Comparison of Nonlocal Nonlinear Wave Equations in the Long-Wave Limit

H. A. Erbay[1], S. Erbay[1], A. Erkip[2]

[1]Department of Natural and Mathematical Sciences, Faculty of Engineering, Ozyegin University, Cekmekoy 34794, Istanbul, Turkey

[2]Faculty of Engineering and Natural Sciences, Sabanci University, Tuzla 34956, Istanbul, Turkey

Abstract

We consider a general class of convolution-type nonlocal wave equations modeling bidirectional nonlinear wave propagation. The model involves two small positive parameters measuring the relative strengths of the nonlinear and dispersive effects. We take two different kernel functions that have similar dispersive characteristics in the long-wave limit and compare the corresponding solutions of the Cauchy problems with the same initial data. We prove rigorously that the difference between the two solutions remains small over a long time interval in a suitable Sobolev norm. In particular our results show that, in the long-wave limit, solutions of such nonlocal equations can be well approximated by those of improved Boussinesq-type equations.

1 Introduction

In the present work we start with the nonlocal wave equation:

$$u_{tt} = \beta_\delta * (u + \epsilon^p u^{p+1})_{xx}, \quad x \in \mathbb{R}, \quad t > 0, \tag{1}$$

$$u(x,0) = u_0(x), \quad u_t(x,0) = u_1(x), \quad x \in \mathbb{R}, \tag{2}$$

where $u = u(x,t)$ is a real-valued function, ϵ and δ are two small positive parameters measuring the effect of nonlinearity and the effect of dispersion, respectively, p is a positive integer, the symbol $*$ denotes convolution in the

---

E-mail: husnuata.erbay@ozyegin.edu.tr, saadet.erbay@ozyegin.edu.tr, albert@sabanciuniv.edu

x-variable and $\beta_\delta(x) = \frac{1}{\delta}\beta(\frac{x}{\delta})$ where β is a kernel function. The two parameter family (1) is obtained from a single equation given with the kernel $\beta(x)$ through a suitable scaling. Our main purpose is to investigate the dependence of the solutions of (1)-(2) on the kernel β in the long-wave limit. To that end, we take two different kernel functions for which the corresponding convolution operators are elliptic (and bounded) and compare the corresponding solutions in the case of the same initial data. We prove that, over a long time interval of length of order $1/\epsilon^p$, the difference between the two solutions is of order δ^{2k} (uniformly in ϵ) where the integer k is determined by the dispersive behaviors of the two kernels.

The nonlocal equation (1) describes the one-dimensional motion of a non-locally and nonlinearly elastic medium and u represents the strain (we refer the reader to [1] for a detailed description of the nonlocally and nonlinearly elastic medium). The dispersive nature of the elastic medium described by (1) is revealed by the kernel β_δ and the parameter δ is a measure of wavenumber so that smaller δ implies longer wavelength. The parameter ϵ assures that the solutions exist over a long time interval of length of order $1/\epsilon^p$ (see [2]). For particular choices of the kernel function, (1) involves many well-known nonlinear wave equations. Two well-known examples that appear as model equations for various physical problems are the improved Boussinesq equation

$$u_{tt} - u_{xx} - \delta^2 u_{xxtt} = \epsilon^p (u^{p+1})_{xx} \tag{3}$$

corresponding to the exponential kernel $\beta_\delta(x) = \frac{1}{2\delta}e^{-|x|/\delta}$ and the classical elasticity equation

$$u_{tt} - u_{xx} = \epsilon^p (u^{p+1})_{xx} \tag{4}$$

corresponding to the Dirac measure. We refer the reader to [1] for other examples of the kernels widely used in continuum mechanics.

A relevant question is how the choice of the kernel function of (1) affects solutions. In that respect one needs a measure for the closeness of two kernels. In this work we show that the dispersive characteristics of the kernel in the long-wave limit provides a suitable measure. We note that in the long-wave limit the Fourier modes are concentrated about the small wave numbers and therefore the dispersive nature of (1) in the small wavenumber regime is related to the Taylor expansion of the Fourier transform of the kernel around zero. As the Taylor coefficients are determined by the moments of the kernel, we can explicitly classify "close" kernels by comparing their moments and investigate how the family of solutions of (1) depend on the kernel. To be more precise we take two kernels $\beta^{(1)}$ and $\beta^{(2)}$ with the same moments up to order $2k$ and consider the corresponding solutions $u_1^{\epsilon,\delta}$ and $u_2^{\epsilon,\delta}$ of (1)-(2) with the same initial values. We then prove that, for a suitable norm, $\|u_1^{\epsilon,\delta}(t) - u_2^{\epsilon,\delta}(t)\|$ is of order δ^{2k} over a long time interval of length of order $1/\epsilon^p$. This shows that for varying kernels with the same dispersive nature, solutions of (1) approximate each other and the approximation errors originate from the dispersive nature of the kernels rather than from their shapes. We note that, in the terminology of some authors, our long-time existence results [2] together with the work presented here are

2

in fact consistency, existence, convergence results for the nonlocal bidirectional approximations of the nonlocal equation (1). We refer to [3–7] and the references therein for a detailed discussion of these concepts.

This work differs from most previous studies available in the literature for several reasons; it does not focus on unidirectional solutions and it compares solutions of two families of equations of the same type. In the literature, there have been a number of works comparing solutions of a parent equation with those of a model equation describing the unidirectional propagation of long waves. The most well-known examples of one-dimensional model equations are the Korteweg-de Vries (KdV) equation [8], the Benjamin-Bona-Mahony (BBM) equation [9], the Camassa-Holm (CH) equation [10] and the Fornberg-Whitham (FW) equation [11]. For a comprehensive treatment of the methods on comparisons of solutions between those unidirectional equations and the parent equations (for instance, in the context of a shallow water approximation) we refer to [12–14] for the case of the KDV and the BBM equations and to [4,7] for the case of the CH equation (see, for instance, [5,15] for comparisons of two-dimensional model equations). In a recent study [16], the present authors have made similar comparisons between (1) in the context of nonlocal elasticity and the CH equation. For emphasis, we remind the reader that all these studies consider unidirectional approximations of nonlinear dispersive equations whereas the present work is about bidirectional approximations. Another strength of this work lies in its level of generality because it poses minimal restrictions on the kernel.

The structure of the paper is as follows. Section 2 is devoted to preliminaries. In Section 3 we state the moment conditions to be satisfied by the kernels, prove the main result that establishes the estimate on the difference between two families of solutions, and discuss how our general result can be applied to certain cases.

Throughout this paper, we use the standard notation for function spaces. The Fourier transform $\widehat{u}$ of u is defined by $\widehat{u}(\xi) = \int_{\mathbb{R}} u(x)e^{-i\xi x}dx$. The L^p $(1 \leq p < \infty)$ norm of u on $\mathbb{R}$ is represented by $\|u\|_{L^p}$. To denote the inner product of u and v in L^2, the symbol $\langle u, v \rangle$ is used. The notation $H^s = H^s(\mathbb{R})$ is used to denote the L^2-based Sobolev space of order s on $\mathbb{R}$, with the norm $\|u\|_{H^s} = \left(\int_{\mathbb{R}} (1 + \xi^2)^s |\widehat{u}(\xi)|^2 d\xi \right)^{1/2}$. All integrals in this paper extend over the whole real line and the limits of integration will not be explicitly written. C is a generic positive constant. Partial differentiations are denoted by D_x etc.

2 Preliminaries

In this section we give a brief discussion of the nonlocal equation and provide some preliminaries. We first note that the family of equations (1) can be obtained from the single equation

$$u_{tt} = \beta * \left(u + \epsilon^p u^{p+1} \right)_{xx}, \tag{5}$$

under the change of variables $(x, t) \to (x/\delta, t/\delta)$. The nonlocal equation (5) with a general even kernel β was proposed in [1] to model longitudinal motions

in nonlinear nonlocal elasticity, in terms of non-dimensional quantities.

In [1], local well-posedness of the Cauchy problem for (5) (and hence for (1)) was proved under the regularization assumption

$$0 \leq \widehat{\beta}(\xi) \leq C \left(1 + \xi^2\right)^{-r/2} \tag{6}$$

for some $r \geq 2$. In this case (5) becomes an H^s-valued ordinary differential equation (ODE). In a recent work [2], this result has been improved in two ways. First, the condition (6) has been replaced by the ellipticity and boundedness condition

$$c_1^2 \leq \widehat{\beta}(\xi) \leq c_2^2 \tag{7}$$

for some $c_1, c_2 > 0$. The condition (7) implies that the convolution operator is invertible. On the other hand, (7) lacks the regularization effect of (6) for $r \geq 2$. Hence the regularity requirement $r \geq 2$ in (6) is replaced by the much weaker condition $r \geq 0$. Secondly, the long-time existence of solutions to the family of initial-value problems of (5) has been established for times up to $\mathcal{O}(1/\epsilon^p)$. We note that when $0 \leq r < 2$, the nonlocal equation (1) is no longer an ODE but the parameter ϵ can be chosen small enough so that it is in the hyperbolic regime. Theorem 5.2 of [2] is as follows:

Theorem 2.1. *Suppose the kernel β satisfies the ellipticity and boundedness condition (7). Let $D > 3$, $P > P_{min}$, $s > \frac{7}{2}$ and $(u_0^\epsilon, w_0^\epsilon)$ be bounded in $H^{s+P} \times H^{s+P}$. Then there exist some $\epsilon_0 > 0$, $T > 0$ and a unique family $(u^\epsilon)_{0 < \epsilon < \epsilon_0}$ bounded in $C\big([0, \frac{T}{\epsilon^p}]; H^{s+D}\big) \cap C^1\big([0, \frac{T}{\epsilon^p}]; H^{s+D-1}\big)$ and satisfying (5) with initial values $u_0 = u_0^\epsilon$, $u_1 = \big(w_0^\epsilon\big)_x$.*

The numbers P and D of this theorem are related to the required extra smoothness and the restrictions on these numbers are due to the Nash-Moser scheme used in the proof of Theorem 2.1 (see also [17] for more details). For a given $D > 3$, the number $P_{\min}$ is determined as $P_{\min} = 3 + \frac{D}{D-3}\left(\sqrt{3} + \sqrt{2D}\right)^2$. The computation in [2] shows that the optimal values of D and P are approximately 7.35 and 55.34, respectively.

Throughout this work we will assume that the kernel $\beta(x)$ is an even function with $\int \beta(x)dx = 1$ and satisfies (7).

In order to write the nonlocal equation (5) as a first-order system we now introduce the pseudo-differential operator

$$Kw(x) = \mathcal{F}^{-1}\Big(k(\xi)\widehat{w}(\xi)\Big), \tag{8}$$

where $k(\xi) = \sqrt{\widehat{\beta}(\xi)}$ and $\mathcal{F}^{-1}$ denotes the inverse Fourier transform. We note that $K^2 w = \beta * w$ and, by (7), K is an invertible bounded operator on H^s. We then convert (5) to

$$u_t = Kv_x, \tag{9}$$

$$v_t = K\left(u + \epsilon^p u^{p+1}\right)_x. \tag{10}$$

4

As mentioned earlier, applying the coordinate transformation $(x,t) \to (x/\delta, t/\delta)$ to (5) yields (1) with $\beta_\delta(x) = \frac{1}{\delta}\beta(\frac{x}{\delta})$ (note that the Fourier domain counterpart of the latter is $\widehat{\beta_\delta}(\xi) = \widehat{\beta}(\delta\xi)$). Similarly, with the same coordinate transformation, the initial-value problem for (9)-(10) reduces to

$$u_t = K_\delta v_x, \qquad\qquad u(x,0) = u_0(x), \qquad\qquad (11)$$

$$v_t = K_\delta \left(u + \epsilon^p u^{p+1}\right)_x, \qquad v(x,0) = v_0(x), \qquad\qquad (12)$$

with the operator $K_\delta w = \mathcal{F}^{-1}\big(k_\delta(\xi)\widehat{w}(\xi)\big)$ and $k_\delta(\xi) = k(\delta\xi)$. Clearly, when $u_1 = (w_0)_x$ with $v_0 = K_\delta^{-1}w_0$ the initial-value problem (1)-(2) becomes equivalent to (11)-(12); hence we have long time existence and uniform bounds for the solutions $\left(u^{\epsilon,\delta}, v^{\epsilon,\delta}\right)$ of (11)-(12). In fact the long-time existence result in [2] was first proved for (11)-(12) and then extended to (1)-(2).

Remark 1. The uniform bounds for the solution in Theorem 2.1 depend only on T, the bounds on the initial data and the bound for the operator K. Due to (7), the family of operators K_δ all have the same bound as K; therefore Theorem 2.1 also holds uniformly in δ when K is replaced by K_δ or β is replaced by β_δ.

Corollary 2.2. *Suppose the kernel β satisfies the ellipticity and boundedness condition (7). Let $D > 3$, $P > P_{min}$, $s > \frac{7}{2}$ and $(\varphi, \psi) \in H^{s+P} \times H^{s+P}$. Then there are some $\epsilon_0 > 0$, $T > 0$ so that for all $0 < \epsilon < \epsilon_0$, the family of Cauchy problems for (1) with initial values $(u_0, u_1) = (\varphi, \psi_x)$ have unique solutions $u^{\epsilon,\delta}$ on the interval $[0, \frac{T}{\epsilon^p}]$, uniformly bounded in $C\big([0, \frac{T}{\epsilon^p}]; H^{s+D}\big) \cap C^1\big([0, \frac{T}{\epsilon^p}]; H^{s+D-1}\big)$.*

3　Comparison of Solutions

In this section we prove the main result of this paper: the difference between the solutions of the Cauchy problems corresponding to two different kernel functions remains small over a long time interval in a suitable Sobolev norm if the kernels have the same dispersive nature in the long-wave limit.

Consider two different kernel functions $\beta^{(1)}$ and $\beta^{(2)}$ satisfying the following three conditions for some $k \geq 1$:

(C1)　$\beta^{(1)}$ and $\beta^{(2)}$ satisfy the ellipticity and boundness condition (7),

(C2)　$\beta^{(1)}$ and $\beta^{(2)}$ have the same first $(2k-1)$-order moments, namely

$$\int x^j \beta^{(1)}(x)dx = \int x^j \beta^{(2)}(x)dx \quad \text{for} \quad 0 \leq j < 2k-1, \qquad (13)$$

(C3)　$x^{2k}\beta^{(i)}(x) \in L^1(\mathbb{R})$ $(i = 1, 2)$.

Clearly in the case when $\beta = \mu$ is a finite measure, the moment integral should be replaced by $\int x^j d\mu$. We consider (1)-(2) with $\beta^{(1)}$ and $\beta^{(2)}$ and denote the

corresponding solutions by $u_1^{\epsilon,\delta}$ and $u_2^{\epsilon,\delta}$, respectively. Our aim is to estimate the difference $u_1^{\epsilon,\delta} - u_2^{\epsilon,\delta}$.

Before proceeding to state the main result, we introduce a lemma which provides certain commutator estimates (see Proposition B.8 of [6]).

Lemma 3.1. *Let $q_0 > 1/2$, $s \geq 0$ and let $[\Lambda^s, w]g = \Lambda^s(wg) - w\Lambda^s g$ with $\Lambda = (1 - D_x^2)^{1/2}$.*

1. If $0 \leq s \leq q_0 + 1$ and $w \in H^{q_0+1}$ then, for all $g \in H^{s-1}$, one has

$$\|[\Lambda^s, w]g\|_{L^2} \leq C\|w_x\|_{H^{q_0}}\|g\|_{H^{s-1}},$$

2. If $-q_0 < r \leq q_0 + 1 - s$ and $w \in H^{q_0+1}$ then, for all $g \in H^{r+s-1}$, one has

$$\|[\Lambda^s, w]g\|_{H^r} \leq C\|w_x\|_{H^{q_0}}\|g\|_{H^{r+s-1}}.$$

We will use this lemma in finding the energy estimates for the system (17)-(18) to be introduced below. We are now ready to prove the main result.

Theorem 3.2. *Let $\beta^{(1)}$ and $\beta^{(2)}$ be two kernels satisfying the conditions (C1), (C2) and (C3) for some $k \geq 1$. Let P, D, s be as in Theorem 2.1 and $\varphi, \psi \in H^{s+P+2k+1}$. Then there are some constants ϵ_0, C and $T > 0$ independent of ϵ ($0 < \epsilon < \epsilon_0$) and δ so that the solutions $u_i^{\epsilon,\delta}$ of the Cauchy problems*

$$u_{tt} = \beta_\delta^{(i)} * (u + \epsilon^p u^{p+1})_{xx}, \qquad x \in \mathbb{R}, \quad t > 0, \qquad (14)$$

$$u(x,0) = \varphi(x), \qquad u_t(x,0) = \psi_x(x), \qquad x \in \mathbb{R}, \qquad (15)$$

for $i = 1, 2$ are defined for all $t \in \left[0, \frac{T}{\epsilon^p}\right]$ and satisfy

$$\|u_1^{\epsilon,\delta}(t) - u_2^{\epsilon,\delta}(t)\|_{H^{s+D}} \leq C\delta^{2k}(1+t) \quad \text{for all } t \leq \frac{T}{\epsilon^p}. \qquad (16)$$

Proof. We will complete the proof in several steps. For the rest of the proof we will drop the superscripts ϵ, δ to simplify the notation.

Step 1 Since $\varphi, \psi \in H^{s+P+2k+1}$, Theorem 2.1 gives the uniform bound

$$\|u_i(t)\|_{H^{s+D+2k+1}} + \|(u_i)_t(t)\|_{H^{s+D+2k}} \leq C \quad \text{for all } t \leq \frac{T}{\epsilon^p}, \quad (i = 1, 2)$$

for both families of solutions.

Step 2 Converting (14)-(15) into the corresponding systems of the form (11)-(12) we obtain

$$(u_i)_t = K_\delta^{(i)}(v_i)_x, \qquad u_i(x,0) = \varphi(x),$$

$$(v_i)_t = K_\delta^{(i)}\left(u_i + \epsilon^p u_i^{p+1}\right)_x, \qquad v_i(0)(x,0) = (K_\delta^{(i)})^{-1}\psi(x)$$

6

for $i = 1, 2$. Then the pair (r, ρ), which are defined by the differences $r = u_1 - u_2$ and $\rho = v_1 - v_2$ between the solutions, satisfy:

$$r_t = K_\delta^{(2)} \rho_x + f_1, \qquad r(x, 0) = 0, \tag{17}$$

$$\rho_t = K_\delta^{(2)} r_x + \epsilon^p K_\delta^{(2)} (wr)_x + f_2, \qquad \rho(x, 0) = g(x), \tag{18}$$

where

$$f_1 = \left(K_\delta^{(1)} - K_\delta^{(2)} \right) (v_1)_x, \tag{19}$$

$$f_2 = \left(K_\delta^{(1)} - K_\delta^{(2)} \right) \left(u_1 + \epsilon^p u_1^{p+1} \right)_x, \tag{20}$$

$$g = \left(\left(K_\delta^{(1)} \right)^{-1} - \left(K_\delta^{(2)} \right)^{-1} \right) \psi, \tag{21}$$

$$w = u_1^p + u_1^{p-1} u_2 + \cdots + u_1 u_2^{p-1} + u_2^p. \tag{22}$$

Step 3 Conditions (C2) and (C3) mean that the Fourier transforms of the kernels satisfy $\widehat{\beta^{(i)}} \in C^{2k}$ for $i = 1, 2$ and

$$\frac{d^j}{d\xi^j} \left(\widehat{\beta^{(2)}}(\xi) - \widehat{\beta^{(1)}}(\xi) \right) \Big|_{\xi=0} = 0 \quad \text{for} \quad 0 \le j < 2k - 1. \tag{23}$$

Then $\widehat{\beta^{(2)}}(\xi) - \widehat{\beta^{(1)}}(\xi) = \mathcal{O}\left(\xi^{2k} \right)$ and $k^{(2)}(\xi) - k^{(1)}(\xi) = \mathcal{O}\left(\xi^{2k} \right)$ near the origin. Thus, we have

$$k^{(2)}(\xi) - k^{(1)}(\xi) = \xi^{2k} m(\xi)$$

for some continuous function m. Moreover, as both $k^{(i)}(\xi)$ $(i = 1, 2)$ are bounded, then so is $m(\xi)$. Then the corresponding operators $K_\delta^{(i)}$ will satisfy

$$K_\delta^{(2)} = K_\delta^{(1)} + (-1)^k \delta^{2k} D_x^{2k} M_\delta,$$

where M_δ is the operator with symbol $m(\delta\xi)$. Since m is bounded, we have the estimate

$$\left\| \left(K_\delta^{(2)} - K_\delta^{(1)} \right) u \right\|_{H^s} = \delta^{2k} \left\| D_x^{2k} M_\delta u \right\|_{H^s} \le C \delta^{2k} \| u \|_{H^{s+2k}}, \tag{24}$$

with the constant C independent of δ.

Step 4 We now estimate the terms g, w, f_1, f_2 appearing in (17)-(18).

1. Noting that

$$\begin{aligned} g &= \left(\left(K_\delta^{(1)} \right)^{-1} - \left(K_\delta^{(2)} \right)^{-1} \right) \psi \\ &= \left(K_\delta^{(1)} \right)^{-1} \left(K_\delta^{(2)} \right)^{-1} \left(K_\delta^{(2)} - K_\delta^{(1)} \right) \psi \end{aligned}$$

and that $\left(K_\delta^{(i)} \right)^{-1}$ $(i = 1, 2)$ are uniformly bounded, we have

$$\begin{aligned} \| g \|_{H^{s+P}} &\le C \left\| \left(K_\delta^{(2)} - K_\delta^{(1)} \right) \psi \right\|_{H^{s+P}} \\ &\le C \delta^{2k} \| \psi \|_{H^{s+P+2k}} \le C \delta^{2k}, \end{aligned}$$

where we have made use of (24).

7

2. From the definition $w = u_1^p + u_1^{p-1} u_2 + \cdots + u_2^p$, we get

$$\|w(t)\|_{H^{s+D}} \le C \sum_{i=0}^{p} \|u_1(t)\|_{H^{s+D}}^{p-i} \|u_2(t)\|_{H^{s+D}}^{i} \le C, \tag{25}$$

$$\|w_t(t)\|_{H^{s+D-1}} \le C \Big(\|(u_1)_t(t)\|_{H^{s+D-1}} + \|(u_2)_t(t)\|_{H^{s+D-1}} \Big) \le C. \tag{26}$$

3. We have

$$
\begin{aligned}
\|f_1(t)\|_{H^{s+D}} &= \left\| \left(K_\delta^{(1)} - K_\delta^{(2)} \right)(v_1)_x(t) \right\|_{H^{s+D}} \\
&\le C\delta^{2k} \|(v_1)_x(t)\|_{H^{s+D+2k}} \\
&\le C\delta^{2k} \|v_1(t)\|_{H^{s+D+2k+1}} \le C\delta^{2k}. \tag{27}
\end{aligned}
$$

4. Similarly,

$$
\begin{aligned}
\|f_2(t)\|_{H^{s+D}} &= \left\| \left(K_\delta^{(1)} - K_\delta^{(2)} \right) \left(u_1 + \epsilon^p u_1^{p+1} \right)_x(t) \right\|_{H^{s+D}} \\
&\le C\delta^{2k} \left\| \left(u_1 + \epsilon^p u_1^{p+1} \right)_x(t) \right\|_{H^{s+D+2k}} \\
&\le C\delta^{2k} \left\| \left(u_1 + \epsilon^p u_1^{p+1} \right)(t) \right\|_{H^{s+D+2k+1}} \le C\delta^{2k}. \tag{28}
\end{aligned}
$$

Step 5 Next we define the energy

$$E^2(t) = \frac{1}{2} \left(\|r(t)\|_{H^{s+D}}^2 + \|\rho(t)\|_{H^{s+D}}^2 + \epsilon^p \langle r(t), w(t) r(t) \rangle_{s+D} \right), \tag{29}$$

with the H^{s+D} inner product $\langle f, g \rangle_{s+D} = \langle \Lambda^{s+D} f, \Lambda^{s+D} g \rangle$. Since

$$\epsilon^p \left| \langle r(t), w(t) r(t) \rangle_{s+D} \right| \le \epsilon^p \|w(t)\|_{H^{s+D}} \|r(t)\|_{H^{s+D}}^2 \le C\epsilon^p \|r(t)\|_{H^{s+D}}^2,$$

there is some $\epsilon_1 \le \epsilon_0$ so that for all $\epsilon < \epsilon_1$,

$$\|r(t)\|_{H^{s+D}}^2 + \|\rho(t)\|_{H^{s+D}}^2 \le CE^2(t) \tag{30}$$

(see Lemma 3.1 of [2] for details). Differentiation both sides of (29) with respect to t gives

$$
\begin{aligned}
\frac{d}{dt} E^2(t) = {}& \langle r, r_t \rangle_{s+D} + \langle \rho, \rho_t \rangle_{s+D} \\
&+ \frac{\epsilon^p}{2} \Big(\langle r_t, wr \rangle_{s+D} + \langle r, w_t r \rangle_{s+D} + \langle r, wr_t \rangle_{s+D} \Big). \tag{31}
\end{aligned}
$$

By making use of the system (17)-(18) in this equation we get

$$
\begin{aligned}
\frac{d}{dt} E^2(t) = {}& \langle r, f_1 \rangle_{s+D} + \langle \rho, f_2 \rangle_{s+D} + \langle \rho, \epsilon^p K_\delta^{(2)}(wr)_x \rangle_{s+D} \\
&+ \frac{\epsilon^p}{2} \Big(\langle r_t, wr \rangle_{s+D} + \langle r, w_t r \rangle_{s+D} + \langle r, wr_t \rangle_{s+D} \Big), \tag{32}
\end{aligned}
$$

where we have made use of the identity $\langle r, K_\delta^{(2)} \rho_x \rangle_{s+D} = -\langle \rho, K_\delta^{(2)} r_x \rangle_{s+D}$. Using (17) the third term on the right-hand side of this equation can be written as

$$
\begin{aligned}
\langle \rho, \epsilon^p K_\delta^{(2)} (wr)_x \rangle_{s+D} &= -\epsilon^p \langle K_\delta^{(2)} \rho_x, wr \rangle_{s+D} \\
&= -\epsilon^p \langle r_t, wr \rangle_{s+D} + \epsilon^p \langle f_1, wr \rangle_{s+D} \quad (33)
\end{aligned}
$$

Substitution of this result into (32) yields

$$
\begin{aligned}
\frac{d}{dt} E^2(t) = &\langle r, f_1 \rangle_{s+D} + \langle \rho, f_2 \rangle_{s+D} + \epsilon^p \langle wr, f_1 \rangle_{s+D} + \frac{\epsilon^p}{2} \langle r, w_t r \rangle_{s+D} \\
&+ \frac{\epsilon^p}{2} \Big(\langle r, wr_t \rangle_{s+D} - \langle r_t, wr \rangle_{s+D} \Big). \quad (34)
\end{aligned}
$$

Noting that $\Lambda^{s+D}(wf) = [\Lambda^{s+D}, w]f + w\Lambda^{s+D} f$ we can rewrite the last term of (34) as

$$
\begin{aligned}
\langle r, wr_t \rangle_{s+D} - \langle r_t, wr \rangle_{s+D} &= \langle \Lambda^{s+D} r, [\Lambda^{s+D}, w] r_t \rangle \\
&\quad - \langle \Lambda^{s+D-1} r_t, \Lambda[\Lambda^{s+D}, w] r \rangle. \quad (35)
\end{aligned}
$$

To estimate this term we start with

$$
\begin{aligned}
\Big| \langle r, wr_t \rangle_{s+D} - \langle r_t, wr \rangle_{s+D} \Big| \leq\ &C \Big(\big\| \Lambda^{s+D} r \big\|_{L^2} \big\| [\Lambda^{s+D}, w] r_t \big\|_{L^2} \\
&+ \big\| \Lambda^{s+D-1} r_t \big\|_{L^2} \big\| \Lambda[\Lambda^{s+D}, w] r \big\|_{L^2} \Big).
\end{aligned}
$$
$$(36)$$

On the other hand, using the commutator estimates given in Lemma 3.1 we get

$$
\begin{aligned}
\big\| [\Lambda^{s+D}, w] r_t \big\|_{L^2} &\leq C \|w\|_{H^{s+D+1}} \|r_t\|_{H^{s+D-1}}, \\
&\leq C \|w\|_{H^{s+D+1}} \Big(\|\rho\|_{H^{s+D}} + \|f_1\|_{H^{s+D-1}} \Big), \\
\big\| \Lambda[\Lambda^{s+D}, w] r \big\|_{L^2} &= \big\| [\Lambda^{s+D}, w] r \big\|_{H^1} \leq C \|w\|_{H^{s+D+1}} \|r\|_{H^{s+D}},
\end{aligned}
$$

and

$$
\begin{aligned}
\big\| \Lambda^{s+D-1} r_t \big\|_{L^2} &\leq C \Big(\big\| \Lambda^{s+D-1} K_\delta^{(2)} \rho_x \big\|_{L^2} + \big\| \Lambda^{s+D-1} f_1 \big\|_{L^2} \Big) \\
&\leq C \Big(\|\rho\|_{H^{s+D}} + \|f_1\|_{H^{s+D-1}} \Big).
\end{aligned}
$$

By making use of these results in (36) we obtain

$$
\begin{aligned}
\Big| \langle r, wr_t \rangle_{s+D} &- \langle r_t, wr \rangle_{s+D} \Big| \\
&\leq C \|r\|_{H^{s+D}} \|w\|_{H^{s+D+1}} \Big(\|\rho\|_{H^{s+D}} + \|f_1\|_{H^{s+D-1}} \Big). \quad (37)
\end{aligned}
$$

Combining (37) with (34) yields

$$\frac{d}{dt}E^2(t) \leq C\Big(\|r\|_{H^{s+D}}\|f_1\|_{H^{s+D}} + \|\rho\|_{H^{s+D}}\|f_2\|_{H^{s+D}}\Big)$$
$$+ C\epsilon^p\|r\|_{H^{s+D}}\Big(\|w\|_{H^{s+D}}\|f_1\|_{H^{s+D}} + \|w\|_{H^{s+D+1}}\|f_1\|_{H^{s+D-1}}\Big)$$
$$+ C\epsilon^p\|r\|_{H^{s+D}}\Big(\|r\|_{H^{s+D}}\|w_t\|_{H^{s+D}} + \|\rho\|_{H^{s+D}}\|w\|_{H^{s+D+1}}\Big).$$

Using the estimates on $\|w\|_{H^{s+D}}$, $\|w_t\|_{H^{s+D}}$, $\|f_i\|_{H^{s+D}}$ $(i=1,2)$ given by (25), (26), (27), (28), respectively, we get

$$\frac{d}{dt}E^2(t) \leq C\delta^{2k}\Big(\|r\|_{H^{s+D}} + \|\rho\|_{H^{s+D}}\Big) + C\epsilon^p\delta^{2k}\|r\|_{H^{s+D}}$$
$$+ C\epsilon^p\|r\|_{H^{s+D}}\Big(\|r\|_{H^{s+D}} + \|\rho\|_{H^{s+D}}\Big).$$

and

$$\frac{d}{dt}E^2(t) \leq C\Big(\delta^{2k}E(t) + \epsilon^p E^2(t)\Big).$$

Gronwall's lemma then implies

$$E(t) \leq e^{C\epsilon^p t}E(0) + \frac{e^{C\epsilon^p t} - 1}{\epsilon^p}\delta^{2k} \leq e^{CT}\big(E(0) + C\delta^{2k}t\big) \qquad (38)$$

for all $t \leq \frac{T}{\epsilon^p}$. Equation (29) shows that the initial energy is

$$E(0) = \frac{1}{\sqrt{2}}\Big(\|r(0)\|^2_{H^{s+D}} + \|\rho(0)\|^2_{H^{s+D}} + \epsilon^p\langle r(0), w(0)r(0)\rangle_{s+D}\Big)^{1/2}$$

from which we get

$$E(0) = \frac{1}{\sqrt{2}}\|g\|_{H^{s+D}} \leq C\|g\|_{H^{s+P}} \leq C\delta^{2k}.$$

By making use of this result in (38) we obtain (16) and this completes the proof of the theorem.

$\square$

Remark 2. In this remark, we show that solutions of (1) for suitable kernels are well approximated by solutions of a high order improved Boussinesq-type equation. Suppose that the kernel β satisfies $x^{2k}\beta \in L^1(\mathbb{R})$ and the ellipticity and boundedness condition (7). We now assume that the $2k$-order Taylor polynomial of $\big(\widehat{\beta}(\xi)\big)^{-1}$ about zero is as follows

$$\big(\widehat{\beta_0}(\xi)\big)^{-1} = 1 + \gamma_1\xi^2 + \gamma_2\xi^4 + \cdots + \gamma_{k-1}\xi^{2(k-1)},$$

10

with suitable constants γ_j ($j = 1, \cdots, k-1$). Clearly $\widehat{\beta_0}$ satisfies the ellipticity and boundedness condition (7). The wave equation corresponding to the kernel β_0 is the high order improved Boussinesq-type equation

$$Lu_{tt} - u_{xx} = \epsilon^p (u^{p+1})_{xx} \tag{39}$$

where

$$L = I - \gamma_1 \delta^2 D_x^2 + \gamma_2 \delta^4 D_x^4 + \cdots + (-1)^{k-1} \gamma_{k-1} \delta^{2k-2} D_x^{2k-2}.$$

We now compare solutions of (1) to solutions of (39). In that respect we note that the Taylor expansions for $\left(\widehat{\beta}(\xi)\right)^{-1}$ and $\left(\widehat{\beta_0}(\xi)\right)^{-1}$ around the origin agree on all terms of degree less than $2k$. Applying Theorem 3.2 for the pair (β, β_0) we conclude that solutions of (1) are well approximated by solutions of (39) with an approximation error of order δ^{2k}. More explicitly, for sufficiently smooth initial conditions the solutions corresponding to the pair (β, β_0) approximate each other at the order δ^{2k} over times $t \le \frac{T}{\epsilon^p}$.

In particular, for the improved Boussinesq equation (3) we have $k = 2$, $\gamma_1 = 1$ and $\left(\widehat{\beta_0}(\xi)\right)^{-1} = 1 + \xi^2$. When we compare solutions of (1) with those of (3), Theorem 3.2 will yield an approximation error of order δ^4.

Remark 3. Here we explain why the equality of the lowest-order moments of the kernels is more important than the similarity of the shapes of the kernels when comparing two nonlocal wave equations in the long-wave limit. To this end, we consider three types of general perturbations of an arbitrary kernel function β_0, and apply Theorem 3.2 to investigate the closeness of the corresponding solutions of (1). With some appropriate kernel function φ, we define the three perturbations in the form

$$\begin{aligned} \beta_1 &= (1-\nu)\beta_0 + \nu\varphi, \\ \beta_2 &= \varphi_\nu * \beta_0, \\ \beta_3 &= \beta_0 + aD_x^{2k}\varphi, \end{aligned}$$

where $\varphi_\nu(x) = \frac{1}{\nu}\varphi(\frac{x}{\nu})$, $\nu \ll 1$ and a are nonzero constants. We denote the corresponding solutions by $u_i^{\epsilon,\delta}$ (i=0,1,2,3). Clearly both β_1 and β_2 are classical small perturbations of β_0 whereas $\beta_0 - \beta_3 = \mathcal{O}(a)$ can be quite large. On the other hand, if $\int x^2 \varphi(x) dx \ne \int x^2 \beta_0(x) dx$, then β_0 and β_1 will have unequal second moments. Then for the corresponding solutions of (1), Theorem 3.2 will only guarantee the estimate $u_0^{\epsilon,\delta} - u_1^{\epsilon,\delta} = \mathcal{O}(\delta^2)$ over long times. Similarly, when $\int x^2 \varphi(x) dx \ne 0$, the kernels β_0 and β_2 will have unequal second moments; hence $u_0^{\epsilon,\delta} - u_2^{\epsilon,\delta} = \mathcal{O}(\delta^2)$. When $\varphi \in W^{2k,1}$, β_0 and β_3 will have equal moments up to order $2k - 1$. For $k \ge 2$, Theorem 3.2 this time implies the much stronger estimate $u_0^{\epsilon,\delta} - u_3^{\epsilon,\delta} = \mathcal{O}(\delta^{2k})$. These three examples show that the behavior of solutions of (1) over large times is determined by the dispersive character of the kernel in the long-wave limit rather than its shape.

11

Remark 4. It is worth to note that, under the change of variable $\widetilde{u} = \epsilon u$, the Cauchy problem (1)-(2) becomes

$$\widetilde{u}_{tt} = \beta_\delta * (\widetilde{u} + \widetilde{u}^{p+1})_{xx}, \quad x \in \mathbb{R}, \quad t > 0,$$
$$\widetilde{u}(x,0) = \epsilon u_0(x), \quad \widetilde{u}_t(x,0) = \epsilon u_1(x), \quad x \in \mathbb{R},$$

which shows that our results can also be interpreted as comparison results for the long-wave approximations of solutions with small initial data.

References

[1] Duruk N, Erbay HA, Erkip A. Global existence and blow-up for a class of nonlocal nonlinear Cauchy problems arising in elasticity. Nonlinearity. 2010;23:107–118.

[2] Erbay HA, Erbay S, Erkip A. Long-time existence of solutions to nonlocal nonlinear bidirectional wave equations. Discrete Contin Dyn Syst. 2018;(To appear).

[3] Bona JL, Colin T, Lannes D. Long wave approximations for water waves. Arch Ration Mech Anal. 2005;178:373–410.

[4] Constantin A, Lannes D. The hydrodynamical relevance of the Camassa-Holm and Degasperis-Procesi equations. Arch Ration Mech Anal. 2009; 192:165–186.

[5] Lannes D, Linares F, Saut JC. The Cauchy problem for the Euler-Poisson system and derivation of the Zakharov-Kuznetsov equation. Prog Nonlinear Differ Equ Appl. 2013;84:181–213.

[6] Lannes D. The water waves problem: Mathematical analysis and asymptotics (ams mathematical surveys and monographs, vol. 188). Providence, RI: American Mathematical Society; 2013.

[7] Duchene V, Israwi S, Talhouk R. A new fully justified asymptotic model for the propagation of internal waves in the Camassa-Holm regime. SIAM J Math Anal. 2015;47:240–290.

[8] Korteweg DJ, de Vries G. On the change of form of long waves advancing in a rectangular channel, and on a new type of long stationary waves. Philos Mag. 1895;39:422–443.

[9] Benjamin TB, Bona JL, Mahony JJ. Model equations for long waves in nonlinear dispersive systems. Philos Trans R Soc Lond Ser A: Math Phys Sci. 1972;272:47–78.

[10] Camassa R, Holm DD. An integrable shallow water equation with peaked solitons. Phys Rev Lett. 1993;71:1661–1664.

[11] Fornberg B, Whitham GB. A numerical and theoretical study of certain nonlinear wave phenomena. Philos Trans R Soc Lond Ser A. 1978;289:373–404.

[12] Craig W. An existence theory for water waves and the Boussinesq and Korteweg-de Vries scaling limits. Comm Partial Differential Equations. 1985;10:787–1003.

[13] Schneider G, Wayne CE. The long-wave limit for the water wave problem. I. The case of zero surface tension. Comm Pure Appl Math. 2000;53:1475–1535.

[14] Alazman AA, Albert JP, Bona JL, et al. Comparisons between the BBM equation and a Boussinesq system. Adv Differential Equations. 2006; 11:121–166.

[15] Gallay T, Schneider G. KP description of unidirectional long waves. The model case. Proc Roy Soc Edinburgh Sect A. 2001;131:885–898.

[16] Erbay HA, Erbay S, Erkip A. The Camassa-Holm equation as the long-wave limit of the improved Boussinesq equation and of a class of nonlocal wave equations. Discrete Contin Dyn Syst. 2016;36:6101–6116.

[17] Alvarez-Samaniego B, Lannes D. A Nash-Moser theorem for singular evolution equations. Application to the Serre and Green-Naghdi equations. Indiana Univ Math Journal. 2008;57:97–131.