

# Long-time existence of solutions to nonlocal nonlinear bidirectional wave equations

H. A. Erbay<sup>1</sup>, S. Erbay<sup>1</sup>, A. Erkip<sup>2</sup>

<sup>1</sup>Department of Natural and Mathematical Sciences, Faculty of Engineering,  
Ozyegin University, Cekmekoy 34794, Istanbul, Turkey

<sup>2</sup>Faculty of Engineering and Natural Sciences, Sabanci University, Tuzla  
34956, Istanbul, Turkey

## Abstract

We consider the Cauchy problem defined for a general class of nonlocal wave equations modeling bidirectional wave propagation in a nonlocally and nonlinearly elastic medium whose constitutive equation is given by a convolution integral. We prove a long-time existence result for the nonlocal wave equations with a power-type nonlinearity and a small parameter. As the energy estimates involve a loss of derivatives, we follow the Nash-Moser approach proposed by Alvarez-Samaniego and Lannes. As an application to the long-time existence theorem, we consider the limiting case in which the kernel function is the Dirac measure and the nonlocal equation reduces to the governing equation of one-dimensional classical elasticity theory. The present study also extends our earlier result concerning local well-posedness for smooth kernels to nonsmooth kernels.

2010 AMS Subject Classification: 35A01, 35L15, 35L70, 35Q74, 74B20

Keywords: Long-time existence, Nonlocal wave equation, Nash-Moser iteration, Improved Boussinesq equation.

## 1 Introduction

In the present paper we prove the long-time existence and uniform estimates of solutions to the Cauchy problem

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

$$u(x, 0) = u_0(x), \quad u_t(x, 0) = u_1(x) \quad (2)$$

---

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

for sufficiently smooth initial data. Here  $u(x, t)$  is a real-valued function,  $\epsilon$  is a small positive parameter measuring the smallness of the initial data, the symbol  $*$  denotes convolution in the  $x$ -variable and  $p$  is a positive integer. We assume that the convolution with the kernel function  $\beta$  is a positive bounded operator on the Sobolev space  $H^s(\mathbb{R})$ . This can be realized by assuming that  $\beta$  is integrable or more generally is a finite measure on  $\mathbb{R}$  with positive Fourier transform.

The nonlocal wave equation (1) describes the one-dimensional motion of a nonlocally and nonlinearly elastic medium and  $u$  represents the elastic strain (we refer the reader to [3] both for a detailed description of the nonlocally and nonlinearly elastic medium and for some examples of the kernels widely used in continuum mechanics). Moreover, (1) involves many well-known nonlinear wave equations for particular choices of the kernel function. One well-known example is the improved Boussinesq equation

$$u_{tt} - u_{xx} - u_{xxtt} - \epsilon^p (u^{p+1})_{xx} = 0 \quad (3)$$

corresponding to the exponential kernel  $\beta(x) = \frac{1}{2}e^{-|x|}$ . On the other hand, if  $\beta$  is taken as the Dirac measure, (1) reduces to the nondispersive nonlinear wave equation

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

of classical elasticity. The local well-posedness of (1)-(2) was proved in [3] under a smoothness assumption on  $\beta$ . This assumption is equivalent to saying that the operator  $\beta'' * (\cdot)$  is bounded on  $H^s(\mathbb{R})$ . In that case, (1) becomes a Banach space-valued ODE and the local existence result holds without any smallness assumption on the initial data.

In the present study, we consider the long-time existence of solutions to (1)-(2) and provide an existence result on time intervals of order  $1/\epsilon^p$ . Additionally, we relax the restriction imposed in [3] on the smoothness of  $\beta$ ; in this case the smallness of  $\epsilon$  guarantees that (1) stays in the hyperbolic regime. At this point, it is worth pointing out that the smallness of the parameter  $\epsilon$  plays an essential role in obtaining the long-time existence result. The key point in our approach is to prove that the bounds can be made uniform in  $\epsilon$ .

To prove our long-time existence result we start by converting (1) into a perturbation of the symmetric hyperbolic linear system and obtain the energy estimates for the corresponding linearized equation. Nevertheless, the energy estimates involve a loss of derivatives. Due to the loss of derivatives one cannot directly pass from the linearized equation to the nonlinear equation via the Picard iteration, and we need a Nash-Moser-type approach. In that respect, the Nash-Moser theorem proved in [1] for a wide class of singular evolution equations is the main technical tool used in the present work.

In the literature, there are a number of studies concerning long-time existence of solutions to PDEs. The long-time existence results of the studies focused on water waves [8, 9, 10] have been also used for the rigorous justification of approximate asymptotical models (such as the Green-Naghdi equations, the shallow water equations and the Boussinesq system) starting from

the Euler equations describing the motion of an inviscid, incompressible fluid. In addition to the studies about asymptotic models of water waves, there are also studies presenting the rigorous derivation of various asymptotic models for nonlinear elastic waves in the long-wave-small-amplitude regime (for instance we refer the reader to [5] where the Camassa-Holm equation and (1) are compared). The present research is motivated by the long-time existence results that were reported for water waves and aims to extend those results to elastic waves propagating in nonlocal elastic solids. In a previous work [5], for smooth kernels and quadratic nonlinearity ( $p = 1$ ) we rigorously established that, in the long-wave-small-amplitude regime, unidirectional solutions of (1) tend to associated solutions of Korteweg-de Vries-type, or Benjamin-Bona-Mahony-type or Camassa-Holm-type equations depending on the balance between dispersion and nonlinearity. Providing a precise control of the approximation error through the justification process, we showed that those unidirectional asymptotic models are good approximations for (1) on time intervals of order  $1/\epsilon$ . There we used the fact that solutions of the unidirectional asymptotic models live over a long-time scale [2]. Hence the corresponding solution of the nonlocal wave equation will live on a sufficiently large time interval. As future work, we plan to investigate similar comparison results between two nonlocal equations. The long-time existence result and uniform bound obtained in this study will be used in a future work to explore comparison of nonlocal equations.

The structure of the paper is as follows. Section 2 is devoted to preliminaries, where (1) is converted into a system of first-order equations, and some function spaces are introduced. In Section 3, we consider a related linear system and derive the estimates to be used in the next section. In Section 4 we consider the Nash-Moser approach of [1] and prove that the required assumptions in the hypothesis of the existence theorem in [1] are satisfied in our case. Finally, in the last section we present our long-time existence result and discuss some particular cases.

Throughout this paper, we use the standard notation for function spaces. The Fourier transform  $\widehat{u}$  of  $u$  is defined by  $\widehat{u}(\xi) = \int_{\mathbb{R}} u(x)e^{-i\xi x} dx$ . We also use  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  to denote the Fourier transform and the inverse Fourier transform. The  $L^p(\mathbb{R})$  norm is denoted by  $\|u\|_{L^p}$  and the symbol  $\langle u, v \rangle_{L^2}$  denotes the inner product of  $u$  and  $v$  in  $L^2$ .  $H^s = H^s(\mathbb{R})$  denotes the  $L^2$ -based Sobolev space of order  $s$  on  $\mathbb{R}$ , with the norm  $\|u\|_{H^s} = \left( \int_{\mathbb{R}} (1 + \xi^2)^s |\widehat{u}(\xi)|^2 d\xi \right)^{1/2}$ .  $C$  is a generic positive constant. Partial differentiations are denoted by  $D_x$  etc.

## 2 Preliminaries

In this section, we recast (1) as an appropriate first-order system, recall the local existence theorem from [3] and introduce function spaces that will be used in the paper. For the rest of this work, we assume that the kernel is an integrable function (or more generally a finite measure) satisfying the nonnegativity and boundedness condition

$$0 \leq \widehat{\beta}(\xi) \leq C. \tag{5}$$

We first convert the Cauchy problem (1)-(2) to

$$u_t = K v_x, \quad u(x, 0) = u_0(x), \quad (6)$$

$$v_t = K u_x + \epsilon^p K (u^{p+1})_x, \quad v(x, 0) = v_0(x) \quad (7)$$

by writing the nonlocal equation (1) as a first-order system and introducing the pseudo-differential operator

$$Kw(x) = \mathcal{F}^{-1} \left( \sqrt{\widehat{\beta}(\xi)} \widehat{w}(\xi) \right), \quad (8)$$

for which  $K^2 w = \beta * w$ . We note that, by (5),  $K$  is a bounded operator on  $H^s$ . Clearly, for the choice  $u_1 = K(v_0)_x$  the initial-value problem (1)-(2) reduces to the first-order system (6)-(7).

The local well-posedness of the Cauchy problem (1)-(2), equivalently the local-well posedness of (6)-(7), was proved in [3] under the regularity assumption

$$0 \leq \widehat{\beta}(\xi) \leq C(1 + \xi^2)^{-r/2}, \quad r \geq 2. \quad (9)$$

In such a case,  $K$  is an operator of order  $-r/2$  and hence maps  $H^s$  into  $H^{s+\frac{r}{2}} \subset H^s$ . For  $r \geq 2$ , due to the regularizing effect, both (1)-(2) and (6)-(7) are the initial-value problems defined for  $H^s$ -valued ODE's. Consequently, the Cauchy problems are locally well-posed with solutions in  $C^1([0, T], H^s)$  for some  $T > 0$ . Obviously, the parameter  $r$  in (9) is a measure of the smoothness of  $\beta$  and hence the regularizing effect of the convolution operator. Namely, the regularizing effect increases as the decay rate  $r$  gets larger. In [3] it was shown that, when  $r \geq 2$ , a possible finite-time blow-up of solutions will be controlled by the  $L^\infty$ -norm. On the other hand, when the smoothing effect of  $\beta$  is weaker, solutions may behave in a nonlinear hyperbolic manner and may evolve to breaking.

In the present work we extend the local well-posedness result proved for the case  $r \geq 2$  in [3] to the case  $r \geq 0$  and show long-time existence in both cases. Indeed, the existence of two different intervals for  $r$  can be attributed to the dual nature of (1). Due to the existence of second-order spatial derivative in (1),  $r = 2$  is the threshold value determining whether (1) behaves like a hyperbolic equation or an ODE. Here we merely state that the well-posedness question of (1)-(2) for the case  $0 \leq r < 2$  requires the techniques that are different from those in [3]. To prove the long-time existence of the solution we need suitable energy estimates. In our problem, even for the regularized case, the energy estimates involve a loss of derivatives. To overcome the loss of derivatives we use a Nash-Moser-type approach. This enables us to extend the existence result to the hyperbolic regime. In particular, we follow the approach given in [1] and prove that a unique solution to (1)-(2) exists over long-time scales of order  $1/\epsilon^p$  in appropriate function spaces.

We now define the function spaces that will be used in this work. For a fixed time  $T$  let

$$H_{\epsilon(0)}^s = C\left([0, \frac{T}{\epsilon^p}]; H^s\right), \quad H_{\epsilon(1)}^s = C\left([0, \frac{T}{\epsilon^p}]; H^s\right) \cap C^1\left([0, \frac{T}{\epsilon^p}]; H^{s-1}\right),$$

with norms

$$\|u\|_{H_{\epsilon(0)}^s} = \sup_{t \in [0, \frac{T}{\epsilon^p}]} \|u(t)\|_{H^s}, \quad \|u\|_{H_{\epsilon(1)}^s} = \sup_{t \in [0, \frac{T}{\epsilon^p}]} (\|u(t)\|_{H^s} + \|u_t(t)\|_{H^{s-1}}).$$

We also introduce the vector-valued counterparts

$$\begin{aligned} X_{\epsilon(0)}^s &= C([0, \frac{T}{\epsilon^p}]; X^s), & X_{\epsilon(1)}^s &= C([0, \frac{T}{\epsilon^p}]; X^s) \cap C^1([0, \frac{T}{\epsilon^p}]; X^{s-1}), \\ \|\mathbf{u}\|_{X_{\epsilon(0)}^s} &= \sup_{t \in [0, \frac{T}{\epsilon^p}]} \|\mathbf{u}(t)\|_{X^s}, & \|\mathbf{u}\|_{X_{\epsilon(1)}^s} &= \sup_{t \in [0, \frac{T}{\epsilon^p}]} (\|\mathbf{u}(t)\|_{X^s} + \|\mathbf{u}_t(t)\|_{X^{s-1}}), \end{aligned}$$

where

$$X^s = H^s \times H^s, \quad \|\mathbf{u}\|_{X^s} = \|(u, v)\|_{X^s} = \|u\|_{H^s} + \|v\|_{H^s}.$$

Note that these spaces depend on two parameters  $T$  and  $\epsilon$ , but, to simplify the notation, we have suppressed the index  $T$ .

### 3 Energy estimates for a related linear system

As a starting point, we consider the following initial-value problem

$$u_t = K v_x + \epsilon^p f_1, \quad u(x, 0) = g_1(x), \quad (10)$$

$$v_t = K u_x + \epsilon^p K (w u)_x + \epsilon^p f_2, \quad v(x, 0) = g_2(x), \quad (11)$$

defined for a nonhomogeneous linear system of differential equations associated with (6)-(7), where  $w$ ,  $f_i$  and  $g_i$  ( $i = 1, 2$ ) are fixed, given and sufficiently smooth functions. In this section we obtain a priori estimates for solutions of (10)-(11) over the long-time intervals  $[0, \frac{T}{\epsilon^p}]$ .

We begin by defining the energy functional for the linear system (10)-(11) as follows

$$E_s^2(t) = \frac{1}{2} (\|u(t)\|_{H^s}^2 + \|v(t)\|_{H^s}^2 + \epsilon^p \langle u(t), w(t)u(t) \rangle_{H^s}) \quad (12)$$

where we have made use of the representation  $\langle u, v \rangle_{H^s} = \langle \Lambda^s u, \Lambda^s v \rangle_{L^2}$  with  $\Lambda^{2s} = (1 - D_x^2)^s$ . The following lemma states that, under some smallness assumption on  $\epsilon$ , the energy functional  $E_s$  is uniformly equivalent to the  $X^s$  norm of the vector  $\mathbf{u} = (u, v)$ .

**Lemma 3.1** *Let  $w \in H_{\epsilon(0)}^s$ . Then there is some  $\epsilon_0 > 0$  so that, for all  $0 < \epsilon \leq \epsilon_0$ ,  $E_s(t)$  is uniformly equivalent to  $\|\mathbf{u}(t)\|_{X^s}$ .*

**Proof.** Recall that, for  $s > 1/2$ ,  $H^s$  is an algebra:

$$|\langle u(t), w(t)u(t) \rangle_{H^s}| \leq C \|u(t)\|_{H^s}^2 \|w(t)\|_{H^s}.$$

For

$$\epsilon^p \leq \frac{1}{2C \|w\|_{H_{\epsilon(0)}^s}} = \epsilon_0^p,$$

we have

$$-\frac{1}{2}\|u(t)\|_{H^s}^2 \leq \epsilon^p \langle u(t), w(t)u(t) \rangle_{H^s} \leq \frac{1}{2}\|u(t)\|_{H^s}^2.$$

Using this and (12) we get

$$\frac{1}{2}(\|u(t)\|_{H^s}^2 + \|v(t)\|_{H^s}^2 - \frac{1}{2}\|u(t)\|_{H^s}^2) \leq E_s^2(t) \leq \frac{1}{2}(\|u(t)\|_{H^s}^2 + \|v(t)\|_{H^s}^2 + \frac{1}{2}\|u(t)\|_{H^s}^2)$$

from which it follows that

$$\frac{1}{2\sqrt{2}}(\|u(t)\|_{H^s} + \|v(t)\|_{H^s}) \leq E_s(t) \leq \frac{\sqrt{3}}{2}(\|u(t)\|_{H^s} + \|v(t)\|_{H^s}). \quad (13)$$

■

**Remark 1** *Lemma 3.1 shows that if  $\epsilon < \epsilon_0$ , then  $\epsilon^p \|w(t)\|_{H^s}$  is small. This in turn implies the usual hyperbolicity condition  $1 + \epsilon^p w > 0$  for (10)-(11).*

Next we state a lemma about the product and commutator estimates that will be needed in the sequel. It corresponds to Lemma 4.6 of [1] and provides a classical Moser tame product estimate and a particular case of Kato-Ponce commutator estimate:

**Lemma 3.2** *Let  $s > s_0 > 1/2$ .*

1. *For all  $f, g \in H^s(\mathbb{R})$ , one has*

$$\|fg\|_{H^s} \leq C(\|f\|_{H^{s_0}} \|g\|_{H^s} + \|f\|_{H^s} \|g\|_{H^{s_0}}). \quad (14)$$

2. *Let  $r \in \mathbb{R}$  be such that  $-s_0 < r \leq s_0 + 1$ . For all  $f \in H^{s_0+1}(\mathbb{R}) \cap H^{s+r}(\mathbb{R})$  and  $u \in H^{s+r-1}(\mathbb{R})$ ,*

$$\|[\Lambda^s, f]u\|_{H^r} \leq C(\|f_x\|_{H^{s_0}} \|u\|_{H^{s+r-1}} + \|f_x\|_{H^{s+r-1}} \|u\|_{H^{s_0}}). \quad (15)$$

In the rest of this work we will always assume that  $s_0 > \frac{1}{2}$ .

We are now ready to state and prove an a priori energy estimate for the linear system (10)-(11).

**Proposition 1** *Let  $s \geq s_0 + 1$ ,  $T > 0$ ,  $w \in H_{\epsilon(1)}^{s+1}$ ,  $\mathbf{f} = (f_1, f_2) \in X_{\epsilon(0)}^s$ ,  $\mathbf{g} = (g_1, g_2) \in X^s$ . Suppose  $\mathbf{u} = (u, v) \in X_{\epsilon(0)}^s$  satisfies the initial-value problem (10)-(11). Then, there is some  $\epsilon_0$  such that for all  $0 < \epsilon < \epsilon_0$  and  $t \in [0, \frac{T}{\epsilon^p}]$*

$$\|\mathbf{u}(t)\|_{X^s} \leq C(T, \|w\|_{H_{\epsilon(1)}^{s_0+2}}) \left( \mathfrak{J}^s(t, \mathbf{f}, \mathbf{g}) + \|w\|_{H_{\epsilon(1)}^{s+1}} \mathfrak{J}^{s_0+1}(t, \mathbf{f}, \mathbf{g}) \right), \quad (16)$$

where  $\epsilon_0$  is determined as in Lemma 3.1 and

$$\mathfrak{J}^s(t, \mathbf{f}, \mathbf{g}) = \|\mathbf{g}\|_{X^s} + \int_0^t \sup_{0 \leq t' \leq t''} \|\mathbf{f}(t')\|_{X^s} dt''. \quad (17)$$

**Proof.** Taking the  $L^2$  inner product of (10) with  $\Lambda^{2s}u$ , (11) with  $\Lambda^{2s}v$  and adding them up yields

$$\begin{aligned}
\frac{d}{dt}E_s^2(t) &= \epsilon^p \left( \langle u, f_1 \rangle_{H^s} + \langle v, f_2 \rangle_{H^s} \right) + \epsilon^p \langle v, K(wu)_x \rangle_{H^s} + \frac{\epsilon^p}{2} \frac{d}{dt} \langle u, wu \rangle_{H^s} \\
&= \epsilon^p \left( \langle u, f_1 \rangle_{H^s} + \langle v, f_2 \rangle_{H^s} \right) - \epsilon^p \langle Kv_x, wu \rangle_{H^s} \\
&\quad + \frac{\epsilon^p}{2} \left( \langle u_t, wu \rangle_{H^s} + \langle u, w_t u \rangle_{H^s} + \langle u, wu_t \rangle_{H^s} \right) \\
&= \epsilon^p \left( \langle u, f_1 \rangle_{H^s} + \langle v, f_2 \rangle_{H^s} \right) + \epsilon^{2p} \langle f_1, wu \rangle_{H^s} + \frac{\epsilon^p}{2} \langle u, w_t u \rangle_{H^s} \\
&\quad + \frac{\epsilon^p}{2} \left( \langle u, wu_t \rangle_{H^s} - \langle u_t, wu \rangle_{H^s} \right), \tag{18}
\end{aligned}$$

where we have used (10) and (12). We now estimate the terms in (18). The first and second terms on the right-hand side of (18) are estimated as

$$|\langle u, f_1 \rangle_{H^s} + \langle v, f_2 \rangle_{H^s}| \leq C(\|f_1\|_{H^s} \|u\|_{H^s} + \|f_2\|_{H^s} \|v\|_{H^s}), \tag{19}$$

and

$$|\langle f_1, wu \rangle_{H^s}| \leq C\|f_1\|_{H^s} \|w\|_{H^s} \|u\|_{H^s}, \tag{20}$$

respectively. Using the Cauchy-Schwarz inequality and the estimate (14), we get

$$|\langle u, w_t u \rangle_{H^s}| \leq C(\|u\|_{H^s}^2 \|w_t\|_{H^{s_0}} + \|u\|_{H^s} \|u\|_{H^{s_0}} \|w_t\|_{H^s}) \tag{21}$$

for the third term on the right-hand side of (18). Similarly, the use of the Cauchy-Schwarz inequality and the commutator  $[\Lambda^s, w]f = \Lambda^s(wf) - w\Lambda^s f$  makes possible to write the last term in (18) as

$$|\langle u, wu_t \rangle_{H^s} - \langle u_t, wu \rangle_{H^s}| \leq C(\|\Lambda^s u\|_{L^2} \|[\Lambda^s, w]u_t\|_{L^2} + \|\Lambda^{s-1} u_t\|_{L^2} \|\Lambda[\Lambda^s, w]u\|_{L^2}). \tag{22}$$

Applying the Kato-Ponce commutator estimate (15) to the terms  $\|\Lambda[\Lambda^s, w]u\|_{L^2}$  and  $\|[\Lambda^s, w]u_t\|_{L^2}$  in (22), we get

$$\|\Lambda[\Lambda^s, w]u\|_{L^2} \leq C(\|u\|_{H^s} \|w\|_{H^{s_0+1}} + \|u\|_{H^{s_0}} \|w_x\|_{H^s}), \tag{23}$$

$$\|[\Lambda^s, w]u_t\|_{L^2} \leq C(\|u_t\|_{H^{s-1}} \|w_x\|_{H^{s_0}} + \|u_t\|_{H^{s_0}} \|w_x\|_{H^{s-1}}). \tag{24}$$

Substituting (23) and (24) into (22) and using (10) to eliminate  $u_t$  in the resulting expression we obtain

$$\begin{aligned}
|\langle u, wu_t \rangle_{H^s} - \langle u_t, wu \rangle_{H^s}| &\leq C \left( \|u\|_{H^s} \|v\|_{H^s} \|w\|_{H^{s_0+1}} + \|u\|_{H^s} \|v\|_{H^{s_0+1}} \|w\|_{H^s} \right. \\
&\quad \left. + \|u\|_{H^{s_0}} \|v\|_{H^s} \|w\|_{H^{s+1}} + \epsilon^p (\|f_1\|_{H^{s_0}} \|u\|_{H^s} \|w\|_{H^s} \right. \\
&\quad \left. + \|f_1\|_{H^s} \|u\|_{H^s} \|w\|_{H^{s_0+1}} + \|f_1\|_{H^s} \|u\|_{H^{s_0}} \|w\|_{H^s}) \right). \tag{25}
\end{aligned}$$

Using (13) and the estimates (19), (20), (21), (25) in (18) we obtain

$$\begin{aligned}
\frac{d}{dt}E_s(t) &\leq C\epsilon^p\left(\|w_t\|_{H^{s_0}}+\|w_x\|_{H^{s_0}}\right)E_s(t)+\|f_1\|_{H^s}+\|f_2\|_{H^s} \\
&\quad +\epsilon^p\|f_1\|_{H^s}\left(\|w\|_{H^s}+\|w\|_{H^{s_0+1}}\right) \\
&\quad +\left(\|w_x\|_{H^s}+\|w_t\|_{H^s}\right)\left(\|u\|_{H^{s_0}}+\|v\|_{H^{s_0+1}}\right) \\
&\leq C\epsilon^p\left(\|w\|_{H_{\epsilon(1)}^{s_0+1}}E_s(t)+\|w\|_{H_{\epsilon(1)}^{s+1}}E_{s_0+1}(t)+(1+\epsilon^p\|w\|_{H_{\epsilon(0)}^s})\|\mathbf{f}(t)\|_{X^s}\right) \\
&\leq C\epsilon^p\left(\|w\|_{H_{\epsilon(1)}^{s_0+1}}E_s(t)+\|w\|_{H_{\epsilon(1)}^{s+1}}E_{s_0+1}(t)+\|\mathbf{f}(t)\|_{X^s}\right), \tag{26}
\end{aligned}$$

where we have used  $\epsilon^p\|w\|_{H_{\epsilon(0)}^s}\leq C$ . Applying the Gronwall inequality,  $E_s(t)$  is estimated as follows

$$E_s(t)\leq e^{C\epsilon^pt\|w\|_{H_{\epsilon(1)}^{s_0+1}}}\left(E_s(0)+C\epsilon^p\|w\|_{H_{\epsilon(1)}^{s+1}}\int_0^tE_{s_0+1}(t')dt'+C\epsilon^p\int_0^t\|\mathbf{f}(t')\|_{X^s}dt'\right). \tag{27}$$

The next step is to eliminate the term  $E_{s_0+1}(t)$  in (27). This is accomplished by getting a similar inequality for  $E_{s_0+1}(t)$ . For  $s=s_0+1$ , the differential inequality (26) takes the form

$$\frac{d}{dt}E_{s_0+1}(t)\leq C\epsilon^p\left(\|w\|_{H_{\epsilon(1)}^{s_0+2}}E_{s_0+1}(t)+\|\mathbf{f}(t)\|_{X^{s_0+1}}\right).$$

Again, by the Gronwall inequality, we get

$$E_{s_0+1}(t)\leq e^{C\epsilon^pt\|w\|_{H_{\epsilon(1)}^{s_0+2}}}\left(E_{s_0+1}(0)+C\epsilon^p\int_0^t\|\mathbf{f}(t')\|_{X^{s_0+1}}dt'\right).$$

We need an estimate for time integral of  $E_{s_0+1}(t)$

$$\begin{aligned}
\int_0^tE_{s_0+1}(t')dt' &\leq\int_0^te^{C\epsilon^pt'\|w\|_{H_{\epsilon(1)}^{s_0+2}}}\left(E_{s_0+1}(0)\right. \\
&\quad \left.+C\epsilon^p\int_0^{t'}\sup_{0\leq t''\leq t'''}\|\mathbf{f}(t'')\|_{X^{s_0+1}}dt''\right)dt' \\
&\leq Cte^{C\epsilon^pt\|w\|_{H_{\epsilon(1)}^{s_0+2}}}\left(E_{s_0+1}(0)+\int_0^t\sup_{0\leq t'\leq t''}\|\mathbf{f}(t')\|_{X^{s_0+1}}dt''\right).
\end{aligned}$$

Using this result in (27) we obtain

$$\begin{aligned}
E_s(t) &\leq Ce^{C\epsilon^pt\|w\|_{H_{\epsilon(1)}^{s_0+1}}}\left(E_s(0)+\int_0^t\sup_{0\leq t'\leq t''}\|\mathbf{f}(t')\|_{X^s}dt''\right. \\
&\quad \left.+\epsilon^pt\|w\|_{H_{\epsilon(1)}^{s+1}}e^{C\epsilon^pt\|w\|_{H_{\epsilon(1)}^{s_0+2}}}\left(E_{s_0+1}(0)+\int_0^t\sup_{0\leq t'\leq t''}\|\mathbf{f}(t')\|_{X^{s_0+1}}dt''\right)\right).
\end{aligned}$$

Then, using the definition in (17), we get

$$\|\mathbf{u}(t)\|_{X^s}\leq C(T,\|w\|_{H_{\epsilon(1)}^{s_0+2}})\left(\mathcal{I}^s(t,\mathbf{f},\mathbf{g})+\|w\|_{H_{\epsilon(1)}^{s+1}}\mathcal{I}^{s_0+1}(t,\mathbf{f},\mathbf{g})\right)$$



for  $t \in [0, \frac{T}{\epsilon^p}]$ . This completes the proof. ■

For convenient reference in the remainder of the paper, it is useful to write our energy estimate in terms of a new scaled time variable instead of  $t$ . Let us do so by introducing the scaled time  $\tau = \epsilon^p t$ , so that we change the time interval  $[0, \frac{T}{\epsilon^p}]$  for  $t$  to  $[0, T]$  for  $\tau$ . Then, the linear initial-value problem (10)-(11) becomes

$$u_\tau = \frac{1}{\epsilon^p} K v_x + f_1, \quad u(x, 0) = g_1(x) \quad (28)$$

$$v_\tau = \frac{1}{\epsilon^p} K u_x + K(wu)_x + f_2, \quad v(x, 0) = g_2(x) \quad (29)$$

in the scaled variable  $\tau$ . In such case, all of the previous arguments used to prove the energy estimate of Proposition 1 still hold for the new system with obvious modifications. Indeed, we now define the spaces  $X_{(0)}^s$  and  $X_{(1)}^s$  as  $X_{(0)}^s = C([0, T]; X^s)$  and  $X_{(1)}^s = C([0, T]; X^s) \cap C^1([0, T]; X^{s-1})$ , respectively. Additionally, the  $X_{(1)}^s$ -norm takes the form

$$\|\mathbf{u}\|_{X_{(1)}^s} = \sup_{\tau \in [0, T]} \left( \|\mathbf{u}(\tau)\|_{X^s} + \epsilon^p \|\mathbf{u}_\tau(\tau)\|_{X^{s-1}} \right).$$

**Corollary 1** *Let  $s \geq s_0 + 1$ ,  $T > 0$ ,  $w \in H_{(1)}^{s+1}$ ,  $\mathbf{f} = (f_1, f_2) \in X_{(0)}^s$ ,  $\mathbf{g} = (g_1, g_2) \in X^s$ . Suppose  $\mathbf{u} = (u, v) \in X_{(0)}^s$  satisfies the initial-value problem (28)-(29). Then, there is some  $\epsilon_0$  such that for all  $0 < \epsilon \leq \epsilon_0$  and  $\tau \in [0, T]$*

$$\|\mathbf{u}(\tau)\|_{X^s} \leq C(T, \|w\|_{H_{(1)}^{s_0+2}}) \left( \mathcal{J}^s(\tau, \mathbf{f}, \mathbf{g}) + \|w\|_{H_{(1)}^{s+1}} \mathcal{J}^{s_0+1}(\tau, \mathbf{f}, \mathbf{g}) \right).$$

We note that the constant  $C$  in Proposition 1 and Corollary 1 also depends on the operator norm of  $K$ .

As seen in Corollary 1, there is a loss of derivative in the energy estimate for the linear system, that is,  $\|\mathbf{u}(\tau)\|_{X^s}$  is controlled by  $\|w\|_{H_{(1)}^{s+1}}$ , the norm of the reference state. This loss of derivative propagates along the iteration scheme and may cause problems in a standard Picard iteration scheme for the nonlinear system (6)-(7). In order to handle the loss of derivative, we will use the Nash-Moser-type approach described in [1] for a general system of evolution equations. The following section serves as preparation for the Nash-Moser scheme.

## 4 Preparation for the Nash-Moser scheme

In preparation for the proof of our main result, we outline in this section certain preliminaries essential for understanding how the approach used in [1] is related to the present case.

In [1], the authors have studied the well-posedness of the following general class of initial-value problems (see equation (1.1) of [1])

$$\partial_t \mathbf{u}^\epsilon + \frac{1}{\epsilon} \mathcal{L}^\epsilon(t) \mathbf{u}^\epsilon + \mathcal{N}^\epsilon[t, \mathbf{u}^\epsilon] = \mathbf{h}^\epsilon, \quad \mathbf{u}^\epsilon(0) = \mathbf{u}_0^\epsilon, \quad (30)$$

where  $\epsilon > 0$  is a small parameter and  $\mathcal{L}^\epsilon(t)$  and  $\mathcal{N}^\epsilon[t, \cdot]$  are linear and nonlinear operators, respectively. By making three simplifying assumptions, they have proved their well-posedness theorem for time intervals  $[0, \bar{T}]$  where  $\bar{T} > 0$  is independent of  $\epsilon$ . While two of the assumptions are concerned with  $\mathcal{L}^\epsilon$  and  $\mathcal{N}^\epsilon$ , the third assumption is about the existence of a tame estimate for the solution of the related linearized system. Their main result is the following general theorem:

**Theorem 4.1** (*Theorem 2.1 of [1]*) *Let  $T > 0$ ,  $s_0, m, d_1$  and  $d'_1$  be such that Assumptions 1.2, 1.3 and 1.5 of [1] are satisfied. Let also  $D > \delta$ ,  $P > P_{\min}$ ,  $s \geq s_0 + m$  and  $(\mathbf{h}^\epsilon, \mathbf{u}_0^\epsilon)_{0 < \epsilon < \epsilon_0}$  be bounded in  $F^{s+P}$ . Then there exist  $0 < \underline{T} \leq T$  and a unique family  $(\mathbf{u}^\epsilon)_{0 < \epsilon < \epsilon_0}$  bounded in  $C([0, \underline{T}], X^{s+D})$  and solving the initial value problems (30) <sub>$0 < \epsilon < \epsilon_0$</sub> .*

The spaces  $\{X^s\}_{s \geq 0}$  form a Banach scale and  $F^s$  is defined as  $F^s = C([0, T]; X^s) \times X^{s+m}$ . The constants  $\delta$  and  $P_{\min}$  appearing in the statement of the theorem are related to certain constants resulting from the above-mentioned three assumptions. For a detailed discussion of these assumptions, we refer the reader to [1].

In the rest of this section we will show that each of the three assumptions of Theorem 4.1, namely each of Assumptions 1.2, 1.3 and 1.5 of [1], also holds for our problem (6)-(7). We first rewrite (6)-(7) in the form

$$\mathbf{u}_\tau + \frac{1}{\epsilon^p} \mathcal{L} \mathbf{u} + \mathcal{N}[\mathbf{u}] = \mathbf{0}, \quad \mathbf{u}(x, 0) = \mathbf{u}_0(x), \quad (31)$$

where  $\tau = \epsilon^p t$ ,  $\mathbf{u} = (u, v)^T$ , and the linear operator  $\mathcal{L}$  and the nonlinear map  $\mathcal{N}[\cdot]$  are given respectively by

$$\mathcal{L} = \begin{pmatrix} 0 & -KD_x \\ -KD_x & 0 \end{pmatrix}, \quad \mathcal{N}[\mathbf{u}] = \begin{pmatrix} 0 \\ -KD_x(u^{p+1}) \end{pmatrix}.$$

We note that our case is simpler than the one in [1] since  $\mathcal{L}$  and  $\mathcal{N}$  are both independent of  $\epsilon$  and  $\tau$  so that  $T$  below can be chosen arbitrarily large. Also, the parameter  $\epsilon$  appearing in (30) is replaced by  $\epsilon^p$  in our case.

For (31) we work with the Banach space  $X^s$  and the smoothing operators  $S_\theta u = \mathcal{F}^{-1}(\chi_{[-\theta, \theta]}(\xi) \widehat{u}(\xi))$  where  $\chi$  is the characteristic function. These choices satisfy the requirements of a Banach scale. We now proceed to show that the three assumptions of [1] also hold for (31).

The first assumption is about the evolution operator generated by  $\mathcal{L}$ .

**Assumption 4.2** *Consider the linear problem*

$$\mathbf{u}_\tau + \frac{1}{\epsilon^p} \mathcal{L} \mathbf{u} = \mathbf{0}, \quad \mathbf{u}(x, 0) = \mathbf{g}(x). \quad (32)$$

1. *The linear operator  $\mathcal{L}$  is uniformly bounded in  $C(\mathbb{R}; \mathfrak{L}(X^{s+1}, X^s))$ ;*
2. *The solution operator  $U^\epsilon(\tau)$  for the linear problem (32) is uniformly bounded in  $C(\mathbb{R}; \mathfrak{L}(X^s, X^s))$ .*

The following discussion shows that this assumption is valid.  $K$  is a bounded operator on  $H^s$  due to (5). Thus  $KD_x$  maps  $H^{s+1}$  into  $H^s$  and so  $\mathcal{L} : X^{s+1} \rightarrow X^s$  is a bounded operator independent of  $\tau$ . The evolution operator  $U^\epsilon$  satisfies  $U^\epsilon(\tau)\mathbf{g} = \mathbf{u}(\tau)$ . In Fourier space the solution of (32) is given by

$$\begin{pmatrix} \widehat{u}(\xi, \tau) \\ \widehat{v}(\xi, \tau) \end{pmatrix} = e^{\tau\mathbf{A}} \begin{pmatrix} \widehat{g}_1(\xi) \\ \widehat{g}_2(\xi) \end{pmatrix}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & -i\frac{\xi}{\epsilon^p}\sqrt{\widehat{\beta}(\xi)} \\ -i\frac{\xi}{\epsilon^p}\sqrt{\widehat{\beta}(\xi)} & 0 \end{bmatrix}.$$

The operator  $U^\epsilon(\tau)$  can be easily computed in the Fourier space as  $U^\epsilon(\tau) = \mathcal{F}^{-1}e^{\tau\mathbf{A}}\mathcal{F}$  with

$$e^{\tau\mathbf{A}} = \begin{bmatrix} \cos\left(\frac{\xi}{\epsilon^p}\sqrt{\widehat{\beta}(\xi)}\tau\right) & -i\sin\left(\frac{\xi}{\epsilon^p}\sqrt{\widehat{\beta}(\xi)}\tau\right) \\ -i\sin\left(\frac{\xi}{\epsilon^p}\sqrt{\widehat{\beta}(\xi)}\tau\right) & \cos\left(\frac{\xi}{\epsilon^p}\sqrt{\widehat{\beta}(\xi)}\tau\right) \end{bmatrix}.$$

Obviously  $U^\epsilon(\tau)$  is uniformly bounded in  $C(\mathbb{R}; \mathfrak{L}(X^s, X^s))$ .

The second assumption is about estimates for the nonlinear term.

**Assumption 4.3** *For  $s \geq s_0$  we have the following nonlinear estimates:*

$$1. \quad \|\mathcal{N}[\mathbf{u}]\|_{X^s} \leq C \|\mathbf{u}\|_{X^{s_0}}^p \|\mathbf{u}\|_{X^{s+1}}; \quad (33)$$

$$2. \quad \|\mathcal{N}_{\mathbf{u}}[\mathbf{u}]\boldsymbol{\phi}\|_{X^s} \leq C \left( \|\mathbf{u}\|_{X^{s_0}}^p + \|\mathbf{u}\|_{X^{s_0}}^{p-1} \right) \left( \|\boldsymbol{\phi}\|_{X^{s+1}} + \|\boldsymbol{\phi}\|_{X^{s_0}} \|\mathbf{u}\|_{X^{s+1}} \right); \quad (34)$$

$$3. \quad \|\mathcal{N}_{\mathbf{u}\mathbf{u}}[\mathbf{u}](\boldsymbol{\phi}, \boldsymbol{\psi})\|_{X^s} \leq C \left( \|\mathbf{u}\|_{X^{s_0}}^{p-1} + \|\mathbf{u}\|_{X^{s_0}}^{p-2} \right) \left( \|\boldsymbol{\phi}\|_{X^{s+1}} \|\boldsymbol{\psi}\|_{X^{s_0}} \right. \\ \left. + \|\boldsymbol{\phi}\|_{X^{s_0}} \|\boldsymbol{\psi}\|_{X^{s+1}} + \|\mathbf{u}\|_{X^{s+1}} \|\boldsymbol{\phi}\|_{X^{s_0}} \|\boldsymbol{\psi}\|_{X^{s_0}} \right). \quad (35)$$

We now show that the above assumption is valid. By repeatedly applying (14) to  $\mathcal{N}[\mathbf{u}] = (0, -KD_x(u^{p+1}))$  where  $\mathbf{u} = (u, v)$ , we get

$$\begin{aligned} \|\mathcal{N}[\mathbf{u}]\|_{X^s} &= \|KD_x(u^{p+1})\|_{H^s} \leq C \|u^{p+1}\|_{H^{s+1}} \\ &\leq C \|u\|_{H^{s_0}}^p \|u\|_{H^{s+1}} \leq C \|\mathbf{u}\|_{X^{s_0}}^p \|\mathbf{u}\|_{X^{s+1}}, \end{aligned}$$

that is, (33) holds. To check (34) we note that  $\mathcal{N}_{\mathbf{u}}[\mathbf{u}]\boldsymbol{\phi} = (0, -(p+1)KD_x(u^p\phi_2))$ , where  $\boldsymbol{\phi} = (\phi_1, \phi_2)$ . Then we have

$$\begin{aligned} \|\mathcal{N}_{\mathbf{u}}[\mathbf{u}]\boldsymbol{\phi}\|_{X^s} &= (p+1) \|KD_x(u^p\phi_2)\|_{H^s} \leq C \|u^p\phi_2\|_{H^{s+1}} \\ &\leq C \left( \|u\|_{H^{s_0}}^p \|\phi_2\|_{H^{s+1}} + \|u\|_{H^{s_0}}^{p-1} \|\phi_2\|_{H^{s_0}} \|u\|_{H^{s+1}} \right) \\ &\leq C \left( \|\mathbf{u}\|_{X^{s_0}}^p + \|\mathbf{u}\|_{X^{s_0}}^{p-1} \right) \left( \|\boldsymbol{\phi}\|_{X^{s+1}} + \|\boldsymbol{\phi}\|_{X^{s_0}} \|\mathbf{u}\|_{X^{s+1}} \right), \end{aligned}$$

that is, (34) holds. Similarly, to check (35) we first note that the only nonzero element of  $\mathcal{N}_{\mathbf{u}\mathbf{u}}[\mathbf{u}](\boldsymbol{\phi}, \boldsymbol{\psi})$  is  $-p(p+1)KD_x(u^{p-1}\phi_2\psi_2)$  where  $\boldsymbol{\psi} = (\psi_1, \psi_2)$ . Then we have

$$\begin{aligned} \|\mathcal{N}_{\mathbf{u}\mathbf{u}}[\mathbf{u}](\boldsymbol{\phi}, \boldsymbol{\psi})\|_{X^s} &\leq p(p+1)\|KD_x(u^{p-1}\phi_2\psi_2)\|_{H^s} \leq C\|u^{p-1}\phi_2\psi_2\|_{H^{s+1}} \\ &\leq C\left(\|u\|_{H^{s_0}}^{p-1}\|\phi_2\|_{H^{s_0}}\|\psi_2\|_{H^{s+1}} + \|u\|_{H^{s_0}}^{p-1}\|\phi_2\|_{H^{s+1}}\|\psi_2\|_{H^{s_0}}\right. \\ &\quad \left. + \|u\|_{H^{s_0}}^{p-2}\|\phi_2\|_{H^{s_0}}\|\psi_2\|_{H^{s_0}}\|u\|_{H^{s+1}}\right) \\ &\leq C\left(\|\mathbf{u}\|_{X^{s_0}}^{p-1} + \|\mathbf{u}\|_{X^{s_0}}^{p-2}\right)(\|\boldsymbol{\phi}\|_{X^{s+1}}\|\boldsymbol{\psi}\|_{X^{s_0}} + \|\boldsymbol{\phi}\|_{X^{s_0}}\|\boldsymbol{\psi}\|_{X^{s+1}} \\ &\quad + \|\mathbf{u}\|_{X^{s+1}}\|\boldsymbol{\phi}\|_{X^{s_0}}\|\boldsymbol{\psi}\|_{X^{s_0}}), \end{aligned}$$

that is, (35) holds.

The third assumption is about the well posedness and a priori estimates for the linearized system:

**Assumption 4.4** Consider the linearized problem around  $\bar{\mathbf{u}} = (\bar{u}, \bar{v})$ :

$$\mathbf{u}_\tau^\epsilon + \frac{1}{\epsilon^p}\mathcal{L}\mathbf{u}^\epsilon + \mathcal{N}_{\mathbf{u}}[\bar{\mathbf{u}}]\mathbf{u}^\epsilon = \mathbf{f}^\epsilon, \quad \mathbf{u}^\epsilon(x, 0) = \mathbf{g}^\epsilon(x). \quad (36)$$

For given  $\bar{\mathbf{u}} \in X_{(1)}^{s+1}$ , and bounded families  $\mathbf{f}^\epsilon \in X_{(0)}^s$ ,  $\mathbf{g}^\epsilon \in X^s$  with  $s \geq s_0 + 3$ , the initial-value problem (36) has a unique solution  $\mathbf{u}^\epsilon \in X_{(0)}^s$  and for all  $0 < \epsilon < \epsilon_0$  and  $\tau \in [0, T]$

$$\|\mathbf{u}^\epsilon\|_{X_{(0)}^s} \leq C(T, \|\bar{\mathbf{u}}\|_{X_{(1)}^{s_0+2}}) \left( \mathcal{J}^s(\tau, \mathbf{f}^\epsilon, \mathbf{g}^\epsilon) + \|\bar{\mathbf{u}}\|_{X_{(1)}^{s+1}} \mathcal{J}^{s_0+1}(\tau, \mathbf{f}^\epsilon, \mathbf{g}^\epsilon) \right).$$

In the remaining part of this section we show that the above assumption holds for our system.

We first note that (36) is just (28)-(29) with  $w = (p+1)\bar{u}^p$ . For the existence proof of the linearized system, we follow the standard hyperbolic approach [11] to obtain a solution of (28)-(29) as a limit of solutions of the regularized system

$$\mathbf{u}_\tau^h + \frac{1}{\epsilon^p}J^h\mathcal{L}\mathbf{u}^h + J^h\mathcal{N}_{\mathbf{u}}[\bar{\mathbf{u}}]\mathbf{u}^h = \mathbf{f}, \quad \mathbf{u}^h(x, 0) = \mathbf{g}(x). \quad (37)$$

where we omit the index  $\epsilon$  for convenience. In (37)  $J^h$  is the Friedrichs mollifier given by

$$J^h z(x) = \frac{1}{h} \int \eta\left(\frac{x-y}{h}\right) z(y) dy$$

with some nonnegative  $\eta \in C_0^\infty(\mathbb{R})$  with  $\int \eta dx = 1$ .

Due to regularizing effect of  $J^h$ , (37) is a system of  $X^s$ -valued ODEs and hence has solution  $\mathbf{u}^h$  in  $X_{(0)}^s$ . Moreover, since  $J^h\mathcal{L}$  and  $J^h\mathcal{N}_{\mathbf{u}}$  satisfy the same bounds as  $\mathcal{L}$  and  $\mathcal{N}_{\mathbf{u}}$ , respectively; Corollary 1 will hold uniformly in  $h$ . Using the inequality

$$\|w\|_{H_{(1)}^{s+1}} = (p+1)\|\bar{u}^p\|_{H_{(1)}^{s+1}} \leq C\|\bar{u}\|_{H_{(1)}^{s_0}}^{p-1}\|\bar{u}\|_{H_{(1)}^{s+1}},$$

we then have, for all  $\epsilon < \epsilon_0$ ,

$$\begin{aligned} \|\mathbf{u}^h(\tau)\|_{X^s} &\leq C(T, \|w\|_{H_{(1)}^{s_0+2}}) \left( \mathfrak{J}^s(\tau, \mathbf{f}, \mathbf{g}) + \|w\|_{H_{(1)}^{s+1}} \mathfrak{J}^{s_0+1}(\tau, \mathbf{f}, \mathbf{g}) \right) \\ &\leq C(T, \|\bar{u}\|_{H_{(1)}^{s_0+2}}) \left( \mathfrak{J}^s(\tau, \mathbf{f}, \mathbf{g}) + \|\bar{u}\|_{H_{(1)}^{s+1}} \mathfrak{J}^{s_0+1}(\tau, \mathbf{f}, \mathbf{g}) \right). \end{aligned}$$

It follows that  $\|\mathbf{u}^h\|_{X_{(0)}^s}$  is uniformly bounded on  $[0, T]$ .

We now show that for any sequence  $(h_n)$  with  $h_n \rightarrow 0$ , the solutions  $\mathbf{u}^{h_n}$  form a Cauchy sequence in  $X_{(0)}^{s-2}$  so that  $\lim \mathbf{u}^{h_n} = \mathbf{u}$  exists in  $X_{(0)}^{s-2}$ . We proceed as follows: We fix  $\epsilon$  and let  $\mathbf{r} = \mathbf{u}^{h_n} - \mathbf{u}^{h_m}$ . Then  $\mathbf{r}$  satisfies

$$\mathbf{r}_\tau + \frac{1}{\epsilon^p} J^{h_n} \mathcal{L} \mathbf{r} + J^{h_n} \mathcal{N}_{\mathbf{u}}[\bar{\mathbf{u}}] \mathbf{r} = (J^{h_m} - J^{h_n}) \left( \frac{1}{\epsilon^p} \mathcal{L} \mathbf{u}^{h_m} + \mathcal{N}_{\mathbf{u}}[\bar{\mathbf{u}}] \mathbf{u}^{h_m} \right), \quad \mathbf{r}(x, 0) = \mathbf{0}.$$

This equation is of the form (28)-(29) with the nonhomogeneous term  $\tilde{\mathbf{f}} = (J^{h_m} - J^{h_n}) \left( \frac{1}{\epsilon^p} \mathcal{L} \mathbf{u}^{h_m} + \mathcal{N}_{\mathbf{u}}[\bar{\mathbf{u}}] \mathbf{u}^{h_m} \right)$  and  $\mathbf{g} = \mathbf{0}$ . Using the mollifier estimate [4]

$$\|J^{h_1} z - J^{h_2} z\|_{H^s} \leq |h_1 - h_2| \|z\|_{H^{s+1}},$$

we have

$$\|\tilde{\mathbf{f}}\|_{X_{(0)}^{s-2}} \leq |h_n - h_m| \left\| \frac{1}{\epsilon^p} \mathcal{L} \mathbf{u}^{h_m} + \mathcal{N}_{\mathbf{u}}[\bar{\mathbf{u}}] \mathbf{u}^{h_m} \right\|_{X_{(0)}^{s-1}} \leq C |h_n - h_m| \|\mathbf{u}^{h_m}\|_{X_{(0)}^s},$$

where we have used the fact that the operator  $\mathcal{L}$  maps  $X^s$  into  $X^{s-1}$ . As  $\|\mathbf{u}^{h_m}\|_{X_{(0)}^s}$  is uniformly bounded, this yields  $\|\tilde{\mathbf{f}}\|_{X_{(0)}^{s-2}} \leq C |h_n - h_m|$ . For  $s \geq s_0 + 3$  this implies that  $\mathfrak{J}^{s-2}(\tau, \tilde{\mathbf{f}}, \mathbf{0})$  and  $\mathfrak{J}^{s_0+1}(\tau, \tilde{\mathbf{f}}, \mathbf{0})$  are both bounded by  $C|h_n - h_m|$ . Then Corollary 1 gives

$$\|\mathbf{u}^{h_n} - \mathbf{u}^{h_m}\|_{X_{(0)}^{s-2}} \leq C |h_n - h_m|.$$

Thus  $(\mathbf{u}^{h_n})$  is a Cauchy sequence in  $X_{(0)}^{s-2}$ . Moreover, it follows from (28)-(29) that  $(\mathbf{u}_\tau^{h_n})$  is also a Cauchy sequence in  $X_{(0)}^{s-3}$ ; in other words we get that  $(\mathbf{u}^{h_n})$  is Cauchy in  $X_{(1)}^{s-2}$ . It is clear that  $\lim \mathbf{u}^{h_n} = \mathbf{u} \in X_{(1)}^{s-2}$  and  $\mathbf{u}$  satisfies (36). Next we show that  $\mathbf{u}$  is in  $X_{(1)}^s$ . We first observe that  $\|\mathbf{u}^{h_n}\|_{X_{(0)}^s}$  is bounded and thus  $(\mathbf{u}^{h_n})$  has a weakly convergent subsequence in  $X_{(0)}^s$ . By uniqueness of the limit, this weak limit should be  $\mathbf{u}$  which gives the regularity result  $\mathbf{u} \in X_{(0)}^s$ . Corollary 1 then gives the required estimate for  $\|\mathbf{u}(\tau)\|_{X^s}$ . Since (36) is linear, uniqueness of the solution follows again from Corollary 1.

Thus we have established that the three assumptions of Theorem 4.1 hold for our problem (31).

## 5 Long-time existence of solutions

We are now ready to apply Theorem 4.1 to our problem (31). We first discuss the parameters appearing in the statement of the theorem. Recall that  $s_0 > \frac{1}{2}$  in our case. As mentioned in Section 4,  $T > 0$  can be taken arbitrarily large. The other three parameters  $m, d_1, d'_1$  appearing in Theorem 4.1 refer to loss of derivatives in the assumptions. We note that, although different values of  $m$  appear in the proofs above, all the assumptions will hold for  $m = 3$ . As a result we have  $m = 3, d_1 = 1$  and  $d'_1 = 0$  in our case. The quantities  $\delta, q$  and  $P_{\min}$  are defined as

$$\delta = \max\{d_1, d'_1 + m\}, \quad q = D - m - d'_1, \quad P_{\min} = \delta + \frac{D}{q}(\sqrt{\delta} + \sqrt{2(\delta + q)})^2 \quad (38)$$

in [1]. Thus, these quantities take the values

$$\delta = 3, \quad q = D - 3, \quad P_{\min} = 3 + \frac{D}{D - 3}(\sqrt{3} + \sqrt{2D})^2. \quad (39)$$

In Theorem 4.1 all assumptions are supposed to hold for  $\epsilon < \epsilon_0$ . In our case, for a given  $w$ , the parameter  $\epsilon_0$  is determined by Lemma 3.1. Considering that  $w$  is replaced by  $(p+1)u^p$  in the linearization (29), we see that  $\epsilon_0$  is determined by the initial data. Thus we obtain the local well-posedness of (31) as follows:

**Corollary 2** *Let  $D > 3, P > P_{\min}, s \geq s_0 + 3$  and  $(\mathbf{u}_0^\epsilon)$  be bounded in  $X^{s+P}$ . Then there exist some  $\epsilon_0 > 0, T > 0$  and a unique family  $(\mathbf{u}^\epsilon)_{0 < \epsilon < \epsilon_0}$  bounded in  $C([0, T]; X^{s+D})$  and satisfying (31) with initial values  $\mathbf{u}_0(x) = \mathbf{u}_0^\epsilon(x)$ .*

We are now in a position to state our main results. Corollary 2 provides the local well-posedness of solutions to (31) on the interval  $[0, T]$  of the scaled time variable  $\tau$ . Setting  $t = \tau/\epsilon^p$ , that is, changing back from the scaled time variable  $\tau$  to the original time variable  $t$  we get the long-time existence result for solutions of the initial value problem (6)-(7):

**Theorem 5.1** *Let  $D > 3, P > P_{\min}, s \geq s_0 + 3$  and  $(\mathbf{u}_0^\epsilon)$  be bounded in  $X^{s+P}$ . Then there exist some  $\epsilon_0 > 0, T > 0$  and a unique family  $(\mathbf{u}^\epsilon)_{0 < \epsilon < \epsilon_0}$  bounded in  $C([0, \frac{T}{\epsilon^p}]; X^{s+D})$  and satisfying (6)-(7) with initial values  $\mathbf{u}_0(x) = \mathbf{u}_0^\epsilon(x)$ .*

This is our first main result and it allows us to say that solutions to (6)-(7) exist over the long time interval  $[0, \frac{T}{\epsilon^p}]$  of  $t$ . Now we want to state an analog of the above theorem for solutions of the initial-value problem (1)-(2). The basic issue is how to choose the initial data  $(u_0, v_0)$  for (6)-(7) when the initial data  $(u_0, u_1)$  for (1)-(2) is given. Obviously,  $u_0$ 's are the same and we have  $u_1 = K(v_0)_x$ . Hence if we have the initial data in the form  $u_1 = (w_0)_x$ , we get  $w_0 = \widetilde{K}v_0$ . When  $K$  is invertible, we have  $v_0 = K^{-1}w_0$ . We note that invertibility of  $K$  is equivalent to the strict positivity (ellipticity) condition

$$0 < C_1 \leq \widehat{\beta}(\xi) \leq C_2 \quad (40)$$

on the kernel. In this case, we have:

**Theorem 5.2** *Suppose the kernel  $\beta$  satisfies the ellipticity condition (40). Let  $D > 3$ ,  $P > P_{min}$ ,  $s \geq s_0 + 3$  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  $C([0, \frac{T}{\epsilon^p}]; H^{s+D}) \times C^1([0, \frac{T}{\epsilon^p}]; H^{s+D-1})$  and satisfying (1)-(2) with initial values  $u_0(x) = u_0^\epsilon(x)$ ,  $u_1(x) = (w_0^\epsilon(x))_x$ .*

**Remark 2** *If  $\beta$  is not elliptic, then  $K$  is not invertible and we cannot perform the above transformation. In this case, if  $u_1$  is of the form  $u_1 = (Kv_0(x))_x$  with  $(u_0^\epsilon, v_0^\epsilon)$  bounded in  $H^{s+P} \times H^{s+P}$ , then Theorem 5.2 will still be valid.*

If we make the transformation  $U = \epsilon u$ , the Cauchy problem (1)-(2) takes the form

$$U_{tt} = \beta * (U + U^{p+1})_{xx}, \quad x \in \mathbb{R}, \quad t > 0, \quad (41)$$

$$U(x, 0) = \epsilon u_0(x), \quad U_t(x, 0) = \epsilon u_1(x), \quad (42)$$

where the small parameter  $\epsilon$  is transferred from the nonlocal equation to the initial conditions. Then we get the following result about the long-time existence for small initial data.

**Theorem 5.3** *Suppose  $\beta$  satisfies the ellipticity condition (40). Let  $D > 3$ ,  $P > P_{min}$ ,  $s \geq s_0 + 3$  and  $(u_0, w_0)$  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  $C([0, \frac{T}{\epsilon^p}]; H^{s+D}) \times C^1([0, \frac{T}{\epsilon^p}]; H^{s+D-1})$  and satisfying (41)-(42) with  $u_1 = (w_0)_x$ .*

**Remark 3** *The extra smoothness requirement  $(u_0^\epsilon) \in X^{s+P}$  (whereas solution is in  $X^{s+D}$  with  $D < P$ ) in the above theorems is due to the technical aspect of the Nash-Moser scheme. Computation shows that the optimal values are obtained when the pair  $(P, D)$  is approximately  $(55.34, 7.35)$ . In certain cases the lower bound for  $P$  may be improved. In particular, we will consider the case when  $\beta$  is more regular, namely when (9) is satisfied with  $r \geq 2$ . In this case both  $K$  and  $KD_x$  are bounded operators on  $H^s$ . Then (31) is an initial-value problem defined for  $H^s$ -valued ODEs, so that we have local existence without loss of derivatives. Moreover, with the simplified energy*

$$\tilde{E}_s^2(\tau) = \frac{1}{2} \left( \|u^\epsilon(\tau)\|_{H^s}^2 + \|v^\epsilon(\tau)\|_{H^s}^2 \right),$$

one directly gets from (31)

$$\frac{d}{d\tau} \tilde{E}_s^2(\tau) = \langle KD_x (u^\epsilon(\tau))^{p+1}, v^\epsilon(\tau) \rangle_{H^s} \leq C \tilde{E}_s^{p+2}(\tau).$$

*This differential inequality gives uniform estimates for the solutions  $(u^\epsilon(\tau), v^\epsilon(\tau))$ . Changing back the variable to  $t$ , we get the following result for (6)-(7) with no extra smoothness requirement. Clearly the same conclusion will be valid for the other theorems, namely when  $r \geq 2$  we can take  $P = D = 0$ . We note that this result is the long-time version of the local existence result given in [3].*

**Theorem 5.4** *Let  $s \geq s_0$ . Suppose the kernel  $\beta$  satisfies (9) with  $r \geq 2$ . Let  $(\mathbf{u}_0^\epsilon)$  be bounded in  $X^s$ . Then there exist some  $\epsilon_0 > 0$ ,  $T > 0$  and a unique family  $(\mathbf{u}^\epsilon)_{0 < \epsilon < \epsilon_0}$  bounded in  $C^1([0, \frac{T}{\epsilon^p}]; X^s)$  and satisfying (6)-(7) with initial values  $\mathbf{u}_0(x) = \mathbf{u}_0^\epsilon(x)$ .*

In the remainder of this section, we provide some examples of the general class (1): the classical elasticity equation, the improved Boussinesq equation and the lattice equation.

**Example 5.5** *(The Classical Elasticity Equation) When  $\beta = \delta$  where  $\delta$  is the Dirac measure, we capture the well-posedness result for the one-dimensional classical (local) elasticity equation (4). Here  $\widehat{\beta} = 1$  so  $K$  is the identity operator and Theorem 5.2 holds.*

**Example 5.6** *(The Improved Boussinesq Equation) When  $\beta(x) = \frac{1}{2}e^{-|x|}$ , the nonlocal equation (1) becomes the improved Boussinesq equation given by (3). The Fourier transform of the kernel is  $\widehat{\beta}(\xi) = (1 + \xi^2)^{-1}$ . So the kernel satisfies both the regularity assumption (9) with  $r = 2$  and the ellipticity condition (40). Then Theorem 5.4 holds.*

**Example 5.7** *(The Lattice Equation) The class (1) involves also the differential-difference equation*

$$u_{tt} = \Delta^d(u + \epsilon^p u^{p+1}) \quad (43)$$

with the discrete Laplacian operator

$$\Delta^d z = z(x-1, t) - 2z(x, t) + z(x+1, t). \quad (44)$$

Equation (43) is widely used as a model in one dimensional non-linear lattice dynamics. Indeed for the triangular kernel,  $\beta(x) = 1 - |x|$  for  $|x| \leq 1$  and  $\beta(x) = 0$  elsewhere, (1) reduces to (43). We have  $\widehat{\beta}(\xi) = \frac{4}{\xi^2} \sin^2(\frac{\xi}{2})$  which satisfies (9) with  $r = 2$  but fails to be elliptic. The operator  $K$  with symbol  $\frac{2}{\xi} \sin(\frac{\xi}{2})$  can be explicitly calculated,  $Ku = \chi_{[-\frac{1}{2}, \frac{1}{2}]} * u$  with the characteristic function  $\chi$ . Then following Remark 2, for initial data of the form  $(u_0^\epsilon, u_1^\epsilon)$  with  $u_1^\epsilon = (Kv_0^\epsilon(x))_x$  and  $(u_0^\epsilon, v_0^\epsilon)$  bounded in  $H^s \times H^s$ , we have uniform bounds and long time existence in  $C^1([0, \frac{T}{\epsilon^p}]; H^s)$ .

Local existence of smooth solutions of (4) is a well-known result. In fact the much more complicated 3D case can be found in [6]. For long-time existence of the 3D case we refer to [7] and the references therein. Our result in Example 5.5 recaptures in fact the simpler 1D case. The wave equations in Examples 5.6 and 5.7 are particular cases of (1) with  $r \geq 2$  and the local well-posedness for those equations follows from the general result in [3]. Our results here yield long-time existence and uniform estimates of the solutions.



## References

- [1] B. Alvarez-Samaniego and D. Lannes, A Nash-Moser theorem for singular evolution equations. Application to the Serre and Green-Naghdi equations, *Indiana Univ. Math. Journal*, **57** (2008), 97–131.
- [2] A. Constantin and D. Lannes, The hydrodynamical relevance of the Camassa-Holm and Degasperis-Procesi equations, *Arch. Rational Mech. Anal.*, **192** (2009), 165–186.
- [3] N. Duruk, H.A. Erbay and A. Erkip, Global existence and blow-up for a class of nonlocal nonlinear Cauchy problems arising in elasticity, *Nonlinearity*, **23** (2010), 107–118.
- [4] M. Ehrnstrom, L. Pei and Y. Wang, A conditional well-posedness result for the bidirectional Whitham equation, preprint, arXiv:1708.04551 [math.AP].
- [5] H. A. Erbay, S. Erbay and A. Erkip, 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.*, **36** (2016), 6101–6116.
- [6] T. J. R. Hughes, T. Kato and J. E. Marsden, Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity, *Arch. Ration. Mech. Anal.*, **63** (1977), 273–294.
- [7] S. Klainerman, Long time behaviour of solutions to nonlinear wave equations, in *Proceedings of the International Congress of Mathematicians (Warsaw, 1983)*, PWN, Warsaw, (1984), 1209-1215.
- [8] M. Ming, J. C. Saut and P. Zhang, Long-time existence of solutions to Boussinesq systems, *SIAM J. Math. Anal.*, **44** (2012), 4078–4100.
- [9] J. C. Saut and L. Xu, The Cauchy problem on large time for surface waves Boussinesq systems, *J. Math. Pures. Appl.*, **97** (2012), 635–662.
- [10] J. C. Saut, C. Wang and L. Xu, The Cauchy problem on large time for surface-waves-type Boussinesq systems II, *SIAM J. Math. Anal.*, **49** (2017), 2321–2386.
- [11] M. E. Taylor, *Partial Differential Equations II. Qualitative Studies of Linear Equations*, 2<sup>nd</sup> edition, Springer, New York, 2011.