0 \tag{0.1}$$

and the RLW-equation

$$u_t + u_x + uu_x - u_{xxt} = 0. \tag{0.2}$$

Bona, Pritchard and Scott showed that solutions of these two evolution equations agree to the neglected order of approximation over a long time scale, if the initial disturbance in question is genuinely of small-amplitude and long-wavelength. The same formal argument that allows one to infer (0.2) from (0.1) in small-amplitude, long-wavelength regimes also produces a third equation, namely

$$u_t + u_x + uu_x + u_{xtt} = 0.$$

* Corresponding author. Department of Mathematics, The University of Texas at Austin, Austin, TX 78712, USA.

AMS(MOS) subject classifications (1985 revision). 35B05, 35E15, 35Q20, 76D33, 76E30

0362-546X/99/$ - see front matter

Kruskal, in a wide-ranging discussion of modelling considerations, pointed to this equation as an example that might not accurately describe water waves. Its status has remained unresolved.

It is our purpose here to show that the initial-value problem for the latter equation is indeed well posed. Moreover, we show that for small-amplitude, long waves, solutions of this model also agree to the neglected order with solutions of either (0.1) or (0.2) provided the initial data is properly imposed.

Keywords: Nonlinear dispersive waves; Korteweg-de Vries equation; Regularized long-wave equation; Comparisons of model equations; Small-amplitude long-wavelength wave motion

1. Introduction

The Korteweg-de Vries equation has been derived as a model for the uni-directional propagation of nonlinear, dispersive waves in an impressive array of physical situations [1,3,10,13,16]. In most cases when it is derived from more complex systems, the Korteweg-de Vries equation (KdV-equation henceforth) appears in the form

$$u_t + u_x + \varepsilon u u_x + \delta u_{xxx} = 0, \tag{1.1}$$

in dimensionless variables, scaled so that the dependent variable $u = u(x,t)$ and its derivatives are order-one. The small positive parameters ε and δ are related to a small-amplitude and a long-wavelength assumption, respectively. The right-hand side of (1.1) is not actually zero in general, but instead is comprised of terms of order ε^2, δ^2 and $\varepsilon\delta$ which are neglected in the KdV-approximation. A further restriction is that the Stokes number, $S = \varepsilon/\delta$ is of order-one, a presumption that formally implies the small nonlinear and dispersive effects to be balanced. It also means that the error terms all have the same, even smaller order of magnitude.

It was observed by Peregrine [12] and Benjamin et al. [4] that whenever (1.1) can be formally justified as a model, the lowest-order relation

$$u_t + u_x = O(\varepsilon, \delta)$$

may be used to alter the higher-order terms without formal loss of accuracy. This point was pursued in the last-quoted references with regard to the dispersive term u_{xxx}. Thus if we write

$$u_{xxx} = -u_{xxt} + O(\varepsilon, \delta), \tag{1.2}$$

then, formally,

$$\delta u_{xxx} = -\delta u_{xxt} + O(\varepsilon\delta, \delta^2).$$

This leads to the regularized long-wave (RLW henceforth) version of the KdV-equation, namely

$$v_t + v_x + \varepsilon v v_x - \delta v_{xxt} = 0. \tag{1.3}$$

It was established by Bona et al. [7] (see also [2,8]) that if the initial-value problem for (1.1) and (1.3) is posed with the same, smooth, order-one, initial data

$$u(x,0) = v(x,0) = f(x), \quad x \in \mathbb{R}, \tag{1.4}$$

then

$$\sup_{x\in\mathbb{R}} |u(x,t) - v(x,t)| \le C\varepsilon^2 t$$

for $0 \le t \le \varepsilon^{-1}$, where C is a constant dependent only on the initial data f, but which is independent of $\varepsilon > 0$ and $t \in [0, \varepsilon^{-1}]$. This result is interpreted as indicating that the solutions of the two initial-value problems agree to the neglected order. More precisely, formal analysis and numerical simulations indicate that nonlinear and dispersive effects may accumulate to have an order-one effect on the solution, which is itself order-one in the present scaling, on a time interval of length ε^{-1} [2, 7, 8, 14]. Thus nonlinear and dispersive effects may alter both u and v significantly by time ε^{-1}, but at this same time, u and v are only order-ε apart. On the other hand, the neglected terms of order ε^2, $\varepsilon\delta$ and δ^2 can, on the time interval ε^{-1}, make an order-ε contribution to the solutions. Thus u and v are seem to agree to the order of resolution of either equation, on the long time scale ε^{-1}.

This result appears to settle the theoretical issue of which of the two evolution equations (1.1) and (1.3) provides a better model. The conclusion one draws from the forgoing is that they have equal predictive power, though when one uses such a model in a practical situation, other facts may come to the fore which would favor one model over the other (see the discussion in [4, Section 2]).

Kruskal [11] pointed out that there is no reason why the general reasoning used to derive (1.3) from (1.1) could not be applied again, thereby leading to the model

$$u_t + u_x + \varepsilon u u_x + \delta u_{xtt} = 0. \tag{1.5a}$$

This somewhat odd-looking differential equation is second-order in time, and therefore appears to require two initial data, say

$$u(x,0) = f(x) \quad \text{and} \quad u_t(x,0) = g(x) \tag{1.5b}$$

to have a chance of determining a unique solution. This seems problematic when considering (1.5) as providing an approximation to uni-directional waves in a nonlinear, dispersive medium since it is generally expected that only a single initial profile is needed to initiate such motion. Bona [5,6] suggested that if small-amplitude, long waves are in question, so that u and its derivatives are order one, then one could use the lowest-order relation $u_t + u_x = 0$ to provide the additional initial condition

$$g(x) = u_t(x,0) = -u_x(x,0) = -f'(x). \tag{1.6}$$

The conjecture was that (1.5) with (1.4) and (1.6) leads to a well-posed initial-value problem whose solutions are close to those of (1.1) or (1.3) with initial value (1.4).

It is our purpose here to investigate this possibility. It will turn out that the general idea leading to the additional initial data (1.6) is well conceived, but the details are a little more subtle. Indeed, it transpires that the initial data imposed upon u_t needs higher-order correction so that solutions of (1.5) resemble those of (1.1) or (1.3) over longer time scales (see Theorem 4.1). An interesting feature of the analysis is that a term that is formally small at a certain order turns out not to be negligible at this order.

It is sometimes convenient in analyzing (1.1) and (1.5) to rescale the dependent and independent variables so that the small-amplitude, long-wavelength assumptions appear applied to the initial disturbance, while the differential equations feature only order-one coefficients. Thus, if we make transforms $U(x,t)=\varepsilon u(\delta^{1/2}x,\delta^{1/2}t)$ in the KdV-equation (1.1) and $V(x,t)=\varepsilon u(\delta^{1/2}x,\delta^{1/2}t)$ in (1.5), then,

$$U_t + U_x + UU_x + U_{xxx} = 0, \tag{1.7}$$

$$V_t + V_x + VV_x + V_{xtt} = 0, \tag{1.8}$$

$$U(x,0) = V(x,0) = \varepsilon f(\delta^{1/2}x), \qquad V_t(x,0) = \varepsilon\delta^{1/2} g(\delta^{1/2}x).$$

It must be kept in mind, however, that when written in this form, the time scale over which nonlinear and dispersive effects may accumulate to have an order-one relative effect is now $\varepsilon^{-3/2}$ and the time scale during which the neglected effects can have an order-one relative effect on U and V is $\varepsilon^{-5/2}$.

The plan of the paper is as follows. Section 2 is concerned with the easier, related question of comparing solutions of the linearized equations (Eqs. (1.1), (1.3) and (1.5) without the uu_x-term). The analysis in this case is straightforward, but the result is illuminating. The nonlinear initial-value problem (1.5) is shown to be well-posed in Section 3. In Section 4, we effect a comparison between suitably initiated solutions of (1.5) and those of (1.1) when the data is genuinely of order one. Ideas from both Sections 2 and 3 prove useful in this comparison.

Throughout, standard notation is used: thus

$$L_p = L_p(\mathbb{R}) = \left\{ f\colon \mathbb{R} \to \mathbb{R}\colon \int_{-\infty}^{\infty} |f(x)|^p \,\mathrm{d}x < \infty \right\}$$

for $1 \le p < \infty$, with the usual modification when $p = \infty$. The standard norm on L_p is denoted $|\cdot|_p$. For $m \ge 0$, H^m is the usual Sobolev space of L_2-functions whose derivatives up to order m also lie in L_2. The norm of a function $f \in H^m$ is $\|f\|_m^2 = \int_{-\infty}^{\infty} (1+k^2)^m |\hat{f}(k)|^2 \,\mathrm{d}k$, where $\hat{f}$ is the Fourier transform of f, defined by

$$\hat{f}(k) = \mathscr{F}(f)(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-ikx} f(x) \,\mathrm{d}x.$$

If $m = 0$, the norm in H^m is that of L_2. We will systematically use an unadorned norm for L_2. Thus, for $f \in L_2$, $\|f\| = \|f\|_0 = |f|_2$. Another standard class of spaces intervening in our analysis is

$$C(0,T;H^m) = \{\text{the continuous functions } f\colon [0,T] \to H^m\},$$

where $T > 0$ and m is a non-negative integer, equipped with the norm

$$\|f\|_{C(0,T;H^m)} = \max_{0 \le t \le T} \|f(\cdot,t)\|_m.$$

2. Analytic comparison of the linearized equations

It is helpful to begin with the associated linearized initial-value problems corresponding to (1.7) and (1.8). The analysis in this case is simple, but the results are not without interest. Moreover, the conclusion points the way to a proper understanding of the nonlinear problem. Thus, in this section, consideration is given to the initial-value problems:

$$U_t + U_x + U_{xxx} = 0,$$
$$U(x,0) = f(x) = \varepsilon F(\varepsilon^{1/2}x), \tag{2.1}$$

and

$$V_t + V_x + V_{xtt} = 0,$$
$$V(x,0) = f(x) = \varepsilon F(\varepsilon^{1/2}x),$$
$$V_t(x,0) = g(x), \tag{2.2}$$

for $x \in \mathbb{R}$ and $t > 0$. Suppose that $\hat{F}$, the Fourier transform of F, has bounded support $[-M, M]$ for some positive number M. While it is not necessary to obtain interesting conclusions, the assumption of bounded support makes the analytical issues especially transparent. For the time being, we leave g independent of f.

Taking the Fourier transform in the spatial variable x in both (2.1) and (2.2) and solving the resulting ordinary differential equations leads to

$$\hat{U}(k,t) = \mathrm{e}^{-\mathrm{i}k(1-k^2)t}\hat{f}(k) \tag{2.3}$$

and

$$\begin{aligned}\hat{V}(k,t) &= \frac{r_+\mathrm{e}^{\mathrm{i}r_-t} - r_-\mathrm{e}^{\mathrm{i}r_+t}}{r_+ - r_-}\hat{f}(k) + \frac{\mathrm{i}\mathrm{e}^{\mathrm{i}r_-t} - \mathrm{i}\mathrm{e}^{\mathrm{i}r_+t}}{r_+ - r_-}\hat{g}(k)\\ &= \frac{r_+\hat{f}(k) + \mathrm{i}\hat{g}(k)}{r_+ - r_-}\mathrm{e}^{\mathrm{i}r_-t} + \frac{-r_-\hat{f} - \mathrm{i}\hat{g}(k)}{r_+ - r_-}\mathrm{e}^{\mathrm{i}r_+t}\\ &= \frac{(1+\sqrt{1+4k^2})\hat{f} + 2\mathrm{i}k\hat{g}}{2\sqrt{1+4k^2}}\mathrm{e}^{\mathrm{i}r_-t} + \frac{(-1+\sqrt{1+4k^2})\hat{f} - 2\mathrm{i}k\hat{g}}{2\sqrt{1+4k^2}}\mathrm{e}^{\mathrm{i}r_+t},\end{aligned} \tag{2.4}$$

where

$$r_+ = r_+(k) = \frac{1+\sqrt{1+4k^2}}{2k} \quad \text{and} \quad r_- = r_-(k) = \frac{1-\sqrt{1+4k^2}}{2k}.$$

Denote by A the function

$$A(k) = \frac{(-1+\sqrt{1+4k^2})\hat{f}(k) - 2\mathrm{i}k\hat{g}(k)}{2\sqrt{1+4k^2}}.$$

Then $\hat{V}$ may be expressed as

$$\hat{V}(k,t) = \{\hat{f}(k) - A(k)\}\mathrm{e}^{\mathrm{i}(1-\sqrt{1+4k^2})/2kt} + A(k)\mathrm{e}^{\mathrm{i}(1+\sqrt{1+4k^2})/2kt}. \tag{2.5}$$

Since $\hat{f}(k)=\varepsilon^{1/2}\hat{F}(\varepsilon^{-1/2}k)$, the support of $\hat{f}$ is $[-\varepsilon^{1/2}M,\varepsilon^{1/2}M]$. If $\varepsilon>0$ is sufficiently small, the quantities $r_{\pm}$ and A can be expressed via a Taylor series which is convergent for $k\in[-\varepsilon^{1/2}M,\varepsilon^{1/2}M]$, namely,

$$r_- = \frac{1-\sqrt{1+4k^2}}{2k} = -k+k^3-4k^5+4k^7+\cdots,$$

$$r_+ = \frac{1+\sqrt{1+4k^2}}{2k} = \frac{1+k^2-k^4+4k^6-4k^8+\cdots}{k},$$

$$A(k) = \frac{\left(k^2-k^4+4k^6-4k^8+\cdots\right)\hat{f}(k)-\mathrm{i}k\hat{g}(k)}{\sqrt{1+4k^2}},$$

$$\mathrm{e}^{\mathrm{i}r_-t} = \mathrm{e}^{\mathrm{i}(1-\sqrt{1+4k^2})/2kt} = \mathrm{e}^{-\mathrm{i}(k-k^3+4k^5-\cdots)t},$$

$$\mathrm{e}^{\mathrm{i}r_+t} = \mathrm{e}^{\mathrm{i}(1+\sqrt{1+4k^2})/2kt} = \mathrm{e}^{\mathrm{i}((1+k^2-k^4+4k^6-4k^8+\cdots)/k)t}.$$

Remark. *From the exact forms* (2.4) *or* (2.5), *it is apparent that the solution of* (2.2) *has two dispersive branches. The branch corresponding to* r_- *is near to the KdV-dispersion for small value of* k, *whereas the branch corresponding to* r_+ *features very high-frequency oscillation for* k *near to* 0.

Subtract (2.3) from (2.5) to reach the formula

$$\begin{aligned}\hat{V}(k,t)-\hat{U}(k,t) &= \{\mathrm{e}^{\mathrm{i}r_-t}-\mathrm{e}^{-\mathrm{i}k(1-k^2)t}\}\hat{f}(k)-A(k)\mathrm{e}^{\mathrm{i}r_-t}+A(k)\mathrm{e}^{\mathrm{i}r_+t}\\ &= \{\mathrm{e}^{\mathrm{i}(-k+k^3-4k^5+\cdots)t}-\mathrm{e}^{\mathrm{i}(-k+k^3)t}\}\hat{f}(k)-A(k)\mathrm{e}^{\mathrm{i}r_-t}+A(k)\mathrm{e}^{\mathrm{i}r_+t}\\ &= \{\mathrm{e}^{\mathrm{i}(-4k^5+\cdots)t}-1\}\mathrm{e}^{-\mathrm{i}(k-k^3)t}\hat{f}(k)-A(k)\mathrm{e}^{\mathrm{i}r_-t}+A(k)\mathrm{e}^{\mathrm{i}r_+t}.\end{aligned} \tag{2.6}$$

It follows readily that as $k\to 0$,

$$|\hat{V}(k,t)-\hat{U}(k,t)| \le O(k^5t)|\hat{f}(k)|+2|A(k)|. \tag{2.7}$$

Two possibilities are distinguished.

(1) If $A\equiv 0$, which is the same as asking that

$$g=\mathscr{F}^{-1}\left\{\frac{-1+\sqrt{1+4k^2}}{2\mathrm{i}k}\hat{f}\right\}, \tag{2.8}$$

then, the oscillatory branch $\mathrm{e}^{\mathrm{i}r_+t}=\mathrm{e}^{\mathrm{i}(1/k)t+O(1)t}$ is of no consequence and it follows that

$$|\hat{V}(\cdot,t)-\hat{U}(\cdot,t)|_1 \le \int_{-\infty}^{\infty}|\hat{U}(\cdot,t)-\hat{V}(\cdot,t)|\,\mathrm{d}k$$

$$= \int_{-\varepsilon^{1/2}M}^{\varepsilon^{1/2}M}|\mathrm{e}^{\mathrm{i}(-4k^5+\cdots)t}-1|\,|\hat{f}(k)|\,\mathrm{d}k$$

$$\leq \int_{-\varepsilon^{1/2}M}^{\varepsilon^{1/2}M} C_1 |k|^5 t\varepsilon^{1/2} |F(\varepsilon^{-1/2}k)| \,\mathrm{d}k$$

$$\leq C\varepsilon^{7/2} t$$

for sufficiently small values of ε in $(0,1)$, where C and C_1 are constants dependent only on F. It follows immediately that

$$|V(\cdot,t) - U(\cdot,t)|_\infty = \max_{x\in\mathbb{R}} \left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}xk} [\hat{V}(k,t) - \hat{U}(k,t)] \,\mathrm{d}k \right|$$

$$\leq \frac{1}{\sqrt{2\pi}} |\hat{V}(\cdot,t) - \hat{U}(\cdot,t)|_1$$

$$\leq C\varepsilon^{7/2} t$$

and, similarly, that

$$\|V(\cdot,t) - U(\cdot,t)\| = \left(\int_{-\infty}^{\infty} |\hat{V}(\cdot,t) - \hat{U}(\cdot,t)|^2 \,\mathrm{d}k \right)^{1/2} \leq C\varepsilon^{13/4} t.$$

Moreover, for $j,m \geq 0$, (2.6) and (2.7) imply there are constants $C_{m,j}$ such that

$$|\widehat{\partial_t^m \partial_x^j V}(k,t) - \widehat{\partial_t^m \partial_x^j U}(k,t)| \leq C_{m,j} k^{5+m+j} t\, |\hat{f}(k)|$$

$$= C_{m,j} k^{5+m+j} t\, |\varepsilon^{1/2} \hat{F}(\varepsilon^{-1/2}k)|,$$

whence,

$$\|\partial_t^m \partial_x^j V(\cdot,t) - \partial_t^m \partial_x^j U(\cdot,t)\| \leq C_{m,j} \varepsilon^{(13/4)+(m/2)+(j/2)} t$$

for $t \geq 0$. This means that models (2.1) and (2.2) agree to the neglected order, at least for $0 \leq t \leq \varepsilon^{-3/2}$.

(2) If $A \neq 0$, consider two sub-cases.

Case I: $g = -f'$ as in (1.6). In this case, $A(k) = O(k^4)\,|\hat{f}(k)|$ as $k \to 0$, so

$$|\hat{V}(k,t) - \hat{U}(k,t)| \leq (O(k^5 t) + O(k^4))|\hat{f}(k)|,$$

and therefore

$$|V(\cdot,t) - U(\cdot,t)|_\infty = \max_{x\in\mathbb{R}} \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}xk} [\hat{V}(k,t) - \hat{U}(k,t)] \,\mathrm{d}k \right|$$

$$\leq C\varepsilon^2 (\varepsilon^{3/2} t) + C\varepsilon^3,$$

$$\|V(\cdot,t) - U(\cdot,t)\| \leq \left(\int_{-\infty}^{\infty} |k^5 t + 2k^4|\, |\varepsilon^{1/2} \hat{F}(\varepsilon^{-1/2}k)|^2 \,\mathrm{d}k \right)^{1/2}$$

$$= \|F^{(5)}\| \varepsilon^{7/4} (\varepsilon^{3/2} t) + \|F^{(4)}\| \varepsilon^{11/4},$$

and, similarly, for $j > 0$,

$$\|\partial_x^j V - \partial_x^j U\| = \left(\int_{-\infty}^{\infty} |(\mathrm{i}k)^j \hat{V}(k,t) - (\mathrm{i}k)^j \hat{V}(k,t)|^2 \,\mathrm{d}k\right)^{1/2}$$
$$\leq C_j \varepsilon^{(7/4)+(j/2)} (\varepsilon^{3/2} t) + C_j \varepsilon^{(11/4)+(j/2)},$$

where the C_j are constants only dependent on F and j. These results are the same as those that obtain when comparing the KdV-equation and the RLW-equation (see [7]). However, if we take the derivative of U and V with respect to t, an interesting phenomenon becomes apparent which is different from the situation arising when comparing the KdV- and RLW-equations. From (2.6) and (2.7) again, it is straightforward to calculate that

$$\begin{aligned}
&\partial_t \hat{V}(k,t) - \partial_t \hat{U}(k,t)\\
&=[\mathrm{i}r_-\{\hat{f}(k) - A(k)\}\mathrm{e}^{\mathrm{i}r_- t} + \mathrm{i}r_+ A(k)\mathrm{e}^{\mathrm{i}r_+ t}] - [-\mathrm{i}(k-k^3)]\hat{f}(k)\mathrm{e}^{-\mathrm{i}(k-k^3)t}\\
&= -\mathrm{i}(k-k^3)\{\mathrm{e}^{\mathrm{i}(-4k^5+\cdots)t} - 1\}\mathrm{e}^{-\mathrm{i}(k-k^3)t}\hat{f}(k)\\
&\quad +\mathrm{i}(-4k^5+\cdots)\mathrm{e}^{\mathrm{i}(-4k^5+\cdots)t}\mathrm{e}^{-\mathrm{i}(k-k^3)t}\hat{f}(k) - \mathrm{i}r_- A(k)\mathrm{e}^{\mathrm{i}r_- t} + \mathrm{i}r_+ A(k)\mathrm{e}^{\mathrm{i}r_+ t},
\end{aligned}$$

so

$$\begin{aligned}
|\partial_t \hat{V}(k,t) - \partial_t \hat{U}(k,t)| &\leq (4k^6 t + 4|k|^5 + |k|^5 + |k|^3)|\hat{f}(k)|\\
&= (4k^6 t + 5|k|^5 + |k|^3)\varepsilon^{1/2}|\hat{F}(\varepsilon^{-1/2}k)|.
\end{aligned}$$

In consequence, it transpires that

$$|\partial_t U(\cdot,t) - \partial_t V(\cdot,t)|_\infty \leq \frac{1}{\sqrt{2\pi}}|\partial_t \hat{U}(\cdot,t) - \partial_t \hat{V}(\cdot,t)|_1 \leq C\varepsilon^{5/2}(\varepsilon^{3/2}t) + C\varepsilon^{5/2},$$

$$\|\partial_t V(\cdot,t) - \partial_t U(\cdot,t)\| \leq C\varepsilon^{9/4}(\varepsilon^{3/2}t) + C\varepsilon^{9/4},$$

and for $j \geq 0$,

$$\|\partial_t\partial_x^j V - \partial_t\partial_x^j U\| \leq C_j\varepsilon^{(9/4)+(j/2)}(\varepsilon^{3/2}t) + C_j\varepsilon^{(9/4)+(j/2)}.$$

More generally, for any $j, m \geq 0$, we see that

$$\begin{aligned}
\|\partial_t^m\partial_x^j V(\cdot,t) - \partial_t^m\partial_x^j U(\cdot,t)\| &\leq C_{j,m}\varepsilon^{(13/4)+(j/2)+(m/2)}t + C_{j,m}\varepsilon^{(11/4)+(j/2)-(m/2)}\\
&= C_{j,m}\varepsilon^{(7/4)+(j/2)+(m/2)}(\varepsilon^{3/2}t) + C_{j,m}\varepsilon^{(11/4)+(j/2)-(m/2)}.
\end{aligned}$$

Thus it becomes clear that for Eq. (2.2), the oscillatory branch plays a non-trivial role when time-derivatives are compared. On the natural, long time interval $[0, \varepsilon^{-3/2}]$ over which nonlinear and dispersive effects are expected to make an order-one relative contribution, spatial derivatives $\partial_x^j u$ and $\partial_x^j v$ of the solution u and v of (1.1) and (1.5), respectively, agree to the neglected order. This is not the case for the difference between temporal derivatives. Moreover, the lack of coherence between u_t and v_t, say, may be traced to the initial condition applied to $v_t(\cdot,0)$.

Case II: Guided by the results in Case I, adjust the initial value for g so that $g = -f' - f'''$. In this case, $A(k) = O(k^6)$, and calculations similar to those above show that

$$\|\partial_t^m \partial_x^j V(\cdot,t) - \partial_t^m \partial_x^j U(\cdot,t)\| \le C_{j,m}\varepsilon^{(13/4)+(j/2)+(m/2)}t + C_{j,m}\varepsilon^{(15/4)+(j/2)-(m/2)},$$

$$= C_{j,m}\varepsilon^{(7/4)+(j/2)+(m/2)}(\varepsilon^{3/2}t) + C_{j,m}\varepsilon^{(15/4)+(j/2)-(m/2)}.$$

Thus $\partial_t^m \partial_x^j V$ is within the neglected order of $\partial_t^m \partial_x^j U$ for $m \le 2$, and therefore, one is indifferent between the two linear models at least on the time interval $[0, \varepsilon^{-3/2}]$.

Remark. *We learn from the foregoing analysis that the proper imposition of initial conditions on Eq.* (1.5) *is crucial for a satisfactory comparison theory. The initial data* $v_t(x,0) = -f'(x) = -\varepsilon^{3/2}F'(\varepsilon^{1/2}x)$ *suggested in* (1.6) *is not accurate enough to obtain the anticipated results.*

3. Well-posedness for the nonlinear problem

Attention is now turned to the nonlinear problem. For suitably restricted initial data f and g, consider the initial-value problem for (1.8), namely

$$v_t + v_x + vv_x + v_{xtt} = 0,$$

$$v(x,0) = f(x), \quad v_t(x,0) = g(x). \tag{3.1}$$

Applying the Fourier transform in the spatial variable x, there appears

$$\mathrm{i}k\hat{v}_{tt} + \hat{v}_t + \mathrm{i}k\hat{v} = -\frac{\mathrm{i}k}{2}\hat{v} * \hat{v},$$

or, what is the same,

$$\hat{v}_{tt} + \frac{1}{\mathrm{i}k}\hat{v}_t + \hat{v} = -\frac{1}{2}\hat{v} * \hat{v},$$

where $v * w$ connotes the convolution of v and w. If r_+ and r_- are defined as before, then the latter equation may be rewritten as

$$(\partial_t - \mathrm{i}r_+)(\partial_t - \mathrm{i}r_-)\hat{v}(k,t) = -\tfrac{1}{2}\hat{v} * \hat{v},$$

$$\hat{v}(k,0) = \hat{f}(k), \quad \hat{v}_t(k,0) = \hat{g}(k).$$

This second-order, ordinary differential equation can be changed to an integral equation by twice solving a first-order, ordinary differential equation, viz.

$$(\partial_t - \mathrm{i}r_-)\hat{v}(k,t) = (\hat{g}(k) - \mathrm{i}r_-\hat{f}(k))\mathrm{e}^{\mathrm{i}r_+t} - \frac{1}{2}\int_0^t \hat{v}^2(k,s)\mathrm{e}^{\mathrm{i}r_+(t-s)}\,\mathrm{d}s,$$

$$\hat{v}(k,t) = \hat{f}(k)\mathrm{e}^{\mathrm{i}r_-t} - \int_0^t \mathrm{e}^{\mathrm{i}r_-(t-s)}(\hat{g}(k) - \mathrm{i}r_-\hat{f}(k))\mathrm{e}^{\mathrm{i}r_+s}\,\mathrm{d}s$$

$$-\frac{1}{2}\int_0^t \mathrm{e}^{\mathrm{i}r_-(t-s)}\int_0^s \mathrm{e}^{\mathrm{i}r_+(s-\tau)}\hat{v}^2(k,\tau)\,\mathrm{d}\tau\,\mathrm{d}s$$

$$= \frac{r_+ e^{ir_- t} - r_- e^{ir_+ t}}{r_+ - r_-} \hat{f}(k) + \frac{e^{ir_+ t} - e^{ir_- t}}{i(r_+ - r_-)} \hat{g}(k)$$

$$- \frac{1}{2} \int_0^t \int_\tau^t e^{ir_-(t-s)} e^{ir_+(s-\tau)} \hat{v}^2(k,\tau) \, ds \, d\tau$$

$$= \frac{r_+ e^{ir_- t} - r_- e^{ir_+ t}}{r_+ - r_-} \hat{f}(k) + \frac{e^{ir_+ t} - e^{ir_- t}}{i(r_+ - r_-)} \hat{g}(k)$$

$$- \frac{1}{2} \frac{1}{i(r_+ - r_-)} \int_0^t [e^{ir_+(t-\tau)} - e^{ir_-(t-\tau)}] \widehat{v^2} \, d\tau$$

$$= \{\hat{f}(k) - A(k)\} e^{ir_- t} + A(k) e^{ir_+ t}$$

$$- \frac{1}{2} \frac{1}{i(r_+ - r_-)} \int_0^t [e^{ir_+(t-\tau)} - e^{ir_-(t-\tau)}] \widehat{v^2} \, d\tau$$

$$= \hat{f}(k) e^{ir_- t} - A(k) e^{ir_- t} + A(k) e^{ir_+ t}$$

$$- \frac{1}{2} \frac{1}{i(r_+ - r_-)} \int_0^t [e^{ir_+(t-\tau)} - e^{ir_-(t-\tau)}] \widehat{v^2} \, d\tau,$$

where $A(k)$ is as it was in Section 2.

Theorem 3.1. (1) *For any initial data $f, g \in H^m$ with integer $m \geq 1$, there is a positive number T which only depends on f and g such that the initial-value problem* (3.1) *has a unique solution in $C^1(0, T; H^m)$.* (2) *If there is a function $\bar{g} \in H^m$ such that $\bar{g}' = g$, then the solution v lies in $C^2(0, T; H^m)$.* (3) *Furthermore, if there are $\bar{f}, \bar{\bar{g}} \in H^m$ such that $\bar{f}' = f$, $\bar{\bar{g}}'' = g$, then $v \in C^3(0, T; H^m)$, and so on.*

Proof. We intend to apply the contraction-mapping principle to prove this theorem. Define an operator $\mathscr{A}$ on $C(0, T; H^m)$ by

$$\widehat{\mathscr{A} v}(k,t) = \frac{r_+ e^{ir_- t} - r_- e^{ir_+ t}}{r_+ - r_-} \hat{f}(k) + \frac{e^{ir_+ t} - e^{ir_- t}}{i(r_+ - r_-)} \hat{g}(k)$$

$$- \frac{1}{2} \frac{1}{i(r_+ - r_-)} \int_0^t [e^{ir_+(t-\tau)} - e^{ir_-(t-\tau)}] \widehat{v^2}(k,\tau) \, d\tau.$$

Let $X = \overline{B_R(0)}$ be the closed ball of radius $R \geq 0$ about the origin in $C(0, T; H^m)$, where $R, T > 0$ are constants to be determined. Then X is a complete metric space because it is a closed subset of $C(0, T; H^m)$. If we can find R and T such that $\mathscr{A}$ maps X into X contractively, then the proof will be essentially complete. To this end, we need a few estimates: for all k,

$$\left| \frac{r_+ e^{ir_- t} - r_- e^{ir_+ t}}{r_+ - r_-} \right| \leq 1, \qquad \left| \frac{e^{ir_+ t} - e^{ir_- t}}{i(r_+ - r_-)} \right| \leq 1,$$

so

$$|\widehat{\mathscr{A}\theta}| \le |\hat{f}| + |\hat{g}|,$$

where $\theta(k,t) \equiv 0$, whence

$$\|\mathscr{A}\theta\|_{C(0,T;H^m)} \le \|f\|_m + \|g\|_m.$$

For any $u, v \in X$, $t \in [0,T]$, and $k \in \mathbb{R}$, we have

$$\begin{aligned}
&\widehat{\mathscr{A}u}(k,t) - \widehat{\mathscr{A}v}(k,t) \\
&\quad = -\frac{1}{2}\frac{1}{\mathrm{i}(r_+ - r_-)} \int_0^t [\mathrm{e}^{\mathrm{i}r_+(t-\tau)} - \mathrm{e}^{\mathrm{i}r_-(t-\tau)}][\widehat{u^2}(k,\tau) - \widehat{v^2}(k,\tau)]\,\mathrm{d}\tau \\
&\quad = -\frac{k}{\mathrm{i}2\sqrt{1+4k^2}} \int_0^t [\mathrm{e}^{\mathrm{i}r_+(t-\tau)} - \mathrm{e}^{\mathrm{i}r_-(t-\tau)}]\mathscr{F}\{(u+v)(u-v)\}(k,\tau)\,\mathrm{d}\tau,
\end{aligned}$$

so that

$$|\widehat{Au}(k,t) - \widehat{Av}(k,t)| \le \frac{1}{2}\int_0^t |\mathscr{F}\{(u+v)(u-v)\}(k,\tau)|\,\mathrm{d}\tau.$$

It follows that for $u, v \in X$, and $t \in [0,T]$,

$$\begin{aligned}
&\|\mathscr{A}u(\cdot,t) - \mathscr{A}v(\cdot,t)\|_m^2 \\
&\quad = \int_{-\infty}^{\infty} (1+k^2)^m |\widehat{Au}(k,t) - \widehat{Av}(k,t)|^2\,\mathrm{d}k \\
&\quad \le \frac{1}{4}\int_{-\infty}^{\infty} (1+k^2)^m \left(\int_0^t |\mathscr{F}\{(u+v)(u-v)\}(k,\tau)|\,\mathrm{d}\tau\right)^2 \mathrm{d}k \\
&\quad = \frac{1}{4}\int_0^t\int_0^t\int_{-\infty}^{\infty} (1+k^2)^m |\mathscr{F}\{(u+v)(u-v)\}(k,\tau)| \\
&\qquad \times |\mathscr{F}\{(u+v)(u-v)\}(k,s)|\,\mathrm{d}k\,\mathrm{d}\tau\,\mathrm{d}s \\
&\quad \le \frac{1}{4}\int_0^t\int_0^t \|(u+v)(u-v)(\cdot,\tau)\|_m \|(u+v)(u-v)(\cdot,s)\|_m\,\mathrm{d}\tau\,\mathrm{d}s \\
&\quad = \frac{1}{4}\left(\int_0^t \|(u+v)(u-v)(\cdot,\tau)\|_m\,\mathrm{d}\tau\right)^2 \\
&\quad \le \frac{1}{4}\left(\int_0^t c_m(\|u\|_{C(0,T;H^m)} + \|v\|_{C(0,T;H^m)})\|u(\cdot,\tau) - v(\cdot,\tau)\|_m\,\mathrm{d}\tau\right)^2 \\
&\quad \le c_m^2 R^2 T^2 \|u - v\|_{C(0,T;H^m)}^2,
\end{aligned}$$

where $c_m = \max\{\|fg\|_m : \ f, g \in H^m, \|f\|_m = \|g\|_m = 1\}$. On the other hand, if $u \in X$, then

$$\begin{aligned}
\|\mathscr{A}u\|_{C(0,T;H^m)} &\le \|\mathscr{A}\theta\|_{C(0,T;H^m)} + \|\mathscr{A}u - \mathscr{A}\theta\|_{C(0,T;H^m)} \\
&\le \|f\|_m + \|g\|_m + c_m RT\|u\|_{C(0,T;H^m)},
\end{aligned}$$

so, if $2\|\mathscr{A}\theta\|_{C(0,T;H^m)} \le 2\{\|f\|_m + \|g\|_m\} = R$ and $T = 1/2c_mR$, then the mapping $\mathscr{A}$ is seen to be contractive from X to X. The contraction-mapping theorem assures that there exists a unique solution $v \in C(0, T; H^m)$ to problem (3.1). To prove $v_t(\cdot,t) \in H^m$, consider the defining relation for v, namely

$$\hat{v}(k,t) = \frac{r_+ e^{ir_- t} - r_- e^{ir_+ t}}{r_+ - r_-}\hat{f}(k) + \frac{e^{ir_+ t} - e^{ir_- t}}{i(r_+ - r_-)}\hat{g}(k)$$

$$-\frac{1}{2}\frac{1}{i(r_+ - r_-)}\int_0^t [e^{ir_+(t-\tau)} - e^{ir_-(t-\tau)}]v^2(k,t)\,d\tau,$$

and take the derivative with respect to t of both sides of this identity to reach the formula

$$\hat{v}_t(k,t) = ir_+r_-\frac{e^{ir_- t} - e^{ir_+ t}}{r_+ - r_-}\hat{f}(k) + \frac{r_+ e^{ir_+ t} - r_- e^{ir_- t}}{(r_+ - r_-)}\hat{g}(k)$$

$$-\frac{1}{2}\frac{1}{(r_+ - r_-)}\int_0^t [r_+ e^{ir_+(t-\tau)} - r_- e^{ir_-(t-\tau)}]v^2(k,\tau)\,d\tau$$

$$= -i\frac{e^{ir_- t} - e^{ir_+ t}}{r_+ - r_-}\hat{f}(k) + \frac{r_+ e^{ir_+ t} - r_- e^{ir_- t}}{(r_+ - r_-)}\hat{g}(k)$$

$$-\frac{1}{2}\frac{1}{(r_+ - r_-)}\int_0^t [r_+ e^{ir_+(t-\tau)} - r_- e^{ir_-(t-\tau)}]v^2(k,\tau)\,d\tau.$$

Since $|r_\pm/(r_+ - r_-)| \le 1$, and $v \in C(0, T; H^m)$, it follows from the preceding equation that

$$|\hat{v}_t(k,t)| \le |\hat{f}(k)| + |\hat{g}(k)| + \int_0^t |\widehat{v^2}(k,\tau)|\,d\tau,$$

and, hence, that

$$\|v_t(\cdot,t)\|_m \le \|f\|_m + \|g\|_m + \int_0^t c_m\|v(\cdot,\tau)\|_m^2\,d\tau.$$

Thus, $v_t \in C(0, T; H^m)$, and part (1) of the theorem is proved.

If $\int_{-\infty}^x g(y)\,dy = \bar{g} \in H^m$, which is to say, $(1+k^2)^{m/2}(1/k)\hat{g} \in L_2$, then differentiating $\hat{v}_t$ leads to

$$\hat{v}_{tt}(k,t) = -\frac{r_+ e^{ir_+ t} - r_- e^{ir_- t}}{r_+ - r_-}\hat{f}(k) + i\frac{r_+^2 e^{ir_+ t} - r_-^2 e^{ir_- t}}{(r_+ - r_-)}\hat{g}(k)$$

$$-\frac{1}{2}\frac{r_+ e^{ir_+ t} - r_- e^{ir_- t}}{r_+ - r_-}\widehat{v^2}(k,0)$$

$$-\frac{1}{2}\frac{1}{(r_+ - r_-)}\int_0^t (r_+ e^{ir_+\tau} - r_- e^{ir_-\tau})\widehat{2vv_t}(k,t-\tau)\,d\tau.$$

Because

$$\left|\frac{r_-^2}{r_+ - r_-}\right| \le 1, \qquad \left|\frac{r_+^2}{r_+ - r_-}\right| \le 1 + \left|\frac{1}{k}\right|,$$

and

$$v \in C^1(0, T; H^m),$$

it follows as above that

$$v_{tt} \in C(0, T; H^m),$$

and part (2) is proved. Part (3) may be established in the same way. □

Proposition 3.2. *Let v be a solution of* (3.1) *corresponding to $f, g \in H^1$, where g is such that there is a $\bar{g} \in H^1$ with $\bar{g}' = g$. Then the following functionals are independent of t:*

$$\int_{-\infty}^{\infty} v(x,t)\,\mathrm{d}x = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x,$$

$$\int_{-\infty}^{\infty} [v^2(x,t) - 2v_x(x,t)v_t(x,t)]\,\mathrm{d}x = \int_{-\infty}^{\infty} [f^2(x) - 2f'(x)g(x)]\,\mathrm{d}x,$$

and

$$\int_{-\infty}^{\infty} \left[v^2(x,t) + v_t^2(x,t) + \frac{1}{3}v^3(x,t)\right]\,\mathrm{d}x = \int_{-\infty}^{\infty} \left[f^2(x) + g^2(x) + \frac{1}{3}f^3(x,t)\right]\,\mathrm{d}x.$$

Proof. These follow from multiplying the equation in (3.1) by $\phi(x) = 1, v$, and $v + \frac{1}{2}v^2 + v_{tt}$, respectively, and integrating the result over a bounded spatial interval $[a, b]$, say. Integrating by parts, integrating over $[0, t]$, and taking the limit as $a \to -\infty$, $b \to \infty$ leads to the advertised result because the boundary values appearing after integration by parts in space are uniformly bounded and tend pointwise to zero at $\pm\infty$. □

Now we return to the case of small initial data. The analysis of the linearized problems in Section 2 indicates one should impose initial data in the following manner:

$$v_t + v_x + vv_x + v_{xtt} = 0,$$

$$v(x,0) = \varepsilon F(\varepsilon^{1/2}x),$$

$$v_t(x,0) = -\varepsilon^{3/2}F'(\varepsilon^{1/2}x) - \varepsilon^{5/2}\{F'''(\varepsilon^{1/2}x) + FF'(\varepsilon^{1/2}x)\}. \tag{3.2}$$

Using the idea put forward in [9] (see also [15]), define η by the change of variables

$$v(x,t) = \varepsilon\eta(\varepsilon^{1/2}(x-t), \varepsilon^{3/2}t),$$

or, what is the same,

$$\eta(x,t) = \varepsilon^{-1}v(\varepsilon^{-1/2}x + \varepsilon^{-3/2}t, \varepsilon^{-3/2}t).$$

Writing (3.2) in terms of η gives

$$\eta_t + \eta\eta_x + \eta_{xxx} - 2\varepsilon\eta_{xxt} + \varepsilon^2\eta_{xtt} = 0,$$

$$\eta(x,0) = F(x), \quad \eta_t(x,0) = -F'''(x) - FF'(x). \tag{3.2a}$$

This formulation looks promising because it appears formally as a perturbed KdV-equation. Of course, η_{xxt} and $\varepsilon\eta_{xtt}$ might not be order-one quantities, in which case the formalism would be misleading. Indeed, because temporal derivatives can feature inverse powers of ε, this point is potentially important. Taking the Fourier transform of (3.2a) with respect to the spatial variable x and simplifying yields

$$\hat{\eta}_{tt} + \frac{1+2\varepsilon k^2}{\varepsilon^2(\mathrm{i}k)}\hat{\eta}_t - \frac{k^2}{\varepsilon^2}\hat{\eta} = -\frac{1}{2\varepsilon^2}\eta^2,$$

$$\hat{\eta}(k,0) = \hat{F}(k), \quad \eta_t(k,0) = \mathrm{i}k^3\hat{F}(k) - \frac{\mathrm{i}k}{2}\widehat{F^2}(k). \tag{3.2b}$$

The system in (3.2b) is an initial-value problem for a one-parameter family of nonlinear ordinary differential equations. As before, these may be rewritten as an integral equation, viz.

$$\begin{aligned}\hat{\eta}(k,t) = {} & \frac{\lambda_+ \mathrm{e}^{\mathrm{i}\lambda_- t} - \lambda_- \mathrm{e}^{\mathrm{i}\lambda_+ t}}{\lambda_+ - \lambda_-}\hat{F}(k) + \frac{\mathrm{e}^{\mathrm{i}\lambda_+ t} - \mathrm{e}^{\mathrm{i}\lambda_- t}}{\mathrm{i}(\lambda_+ - \lambda_-)}\left\{\mathrm{i}k^3\hat{F}(k) - \frac{\mathrm{i}k}{2}\widehat{F^2}(k)\right\} \\ & + \frac{1}{\mathrm{i}(\lambda_+ - \lambda_-)}\int_0^t [\mathrm{e}^{\mathrm{i}\lambda_+(t-\tau)} - \mathrm{e}^{\mathrm{i}\lambda_-(t-\tau)}]\left[-\frac{1}{2\varepsilon^2}\eta^2(k,\tau)\right] \mathrm{d}\tau, \end{aligned} \tag{3.3}$$

where

$$\begin{aligned} \lambda_+ &= \frac{1+2\varepsilon k^2 + \sqrt{1+4\varepsilon k^2}}{2\varepsilon^2 k} = \frac{1+2\varepsilon k^2 - \varepsilon^2 k^4 + \cdots}{\varepsilon^2 k}, \\ \lambda_- &= \frac{1+2\varepsilon k^2 - \sqrt{1+4\varepsilon k^2}}{2\varepsilon^2 k} = k^3 - 4\varepsilon k^5 + \cdots. \end{aligned} \tag{3.4}$$

(The series in (3.4) are uniformly convergent if, for example, $|k| \le M$ and $0 < \varepsilon < 1$ is sufficiently small.) Substituting (3.4) into (3.3) gives

$$\begin{aligned} \hat{\eta}(k,t) = {} & \frac{(\lambda_+ - k^3)\hat{F}(k) + (k/2)\widehat{F^2}(k)}{\lambda_+ - \lambda_-}\mathrm{e}^{\mathrm{i}\lambda_- t} - \frac{(\lambda_- - k^3)\hat{F}(k) + (k/2)\widehat{F^2}(k)}{\lambda_+ - \lambda_-}\mathrm{e}^{\mathrm{i}\lambda_+ t} \\ & + \frac{1}{\mathrm{i}(\lambda_+ - \lambda_-)}\int_0^t [\mathrm{e}^{\mathrm{i}\lambda_+(t-\tau)} - \mathrm{e}^{\mathrm{i}\lambda_-(t-\tau)}]\left[-\frac{1}{2\varepsilon^2}\eta^2(k,\tau)\right] \mathrm{d}\tau \\ = {} & \frac{(1+2\varepsilon k^2 + \sqrt{1+4\varepsilon k^2} - 2\varepsilon^2 k^4)\hat{F}(k) + \varepsilon^2 k^2 \widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}\mathrm{e}^{\mathrm{i}\lambda_- t} \\ & - \frac{(1+2\varepsilon k^2 - \sqrt{1+4\varepsilon k^2} - 2\varepsilon^2 k^4)\hat{F}(k) + \varepsilon^2 k^2 \widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}\mathrm{e}^{\mathrm{i}\lambda_+ t} \\ & + \frac{\mathrm{i}k}{2\sqrt{1+4\varepsilon k^2}}\int_0^t \left[\mathrm{e}^{\mathrm{i}\lambda_+(t-\tau)} - \mathrm{e}^{\mathrm{i}\lambda_-(t-\tau)}\right]\eta^2(k,\tau)\,\mathrm{d}\tau, \end{aligned} \tag{3.5a}$$

and this is formally of the form

$$\hat{\eta}(k,t) = \{(1 - 4\varepsilon^3 k^6 + O(\varepsilon^4 k^8))\hat{F}(k) + \tfrac{1}{2}\left(\varepsilon^2 k^2 - 2\varepsilon^3 k^4 + O(\varepsilon^4 k^6)\right)\widehat{F^2}(k)\}\mathrm{e}^{\mathrm{i}\lambda_- t}$$

$$+\{(4\varepsilon^3k^6+O(\varepsilon^4k^8))\hat{F}(k)-\tfrac{1}{2}(\varepsilon^2k^2-2\varepsilon^3k^4+O(\varepsilon^4k^6))\widehat{F^2}(k)\}e^{i\lambda_+t}$$

$$+\frac{ik}{2\sqrt{1+4\varepsilon k^2}}\int_0^t\left[e^{i\lambda_+(t-\tau)}-e^{i\lambda_-(t-\tau)}\right]\eta^2(k,\tau)\,d\tau, \tag{3.5b}$$

as $\varepsilon\to 0$.

Theorem 3.2. *Let $F\in H^{m+3}$ for $m\geq 3/2$. Then there exists a positive number T of order at least $\varepsilon^{1/2}$ such that (3.2a) has a unique solution $\eta\in C(0,T;H^m)\cap C^1(0,T;H^{m-1})\cap C^2(0,T;H^{m-2})$. Furthermore, $\eta(\cdot,t)$ and $\eta_t(\cdot,t)$ are of order one, and $\eta_{tt}(\cdot,t)$ is of order ε^{-1} for $t\in[0,T]$.*

Proof. The contraction-mapping principle will be applied. Fix $T>0$ and define a mapping $\mathscr{A}$ on $C(0,T;H^m)$ by

$$\widehat{\mathscr{A}\eta}(k,t)=\frac{(1+2\varepsilon k^2+\sqrt{1+4\varepsilon k^2}-2\varepsilon^2k^4)\hat{F}(k)+\varepsilon^2k^2\widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}e^{i\lambda_-t}$$

$$-\frac{(1+2\varepsilon k^2-\sqrt{1+4\varepsilon k^2}-2\varepsilon^2k^4)\hat{F}(k)+\varepsilon^2k^2\widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}e^{i\lambda_+t}$$

$$+\frac{ik}{2\sqrt{1+4\varepsilon k^2}}\int_0^t\left[e^{i\lambda_+(t-\tau)}-e^{i\lambda_-(t-\tau)}\right]\eta^2(k,\tau)\,d\tau.$$

For $R,T>0$, let X be the closed ball

$$X=X_{R,T}=\{h\in C(0,T;H^m):\ \|h(\cdot,t)\|_m\leq R,\ \text{for } 0\leq t\leq T\}$$

in $C(0,T;H^m)$. We plan to choose R and T so that $\mathscr{A}$ is a contraction mapping of X to itself.

First, notice that at the zero-function $\theta(x,t)\equiv 0$

$$\widehat{\mathscr{A}\theta}(k,t)=\frac{(1+2\varepsilon k^2+\sqrt{1+4\varepsilon k^2}-2\varepsilon^2k^4)\hat{F}(k)+\varepsilon^2k^2\widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}e^{i\lambda_-t},$$

$$-\frac{(1+2\varepsilon k^2-\sqrt{1+4\varepsilon k^2}-2\varepsilon^2k^4)\hat{F}(k)+\varepsilon^2k^2\widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}e^{i\lambda_+t},$$

whence

$$|\widehat{\mathscr{A}\theta}(k,t)|\leq(1+\varepsilon k^2+\varepsilon^{3/2}|k|^3)|\hat{F}(k)|+\tfrac{1}{2}\varepsilon^{3/2}|k||\widehat{F^2}(k)|,$$

and thus

$$\|\mathscr{A}\theta\|_m\leq C(\|F\|_{m+3}),$$

where $C(\|F\|_{m+3})$ is a constant only dependent on $\|F\|_{m+3}$. Second, for any $\eta,\xi\in X$,

$$|\widehat{\mathscr{A}\eta}(k,t)-\widehat{\mathscr{A}\xi}(k,t)|\leq\frac{|k|}{\sqrt{1+4\varepsilon k^2}}\int_0^t|\widehat{\eta^2}(k,\tau)-\widehat{\xi^2}(k,\tau)|\,d\tau$$

$$\leq\frac{1}{2\varepsilon^{1/2}}\int_0^t|\widehat{\eta^2}(k,\tau)-\widehat{\xi^2}(k,\tau)|\,d\tau,$$

so,

$$\|\mathscr{A}\eta(\cdot,t)-\mathscr{A}\xi(\cdot,t)\|_m=\left(\int_{-\infty}^{\infty}(1+k^2)^m|\widehat{\mathscr{A}\eta}(k,t)-\widehat{\mathscr{A}\xi}(k,t)|^2\,\mathrm{d}k\right)^{1/2}$$
$$\le\frac{c_m}{2\varepsilon^{1/2}}\int_0^t\|\eta(\cdot,\tau)+\xi(\cdot,\tau)\|_m\|\eta(\cdot,\tau)-\xi(\cdot,\tau)\|_m\,\mathrm{d}\tau$$
$$\le c_mR\varepsilon^{-1/2}t\|\eta-\xi\|_{C(0,T;H^m)}.$$

Taking the maximum over $t\in[0,T]$ of both sides of this inequality gives

$$\|\mathscr{A}\eta-\mathscr{A}\xi\|_{C(0,T;H^m)}\le c_mRT\varepsilon^{-1/2}\|\eta-\xi\|_{C(0,T;H^m)},$$

and therefore

$$\|\mathscr{A}\eta\|_{C(0,T;H^m)}\le\|\mathscr{A}\theta\|_{C(0,T;H^m)}+\|\mathscr{A}\eta-\mathscr{A}\theta\|_{C(0,T;H^m)}$$
$$\le C(\|F\|_{m+3})+\varepsilon^{-1/2}c_mRT\|\eta\|_{C(0,T;H^m)}.$$

Hence, if we let $R=2\|\mathscr{A}\theta\|_{C(0,T;H^m)}\le 2C(\|F\||_{m+3})$ and $T=\varepsilon^{1/2}/4c_mR$, then the mapping $\mathscr{A}$ is a contraction from X to X. It follows there is a fixed point of $\mathscr{A}$ in X which is the unique solution η of (3.2a) in X. To prove the stated properties of η_t and η_{tt}, use is made of the identity

$$\hat\eta(k,t)=\widehat{\mathscr{A}\eta}(k,t). \tag{3.6}$$

The right-hand side of (3.6) is plainly differentiable with respect to t. Hence so is the left-hand side, and

$$\hat\eta_t(k,t)=(\mathscr{A}\widehat{\eta(k},t))_t$$
$$=\mathrm{i}\lambda_-\frac{(1+2\varepsilon k^2+\sqrt{1+4\varepsilon k^2}-2\varepsilon^2k^4)\hat F(k)+\varepsilon^2k^2\widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}\mathrm{e}^{\mathrm{i}\lambda_-t}$$
$$-\mathrm{i}\lambda_+\frac{(1+2\varepsilon k^2-\sqrt{1+4\varepsilon k^2}-2\varepsilon^2k^4)\hat F(k)+\varepsilon^2k^2\widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}\mathrm{e}^{\mathrm{i}\lambda_+t}$$
$$+\frac{\mathrm{i}k}{2\sqrt{1+4\varepsilon k^2}}\left\{\widehat{F^2}(k)\mathrm{e}^{\mathrm{i}\lambda_+t}+\int_0^t\mathrm{e}^{\mathrm{i}\lambda_+\tau}\widehat{2\eta\eta_t}(k,t-\tau)\,\mathrm{d}\tau\right\}$$
$$+\frac{\mathrm{i}k}{2\sqrt{1+4\varepsilon k^2}}\left\{-\widehat{F^2}(k)\mathrm{e}^{\mathrm{i}\lambda_-t}-\int_0^t\mathrm{e}^{\mathrm{i}\lambda_-\tau}\widehat{2\eta\eta_t}(k,t-\tau)\,\mathrm{d}\tau\right\}, \tag{3.7}$$

whence

$$|\hat\eta_t(k,t)|\le(3|k|^3+\varepsilon^{1/2}k^4)|\hat F(k)|+\frac12(|k|+\varepsilon^{1/2}k^2)|\widehat{F^2}(k)|+\frac{1}{\sqrt\varepsilon}\int_0^t|\widehat{\eta\eta_t}(\cdot,\tau)|\,\mathrm{d}\tau$$

and, thus,

$$\|\eta_t(\cdot,t)\|_{m-1} \le C(\|F\|_{m+3}) + \frac{1}{\varepsilon^{1/2}}\int_0^t \|\eta\eta_t(\cdot,t-\tau)\|_{m-1}\,\mathrm{d}\tau$$

$$\le C(\|F\|_{m+3}) + \frac{c_{m-1}\|\eta\|_{C(0,T;H^m)}}{\varepsilon^{1/2}}\int_0^t \|\eta_t(\cdot,\tau)\|_{m-1}\,\mathrm{d}\tau.$$

Gronwall's lemma then yields

$$\|\eta_t(\cdot,t)\|_{m-1} \le C(\|F\|_{m+3})\mathrm{e}^{c_{m-1}\|\eta(\cdot,t)\|_{C(0,T;H^m)}\varepsilon^{-1/2}t}. \tag{3.8}$$

The quantity $\|\eta\|_{C(0,T;H^m)}$ is order of one, so if t is order of $\varepsilon^{1/2}$, then η_t is also of order one. Since η_t exists and lies in $C(0,T;H^{m-1})$, it follows from (3.7) that η_{tt} also exists, and, moreover,

$$\hat{\eta}_{tt}(k,t) = (\widehat{\mathscr{A}\eta(k,t)})_{tt}$$

$$= -\lambda_-^2\frac{(1+2\varepsilon k^2+\sqrt{1+4\varepsilon k^2}-2\varepsilon^2k^4)\hat{F}(k)+\varepsilon^2k^4\widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}\mathrm{e}^{\mathrm{i}\lambda_- t}$$

$$+\lambda_+^2\frac{(1+2\varepsilon k^2-\sqrt{1+4\varepsilon k^2}-2\varepsilon^2k^4)\hat{F}(k)+\varepsilon^2k^2\widehat{F^2}(k)}{2\sqrt{1+4\varepsilon k^2}}\mathrm{e}^{\mathrm{i}\lambda_+ t}$$

$$+\frac{\mathrm{i}k}{2\sqrt{1+4\varepsilon k^2}}\Big\{\mathrm{i}\lambda_+\widehat{F^2}(k)\mathrm{e}^{\mathrm{i}\lambda_+ t}+2\widehat{\eta\eta_t}(k,0)\mathrm{e}^{\mathrm{i}\lambda_+ t}$$

$$+\,2\int_0^t \mathrm{e}^{\mathrm{i}\lambda_+\tau}(\widehat{\eta_t^2}+\widehat{\eta\eta_{tt}}(k,t-\tau))\,\mathrm{d}\tau\Big\}$$

$$+\frac{\mathrm{i}k}{2\sqrt{1+4\varepsilon k^2}}-\Big\{\mathrm{i}\lambda_-\widehat{F^2}(k)\mathrm{e}^{\mathrm{i}\lambda_- t}-2\widehat{\eta\eta_t}(k,0)\mathrm{e}^{\mathrm{i}\lambda_- t}$$

$$-\,2\int_0^t \mathrm{e}^{\mathrm{i}\lambda_- t}(\widehat{\eta_t^2}-\widehat{\eta\eta_{tt}}(k,t-\tau))\,\mathrm{d}\tau\Big\}.$$

Collecting terms in $\mathrm{e}^{\mathrm{i}\lambda_- t}$ and $\mathrm{e}^{\mathrm{i}\lambda_+ t}$, respectively, (for simplicity, suppose that the support of $\hat{F}$ is finite, so we can make use of the Taylor expansion) yields

$$\hat{\eta}_{tt} = \{-k^6\hat{F}(k)+\tfrac{1}{2}k^4\widehat{F^2}(k)-\mathrm{i}k\widehat{FF'''}-\mathrm{i}k\widehat{F^2F'}+O(\varepsilon)\}\mathrm{e}^{\mathrm{i}\lambda_- t}$$

$$+\Big\{-\lambda_+^2\,4\varepsilon^3k^6\hat{F}+\lambda_+^2(\tfrac{1}{2}\varepsilon^2k^2-\varepsilon^3k^3)\widehat{F^2}(k)+O(1)$$

$$-\frac{k\lambda_+}{2\sqrt{1+4\varepsilon k^2}}\widehat{F^2}+\frac{\mathrm{i}k}{2\sqrt{1+4\varepsilon k^2}}2(-\widehat{FF'''}-\widehat{F^2F'})\Big\}\,\mathrm{e}^{\mathrm{i}\lambda_+ t}$$

$$+\frac{\mathrm{i}k}{\sqrt{1+4\varepsilon k^2}}\int_0^t \mathrm{e}^{\mathrm{i}\lambda_+\tau}(\widehat{\eta_t^2}+\widehat{\eta\eta_{tt}}(k,t-\tau))\,\mathrm{d}\tau\Bigg\}$$

$$-\frac{\mathrm{i}k}{\sqrt{1+4\varepsilon k^2}}\int_0^t \mathrm{e}^{\mathrm{i}\lambda_-t}(\widehat{\eta_t^2}-\widehat{\eta\eta_{tt}}(k,t-\tau))\,\mathrm{d}\tau\Bigg\}$$

$$=O(1)\mathrm{e}^{\mathrm{i}\lambda_- t}+\Bigg\{-\frac{(1+2\varepsilon k^2-\varepsilon^2k^4)^2}{\varepsilon^4k^2}4\varepsilon^3k^6\hat{F}$$

$$+\frac{(1+2\varepsilon k^2-\varepsilon^2k^4)^2}{2\varepsilon^4k^2}(\varepsilon^2k^2-2\varepsilon^3k^3)\widehat{F^2}$$

$$-\frac{k(1+2\varepsilon k^2-\varepsilon^2k^4)}{2\varepsilon^2k\sqrt{1+4\varepsilon k^2}}\widehat{F^2}+O(1)\Bigg\}\,\mathrm{e}^{\mathrm{i}\lambda_+t}$$

$$+\frac{\mathrm{i}k}{\sqrt{1+4\varepsilon k^2}}\int_0^t[(\mathrm{e}^{\mathrm{i}\lambda_+(t-\tau)}-\mathrm{e}^{\mathrm{i}\lambda_-(t-\tau)})\widehat{\eta_t^2}(k,\tau)$$

$$+(\mathrm{e}^{\mathrm{i}\lambda_+(t-\tau)}+\mathrm{e}^{\mathrm{i}\lambda_-(t-\tau)})]\widehat{\eta\eta_{tt}}(k,\tau)\,\mathrm{d}\tau.$$

Simplifying the above gives

$$\hat{\eta}_{tt}=O(1)\mathrm{e}^{\mathrm{i}\lambda_-t}+\Bigg\{-\frac{4k^4}{\varepsilon}\hat{F}+O(1)\Bigg\}\,\mathrm{e}^{\mathrm{i}\lambda_+t}$$

$$+\frac{\mathrm{i}k}{\sqrt{1+4\varepsilon k^2}}\int_0^t(\mathrm{e}^{\mathrm{i}\lambda_+(t-\tau)}-\mathrm{e}^{\mathrm{i}\lambda_-(t-\tau)})\widehat{\eta_t^2}(k,\tau)\,\mathrm{d}\tau$$

$$+\frac{\mathrm{i}k}{\sqrt{1+4\varepsilon k^2}}\int_0^t(\mathrm{e}^{\mathrm{i}\lambda_+(t-\tau)}+\mathrm{e}^{\mathrm{i}\lambda_-(t-\tau)})\widehat{\eta\eta_{tt}}(k,\tau))\,\mathrm{d}\tau,$$

as $\varepsilon\to 0$. It follows readily that

$$\|\eta_{tt}(\cdot,t)\|_{m-2}\le C(\|F\|_{m+3})\frac{1}{\varepsilon}$$

$$+\frac{1}{\varepsilon^{1/2}}\int_0^t\|\eta_t^2(\cdot,\tau)\|_{m-1}\mathrm{d}\tau+\frac{1}{\varepsilon^{1/2}}\int_0^t\|\eta\eta_{tt}(\cdot,t-\tau)\|_{m-2}\,\mathrm{d}\tau.$$

Because η and η_t are order-one, at least on the time interval $[0,\varepsilon^{1/2}]$, Gronwall's principle brings out the inequality

$$\|\eta_{tt}(\cdot,t)\|_{m-2}\le C(\|F\|_{m+3})\varepsilon^{-1}\mathrm{e}^{\varepsilon^{-1/2}t},$$

or

$$\|\varepsilon\eta_{tt}(\cdot,t)\|_{m-2}\le C(\|F\|_{m+3})\mathrm{e}^{\varepsilon^{-1/2}t}.$$

The theorem is proved. □

Remark. *By now, it becomes clear that the difference and similarity between system (3.1) and the KdV-equation is somewhat subtle. From our analysis, it is seen that formal order-one quantities are not necessarily order-one in fact.*

Corollary 3.3. *Under the assumptions in Theorem 3.2, the initial-value problem (3.2) has a solution in* $C^2(0,T_0;H^m)$ *for some* $T_0 \geq \varepsilon^{-1}/C\|F\|_{m+3}$, *where the constant* C *is independent of* ε *and* F, *and for* $0 \leq t \leq \min\{T_0, O(\varepsilon^{-3/2})\}$, *we have the following estimates*:

$$\|v_{(j)}(\cdot,t)\| \leq E_j \varepsilon^{(j/2)+(3/4)},$$

$$\sup_{x\in\mathbb{R}} |v_{(j)}(x,t)| \leq N_j \varepsilon^{(j/2)+1}$$

for $j=1,2,\ldots$, *where* E_j *and* N_j *are order-one constants only dependent on* F *and* j.

Proof. This is simply a reinterpretation of Theorem 3.2 in the dependent variable v instead of the variable η. □

4. Comparison to the KdV-equation

According to the view put forward in the previous sections, the initial-value problem for (1.8) needs to be imposed in the form

$$\begin{aligned} &v_t + v_x + vv_x + v_{ttx} = 0,\\ &v(x,0) = \varepsilon F(\varepsilon^{1/2}x),\\ &v_t(x,0) = -\varepsilon^{3/2}F'(\varepsilon^{1/2}x) - \varepsilon^{5/2}(F'''(\varepsilon^{1/2}x) + F(\varepsilon^{1/2}x)F'(\varepsilon^{1/2}x)), \end{aligned} \tag{4.1}$$

to model small-amplitude long waves at the same level of approximation as does the KdV-equation. To compare the new equation (4.1) with KdV, as in (3.2a), define

$$\begin{aligned} &u(x,t) = \varepsilon\xi(\varepsilon^{1/2}(x-t), \varepsilon^{3/2}t),\\ &v(x,t) = \varepsilon\eta(\varepsilon^{1/2}(x-t), \varepsilon^{3/2}t), \end{aligned} \tag{4.2}$$

or

$$\begin{aligned} &\xi(x,t) = \varepsilon^{-1}u(\varepsilon^{-1/2}x + \varepsilon^{-3/2}t, \varepsilon^{-3/2}t),\\ &\eta(x,t) = \varepsilon^{-1}v(\varepsilon^{-1/2}x + \varepsilon^{-3/2}t, \varepsilon^{-3/2}t). \end{aligned}$$

Then, problems (4.1) and (4.2) take the form

$$\begin{aligned} &\eta_t + \eta\eta_x + \eta_{xxx} - 2\varepsilon\eta_{xxt} + \varepsilon^2\eta_{xtt} = 0,\\ &\xi_t + \xi\xi_x + \xi_{xxx} = 0,\\ &\eta(x,0) = \xi(x,0) = F(x), \quad \eta_t(x,0) = -F'''(x) - F(x)F'(x). \end{aligned} \tag{4.3}$$

Set $w=\eta-\xi$, so that $\eta=w+\xi$. Straightforward calculation shows that w satisfies the initial-value problem

$$\begin{aligned} &w_t+ww_x+w_{xxx}-2\varepsilon w_{xxt}=-(\xi w)_x+2\varepsilon\xi_{xxt}-\varepsilon^2\eta_{xtt},\\ &w(x,0)=0,\quad w_t(x,0)=0. \end{aligned} \tag{4.4}$$

Multiply (4.4) by $2w$, then integrate over $\mathbb{R}$ with respect to x: simple estimates lead to the inequality

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}(w^2+2\varepsilon w_x^2)\,\mathrm{d}x &= \int_{-\infty}^{\infty}(-(\xi w)_x+2\varepsilon\xi_{xxt}-\varepsilon^2\eta_{xtt})2w\,\mathrm{d}x\\ &= \int_{-\infty}^{\infty} -\xi_x w^2+2(\varepsilon\xi_{ttx}-2\varepsilon^2\eta_{xxt})w\\ &\le |\xi_x|_\infty\|w\|^2+2\varepsilon\|2\xi_{xxt}-\varepsilon\eta_{xtt}\|\,\|w\|. \end{aligned} \tag{4.5}$$

Integrating (4.5) over $[0,t]$ yields

$$\|w(\cdot,t)\|^2+2\varepsilon\|w_x^2(\cdot,t)\|\le\int_0^t(|\xi_x|_\infty\|w(\cdot,\tau)\|^2+2\varepsilon\|2\xi_{xxt}-\varepsilon\eta_{xtt}\|\,\|w(\cdot,\tau)\|)\,\mathrm{d}\tau.$$

As we have already derived that $\|\varepsilon\eta_{tt}\|_m=O(1)$ at least for $0\le t\le T_0$, where T_0 is as in Corollary 3.3, we may apply a variation of Gronwall's lemma to come to

$$\|w(\cdot,t)\|\le\frac{\varepsilon\|2\xi_{xxt}-\varepsilon\eta_{xtt}\|}{|\xi_x|_\infty}(\mathrm{e}^{|\xi_x|_\infty t}-1)\le C\varepsilon t. \tag{4.6}$$

Define $w_{(k)}$ to be $\partial_x^k w$, for $k=0,1,2,\ldots$, and note $w_{(0)}=w$. Suppose it is true that for $0\le t\le T_0$,

$$\|w_{(j)}(\cdot,t)\|\le C_j\varepsilon t, \tag{4.7}$$

for all integers from 0 to $n-1$. To estimate $\|w_{(n)}(\cdot,t)\|$, multiply (4.3) by $(-1)^n 2w_{(2n)}(x,t)$ and integrate the result over $\mathbb{R}$ with respect to x, to obtain

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\int(w_{(n)}^2(x,t)+2\varepsilon w_{(n+1)}^2(x,t))\,\mathrm{d}x\\ &=-(-1)^n\int ww_xw_{(2n)}(x,t)\,\mathrm{d}x-(-1)^n\int(\xi w)_xw_{(2n)}(x,t)\,\mathrm{d}x\\ &\quad+(-1)^n\varepsilon\int(2\xi_{xxt}-\varepsilon\eta_{xtt})w_{(2n)}(x,t)\,\mathrm{d}x\\ &=-\int(ww_x)_{(n)}w_{(n)}(x,t)\,\mathrm{d}x-\int(\xi w)_{(n+1)}w_{(n)}(x,t)\,\mathrm{d}x\\ &\quad+\int(2\xi_{xxt}-\varepsilon\eta_{xtt})_{(n)}w_{(n)}(x,t)\,\mathrm{d}x\\ &=-\sum_{j=0}^{n}\binom{n}{j}\int w_{(j)}w_{(n+1-j)}w_{(n)}(x,t)\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&-\sum_{j=0}^{n}\binom{n+1}{j}\int \xi_j w_{(n+1-j)} w_{(n)}(x,t)\,\mathrm{d}x \\
&+\varepsilon\int (2\partial_t \xi_{(n+2)} - \varepsilon\partial_t^2 \eta_{(n+1)}) w_{(n)}(x,t)\,\mathrm{d}x \\
=&\frac{1}{2}\int \Big(w_x w_{(n)}^2(x,t) + \sum_{j=0}^{\lfloor n/2 \rfloor}\binom{n}{j} w_{(2j-1)} w_{(n-j+1)}^2(x,t) \\
&+ \sum_{j=0}^{\lfloor (n+1)/2 \rfloor}\binom{n}{j} \xi_{(2j-1)} w_{(n-j+1)}^2(x,t)\Big)\,\mathrm{d}x \\
&+\varepsilon\int (2\partial_t \xi_{(n+2)} - \varepsilon\partial_t^2 \eta_{(n+1)}) w_{(n)}(x,t)\,\mathrm{d}x \\
\le&\; C_n \|w_{(n)}(\cdot,t)\|^2 + C_n\varepsilon \|w_{(n)}(\cdot,t)\|, \qquad (4.8)
\end{aligned}
$$

where $\lfloor s \rfloor$ is the biggest integer which is less than or equal to s. Applying Gronwall's inequality again leads to

$$\|w_{(n)}(\cdot,t)\| \le C_n \varepsilon t, \qquad (4.9)$$

for $0 \le t \le T_0$. Thus inequality (4.9) is true for all $n = 0, 1, \ldots, m$. These ruminations are recorded in the following theorem.

Theorem 4.1. *Let $F \in H^{m+5}$ where $m \ge 0$. Let $\varepsilon > 0$ and let u be the solution of the initial-value problem* (1.1) *for the KdV-equation with initial data F and let v be the solution of the initial-value problem* (3.2). *Then there exists an $\varepsilon_0 > 0$ and order-one constants C_j such that if $0 \le \varepsilon \le \varepsilon_0$, then for $0 \le j \le m$,*

$$
\begin{aligned}
&\|u(\cdot,t) - v(\cdot,t)\| \le C\varepsilon^{7/4}(\varepsilon^{3/2} t), \\
&\|u_{(j)}(\cdot,t) - v_{(j)}(\cdot,t)\| \le C_j \varepsilon^{(7/4)+(j/2)}(\varepsilon^{3/2} t),
\end{aligned} \qquad (4.10)
$$

for $0 \le t \le \min\{\varepsilon^{-3/2}, T_0\}$, where T_0 is as in Corollary 3.3.

Proof. To obtain the result (4.10), we need only use the fact that

$$u(x,t) = \varepsilon\xi(\varepsilon^{1/2}(x-t), \varepsilon^{3/2} t) \quad \text{and} \quad v(x,t) = \varepsilon\eta(\varepsilon^{1/2}(x-t), \varepsilon^{3/2} t).$$

Substituting these relationships into (4.9) gives (4.10). □

Appendix

In 1871, Boussinesq derived the original system of equations

$$
\begin{cases}
h_t + (uh)_x = 0, \\
u_t + u u_x + g h_x + \dfrac{1}{3} h_0 h_{xtt} = 0,
\end{cases} \qquad (\mathrm{B})
$$

for the two-way propagation of long-crested surface water waves in a channel that bears his name. Here, the variables are dimensional, h_0 is the undisturbed water depth, h=$h(x,t)$ is the depth of water at the point in the channel corresponding to x at time t, g is the gravitational constant and $u = u(x,t)$ is the horizontal velocity of the free surface corresponding to x at time t. If the dispersive term h_{xtt} is dropped in (B), the resulting system is valid for waves of extreme length and are commonly called the shallow-water equations (cf. Whitham [16] formula (13.92)). As is well known, the characteristic velocities of this hyperbolic system are $u \pm \sqrt{gh}$ and the Riemann invariants are $u \pm 2\sqrt{gh}$. If attention is focussed upon waves moving only to the right, say, these therefore satisfy

$$u = 2\sqrt{g(h_0+\eta)} - 2\sqrt{gh_0}, \tag{R}$$

where $\eta = h - h_0$ is the deviation of the free surface from its rest position at the point x at time t (see again, Whitham [16] formula (13.80)).

Rescale the independent variables x and t by λ and $\lambda/\sqrt{gh_0}$, respectively, and the dependent variables η and u by a and $\sqrt{gh_0}$, respectively, where a is the maximum amplitude of the wavefield being modelled and λ is a typical wavelength. The Boussinesq system and the relation (R) both subsist on a small-amplitude assumption wherein the dimensionless parameter $\alpha = a/h_0$ is taken to be small. Rewrite h as

$$h = h_0 + h_0\alpha\eta', \tag{A$_1$}$$

and the velocity u in the form

$$u = 2\sqrt{gh_0}(\sqrt{1+\alpha\eta'} - 1) = 2\sqrt{gh_0}\left(\frac{\alpha}{2}\eta' - \frac{\alpha^2}{8}\eta'^2\right) + O(\alpha^3)$$
$$= \alpha\sqrt{gh_0}(\eta' - \frac{\alpha}{4}\eta'^2) + O(\alpha^3) \simeq \alpha\sqrt{gh_0}\left(\eta' - \frac{\alpha}{4}\eta'^2\right), \tag{A$_2$}$$

as $\alpha \to 0$. Substitute these consequences of a small-amplitude assumption and unidirectionality into the second equation in (B) and remove the prime to obtain

$$\alpha\sqrt{gh_0}\left(\eta - \frac{\alpha}{4}\eta^2\right)_t \frac{\sqrt{gh_0}}{\lambda} + (\alpha\sqrt{gh_0})^2\left(\eta - \frac{\alpha}{4}\eta^2\right)\left(\eta - \frac{\alpha}{4}\eta^2\right)_x \frac{1}{\lambda}$$
$$+ \alpha g h_0 \eta_x \frac{1}{\lambda} + \frac{1}{3}h_0^2\alpha\eta_{xtt}\frac{gh_0}{P\lambda^3} = O(\alpha^3).$$

Dividing both sides of the last equation by $\alpha gh_0/\lambda$ yields

$$\left(\eta - \frac{\alpha}{4}\eta^2\right)_t + \alpha\left(\eta - \frac{\alpha}{4}\eta^2\right)\left(\eta - \frac{\alpha}{4}\eta^2\right)_x + \eta_x + \frac{1}{3}\frac{h_0^2}{\lambda^2}\eta_{xtt} = O(\alpha^2).$$

Denote h_0^2/λ^2 by β. Then the last equation may be rewritten as

$$\eta_t - \frac{\alpha}{2}\eta\eta_t + \alpha\eta\eta_x + \eta_x + \frac{1}{3}\beta\eta_{xtt} = O(\alpha^2),$$

or, what is the same,

$$\left(1 - \frac{\alpha}{2}\eta\right)\eta_t + \eta_x + \alpha\eta\eta_x + \frac{1}{3}\beta\eta_{xtt} = O(\alpha^2).$$