Multiscale Model. Simul.
Vol. 12, No. 3, pp. 953–995

APPROXIMATION OF POLYATOMIC FPU LATTICES BY KdV EQUATIONS*

JEREMY GAISON†, SHARI MOSKOW†, J. DOUGLAS WRIGHT†, AND QIMIN ZHANG†

Abstract. We consider the evolution of small amplitude, long wavelength initial data by a polyatomic Fermi–Pasta–Ulam lattice differential equation whose material properties vary periodically. Using the methods of homogenization theory, we prove rigorous estimates that show that the solution breaks up into the linear superposition of two appropriately scaled and modulated counter-propagating waves, each of which solves a Korteweg–de Vries equation, plus a small error. The estimates are valid over very long time scales.

Key words. Fermi–Pasta–Ulam (FPU), Korteweg–de Vries (KdV), solitary waves, homogenization, polyatomic lattice

AMS subject classifications. 37L60, 37L50, 74J35

DOI. 10.1137/130941638

1. Introduction.

Newton's law for an infinite chain of oscillators with nearest neighbor (conservative) interactions is

$$m(j)\ddot{u}(j) = \mathcal{V}_j'(u(j+1) - u(j) - l) - \mathcal{V}_{j-1}'(u(j) - u(j-1) - l). \tag{1.1}$$

Here $j \in \mathbf{Z}$ and $u(j)$ is the position of the jth mass. The constant $l \geq 0$ is the relaxation length of the nonlinear spring and can be taken, without loss of generality, to be zero. The mass of the jth particle is

$$m(j) > 0. \tag{1.2}$$

$\mathcal{V}_j$ is the potential function for the nonlinear spring which is between the jth and $j+1$st particles. We assume the intersite potential $\mathcal{V}_j$ is smooth and

$$\mathcal{V}_j(h) = \frac{1}{2}\kappa(j)h^2 + \frac{1}{3}\beta(j)h^3 + O(h^3). \tag{1.3}$$

Note that

$$\kappa(j) > 0 \tag{1.4}$$

is Hooke's constant for the springs. We take the masses $m(j)$ and the potentials $\mathcal{V}_j$ to vary periodically. That is, there is $N \in \mathbf{N}$ such that

$$m(j+N) = m(j) \quad \text{and} \quad \mathcal{V}_j(h) = \mathcal{V}_{j+N}(h) \tag{1.5}$$

for all $j \in \mathbf{Z}$ and $h \in \mathbf{R}$. In the case when the masses and potentials do not depend on j this system is a classic Fermi–Pasta–Ulam (FPU) lattice (see [7]).

*Received by the editors October 17, 2013; accepted for publication (in revised form) May 7, 2014; published electronically July 1, 2014. This research was partially supported by the National Science Foundation through grants DMS-1105635, DMS-0908299, and DMS-1108858. The work of the first and fourth authors was also partially supported by Drexel's Steinbright Career Development Center co-op program.

http://www.siam.org/journals/mms/12-3/94163.html

†Department of Mathematics, Drexel University, 3141 Chestnut St., Philadelphia, PA 19104 (jkg65@drexel.edu, moskow@math.drexel.edu, jdoug@math.drexel.edu, q243@drexel.edu).

There is a long history of computing and validating macroscopic effective partial differential equations for solutions of the homogeneous FPU lattice. Depending on the scaling regime under investigation, these effective equations include nonlinear Schrödinger equations (see, e.g., [10], [20]), Boussinesq equations (e.g., [2]), or hyperbolic conservation laws (e.g., [11]). Our interest is in the propagation of long waves which are also of small amplitude, the so-called long wave limit. In the homogeneous case, it is well known that such solutions break up into the linear superposition of counterpropagating solutions to a pair of Korteweg–de Vries (KdV) equations plus a small remainder. See [24], [13] for the earliest formal derivations. Various aspects of the formal arguments were first made precise in the articles [8], [9], [21]. Of these, [21] is most closely related to the work here. It concerns general solutions of the initial value problem for FPU in the long wave limit; see below for a more thorough discussion of this article. The articles [8] and [9] use the KdV approximation as a jumping off point for rigorous proofs of the existence and stability of small amplitude solitary wave solutions to FPU.

While there is a wide array of formal asymptotics that demonstrate that long waves behave in polyatomic FPU much as they do in the homogeneous case (see, for instance, [5], [14], [15], [18], [19], [22]) and there is a large collection of results concerning breather solutions in this setting (see [12] and the references therein), at this time the rigorous validation of long wave limits is complete only in the case when the system is linear (see [16]) or $N = 2$ (see [3]). In this article we prove a result which is similar to those in [21] and [3]. Of particular interest here is that the long wave analysis follows from the methods of homogenization theory; in short, since we are interested in long waves, the material coefficients are essentially "rapidly varying" in comparison. As such, the classic tools of homogenization are perfectly suited for the analysis (see, for instance, [4]). A complicating feature is that we need to carry out the homogenization asymptotics to relatively high order so that the "weakly nonlinear" effects which give rise to KdV dynamics manifest.

2. Formulation and main result. We denote a sequence $\{x(j)\}$ by $x = \{x(j)\}$. Similarly, when we write $\mathcal{V}(x)$ we mean $\mathcal{V}(x)(j) = \mathcal{V}_j(x(j))$. Let $S^{\pm}$ be the shift operators which act on sequences $f = \{f(j)\}$ as

$$(S^{\pm} f)(j) := f(j \pm 1),$$

and likewise the operators δ^+ and δ^- are the left and right difference operators given by

$$(\delta^+ f)(j) := f(j+1) - f(j) \quad \text{and} \quad (\delta^- f)(j) := f(j) - f(j-1).$$

Defining

$$r := \delta^+ u \quad \text{and} \quad p := \dot{u},$$

we convert our second order equation (1.1) to the system

$$\begin{aligned} \dot{r} &= \delta^+ p, \\ \dot{p} &= \frac{1}{m}\delta^-(\mathcal{V}'(r)). \end{aligned} \tag{2.1}$$

Here is our main result.

Theorem 2.1 (long wave solutions of (2.1) are approximated by KdV equations). *Fix m and $\mathcal{V}$ subject to the positivity and periodicity conditions in* (1.2), (1.4)*, and*

(1.5). *Define the constants* a, b, c, $\bar{m}$, $\breve{\kappa}$ *from* m *and* $\mathcal{V}$ *according to the formulas found immediately below in Definition* 2.2. *Suppose that* $a \neq 0$. *Fix* $\phi, \psi \in H^5$ *and* $T_0 > 0$. *Let*

$$\Phi(X) := \int_0^X \phi(y)dy \quad \textit{and} \quad \Psi(X) := \int_0^X \psi(y)dy \tag{2.2}$$

and suppose $\Phi, \Psi \in L^\infty$. *Then there exists* $\epsilon_0 = \epsilon_0(\|\phi\|_{H^5}, \|\psi\|_{H^5}, \|\Phi\|_{L^\infty}, \|\Psi\|_{L^\infty}, T_0, m, \mathcal{V}) > 0$ *and* $C_\star = C_\star(\|\phi\|_{H^5}, \|\psi\|_{H^5}, \|\Phi\|_{L^\infty}, \|\Psi\|_{L^\infty}, T_0, m, \mathcal{V}) > 0$ *such that the following is true for all* $0 < \epsilon < \epsilon_0$.

Let $\mathbf{r}(j,t) := (r(j,t), p(j,t)) \in C^1(\mathbf{R}; \ell^2 \times \ell^2)$ *be the solution of* (2.1) *with initial conditions*

$$r(j,0) = \frac{\epsilon^2}{\kappa(j)}\phi(\epsilon j) \quad \textit{and} \quad p(j,0) = \epsilon^2 \psi(\epsilon j). \tag{2.3}$$

Suppose that $A(w,T)$ *and* $B(l,T)$ *solve the KdV equations*

$$\frac{1}{c}A_T + aA_{www} + bAA_w = 0 \quad \textit{and} \quad \frac{1}{c}B_T - aB_{lll} - bBB_l = 0 \tag{2.4}$$

with

$$A(w,0) = \frac{1}{2}\left(\phi(w) - \sqrt{\overline{m}\breve{\kappa}}\psi(w)\right) \quad \textit{and} \quad B(l,0) = \frac{1}{2}\left(\phi(l) + \sqrt{\overline{m}\breve{\kappa}}\psi(l)\right). \tag{2.5}$$

Then

$$\sup_{|t| \leq T_0\epsilon^{-3}} \|\mathbf{r}(t) - \epsilon^2 \mathbf{A}_\epsilon(t)\|_{\ell^2 \times \ell^2} \leq C_\star \epsilon^{5/2}, \tag{2.6}$$

where

$$\begin{aligned} \mathbf{A}_\epsilon(j,t) := \Bigg(&\frac{1}{\kappa(j)}\left[A(\epsilon(j-ct), \epsilon^3 t) + B(\epsilon(j+ct), \epsilon^3 t)\right], \\ &\frac{1}{\sqrt{\overline{m}\breve{\kappa}}}\left[-A(\epsilon(j-ct), \epsilon^3 t) + B(\epsilon(j+ct, \epsilon^3 t))\right]\Bigg). \end{aligned} \tag{2.7}$$

Definition 2.2 (homogenized coefficients). *The constants above are determined from* $m(j)$, $\kappa(j)$, *and* $\beta(j)$ *as follows. Set*

$$\overline{m} := \frac{1}{N}\sum_{j=1}^N m(j), \quad \breve{\kappa} := \frac{N}{\sum_{j=1}^N \kappa^{-1}(j)}, \quad b := \frac{\breve{\kappa}}{N}\sum_{j=1}^N \frac{\beta(j)}{\kappa^3(j)}, \quad \textit{and} \quad c := \sqrt{\frac{\breve{\kappa}}{\overline{m}}}.$$

Let

$$\chi_1(1) := -\frac{1}{N}\sum_{k=2}^N\sum_{j=1}^{k-1}\left(\frac{\breve{\kappa}}{\kappa(j)} - 1\right) \quad \textit{and} \quad \chi_2(1) := -\frac{1}{N}\sum_{k=2}^N\sum_{j=1}^{k-1}\left(\frac{m(j+1)}{\overline{m}} - 1\right)$$

and, for $k = 2, \ldots, N$,

$$\chi_1(k) := \chi_1(1) + \sum_{j=1}^{k-1}\left(\frac{\breve{\kappa}}{\kappa(j)} - 1\right) \quad \textit{and} \quad \chi_2(k) := \chi_2(1) + \sum_{j=1}^{k-1}\left(\frac{m(j+1)}{\overline{m}} - 1\right).$$

For N-periodic sequences $f(j)$ and $g(j)$ let

$$\langle f,g\rangle := \frac{1}{N}\sum_{j=1}^{N} f(j)g(j).$$

With this, set

$$\gamma_0 := \left\langle \frac{\breve{\kappa}}{\kappa},\chi_2\right\rangle, \quad \gamma_1 := \langle \chi_1,\chi_2\rangle, \quad \gamma_2 := \left\langle \chi_1^2, \frac{m}{\overline{m}}\right\rangle, \quad \gamma_3 := \left\langle \chi_2^2, \frac{\breve{\kappa}}{\kappa}\right\rangle,$$

and

$$a := \frac{1}{24}\left[1-12\gamma_0-12\gamma_0^2-24\gamma_1+12\gamma_2+12\gamma_3\right].$$

REMARK 1. *Notice that since $m(j)>0$ and $\kappa(j)>0$, it is clear that $c>0$. It is not clear whether the same can be said for a. Nevertheless, we have numerically computed values of a for a large number of randomly selected positive choices of $m(j)$ and $\kappa(j)$ and have always found that $a>0$. If $a=0$, then notice that the approximation would be by solutions of Burgers' equations as opposed to KdV equations and, as such, would require different techniques for validation. Since we do not expect this situation to be possible, we exclude it, hence the assumption in Theorem* 2.1 *that $a\neq 0$. Finally it is entirely possible for b to vanish, provided $\beta(j)$ oscillates about zero just so. This is not an obstruction for our method, as the modulation equations are then Airy's equations, which are similar enough to KdV equations that there is no additional difficulty in including this case.*

As stated above, Theorem 2.1 is in many ways an extension of the main results of [21] (which applies only to chains with constant material coefficients) and that of [3] (which handles the $N=2$ case), to the case of arbitrary N.[1] There are other major differences here, however. The first, and most obvious, is that our error estimate is a full power of ϵ less accurate than in those works. We stress here that this is not a limitation of our method, but is in fact a natural byproduct of the homogenization process. In particular, we could improve our error estimate to $\epsilon^{7/2}$, but this would require substantive restrictions on the form of the initial data—a technical point we discuss below in Remark 5. Such restrictions are employed in [3] to get their improved error estimate. Our point of view is that the $1/\kappa(j)$ term that appears in the initial data for r is restrictive enough.

There are other notable distinctions between our result and those in [21] and [3]. In [21] the initial data is required to lie in $H^s\cap H^7(3)$, where $s>14$. (Here, $H^s(m) := \left\{f(X) : \sqrt{1+X^{2m}}f(X)\in H^s\right\}$.) Such burdensome regularity and decay conditions on initial conditions is typical in rigorous approximation results (see [23] for a particularly egregious example). In our theorem, we require only H^5 regularity for the initial conditions. Moreover, we have replaced the algebraic decay condition with the much weaker condition that the antiderivatives Φ and Ψ are in L^∞. Much of the reduction in the needed regularity is attributable to nothing more than careful bookkeeping when we estimate the "residuals" (see (3.1) below for a precise definition). Such bookkeeping would allow us to reduce the regularity to H^7, which is the same as the regularity required in [3]. The final two derivatives we eliminate using the fact that

[1] The results of [16] primarily concern various weak limits for linear problems in high dimensions and, as such, are not as comparable to the work done here.

the least regular terms in the residuals appear with extra powers of ϵ on them. Taking ϵ-dependent truncations of the Fourier transforms of the initial data, we are able to exchange these additional powers of ϵ for smoothness (see subsection 5.2). As for the algebraic decay condition, we eliminate this condition by a technical rearrangement of terms in the approximation (see subsection 3.5, specifically Remark 3).

The paper is organized as follows. In section 3 we derive the KdV equation semirigorously. In section 4 we provide precise estimates that demonstrate that the approximation very nearly solves (2.1); that is, it contains estimates on the residuals. In section 5 we prove Theorem 2.1. In section 6 we carry out numerical simulations demonstrating various aspects of our results.

3. Homogenization via multiscale asymptotics.

3.1. Preliminaries. For any functions $\tilde{r}(j,t)$ and $\tilde{p}(j,t)$ define the *residuals* as

$$\mathrm{Res}_1(\tilde{r},\tilde{p}) := \delta^+\tilde{p} - \partial_t\tilde{r} \quad \text{and} \quad \mathrm{Res}_2(\tilde{r},\tilde{p}) := \frac{1}{m}\delta^-\left(\mathcal{V}(\tilde{r})\right) - \partial_t\tilde{p}. \tag{3.1}$$

If (r,p) is a solution of (2.1), then $\mathrm{Res}_1(r,p)$ and $\mathrm{Res}_2(r,p)$ are identically zero. The goal of our asymptotic expansion is to find $(\tilde{r},\tilde{p})$ so that the residuals are sufficiently small and $(\tilde{r},\tilde{p})$ is (in some sense) easier to compute than a true solution.

In our case, we follow the prescriptions of homogenization theory and look for (approximate) solutions of the form

$$\tilde{r}(j,t) = \tilde{r}_\epsilon(j,t) := \epsilon^2 R(j,\epsilon j,\epsilon t,\epsilon^3 t) \quad \text{and} \quad \tilde{p}(j,t) = \tilde{p}_\epsilon(j,t) := \epsilon^2 P(j,\epsilon j,\epsilon t,\epsilon^3 t), \tag{3.2}$$

where R and P are maps

$$\mathbf{Z}\times\mathbf{R}\times\mathbf{R}\times\mathbf{R}\to\mathbf{R}.$$

A critical part of our assumption is that

$$R(j+N,X,\tau,T) = R(j,X,\tau,T) \quad \text{and} \quad P(j+N,X,\tau,T) = P(j,X,\tau,T) \tag{3.3}$$

for all (j,X,τ,T). That is, the functions are periodic in their first slot with the same period as m and $\mathcal{V}$.

We must understand how $\delta^\pm$ act on functions of this type. It is a straightforward exercise to use Taylor's theorem to show that if $u(j) = U(j,\epsilon j)$, then

$$\delta^\pm u(j) = \sum_{n\geq 0}\epsilon^n \delta_n^\pm U,$$

where

$$\delta_0^\pm := \check{\delta}^\pm \quad \text{and} \quad \delta_n^\pm := \frac{(\pm 1)^{n+1}}{n!}\check{S}^\pm\partial_X^n. \tag{3.4}$$

Here $\check{\delta}^\pm$ and $\check{S}^\pm$ act only on the first slot. That is, they are analogous to partial derivatives with respect to j. Precisely,

$$\begin{aligned}
(\check{S}^+U)(j,X) &:= U(j+1,X),\\
(\check{S}^-U)(j,X) &:= U(j-1,X),\\
(\check{\delta}^+U)(j,X) &:= U(j+1,X) - U(j,X), \quad \text{and}\\
(\check{\delta}^-U)(j,X) &:= U(j,X) - U(j-1,X).
\end{aligned}$$

Let

$$(E_M^{\pm} u)(j) := (\delta^{\pm} u)(j) - \sum_{j=0}^{M} \epsilon^n (\delta_n^{\pm} U)(j, \epsilon j)$$

be the error made by truncating the series expansion of $\delta^{\pm} u$ after M terms. On the formal level, we have

$$E_M^{\pm} = O(\epsilon^{M+1}). \tag{3.5}$$

(A rigorous estimate on $E_M^{\pm}$ is found in Lemma 4.3 below.)

Before carrying out the expansion, we make one further refinement to our Ansatz (3.2):

$$\tilde{r}_\epsilon(j,t) := \sum_{n=0}^{3} \epsilon^{n+2} R_n(j, \epsilon j, \epsilon t, \epsilon^3 t) \quad \text{and} \quad \tilde{p}_\epsilon(j,t) := \sum_{n=0}^{3} \epsilon^{n+2} P_n(j, \epsilon j, \epsilon t, \epsilon^3 t). \tag{3.6}$$

That is to say, R and P themselves have an expansion in ϵ. The functions R_n and P_n have the same dependencies and periodic behavior as R and P do in (3.3). Our decision to take this expansion only to $O(\epsilon^5)$ is perhaps not obvious; as the reader shall see, it is at this level that the KdV dynamics appear.

If we insert these into the definition of Res_1 above, we have

$$\mathrm{Res}_1(\tilde{r}_\epsilon, \tilde{p}_\epsilon) = \sum_{n=0}^{3} \epsilon^{n+2} \delta^+ P_n - \sum_{n=0}^{3} \epsilon^{n+2} \partial_t R_n.$$

If we use the definition of E_M^+, this can be rewritten as

$$\mathrm{Res}_1(\tilde{r}_\epsilon, \tilde{p}_\epsilon) = \sum_{n=0}^{3} \epsilon^{n+2} \sum_{k=0}^{3-n} \epsilon^k \delta_k^+ P_n + \sum_{n=0}^{3} \epsilon^{n+2} E_{3-n}^+ P_n - \sum_{n=0}^{3} \epsilon^{n+2} \partial_t R_n. \tag{3.7}$$

In light of (3.5), we expect $\epsilon^{n+2} E_{3-n}^+ P_n = O(\epsilon^6)$ for $n = 0, \ldots, 3$, provided the P_n are well behaved. This is, as we see below, the appropriate power of ϵ the residuals should satisfy to prove our main theorem.

Next we observe that $R_n = R_n(j, \epsilon j, \epsilon t, \epsilon^3 t)$ implies $\partial_t R_n = \epsilon \partial_\tau R_n + \epsilon^3 \partial_T R_n$. This, together with some reorganization, gives

$$\begin{aligned} \mathrm{Res}_1(\tilde{r}_\epsilon, \tilde{p}_\epsilon) = &\sum_{n=0}^{3} \epsilon^{n+2} \sum_{k=0}^{n} \delta_{n-k}^+ P_k - \sum_{n=1}^{3} \epsilon^{n+2} \partial_\tau R_{n-1} - \epsilon^5 \partial_T R_0 \\ &- \epsilon^6 \partial_\tau R_3 - \sum_{n=1}^{3} \epsilon^{n+5} \partial_T R_n + \sum_{n=0}^{3} \epsilon^{n+2} E_{3-n}^+ P_n. \end{aligned} \tag{3.8}$$

Note that the first row contains only terms which are (formally) $O(\epsilon^5)$ and the second row contains only terms which are higher than $O(\epsilon^6)$.

We now carry out a similar calculation for Res_2. This is slightly more complicated because $\mathcal{V}$ is nonlinear. Towards this end, let

$$N_1(\tilde{r}_\epsilon) := \frac{1}{m} \delta^- \left\{ \mathcal{V}'(\tilde{r}_\epsilon) - \kappa \tilde{r}_\epsilon - \beta \tilde{r}_\epsilon^2 \right\}.$$

Given the expansion in (1.3), Taylor's theorem implies that this is formally cubic in $\tilde{r}_\epsilon$. Thus, since $\tilde{r}_\epsilon$ is $O(\epsilon^2)$, $N_1(\tilde{r}_\epsilon)$ is formally $O(\epsilon^6)$. Similarly, if we let

$$N_2(\tilde{r}_\epsilon) := \frac{1}{m}\delta^- \left(\beta\tilde{r}^2 - \epsilon^4\beta R_0^2 - 2\epsilon^5\beta R_0 R_1\right),$$

then (3.6) implies that this is $O(\epsilon^6)$.

These definitions and steps, completely like those used to arrive at (3.8), yield

$$\begin{aligned}\mathrm{Res}_2(\tilde{r}_\epsilon, \tilde{p}_\epsilon) = &\frac{1}{m}\sum_{n=0}^{3}\epsilon^{n+2}\sum_{k=0}^{n}\delta_{n-k}^-\left(\kappa R_k\right) - \sum_{n=1}^{3}\epsilon^{n+2}\partial_\tau P_{n-1} - \epsilon^5\partial_T P_0\\ &+ \epsilon^4\frac{1}{m}\delta_0^-\left(\beta R_0^2\right) + \epsilon^5\frac{2}{m}\delta_0^-\left(\beta R_0 R_1\right) + \epsilon^5\frac{1}{m}\delta_1^-\left(\beta R_0^2\right)\\ &- \epsilon^6\partial_\tau P_3 - \sum_{n=1}^{3}\epsilon^{n+5}\partial_T P_n + \sum_{n=0}^{3}\epsilon^{n+2}E_{3-n}^-(\kappa R_n)\\ &+ \epsilon^4\frac{1}{m}E_1^-\left(\beta R_0^2\right) + \epsilon^5\frac{2}{m}E_0^-\left(\beta R_0 R_1\right) + N_1(\tilde{r}_\epsilon) + N_2(\tilde{r}_\epsilon).\end{aligned}$$

The first two lines have terms which are $O(\epsilon^5)$ or lower. The last two consist of terms which are formally of order higher than $O(\epsilon^6)$.

As mentioned above, the goal of our asymptotics is to choose the R and P functions so that the residuals are formally $O(\epsilon^6)$. We can accomplish this, provided we have

$$0 = \sum_{n=0}^{3}\epsilon^{n+2}\sum_{k=0}^{n}\delta_{n-k}^+ P_k - \sum_{n=1}^{3}\epsilon^{n+2}\partial_\tau R_{n-1} - \epsilon^5\partial_T R_0 \tag{3.9}$$

and

$$\begin{aligned}0 = &\sum_{n=0}^{3}\epsilon^{n+2}\sum_{k=0}^{n}\delta_{n-k}^-\left(\kappa R_k\right) - \sum_{n=1}^{3}\epsilon^{n+2}m\partial_\tau P_{n-1} - \epsilon^5 m\partial_T P_0\\ &+ \epsilon^4\delta_0^-\left(\beta R_0^2\right) + \epsilon^5 2\delta_0^-\left(\beta R_0 R_1\right) + \epsilon^5\delta_1^-\left(\beta R_0^2\right),\end{aligned} \tag{3.10}$$

in which case we have

$$\mathrm{Res}_1(\tilde{r}_\epsilon, \tilde{p}_\epsilon) = -\epsilon^6\partial_\tau R_3 - \sum_{n=1}^{3}\epsilon^{n+5}\partial_T R_n + \sum_{n=0}^{3}\epsilon^{n+2}E_{3-n}^+ P_n \tag{3.11}$$

and

$$\begin{aligned}\mathrm{Res}_2(\tilde{r}_\epsilon, \tilde{p}_\epsilon) = &-\epsilon^6\partial_\tau P_3 - \sum_{n=1}^{3}\epsilon^{n+5}\partial_T P_n + \sum_{n=0}^{3}\epsilon^{n+2}E_{3-n}^-(\kappa R_n)\\ &+ \epsilon^4\frac{1}{m}E_1^-\left(\beta R_0^2\right) + \epsilon^5\frac{2}{m}E_0^-\left(\beta R_0 R_1\right) + N_1(\tilde{r}_\epsilon) + N_2(\tilde{r}_\epsilon).\end{aligned} \tag{3.12}$$

We can make (3.9) and (3.10) happen by setting the right-hand sides of these equations to zero at $O(\epsilon^2)$ to $O(\epsilon^5)$. Doing this, and a little algebra, gives the following equations. At $O(\epsilon^2)$ we have

$$\delta_0^+ P_0 = 0 \quad \text{and} \quad \delta_0^- Q_0 = 0. \tag{3.13}$$

At $O(\epsilon^3)$,

$$\delta_0^+ P_1 = \frac{1}{\kappa}\partial_\tau Q_0 - \delta_1^+ P_0 \quad \text{and} \quad \delta_0^- Q_1 = m\partial_\tau P_0 - \delta_1^- Q_0. \tag{3.14}$$

At $O(\epsilon^4)$,

$$\delta_0^+ P_2 = \frac{1}{\kappa}\partial_\tau Q_1 - \delta_1^+ P_1 - \delta_2^+ P_0 \quad \text{and} \quad \delta_0^- Q_2 = m\partial_\tau P_1 - \delta_1^- Q_1 - \delta_2^- Q_0 - \delta_0^- \left(\frac{\beta}{\kappa^2} Q_0^2\right). \tag{3.15}$$

And at $O(\epsilon^5)$,

$$\begin{aligned} &\delta_0^+ P_3 = \frac{1}{\kappa}\partial_T Q_0 + \frac{1}{\kappa}\partial_\tau Q_2 - \delta_1^+ P_2 - \delta_2^+ P_1 - \delta_3^+ P_0 \quad \text{and} \\ &\delta_0^- Q_3 = m\partial_T P_0 + m\partial_\tau P_2 - \delta_1^- Q_2 - \delta_2^- Q_1 - \delta_3^- Q_0 - \delta_1^- \left(\frac{\beta}{\kappa^2} Q_0^2\right) - 2\delta_0^- \left(\frac{\beta}{\kappa^2} Q_0 Q_1\right). \end{aligned} \tag{3.16}$$

In the above, to make our calculations less cluttered, we have put

$$Q_n(j, X, \tau, T) := \kappa(j) R_n(j, X, \tau, T). \tag{3.17}$$

Note that all of the equations in (3.13) through (3.16) are of the form "$\delta_0^\pm F = G$." The following lemma tells us about solving such equations.

LEMMA 3.1. *Suppose that $g(j)$ is an N-periodic sequence. Then there exists an N-periodic sequence $f(j)$ which satisfies*

$$\delta^\pm f = g$$

for all j if and only if

$$\frac{1}{N}\sum_{j=1}^{N} g(j) = 0.$$

Moreover, if $f_1(j, X)$ and $f_2(j, X)$ are two such solutions, then

$$f_1(j) - f_2(j) = \textit{constant}$$

for all $j \in \mathbf{Z}$. Lastly, if we impose the additional condition that

$$\sum_{j=1}^{N} f(j) = 0,$$

then we have

$$\max_j |f(j)| \le C \max_j |g(j)|.$$

Here $C > 0$ depends only on N.

As the space of N-periodic sequences is really just $\mathbf{R}^N$ in disguise, the proof consists only of elementary linear algebra, so we omit it. Our equations (3.13) through

(3.16) depend on X as well as j, and so we will actually be using the following corollary (whose proof is omitted) of the lemma.

COROLLARY 3.2. *Suppose that* $G : \mathbf{Z} \times \mathbf{R} \to \mathbf{R}$ *has* $G(j+N, X) = G(j, X)$ *for all* $j \in \mathbf{Z}$ *and* $X \in \mathbf{R}$. *Then there exists* $F : \mathbf{Z} \times \mathbf{R} \to \mathbf{R}$ *with* $F(j+N, X) = F(j, X)$ *for all* $j \in \mathbf{Z}$ *and* $X \in \mathbf{R}$ *which satisfies*

$$\delta_0^{\pm} F = G$$

for all j *and* X *if and only if*

$$\frac{1}{N} \sum_{j=1}^{N} G(j, X) = 0 \text{ for all } X \in \mathbf{R}. \tag{3.18}$$

Moreover, if $F_1(j, X)$ *and* $F_2(j, X)$ *are two such solutions, then there exists* $\bar{F} : \mathbf{R} \to \mathbf{R}$ *such that*

$$F_1(j, X) - F_2(j, X) = \bar{F}(X)$$

for all $j \in \mathbf{Z}$. *That is to say,* $F_1 - F_2$ *is constant with respect to* j. *Lastly, if we impose the additional condition that*

$$\sum_{j=1}^{N} F(j, X) = 0 \text{ for all } X \in \mathbf{R},$$

then we have

$$\max_j \|F(j, \cdot)\|_{W^{k,p}} \leq C \max_j \|G(j, \cdot)\|_{W^{k,p}}.$$

Here $C > 0$ *depends only on* N.

3.2. Solving the $O(\epsilon^2)$ equations, (3.13). These are

$$\delta_0^{+} P_0 = 0 \quad \text{and} \quad \delta_0^{-} Q_0 = 0.$$

These imply that P_0 and Q_0 do not depend on j. That is,

$$P_0 = \bar{P}_0(X, \tau, T) \quad \text{and} \quad Q_0 = \bar{Q}_0(X, \tau, T). \tag{3.19}$$

REMARK 2. *In what follows, any function wearing a "$^-$" will not depend on* j, *but only on* (X, τ, T).

3.3. Solving the $O(\epsilon^3)$ equations, (3.14). These are

$$\delta_0^{+} P_1 = \frac{1}{\kappa} \partial_\tau Q_0 - \delta_1^{+} P_0 \quad \text{and} \quad \delta_0^{-} Q_1 = m \partial_\tau P_0 - \delta_1^{-} Q_0.$$

If we are to be able to solve these, Corollary 3.2 and (3.19) tell us we must have

$$\frac{1}{N} \sum_{j=1}^{N} \left(\frac{1}{\kappa} \partial_\tau \bar{Q}_0 - \delta_1^{+} \bar{P}_0 \right) = 0 \quad \text{and} \quad \frac{1}{N} \sum_{j=1}^{N} \left(m \partial_\tau \bar{P}_0 - \delta_1^{-} \bar{Q}_0 \right) = 0. \tag{3.20}$$

Using the expression for $\delta_1^{\pm}$ from (3.4) in (3.20) and then summing gives

$$\partial_\tau \bar{Q}_0 = \breve{\kappa} \partial_X \bar{P}_0 \quad \text{and} \quad \partial_\tau \bar{P}_0 = \frac{1}{\overline{m}} \partial_X \bar{Q}_0. \tag{3.21}$$

In the above, we have set $\breve{\kappa} := N\left(\sum_{j=1}^{N} \kappa^{-1}(j)\right)^{-1}$ and $\overline{m} := \frac{1}{N}\sum_{j=1}^{N} m(j)$, as in Definition 2.2.

Notice that (3.21) is basically the wave equation, since it implies $\partial_\tau^2 \bar{Q}_0 = c^2 \partial_X^2 \bar{Q}_0$, where

$$c := \sqrt{\frac{\breve{\kappa}}{\overline{m}}}.$$

Of course, $\bar{Q}_0$ and $\bar{P}_0$ depend on (X, τ, T), and (3.21) says nothing about the T dependence. Nevertheless d'Alembert's formula implies

$$\begin{aligned} \bar{Q}_0(X,\tau,T) =& \bar{A}(X - c\tau, T) + \bar{B}(X + c\tau, T), \\ \bar{P}_0(X,\tau,T) =& \frac{1}{\sqrt{\overline{m}\breve{\kappa}}} \left\{ -\bar{A}(X - c\tau, T) + \bar{B}(X + c\tau, T) \right\}, \end{aligned} \tag{3.22}$$

where $\bar{A}$ and $\bar{B}$ are as yet unspecified; eventually we will show that they should satisfy KdV equations. Note that the functions $\bar{A}$ and $\bar{B}$ differ from one another in how they depend on X and τ. Thus we will write $\bar{A} = \bar{A}(w, T)$ and $\bar{B} = \bar{B}(l, T)$, where

$$l := X + c\tau \quad \text{and} \quad w := X - c\tau$$

are used to denote waves which move "left" and "right," respectively.

Using (3.21) in (3.14), along with (3.4), gives

$$\delta_0^+ P_1 = \left(\frac{\breve{\kappa}}{\kappa} - 1\right) \partial_X \bar{P}_0 \quad \text{and} \quad \delta_0^- Q_1 = \left(\frac{m}{\overline{m}} - 1\right) \partial_X \bar{Q}_0. \tag{3.23}$$

Define $\chi_1(j)$ to be the unique solution of

$$\delta_0^+ \chi_1 = \frac{\breve{\kappa}}{\kappa} - 1 \quad \text{and} \quad \sum_{j=1}^{N} \chi_1(j) = 0. \tag{3.24}$$

Define $\chi_2(j)$ to be the unique solution of

$$\delta_0^- \chi_2 = \frac{m}{\overline{m}} - 1 \quad \text{and} \quad \sum_{j=1}^{N} \chi_2(j) = 0. \tag{3.25}$$

(Note that the expressions for χ_1 and χ_2 in Definition 2.2 give exact formulas for these in terms of m and κ.)

Then (3.23) implies

$$Q_1 = \chi_2 \partial_X \bar{Q}_0 + \bar{Q}_1 \quad \text{and} \quad P_1 = \chi_1 \partial_X \bar{P}_0 + \bar{P}_1, \tag{3.26}$$

where

$$\bar{Q}_1 = \bar{Q}_1(X, \tau, T) \quad \text{and} \quad \bar{P}_1 = \bar{P}_1(X, \tau, T)$$

are as yet undetermined. They do not depend on j.

3.4. Solving the $O(\epsilon^4)$ equations, (3.15). These are

$$\delta_0^+ P_2 = \frac{1}{\kappa}\partial_\tau Q_1 - \delta_1^+ P_1 - \delta_2^+ P_0 \quad \text{and} \quad \delta_0^- Q_2 = m\partial_\tau P_1 - \delta_1^- Q_1 - \delta_2^- Q_0 - \delta_0^- \left(\frac{\beta}{\kappa^2} Q_0^2\right).$$

If we are to be able to solve (3.15), Corollary 3.2 tells us we must have

$$\begin{aligned} &\frac{1}{N}\sum_{j=1}^{N}\left(\frac{1}{\kappa}\partial_\tau Q_1 - \delta_1^+ P_1 - \delta_2^+ P_0\right) = 0 \quad \text{and} \\ &\frac{1}{N}\sum_{j=1}^{N}\left(m\partial_\tau P_1 - \delta_1^- Q_1 - \delta_2^- Q_0 - \delta_0^-\left(\frac{\beta}{\kappa^2}Q_0^2\right)\right) = 0. \end{aligned} \tag{3.27}$$

First of all, note that

$$\sum_{j=1}^{N} \delta_0^{\pm} F = 0 \tag{3.28}$$

for any periodic function $F(j)$, and so $\sum_{j=1}^N \delta_0^-\left(\frac{\beta}{\kappa^2}Q_0^2\right) = 0$ in the second of these equations. Using (3.19), (3.26), the formulas for the δ_n in (3.4), and the definitions of χ_1, χ_2, $\overline{m}$, and $\breve{\kappa}$, a lengthy but routine computation shows that (3.27) is equivalent to

$$\begin{aligned} &\partial_\tau \bar{Q}_1 = \breve{\kappa}\left(\partial_X \bar{P}_1 + \left(\frac{1}{2} - \gamma\right)\partial_X^2 \bar{P}_0\right) \quad \text{and} \\ &\partial_\tau \bar{P}_1 = \frac{1}{\overline{m}}\left(\partial_X \bar{Q}_1 + \left(\tilde{\gamma} - \frac{1}{2}\right)\partial_X^2 \bar{Q}_0\right), \end{aligned} \tag{3.29}$$

where

$$\gamma := \frac{1}{N}\sum_{j=1}^{N}\frac{\breve{\kappa}}{\kappa(j)}\chi_2(j) \quad \text{and} \quad \tilde{\gamma} := \frac{1}{N}\sum_{j=1}^{N}\frac{m(j)}{\overline{m}}\chi_2(j).$$

As it happens, $\tilde{\gamma} = -\gamma$. Here is the calculation:

$$\begin{aligned} \gamma &= \frac{1}{N}\sum_{j=1}^{N}\frac{\breve{\kappa}}{\kappa}\chi_2(j) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{\breve{\kappa}}{\kappa} - 1\right)\chi_2(j) = \frac{1}{N}\sum_{j=1}^{N}\left(\delta_0^+\chi_1(j)\right)\chi_2(j) \\ &= -\frac{1}{N}\sum_{j=1}^{N}\chi_1(j)\left(\delta_0^-\chi_2(j)\right) = -\frac{1}{N}\sum_{j=1}^{N}\chi_1(j)\left(\frac{m}{\overline{m}} - 1\right) = -\frac{1}{N}\sum_{j=1}^{N}\frac{m}{\overline{m}}\chi_1(j) = -\tilde{\gamma}. \end{aligned}$$

(We used (3.24), (3.25), and summation by parts above.)

A computation using d'Alembert's solution shows that the following is a solution of (3.29), given (3.22):

$$\begin{aligned} &\bar{Q}_1(X,\tau,T) = (1/2 - \gamma)\left(\bar{A}_w(X - c\tau, T) + \bar{B}_l(X + c\tau, T)\right), \\ &\bar{P}_1(X,\tau,T) = 0. \end{aligned} \tag{3.30}$$

While this is not the general solution to (3.29), it turns out that all we require is this particular solution.

Now we compute P_2. Using (3.4), (3.19), (3.24), (3.25), (3.26), and (3.30) in the first equation in (3.15) gives

$$\delta_0^+ P_2 = \left(\frac{\breve{\kappa}}{\kappa} (\chi_2 - \gamma) - \check{S}^+ \chi_1 \right) \partial_X^2 \bar{P}_0 + \frac{1}{2} \left(\delta_0^+ \chi_1 \right) \partial_X^2 \bar{P}_0.$$

Define $\chi_3(j)$ to be the unique solution of

$$\delta_0^+ \chi_3 = \left(\frac{\breve{\kappa}}{\kappa} (\chi_2 - \gamma) - \check{S}^+ \chi_1 \right) \quad \text{and} \quad \sum_{j=1}^{N} \chi_3(j) = 0. \tag{3.31}$$

Such a solution exists by our selection of γ. Then

$$P_2 = \left(\chi_3 + \frac{1}{2} \chi_1 \right) \partial_X^2 \bar{P}_0 + \bar{P}_2, \tag{3.32}$$

where

$$\bar{P}_2 = \bar{P}_2(X, \tau, T)$$

is as yet unspecified.

Now we compute Q_2. Using (3.4), (3.19), (3.24), (3.25), and (3.26) in the second equation in (3.15) gives

$$\delta_0^- Q_2 = \left(\frac{m}{\overline{m}} (\chi_1 + \gamma) - \check{S}^- \chi_2 \right) \partial_X^2 \bar{Q}_0 - \frac{1}{2} \left(\delta_0^- \chi_2 \right) \partial_X^2 \bar{Q}_0 + \left(\delta_0^- \chi_2 \right) \partial_X \bar{Q}_1 - \delta_0^- \left(\frac{\beta}{\kappa^2} \bar{Q}_0^2 \right).$$

Define $\chi_4(j)$ to be the unique solution of

$$\delta_0^- \chi_4 = \left(\frac{m}{\overline{m}} (\chi_1 + \gamma) - \check{S}^- \chi_2 \right) \quad \text{and} \quad \sum_{j=1}^{N} \chi_4(j) = 0. \tag{3.33}$$

Such a solution exists by our selection of γ. Then

$$Q_2 = \left(\chi_4 - \frac{1}{2} \chi_2 \right) \partial_X^2 \bar{Q}_0 + \chi_2 \partial_X \bar{Q}_1 - \frac{\beta}{\kappa^2} \bar{Q}_0^2 + \bar{Q}_2, \tag{3.34}$$

where

$$\bar{Q}_2 = \bar{Q}_2(X, \tau, T)$$

is as yet unspecified.

3.5. Solving the $O(\epsilon^5)$ equations, (3.16). The equations are

$$\delta_0^+ P_3 = \frac{1}{\kappa} \partial_T Q_0 + \frac{1}{\kappa} \partial_\tau Q_2 - \delta_1^+ P_2 - \delta_2^+ P_1 - \delta_3^+ P_0,$$
$$\delta_0^- Q_3 = m \partial_T P_0 + m \partial_\tau P_2 - \delta_1^- Q_2 - \delta_2^- Q_1 - \delta_3^- Q_0 - \delta_1^- \left(\frac{\beta}{\kappa^2} Q_0^2 \right) - 2\delta_0^- \left(\frac{\beta}{\kappa^2} Q_0 Q_1 \right).$$

To be able to solve the first equation of these for P_3 requires

$$\frac{1}{N} \sum_{j=1}^{N} \left(\frac{1}{\kappa} \partial_T Q_0 + \frac{1}{\kappa} \partial_\tau Q_2 - \delta_1^+ P_2 - \delta_2^+ P_1 - \delta_3^+ P_0 \right) = 0.$$

Using (3.4), (3.19), (3.21), (3.26), (3.29), (3.30), (3.32), (3.34), and the definitions of $\overline{m}$, $\breve{\kappa}$, and γ converts this to

$$\frac{1}{\breve{\kappa}}\partial_T\bar{Q}_0 + \left(-\frac{1}{6} - \gamma^2 + \frac{1}{N}\sum_{j=1}^{N}\frac{\breve{\kappa}}{\kappa}\chi_4\right)\partial_X^3\bar{P}_0 - \left(\frac{2\breve{\kappa}}{N}\sum_{j=1}^{N}\frac{\beta}{\kappa^3}\right)\bar{Q}_0\partial_X\bar{P}_0 + \frac{1}{\breve{\kappa}}\partial_\tau\bar{Q}_2 - \partial_X\bar{P}_2 = 0. \tag{3.35}$$

To be able to solve the second equation in (3.16) for Q_3 requires

$$\frac{1}{N}\sum_{j=1}^{N}\left[m\partial_T P_0 + m\partial_\tau P_2 - \delta_1^- Q_2 - \delta_2^- Q_1 - \delta_3^- Q_0 - \delta_1^-\left(\frac{\beta}{\kappa^2}Q_0^2\right) - 2\delta_0^-\left(\frac{\beta}{\kappa^2}Q_0Q_1\right)\right] = 0.$$

Using (3.4), (3.19), (3.21), (3.26), (3.28), (3.29), (3.32), (3.34), and the definitions of $\overline{m}$, $\breve{\kappa}$, and γ converts this to

$$\overline{m}\partial_T\bar{P}_0 + \left(-\frac{1}{6} - \gamma^2 + \frac{1}{N}\sum_{j=1}^{N}\frac{m}{\overline{m}}\chi_3\right)\partial_X^3\bar{Q}_0 + \left(\frac{1}{2} - \gamma\right)\partial_X^2\bar{Q}_1 + \overline{m}\partial_\tau\bar{P}_2 - \partial_X\bar{Q}_2 = 0. \tag{3.36}$$

Given (3.22), we see that (3.35) and (3.36) can be solved for $\bar{A}_T$ and $\bar{B}_T$. To do so is a messy but essentially simple process and relies primarily on (3.22) and (3.30). The result is

$$\begin{aligned}\bar{A}_T = &-\frac{c}{24}\left(1 + 12\gamma + 12\gamma^2 - 12c_1 - 12c_2\right)\bar{A}_{www} - \frac{c}{2}c_3\bar{A}\bar{A}_w\\ &-\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{B}_{lll}\\ &-\frac{c}{2}\left(\sqrt{\overline{m}\breve{\kappa}}J_1 - J_2\right) - \frac{c}{2}c_3\bar{B}\left(\bar{A}_w - \bar{B}_l\right) - \frac{c}{2}c_3\bar{A}\bar{B}_l\end{aligned} \tag{3.37}$$

and

$$\begin{aligned}\bar{B}_T = &\frac{c}{24}\left(1 + 12\gamma + 12\gamma^2 - 12c_1 - 12c_2\right)\bar{B}_{lll} + \frac{c}{2}c_3\bar{B}\bar{B}_l\\ &+\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{A}_{www}\\ &+\frac{c}{2}\left(-\sqrt{\overline{m}\breve{\kappa}}J_1 - J_2\right) + \frac{c}{2}c_3\bar{A}\left(\bar{B}_l - \bar{A}_w\right) - \frac{c}{2}c_3\bar{B}\bar{A}_w.\end{aligned} \tag{3.38}$$

Here we have defined the following constants:

$$c_1 := \frac{1}{N}\sum_{j=1}^{N}\frac{\breve{\kappa}}{\kappa}\chi_4, \quad c_2 := \frac{1}{N}\sum_{j=1}^{N}\frac{m}{\overline{m}}\chi_3, \quad \text{and} \quad c_3 := \frac{2\breve{\kappa}}{N}\sum_{j=1}^{N}\frac{\beta}{\kappa^3}.$$

Also, we define

$$J_1 := \frac{1}{\breve{\kappa}}\partial_\tau\bar{Q}_2 - \partial_X\bar{P}_2 \quad \text{and} \quad J_2 := \overline{m}\partial_\tau\bar{P}_2 - \partial_X\bar{Q}_2.$$

We select $\bar{A}$ and $\bar{B}$ to solve the KdV equations

$$\bar{A}_T = -ca\bar{A}_{www} - cb\bar{A}\bar{A}_w \quad \text{and} \quad \bar{B}_T = ca\bar{B}_{lll} + cb\bar{B}\bar{B}_l. \tag{3.39}$$

Here

$$a := \frac{1}{24}\left(1 + 12\gamma + 12\gamma^2 - 12c_1 - 12c_2\right) \quad \text{and} \quad b := \frac{c_3}{2}.$$

Note that careful unraveling the various definitions for γ, c_1, c_2, c_3, and the various χ functions will convert these definitions for a and b into those given in Definition 2.2.

Our choices for $\bar{A}$ and $\bar{B}$ reduce (3.37) and (3.38) to

$$\begin{aligned} 0 = &-\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{B}_{lll} \\ &- \frac{c}{2}\left(\sqrt{\bar{m}}\breve{\kappa}J_1 - J_2\right) - \frac{c}{2}c_3\bar{B}\left(\bar{A}_w - \bar{B}_l\right) - \frac{c}{2}c_3\bar{A}\bar{B}_l \end{aligned} \tag{3.40}$$

and

$$\begin{aligned} 0 = &\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{A}_{www} \\ &+ \frac{c}{2}\left(-\sqrt{\bar{m}}\breve{\kappa}J_1 - J_2\right) + \frac{c}{2}c_3\bar{A}\left(\bar{B}_l - \bar{A}_w\right) - \frac{c}{2}c_3\bar{B}\bar{A}_w. \end{aligned} \tag{3.41}$$

Surprisingly, it happens that we can solve (3.40) and (3.41) for $\bar{P}_2$ and $\bar{Q}_2$ exactly in terms of $\bar{A}$ and $\bar{B}$. If we make a change of variables $\bar{Q}_2(X,\tau,T) = \mathcal{Q}(X - c\tau, X + c\tau, T)$ and $\bar{P}_2(X,\tau,T) = \mathcal{P}(X - c\tau, X + c\tau, T)$, then we have

$$\begin{aligned} 0 = &-\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{B}_{lll} \\ &- c\mathcal{Q}_l + \breve{\kappa}\mathcal{P}_l - \frac{c}{2}c_3\bar{B}\left(\bar{A}_w - \bar{B}_l\right) - \frac{c}{2}c_3\bar{A}\bar{B}_l \end{aligned}$$

and

$$\begin{aligned} 0 = &\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{A}_{www} \\ &+ c\mathcal{Q}_w + \breve{\kappa}\mathcal{P}_w + \frac{c}{2}c_3\bar{A}\left(\bar{B}_l - \bar{A}_w\right) - \frac{c}{2}c_3\bar{B}\bar{A}_w. \end{aligned}$$

Let

$$\bar{\mathcal{A}}(w,T) := \int_0^w \bar{A}(s,T)ds \quad \text{and} \quad \bar{\mathcal{B}}(l,T) := \int_0^l \bar{B}(s,T)ds.$$

That is, $\partial_w\bar{\mathcal{A}} = \bar{A}$ and $\partial_l\bar{\mathcal{B}} = \bar{B}$. Then we have

$$\begin{aligned} 0 = &-\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{B}_{lll} \\ &- c\mathcal{Q}_l + \breve{\kappa}\mathcal{P}_l - \frac{c}{2}c_3\bar{\mathcal{B}}_l\bar{A}_w + \frac{c}{4}c_3(\bar{B}^2)_l - \frac{c}{2}c_3\bar{A}\bar{B}_l \end{aligned}$$

and

$$\begin{aligned} 0 = &\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{A}_{www} \\ &+ c\mathcal{Q}_w + \breve{\kappa}\mathcal{P}_w + \frac{c}{2}c_3\bar{\mathcal{A}}_w\bar{B}_l - \frac{c}{4}c_3\left(\bar{A}^2\right)_w - \frac{c}{2}c_3\bar{B}\bar{A}_w. \end{aligned}$$

Then we can antidifferentiate these equations with respect to l and w, respectively, to get

$$\begin{aligned} 0 = &-\frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{B}_{ll} \\ &- c\mathcal{Q} + \breve{\kappa}\mathcal{P} - \frac{c}{2}c_3\bar{\mathcal{B}}\bar{A}_w + \frac{c}{4}c_3(\bar{B}^2) - \frac{c}{2}c_3\bar{A}\bar{B} \end{aligned}$$

and

$$0 = \frac{c}{24}\left(-3 + 12\gamma - 12\gamma^2 + 12c_1 - 12c_2\right)\bar{A}_{ww} + c\mathcal{Q} + \breve{\kappa}\mathcal{P} + \frac{c}{2}c_3\bar{\mathcal{A}}\bar{B}_l - \frac{c}{4}c_3\left(\bar{A}^2\right) - \frac{c}{2}c_3\bar{B}\bar{A}.$$

Note that we have taken the constants of integration to be zero. We solve these for $\mathcal{Q}$ and $\mathcal{P}$ (and therefore for $\bar{Q}_2$ and $\bar{P}_2$) to get

$$\begin{aligned}\bar{Q}_2(X,\tau,T) &= g_1\bar{A}_{ww} + g_2\bar{B}_{ll} + g_3\bar{\mathcal{A}}\bar{B}_l + g_4\bar{\mathcal{B}}\bar{A}_w + g_5\bar{A}^2 + g_6\bar{B}^2 + g_7\bar{A}\bar{B},\\ \bar{P}_2(X,\tau,T) &= h_1\bar{A}_{ww} + h_2\bar{B}_{ll} + h_3\bar{\mathcal{A}}\bar{B}_l + h_4\bar{\mathcal{B}}\bar{A}_w + h_5\bar{A}^2 + h_6\bar{B}^2 + h_7\bar{A}\bar{B}\end{aligned} \tag{3.42}$$

for some collection of constants $g_1, \ldots, h_7$.

REMARK 3. *In previous works (for instance,* [21], [23]*), equations analogous to* (3.40) *and* (3.41) *are viewed as evolution equations for $\bar{P}_2$ and $\bar{Q}_2$. A little rearranging shows that these are equivalent to a forced wave equation, where a typical driving term would be something like $\bar{B}(l,T)\bar{A}(w,T)$. The algebraic decay conditions on the initial data in those articles are put in place precisely so that these driving terms do not cause their terms analogous to $\bar{P}_2$ and $\bar{Q}_2$ to grow secularly with t and, consequently, ruin the error estimates. In short, if the initial data is decaying at infinity, then $\bar{A}$ and $\bar{B}$ do as well. Since $\bar{A}$ moves to the right and $\bar{B}$ moves to the left, their product takes place mostly through their tails and is thus small. By showing that we can solve these for $\bar{P}_2$ and $\bar{Q}_2$ explicitly allows us to replace the algebraic decay condition with the weaker condition* (2.2)*.*

Our selection of $\bar{A}$, $\bar{B}$, Q_2, and P_2 allows us to use Corollary 3.2 to solve (3.16) for Q_3 and P_3. Though possible, it is not necessary for us to have exact formulas for these. Instead we presume only that

$$\sum_{j=1}^{N} P_3(j,X,\tau,T) = \sum_{j=1}^{N} Q_3(j,X,\tau,T) = 0 \tag{3.43}$$

for all X, τ, and T.

4. The size of the approximation and its residual. Now that we have explicitly determined $Q_0, Q_1, Q_2, P_0, P_1, P_2$ in terms of $\bar{A}$ and $\bar{B}$, and implicitly defined Q_3 and P_3 with (3.16) and (3.43), we can rigorously estimate their sizes and the size of the residual in terms of $\bar{A}$ and $\bar{B}$. Specifically we will provide estimates which control $\tilde{r}_\epsilon(t)$ and $\tilde{p}_\epsilon(t)$ and the residual for $|t| \leq T_0\epsilon^{-3}$ in terms of T_0 and

$$\bar{A}_0(w) := \bar{A}(w,0), \quad \bar{B}_0(l) := \bar{B}(l,0), \quad \bar{\mathcal{A}}_0(w) := \bar{\mathcal{A}}(w,0), \quad \text{and} \quad \bar{\mathcal{B}}_0(l) := \bar{\mathcal{B}}(l,0). \tag{4.1}$$

We are going to pay more attention to regularity issues here than is typically done when justifying modulation equations. Note that in Theorem 2.1 the initial conditions are sampled from functions in H^5. As we shall see, to estimate the residual requires that $\bar{A}$ and $\bar{B}$ be in H^7. We discuss how to close this gap in the next section. To help us organize the frequently tedious calculations we carry out here, we introduce the following notation.

DEFINITION 4.1. *We write*

$$Z \leq K_5 W$$

if and only if

$$Z \le W\Pi\left(T_0 + \|\bar{A}_0\|_{H^5} + \|\bar{B}_0\|_{H^5} + \|\bar{\mathcal{A}}_0\|_{L^\infty} + \|\bar{\mathcal{B}}_0\|_{L^\infty}\right)$$

for some function $\Pi(h)$ which is continuous, nondecreasing, has $\Pi(0) = 0$, and is determined entirely by m and $\mathcal{V}$. Note that the subscript "5" is put in place to reinforce that this notation involves the H^5 regularity of $\bar{A}_0$ and $\bar{B}_0$.

The estimates we require are contained in the following.

Proposition 4.2 (rigorous estimates on the approximation). *Fix $T_0 > 0$ and suppose that $\bar{A}(w,T)$ and $\bar{B}(l,T)$ are solutions of the KdV equations* (3.39) *with $a \neq 0$. Let $\bar{A}_0, \bar{B}_0, \bar{\mathcal{A}}_0, \bar{\mathcal{B}}_0$ be as in* (4.1)*. Construct $R_0, \dots, R_3$, $P_0, \dots, P_3$, $\tilde{r}_\epsilon$, and $\tilde{p}_\epsilon$ from $\bar{A}$ and $\bar{B}$ as described above in section* 3*. Define $\mathrm{Res}_1(\tilde{r}_\epsilon, \tilde{p}_\epsilon)$ and $\mathrm{Res}_2(\tilde{r}_\epsilon, \tilde{p}_\epsilon)$ as in* (3.1)*. Let*[2]

$$\breve{r}_\epsilon(j,t) := \sum_{n=1}^{3} \epsilon^{n+2} R_n(j, \epsilon j, \epsilon t, \epsilon^3 t) \quad \textit{and} \quad \breve{p}_\epsilon(j,t) := \sum_{n=1}^{3} \epsilon^{n+2} P_n(j, \epsilon j, \epsilon t, \epsilon^3 t).$$

Then

$$\sup_{|t| \le T_0 \epsilon^{-3}} \|(\tilde{r}_\epsilon(t), \tilde{p}_\epsilon(t))\|_{\ell^2 \times \ell^2} \le K_5 \epsilon^{3/2}, \tag{4.2}$$

$$\sup_{|t| \le T_0 \epsilon^{-3}} \|(\tilde{r}_\epsilon(t), \tilde{p}_\epsilon(t))\|_{\ell^\infty \times \ell^\infty} \le K_5 \epsilon^{4}, \tag{4.3}$$

$$\sup_{|t| \le T_0 \epsilon^{-3}} \|(\breve{r}_\epsilon(t), \breve{p}_\epsilon(t))\|_{\ell^2 \times \ell^2} \le K_5 \epsilon^{5/2}, \tag{4.4}$$

$$\begin{aligned} &\sup_{|t| \le T_0 \epsilon^{-3}} \|(\partial_t \tilde{r}_\epsilon(t), \partial_t \tilde{p}_\epsilon(t))\|_{\ell^\infty \times \ell^\infty} \\ &\qquad \le K_5 \epsilon^3 + K_5 \epsilon^7 \left(\|\bar{A}_0\|_{H^6} + \|\bar{B}_0\|_{H^6}\right) + K_5 \epsilon^8 \left(\|\bar{A}_0\|_{H^7} + \|\bar{B}_0\|_{H^7}\right), \end{aligned} \tag{4.5}$$

and

$$\begin{aligned} &\sup_{|t| \le T_0 \epsilon^{-3}} \|(\mathrm{Res}_1(\tilde{r}_\epsilon(t), \tilde{p}_\epsilon(t)), \mathrm{Res}_2(\tilde{r}_\epsilon(t), \tilde{p}_\epsilon(t)))\|_{\ell^2 \times \ell^2} \\ &\quad \le K_5 \epsilon^{11/2} + K_5 \epsilon^{13/2} \left(\|\bar{A}_0\|_{H^6} + \|\bar{B}_0\|_{H^6}\right) + K_5 \epsilon^{15/2} \left(\|\bar{A}_0\|_{H^7} + \|\bar{B}_0\|_{H^7}\right). \end{aligned} \tag{4.6}$$

The main tools for proving the proposition are rigorous estimates on the error functions $E_M^\pm$ and well-posedness results for KdV equations.

4.1. Long wave approximations. The main estimates we need are collected in the following.

Lemma 4.3 (error estimates for long wave approximations). *Suppose that*

$$u(j) = U(j, \epsilon j),$$

[2] Note that $\breve{r}_\epsilon$ and $\breve{p}_\epsilon$ are simply the $O(\epsilon^3)$ and higher terms of $\tilde{r}_\epsilon$ and $\tilde{p}_\epsilon$.

where $U(j+N,X)=U(j,X)$ *for all* $j\in\mathbf{Z}$ *and* $X\in\mathbf{R}$. *Then*

$$\|u\|_{\ell^\infty}\le C\sup_{j\in\mathbf{Z}}\|U(j,\cdot)\|_{L^\infty}, \tag{4.7}$$

$$\|u\|_{\ell^2}\le C\epsilon^{-1/2}\sup_{j\in\mathbf{Z}}\|U(j,\cdot)\|_{H^1}, \tag{4.8}$$

and

$$\|E_M^{\pm}u\|_{\ell^2}\le C\epsilon^{M+1/2}\sup_{j\in\mathbf{Z}}\|U(j,\cdot)\|_{H^{M+1}}. \tag{4.9}$$

The constant $C>0$ *depends only on* N *and* M.

Note that these estimates are sharp with respect to the powers of ϵ. The loss of a half power of ϵ appearing in the ℓ^2-based estimates over the formally expected error is due to the "long wave scaling," $X=\epsilon j$. Proving these results requires the following.

LEMMA 4.4. *Suppose that* $f\in H^1$. *Select* $x_j\in[j,j+1]$ *for all* $j\in\mathbf{Z}$. *Then*

$$\sum_{j\in\mathbf{Z}}f^2(x_j)\le 2\|f\|_{H^1}^2.$$

Proof. Since $f\in H^1$ by the Sobolev embedding theorem we know that $f(x)$ is continuous. Since $f(x)$ is continuous, so is $f^2(x)$. Thus, for each $j\in\mathbf{Z}$, there exists $\underline{x}_j\in[j,j+1]$ with $f^2(\underline{x}_j)=\min_{x\in[j,j+1]}f^2(x)$. Clearly, then, $f^2(\underline{x}_j)\le\int_j^{j+1}f^2(x)dx$. Consequently

$$\sum_{j\in\mathbf{Z}}f^2(\underline{x}_j)\le\|f\|_{L^2}^2. \tag{4.10}$$

Then we use the fundamental theorem of calculus:

$$f^2(x_j)=f^2(\underline{x}_j)+\int_{\underline{x}_j}^{x_j}\partial_x(f^2(x))dx.$$

The chain rule and the Cauchy–Schwarz inequality give

$$\int_{\underline{x}_j}^{x_j}\partial_x(f^2(x))dx=2\int_{\underline{x}_j}^{x_j}f(x)f_x(x)dx\le 2\sqrt{\left|\int_{\underline{x}_j}^{x_j}f^2(x)dx\right|}\sqrt{\left|\int_{\underline{x}_j}^{x_j}f_x^2(x)dx\right|}.$$

Since $\underline{x}_j,x_j\in[j,j+1]$ we get

$$\int_{\underline{x}_j}^{x_j}\partial_x(f^2(x))dx\le 2\sqrt{\int_j^{j+1}f^2(x)dx}\sqrt{\int_j^{j+1}f_x^2(x)dx}.$$

Using $ab\le a^2/2+b^2/2$ we have

$$\int_{\underline{x}_j}^{x_j}\partial_x(f^2(x))dx\le\int_j^{j+1}f^2(x)dx+\int_j^{j+1}f_x^2(x)dx.$$

Using this and (4.10) gives

$$\sum_{j\in\mathbf{Z}}f^2(x_j)\le\sum_{j\in\mathbf{Z}}f^2(\underline{x}_j)+\sum_{j\in\mathbf{Z}}\int_j^{j+1}f^2(x)dx+\sum_{j\in\mathbf{Z}}\int_j^{j+1}f_x^2(x)dx\le 2\|f\|_{L^2}^2+\|f_x\|_{L^2}^2. \qquad\square$$

Proof of Lemma 4.3. The first estimate is trivial. The second estimate follows from Lemma 4.4 and naive estimates. For the third estimate, we prove the "+" case.

Suppose that $f \in H^{M+1}$. Since $f^{(M+1)} \in L^2$, this implies $f^{(M+1)} \in L^1_{loc}$. Taylor's theorem with remainder tells us that

$$f(x) - \sum_{n=0}^{M} \frac{1}{n!} f^{(n)}(y)(x-y)^n = \int_x^y \frac{1}{M!} f^{(M+1)}(s)(x-s)^M ds$$

for any $-\infty < x \leq y < \infty$. A naive estimate gives

$$\left| f(x) - \sum_{n=0}^{M} \frac{1}{n!} f^{(n)}(y)(x-y)^n \right| \leq \frac{1}{M!} |x-y|^M \int_y^x \left| f^{(M+1)}(s) \right| ds.$$

Using the Cauchy–Schwarz inequality on the last term gives

$$\left| f(x) - \sum_{n=0}^{M} \frac{1}{n!} f^{(n)}(y)(x-y)^n \right| \leq \frac{1}{M!} |x-y|^{M+1/2} \| f^{(M+1)} \|_{L^2([x,y])}. \tag{4.11}$$

Now consider $E^+ u$. By definition of E^+ and the δ_n we have

$$\begin{aligned}(E_M^+ u)(j) &= (U(j+1, \epsilon j + \epsilon) - U(j, \epsilon j)) \\ &\quad - (U(j+1, \epsilon j) - U(j, \epsilon j)) - \sum_{n=1}^{M} \frac{\epsilon^n}{n!} \partial_X^n U(j+1, \epsilon j) \\ &= U(j+1, \epsilon j + \epsilon) - \sum_{n=0}^{M} \frac{\epsilon^n}{n!} \partial_X^n U(j+1, \epsilon j).\end{aligned}$$

Using (4.11) with $f(\cdot) = U(j+1, \cdot)$, $x = \epsilon j$, and $y = \epsilon j + \epsilon$, we see

$$\left| (E_M^+ u)(j) \right| \leq \frac{1}{M!} \epsilon^{M+1/2} \| \partial_X^{M+1} U(j, \cdot) \|_{L^2([\epsilon j, \epsilon j + \epsilon])}.$$

Squaring and summing this over $j \in \mathbf{Z}$ gives

$$\| E_M^+ u \|_{\ell^2}^2 \leq \frac{1}{(M!)^2} \epsilon^{2M+1} \sum_{j \in \mathbf{Z}} \| \partial_X^{M+1} U(j, \cdot) \|_{L^2([\epsilon j, \epsilon j + \epsilon])}^2.$$

Adding positive terms on the right-hand side can only make things larger, and thus we have

$$\| E_M^+ u \|_{\ell^2}^2 \leq \frac{1}{(M!)^2} \epsilon^{2M+1} \sum_{k=1}^{N} \sum_{j \in \mathbf{Z}} \| \partial_X^{M+1} U(k, \cdot) \|_{L^2([\epsilon j, \epsilon j + 1])}^2.$$

Of course,

$$\sum_{j \in \mathbf{Z}} \| \partial_X^{M+1} U(k, \cdot) \|_{L^2([\epsilon j, \epsilon j + \epsilon])}^2 = \| \partial_X^{M+1} U(k, \cdot) \|_{L^2}^2 \leq \sup_k \| U(k, \cdot) \|_{H^{M+1}}^2.$$

Thus

$$\| E_M^+ u \|_{\ell^2}^2 \leq \frac{N}{(M!)^2} \epsilon^{2M+1} \sup_k \| U(k, \cdot) \|_{H^{M+1}}^2$$

and the proof is complete. ∎

4.2. Existence and estimates for solutions of KdV equations. The following well-known results concerning solutions of the KdV equation will be used (see [1] and [17]).

THEOREM 4.5 (global existence for solutions of KdV equations). *Suppose that* $U_0 \in H^s$ *for* $s \geq 2$, $s \in \mathbf{Z}$ *and consider the partial differential equation*

$$U_T = aU_{yyy} + b(U^2)_y, \tag{4.12}$$

where $a \neq 0$ *and* $y \in \mathbf{R}$. *Then there exists unique*

$$U \in C(\mathbf{R}; H^s)$$

with

$$\partial_T^n U \in C(\mathbf{R}; H^{s-3n})$$

for all $n \in \mathbf{N}$ *that has* $U(y,0) = U_0$ *and which solves* (4.12) *for all* $T \in \mathbf{R}$.

Additionally, one has, for all $0 \leq k \leq s$, $k \in \mathbf{Z}$,

$$\sup_{T\in\mathbf{R}} \|U(T)\|_{H^k} \leq \|U_0\|_{H^k} + C\|U_0\|_{H^1}\|U_0\|_{H^{k-1}} + \Pi(\|U_0\|_{H^{k-2}}). \tag{4.13}$$

The constant $C > 0$ *depends only on* a, b, *and* k. *Here* Π *is a nondecreasing continuous function with* $\Pi(0) = 0$*; it is wholly determined by* a, b, *and* k.

We also need to control $\bar{\mathcal{A}}$ and $\bar{\mathcal{B}}$, the antiderivatives. Lemma 4.6 below does so.

LEMMA 4.6 (estimates on antiderivatives of solutions of KdV equations). *Suppose that* $U(y,T)$ *is the unique global-in-time function (whose existence is asserted by Theorem* 4.5*) which solves* (4.12) *with* $a \neq 0$ *and for which* $U_0(y) \in H^4$. *Assume*

$$\left|\int_{\mathbf{R}} U_0(s)ds\right| < \infty. \tag{4.14}$$

Let

$$\mathcal{U}(y,T) := \int_0^y U(s,T)ds.$$

Then for all $T_0 > 0$

$$\sup_{|T|<T_0} \|\mathcal{U}(T)\|_{L^\infty} \leq \|\mathcal{U}_0\|_{L^\infty} + T_0\Pi(\|U_0\|_{H^4}).$$

Here Π *is a nondecreasing continuous function with* $\Pi(0) = 0$*; it is wholly determined by* a *and* b. *Of course,* $\mathcal{U}_0(y) := \mathcal{U}(y,0)$.

Proof. Without loss of generality, $a = 1$, due to well-known scaling invariances of the KdV equation. Since $U_0 \in H^4$, Sobolev embedding implies it is in C^3. Thus $U(X,0)$, which is an antiderivative of $U_0(x)$, is in C^4. Condition (4.14) implies that $\lim_{X\to\pm\infty} \mathcal{U}(X,0) < \infty$ and so $\mathcal{U}(X,0) \in L^\infty$.

Next, Theorem 4.5 implies that $U(y,T)$ is in H^4 for all T and, again, Sobolev embedding tells us then that $U \in C^3$ and $W^{3,\infty}$; that is to say, $U(y,T)$ is a classical solution. If we integrate the KdV equation in time, we get

$$U(y,T) = U_0(y) + \int_0^T \left(U_{yyy}(y,t') + b(U^2)_y(y,t')\right) dt'.$$

Next we integrate from 0 in y to get

$$\mathcal{U}(y,T) = \mathcal{U}(y,0) + \int_0^y \int_0^T \left(U_{yyy}(y',t') + b(U^2)_y(y',t')\right) dt' dy'.$$

Since U_{yyy} and U are continuous functions and we are integrating over a compact set, we are free to exchange the order of integration:

$$\mathcal{U}(y,T) = \mathcal{U}(y,0) + \int_0^T \int_0^y \left(U_{yyy}(y',t') + b(U^2)_y(y',t')\right) dy' dt'.$$

Now we use the fundamental theorem of calculus:

$$\mathcal{U}(y,T) = \mathcal{U}(y,0) + \int_0^T \left(U_{yy}(y,t') - U_{yy}(0,t') + bU^2(y,t') - bU^2(0,t')\right) dt'.$$

Thus, if $|T| \leq T_0$,

$$\|\mathcal{U}(T)\|_{L^\infty} \leq \|\mathcal{U}(0)\|_{L^\infty} + 2T_0 \sup_{|t|\leq T_0} \left(\|U_{yy}(t)\|_{L^\infty} + b\|U^2(t)\|_{L^\infty}\right).$$

The Sobolev embedding theorem and Theorem 4.5 then imply

$$\|\mathcal{U}(T)\|_{L^\infty} \leq \|\mathcal{U}(0)\|_{L^\infty} + 2T_0 \sup_{|t|\leq T_0} \left(\|U\|_{H^4} + b\|U\|_{H^4}^2\right) \leq \|\mathcal{U}(0)\|_{L^\infty} + T_0\Pi(\|U_0\|_{H^4}).$$

This estimate finishes the proof. ∎

4.3. The proof of Proposition 4.2. Proving estimates (4.2), (4.3), and (4.4) is "easy" because of the explicit formulas for Q_n and P_n in terms of $\bar{A}$, $\bar{B}$, and their antiderivatives. More or less all we do is search through the formulas (3.16), (3.19), (3.22), (3.26), (3.30), (3.32), (3.34), (3.42), (3.43), count derivatives, and apply estimates from Corollary 3.2, Lemma 4.3, Theorem 4.5, and Lemma 4.6 where appropriate. We omit the particulars; most of the critical ideas will be presented below when we treat (4.6).

Proving (4.5) is much the same. The extra power of ϵ in this estimate over that in (4.3) is due to the fact that all of our functions Q_n and P_n depend on t only through $\tau = \epsilon t$ and $T = \epsilon^3 t$, which is to say that the chain rule automatically produces at least one extra power of ϵ. Application of ∂_t to $(\tilde{r}_\epsilon, \tilde{p}_\epsilon)$ produces terms like $\partial_\tau Q_n$ and $\partial_T Q_n$. These can always be eliminated using the wave equations (3.21), (3.29) and the KdV equations (3.39). After that is complete, we count derivatives and apply Theorem 4.5 and Lemma 4.6. Doing so is straightforward and uninteresting, so we omit it.

Of the estimates in (4.6), the one for Res_2 is the more complicated, and so we present details for it. Recall that our selection of the Q and P functions was made so that Res_2 was given by (3.12), which we recopy here, recalling that $Q_n := \kappa R_n$:

$$\begin{aligned}\mathrm{Res}_2(\tilde{r}_\epsilon, \tilde{p}_\epsilon) = {} & -\epsilon^6 \partial_\tau P_3 - \sum_{n=1}^{3} \epsilon^{n+5} \partial_T P_n + \sum_{n=0}^{3} \epsilon^{n+2} E_{3-n}^- Q_n \\ & + \epsilon^4 \frac{1}{m} E_1^- \left(\frac{\beta}{\kappa^2} Q_0^2\right) + \epsilon^5 \frac{2}{m} E_0^- \left(\frac{\beta}{\kappa^2} Q_0 Q_1\right) + N_1(\tilde{r}_\epsilon) + N_2(\tilde{r}_\epsilon).\end{aligned}$$

If we apply (4.8) and (4.9) to this, we get, with some rearrangement,

$$
\begin{aligned}
&\|\mathrm{Res}_2(\tilde{r}_\epsilon(t), \tilde{p}_\epsilon(t))\|_{\ell^2} \\
&\leq C\epsilon^{11/2} \sup_{j\in\mathbf{Z}} \left(\|Q_0(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^4} + \|Q_1(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^3}\right) \\
&\quad + C\epsilon^{11/2} \sup_{j\in\mathbf{Z}} \left(\|Q_0(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^1}^2 + \|Q_0(j,\cdot,\epsilon t,\epsilon^3 t)Q_1(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^1}\right) \\
&\quad + C\epsilon^{11/2} \sup_{j\in\mathbf{Z}} \left(\|Q_2(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^2} + \|\partial_T P_1(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^1}\right) \\
&\quad + C\epsilon^{13/2} \sup_{j\in\mathbf{Z}} \|\partial_T P_2(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^1} \\
&\quad + C\epsilon^{11/2} \sup_{j\in\mathbf{Z}} \left(\|\partial_\tau P_3(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^1} + \|Q_3(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^1}\right) \\
&\quad + C\epsilon^{15/2} \sup_{j\in\mathbf{Z}} \|\partial_T P_3(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^1} \\
&\quad + \|N_1(\tilde{r}_\epsilon)\|_{\ell^2} + \|N_2(\tilde{r}_\epsilon)\|_{\ell^2}.
\end{aligned}
\tag{4.15}
$$

The most complicated of all of these terms is (unsurprisingly) the one involving $\partial_T P_3$. We will present many of the details for this term and also those for N_1, as it (along with N_2) is handled a bit differently than the rest. As a byproduct, we will see all the tricks for handling every other term along the way.

Estimating $\boldsymbol{\partial_T P_3}$. Differentiation with respect to T of the "P_3" equations (3.16) and (3.43) shows that $\partial_T P_3$ satisfies

$$\delta_0^+ \partial_T P_3 = \frac{1}{\kappa}\partial_T^2 Q_0 + \frac{1}{\kappa}\partial_\tau \partial_T Q_2 - \delta_1^+ \partial_T P_2 - \delta_2^+ \partial_T P_1 - \delta_3^+ \partial_T P_0$$

subject to

$$\sum_{j=1}^{N} \partial_T P_3(j, X, \tau, T) = 0.$$

The estimate in Corollary 3.2 therefore gives

$$
\begin{aligned}
&\sup_{j\in\mathbf{Z}} \|\partial_T P_3(j,\cdot,\tau,T)\|_{H^1} \leq C \sup_{j\in\mathbf{Z}} \left(\|\partial_T^2 Q_0(j,\cdot,\tau,T)\|_{H^1} + \|\partial_\tau\partial_T Q_2(j,\cdot,\tau,T)\|_{H^1}\right. \\
&\qquad \left. + \|\delta_1^+\partial_T P_2(j,\cdot,\tau,T)\|_{H^1} + \|\delta_2^+\partial_T P_1(j,\cdot,\tau,T)\|_{H^1} + \|\delta_3^+\partial_T P_0(j,\cdot,\tau,T)\|_{H^1}\right).
\end{aligned}
\tag{4.16}
$$

Of these five terms, the most complicated is the one involving $\delta_1^+\partial_T P_2$. We present the details for this term, as the others are handled with the same techniques and are no worse in terms of regularity or difficulty.

Given (3.4), we have

$$\sup_{j\in\mathbf{Z}} \|\delta_1^+\partial_T P_2(j,\cdot,\tau,T)\|_{H^1} \leq \sup_{j\in\mathbf{Z}} \|\partial_T P_2(j,\cdot,\tau,T)\|_{H^2}.$$

P_2 is given explicitly by (3.32) and (3.42). Differentiation of these formulas with respect to T, the triangle inequality, and the fact that H^2 is an algebra gives us the following crude estimate:

$$
\begin{aligned}
(4.17)\quad \sup_{j\in\mathbf{Z}} \|\partial_T P_2(j,\cdot,\tau,T)\|_{H^2} \le{}& C\|\partial_T \bar{A}_{ww}(T)\|_{H^2} + C\|\partial_T \bar{B}_{ll}(T)\|_{H^2}\\
&+ C(\|\bar{A}(T)\|_{H^2}\|\bar{A}_T(T)\|_{H^2} + \|\bar{A}(T)\|_{H^2}\|\bar{B}_T(T)\|_{H^2})\\
&+ C(\|\bar{B}(T)\|_{H^2}\|\bar{A}_T(T)\|_{H^2} + \|\bar{B}(T)\|_{H^2}\|\bar{B}_T(T)\|_{H^2})\\
&+ C\|\partial_T(\bar{\mathcal{A}}(\cdot - c\tau, T)\bar{B}_l(\cdot + c\tau, T))\|_{H^2}\\
&+ C\|\partial_T(\bar{\mathcal{B}}(\cdot + c\tau, T)\bar{A}_w(\cdot - c\tau, T))\|_{H^2}.
\end{aligned}
$$

To handle the first line on the right-hand side, we use the fact that $\bar{A}$ and $\bar{B}$ are assumed to solve (3.39). Thus

$$\partial_T \bar{A}_{ww} = -ca\bar{A}_{5w} - cb(\bar{A}\bar{A}_w)_{ww}.$$

And so we have, for all T,

$$\|\partial_T \bar{A}_{ww}(T)\|_{H^2} \le C\|\bar{A}(T)\|_{H^7} + C\|\bar{A}(T)\|^2_{H^5}.$$

Using the estimate (4.13) in Theorem 4.5 gives

$$
\begin{aligned}
\sup_{|T|\le T_0} \|\partial_T \bar{A}_{ww}(T)\|_{H^2} \le{}& \|\bar{A}_0\|_{H^7} + C\|\bar{A}_0\|_{H^1}\|\bar{A}_0\|_{H^6} + \Pi(\|\bar{A}_0\|_{H^5})\\
&+ C\left(\|\bar{A}_0\|_{H^5} + C\|\bar{A}_0\|_{H^1}\|\bar{A}_0\|_{H^4} + \Pi(\|\bar{A}_0\|_{H^3})\right)^2.
\end{aligned}
$$

Using the "K_5" definition (Definition 4.1), we have

$$\sup_{|T|\le T_0} \|\partial_T \bar{A}_{ww}(T)\|_{H^1} \le K_5\|\bar{A}_0\|_{H^7} + K_5.$$

The other three terms in the first three lines of (4.17) are handled using exactly these same ideas.

Dealing with the terms in the final two lines of (4.17) is a bit different. Consider the one on the fourth line which requires control of $\partial_T(\bar{\mathcal{A}}\bar{B}_l)$. The product rule and triangle inequality give

$$
\begin{aligned}
&\|\partial_T(\bar{\mathcal{A}}(\cdot - c\tau, T)\bar{B}_l(\cdot + c\tau, T))\|_{H^2}\\
&\qquad\le \|\partial_T\bar{\mathcal{A}}(\cdot - c\tau, T)\bar{B}_l(\cdot + c\tau, T)\|_{H^2} + \|\bar{\mathcal{A}}(\cdot - c\tau, T)\partial_T\bar{B}_l(\cdot + c\tau, T)\|_{H^2}.
\end{aligned}
$$

Then we use the estimate $\|fg\|_{H^2} \le C\|f\|_{W^{2,\infty}}\|g\|_{H^2}$ and the fact that $W^{k,p}$ norms are shift invariant to get

$$
\begin{aligned}
(4.18)\quad &\|\partial_T(\bar{\mathcal{A}}(\cdot - c\tau, T)\bar{B}_l(\cdot + c\tau, T))\|_{H^2}\\
&\qquad\le C\|\partial_T\bar{\mathcal{A}}(T)\|_{W^{2,\infty}}\|\bar{B}_l(T)\|_{H^2} + C\|\bar{\mathcal{A}}(T)\|_{W^{2,\infty}}\|\partial_T\bar{B}_l(T)\|_{H^2}.
\end{aligned}
$$

Theorem 4.5 implies immediately that $\sup_{|T|\le T_0}\|B_l(T)\|_{H^2} \le K_5$. Using the same sort of steps as were used to estimate $\partial_T A_{ww}$ above, we have

$$\sup_{|T|\le T_0} \|\partial_T B_l(T)\|_{H^2} \le K_5\|\bar{B}_0\|_{H^6} + K_5.$$

Controlling the terms in (4.18) with $\bar{\mathcal{A}}$ proceeds as follows. First, since $\bar{\mathcal{A}}$ is an antiderivative of $\bar{A}$, we have

$$\|\bar{\mathcal{A}}\|_{W^{2,\infty}} = \|\bar{\mathcal{A}}\|_{L^\infty} + \|\bar{A}\|_{W^{1,\infty}} \le \|\bar{\mathcal{A}}\|_{L^\infty} + C\|\bar{A}\|_{H^2}.$$

Here, we used Morrey's inequality, $\|f\|_{W^{s,\infty}} \le C\|f\|_{H^{s+1}}$. Then Lemma 4.6 and Theorem 4.5 give us

$$\sup_{|T|\le T_0} \|\bar{\mathcal{A}}(T)\|_{W^{1,\infty}} \le K_5.$$

Second, since $\bar{A}$ solves (3.39), it follows from the fundamental theorem of calculus that

$$\partial_T \bar{\mathcal{A}}(w,T) = -ca\bar{A}_{ww}(w,T) + ca\bar{A}_{ww}(0,T) - cb\bar{A}^2(w,T) + cb\bar{A}^2(0,T).$$

Morrey's inequality implies

$$\sup_{|T|\le T_0} \|\partial_T \bar{\mathcal{A}}(T)\|_{W^{2,\infty}} \le K_5 \sup_{|T|\le T_0} \|\bar{A}(T)\|_{H^5} + C \sup_{|T|\le T_0} \|\bar{A}(T)\|_{H^3}^2 \le K_5.$$

And so (4.18) becomes

$$\sup_{|T|\le T_0} \|\partial_T(\bar{\mathcal{A}}(\cdot - c\tau, T)\bar{B}_l(\cdot + c\tau, T))\|_{H^2} \le K_5\|\bar{B}_0\|_{H^6} + K_5.$$

With this we can establish

$$\sup_{|t|\le T_0\epsilon^{-3}} \sup_{j\in\mathbf{Z}} \|\partial_T P_2(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^2} \le K_5\left(\|\bar{A}_0\|_{H^7} + \|\bar{B}_0\|_{H^7}\right) + K_5,$$

which in turn implies

$$\sup_{|t|\le T_0\epsilon^{-3}} \sup_{j\in\mathbf{Z}} \|\partial_T P_3(j,\cdot,\epsilon t,\epsilon^3 t)\|_{H^1} \le C\left(\|\bar{A}_0\|_{H^7} + \|\bar{B}_0\|_{H^7}\right) + K_5.$$

Estimating $\boldsymbol{N_1(\tilde{r}_\epsilon)}$. Recall that $N_1(\tilde{r}_\epsilon) = \frac{1}{m}\delta^-\left\{\mathcal{V}'(\tilde{r}_\epsilon) - \kappa\tilde{r}_\epsilon - \beta\tilde{r}_\epsilon^2\right\}$. Thus, since δ^- is a bounded operator on ℓ^2,

$$\|N_1(\tilde{r}_\epsilon)\|_{\ell^2} \le C\|\tilde{N}_1\|_{\ell^2}$$

with

$$\tilde{N}_1 := \mathcal{V}'(\tilde{r}_\epsilon) - \kappa\tilde{r}_\epsilon - \beta\tilde{r}_\epsilon^2.$$

Since $\mathcal{V}(\rho)$ is a smooth function, Taylor's theorem implies

$$\tilde{N}_1 = \frac{1}{2}\int_0^{\tilde{r}_\epsilon} \mathcal{V}^{(4)}(s)\left(\tilde{r}_\epsilon - s\right)^2 ds.$$

Thus crude estimates give

$$|\tilde{N}_1| \le C|\tilde{r}_\epsilon|^3 G(|\tilde{r}_\epsilon|),$$

where

$$G(\rho) := \max_{|s|\le|\rho|} \left|\mathcal{V}^{(4)}(s)\right|.$$

Clearly G is nondecreasing. Moreover, since $\mathcal{V}$ is smooth, G is continuous. Therefore

$$|\tilde{N}_1| \leq C\|\tilde{r}_\epsilon\|_{\ell^\infty}^2 G(\|\tilde{r}\|_{\ell^\infty})\,|\tilde{r}_\epsilon|\,.$$

And so

$$\|N_1(\tilde{r}_\epsilon)\|_{\ell^2} \leq C\|\tilde{r}_\epsilon\|_{\ell^\infty}^2 G(\|\tilde{r}\|_{\ell^\infty})\|\tilde{r}_\epsilon\|_{\ell^2}.$$

The estimates in (4.3) and (4.2) thus give

$$\|N_1(\tilde{r}_\epsilon)\|_{\ell^2} \leq K_5\epsilon^{11/2}.$$

This completes our proof of Proposition 4.2. □

5. Error estimates and the proof of Theorem 2.1. We are now in a position to prove the main result. We quickly recapitulate the hypotheses and conclusions. Fix $T_0 > 0$ and $\phi, \psi \in H^5$. These will be the functions from which we sample the initial conditions for (2.1). Let $\Phi(X) := \int_0^X \phi(y)dy$ and $\Psi(X) := \int_0^X \psi(y)dy$. These are assumed to be in L^∞.

Definition 5.1. *We write*

$$Z \leq C_5 W$$

if and only if

$$Z \leq W\Pi\left(T_0 + \|\phi\|_{H^5} + \|\psi\|_{H^5} + \|\Phi\|_{L^\infty} + \|\Psi\|_{L^\infty}\right)$$

for some function $\Pi(h)$ *which is continuous, nondecreasing, has* $\Pi(0) = 0$, *and is determined entirely by* m *and* $\mathcal{V}$.

The initial data for (2.1) are taken as described in (2.3): $r_\epsilon(j,0) = \frac{\epsilon^2}{\kappa(j)}\phi(\epsilon j)$ and $p_\epsilon(j,0) = \epsilon^2\psi(\epsilon j)$. We denote the corresponding solution of (2.1) by $(r_\epsilon(t), p_\epsilon(t))$.

Given (2.5), we set

$$A_0 := \frac{1}{2}\left(\phi - \sqrt{\overline{m}}\breve{\kappa}\psi\right) \quad \text{and} \quad B_0 := \frac{1}{2}\left(\phi + \sqrt{\overline{m}}\breve{\kappa}\psi\right)$$

with $\overline{m}$ and $\breve{\kappa}$ defined as in Definition 2.2. It is clear that $A_0, B_0 \in H^5$ and in particular

$$\|A_0\|_{H^5} + \|B_0\|_{H^5} \leq C_5.$$

Let $A(w,T)$ and $B(l,T)$ be the unique global-in-time solutions of (3.39) with these initial conditions which are guaranteed by Theorem 4.5. Our goal is to estimate

$$e := \sup_{|t|\leq T_0\epsilon^{-3}} \left\| r_\epsilon(\cdot,t) - \frac{1}{\kappa(\cdot)}\left[\epsilon^2 A(\epsilon(\cdot - ct), \epsilon^3 t) + \epsilon^2 B(\epsilon(\cdot + ct), \epsilon^3 t)\right] \right\|_{\ell^2} \tag{5.1}$$

and

$$f := \sup_{|t|\leq T_0\epsilon^{-3}} \left\| p_\epsilon(\cdot,t) - \sqrt{\overline{m}}\breve{\kappa}\left[\epsilon^2 A(\epsilon(\cdot - ct), \epsilon^3 t) - \epsilon^2 B(\epsilon(\cdot + ct), \epsilon^3 t)\right] \right\|_{\ell^2}. \tag{5.2}$$

Namely, we will have proven Theorem 2.1 if we show

$$e + f \leq C_5\epsilon^{5/2}.$$

5.1. Long time existence for solution of (2.1). The first thing we need to do is ensure that the solutions of (2.1) exist over the very long time scale $|t| \leq T_0\epsilon^{-3}$. The following theorem does just this.

THEOREM 5.2 (small data implies global existence for (2.1)). *There exists $\rho_*, C_* > 0$, determined wholly by m and $\mathcal{V}$, such that if*

$$\|(r_0, p_0)\|_{\ell^2\times\ell^2} \leq \rho_*,$$

then there exists a unique function

$$(r(t), p(t)) \in C^1(\mathbf{R}; \ell^2 \times \ell^2)$$

which solves (2.1) *and for which* $(r(0), p(0)) = (r_0, p_0)$. *Moreover*

$$\sup_{t\in\mathbf{R}} \|(r(t), p(t))\|_{\ell^2\times\ell^2} \leq C_*\|(r_0, p_0)\|_{\ell^2\times\ell^2}.$$

Proof. Define

$$H(t) := \sum_{j\in\mathbf{Z}} \left[\frac{1}{2}m(j)p_j^2 + \mathcal{V}_j(r(j))\right].$$

This quantity is the total mechanical energy of (2.1) and thus

$$H(t) = H(0) \tag{5.3}$$

so long as the solution exists. Moreover, when $\|(r, p)\|_{\ell\times\ell^2}$ is sufficiently small, (1.2), (1.3), and (1.4) imply that $\sqrt{H(t)}$ is equivalent to the usual norm on the $\ell^2 \times \ell^2$ norm. That is, there exist $\rho > 0$ and $C_{**} > 1$ such that $\|(r, p)\|_{\ell^2\times\ell^2} < \rho$ implies

$$\frac{1}{C_{**}}\sqrt{H(t)} \leq \|(r, p)\|_{\ell^2\times\ell^2} \leq C_{**}\sqrt{H(t)}. \tag{5.4}$$

To see this, note that Taylor's theorem, (1.3), and (1.4) imply that there exists $\rho > 0$ such that $|y| \leq \rho$ implies

$$\frac{1}{4}\min_{k\in\mathbf{Z}} \kappa(k)y^2 \leq \mathcal{V}_j(y) \leq \frac{3}{4}\max_{k\in\mathbf{Z}} \kappa(k)y^2 \tag{5.5}$$

for all j. Now suppose that $\|(r, p)\|_{\ell^2\times\ell^2} < \rho$. This implies $\|r\|_{\ell^\infty} \leq \rho$, and thus $\|(r, p)\|_{\ell^2\times\ell^2} \leq \rho$:

$$\frac{1}{4}\min_{k\in\mathbf{Z}} \kappa(k)\|r\|_{\ell^2}^2 \leq \sum_{j\in\mathbf{Z}} \mathcal{V}_j(r(j)) \leq \frac{3}{4}\max_{k\in\mathbf{Z}} \kappa(k)\|r\|_{\ell^2}^2. \tag{5.6}$$

Similarly (1.2) implies

$$\frac{1}{2}\min_{k\in\mathbf{Z}} m(k)\|p\|_{\ell^2}^2 \leq \sum_{j\in\mathbf{Z}} \frac{1}{2}m(j)p^2(j) \leq \frac{1}{2}\max_{k\in\mathbf{Z}} m(k)\|p\|_{\ell^2}^2, \tag{5.7}$$

and thus we have (5.4) for an appropriately defined constant C_{**}.

Let $\rho_* := \rho/2C_{**}^2$ and suppose that $\|(r_0, p_0)\|_{\ell^2\times\ell^2} \leq \rho_*$. Since $C_{**} > 1$, we have $\rho_* < \rho/2$. The right-hand side of (2.1) is a smooth and bounded map from $\ell^2 \times \ell^2$ into itself, and so Picard's theorem provides $t_0 > 0$ and $(r(t), p(t)) \in C^1([-t_0, t_0]; \mathfrak{B}_\rho)$, which uniquely solves (2.1) with $(r(0), p(0)) = (r_0, p_0)$. Here $\mathfrak{B}_\rho := \big\{(r,p) \in \ell^2 \times \ell^2 : \|(r,p)\|_{\ell^2\times\ell^2} < \rho\big\}$. Thus we can employ (5.4) and (5.3) to find

$$\sup_{|t|\le t_0} \|(r(t),p(t))\|_{\ell^2\times\ell^2} \le C_{**}\sqrt{H(t)} = C_{**}\sqrt{H(0)} \le C_{**}^2\|(r(0),p(0))\|_{\ell^2\times\ell^2} \le \rho/2.$$

This implies via a straightforward bootstrap argument that $t_0 = +\infty$. Putting $C_* = C_{**}^2$ finishes the proof. ∎

Now, if we use Lemma 4.3, estimate (4.8), we see that

$$\|(r_\epsilon(0), p_\epsilon(0))\|_{\ell^2\times\ell^2} \le C\epsilon^{3/2}\left(\|\phi\|_{H^1} + \|\psi\|_{H^1}\right) \le C_5\epsilon^{3/2}.$$

As such, there exist $\epsilon_1 > 0$, which depends on $(\|\phi\|_{H^1} + \|\psi\|_{H^1})$ such that

$$\|(r_\epsilon(0), p_\epsilon(0))\|_{\ell^2\times\ell^2} \le \rho_*$$

when $0 \le \epsilon \le \epsilon_1$. Thus $(r_\epsilon(j,t), p_\epsilon(j,t))$ exists for all $t \in \mathbf{R}$. Note that the estimate in Theorem 5.2 shows that

$$\sup_{t\in\mathbf{R}} \|(r_\epsilon(t), p_\epsilon(t))\|_{\ell^2\times\ell^2} \le C_*\|(r_\epsilon(0), p_\epsilon(0))\|_{\ell^2\times\ell^2} \le C_5\epsilon^{3/2}. \tag{5.8}$$

5.2. Smoothing. Now that we have established the existence of the solution of (2.1) for all t, we turn our attention to the approximation. Note that the functions A and B are only in H^5. To use Proposition 4.2 we need functions which are in H^7, and so we will "smooth" our initial conditions using the following result.

LEMMA 5.3 (smooth approximation with ϵ dependent estimates). *Define the map $\mathcal{T}_\epsilon$ as a Fourier multiplier operator $\widehat{\mathcal{T}_\epsilon U}(K) = \widehat{\mathcal{T}}_\epsilon(K)\widehat{U}(K)$, where $\widehat{\mathcal{T}}_\epsilon(K) = 1$ when $|K| \le \epsilon^{-1}$ and is zero otherwise. Then, for $0 \le \epsilon \le 1$, we have*

$$\|U - \mathcal{T}_\epsilon U\|_{H^s} \le \epsilon^{(5-s)}\|U\|_{H^5} \tag{5.9}$$

and

$$\|\mathcal{T}_\epsilon U\|_{H^s} \le \|U\|_{H^s} \tag{5.10}$$

when $0 \le s \le 5$. Also

$$\|\mathcal{T}_\epsilon U\|_{H^s} \le C\epsilon^{(5-s)}\|U\|_{H^5} \tag{5.11}$$

for $s \ge 5$. Here, $C = C(s) > 0$ does not depend on ϵ. Furthermore

$$\int_{\mathbf{R}} \mathcal{T}_\epsilon U(X)dX = \int_{\mathbf{R}} U(X)dX \tag{5.12}$$

and

$$\|\mathcal{U} - \tilde{\mathcal{U}}\|_{L^\infty} \le C\epsilon^{11/2}\|U\|_{H^5}, \tag{5.13}$$

where

$$\mathcal{U}(X) := \int_0^X U(X')dX' \quad \textit{and} \quad \tilde{\mathcal{U}}(X) := \int_0^X \mathcal{T}_\epsilon U(X')dX'.$$

Proof. The proofs of the first four estimates are routine and so we omit them. The proof of (5.13) is less typical. First, note that

$$\mathcal{U}(X) - \tilde{\mathcal{U}}(X) = \int_0^X [U(X') - \mathcal{T}_\epsilon U(X')] dX'.$$

Using the Fourier inversion theorem gives

$$\mathcal{U}(X) - \tilde{\mathcal{U}}(X) = \int_0^X \int_{\mathbf{R}} e^{iKX'} \left[\widehat{U}(K) - \widehat{\mathcal{T}_\epsilon U}(K)\right] dK dX'.$$

The definition of $\mathcal{T}_\epsilon$ gives

$$\mathcal{U}(X) - \tilde{\mathcal{U}}(X) = \int_0^X \int_{|K| \geq \epsilon^{-1}} e^{iKX'} \widehat{U}(K) dK dX'.$$

Now, $U \in H^5$ implies that $\widehat{U} \in L^1$. Thus we may exchange the order of integration above to get

$$\mathcal{U}(X) - \tilde{\mathcal{U}}(X) = \int_{|K| \geq \epsilon^{-1}} \int_0^X e^{iKX'} \widehat{U}(K) dX' dK = \int_{|K| \geq \epsilon^{-1}} \widehat{U}(K) \left(\int_0^X e^{iKX'} dX' \right) dK.$$

We evaluate the X'-integral:

$$\mathcal{U}(X) - \tilde{\mathcal{U}}(X) = \int_{|K| \geq \epsilon^{-1}} \widehat{U}(K) \frac{e^{iKX} - 1}{iK} dK.$$

Thus

$$\left|\mathcal{U}(X) - \tilde{\mathcal{U}}(X)\right| \leq 2 \int_{|K| \geq \epsilon^{-1}} |\widehat{U}(K)| |K|^{-1} dK.$$

Multiplication by 1 gives

$$\left|\mathcal{U}(X) - \tilde{\mathcal{U}}(X)\right| \leq 2 \int_{|K| \geq \epsilon^{-1}} (1 + K^2)^{5/2} |\widehat{U}(K)| (1 + K^2)^{-5/2} |K|^{-1} dK.$$

Then Cauchy–Schwarz gives

$$\left|\mathcal{U}(X) - \tilde{\mathcal{U}}(X)\right| \leq 2 \|U\|_{H^5} \sqrt{\int_{|K| \geq \epsilon^{-1}} (1 + K^2)^{-5} |K|^{-2} dK} \leq C \epsilon^{11/2} \|U\|_{H^5}.$$

Since X was arbitrary, we have (5.13). ∎

Now let

$$\bar{A}_0 := \mathcal{T}_\epsilon A_0 \quad \text{and} \quad \bar{B}_0 := \mathcal{T}_\epsilon B_0.$$

Then Lemma 5.3 implies

$$\|\bar{A}_0 - A_0\|_{H^1} \leq \epsilon^4 \|A_0\|_{H^5} \leq C_5 \epsilon^4 \quad \text{and} \quad \|\bar{B}_0 - B_0\|_{H^1} \leq \epsilon^4 \|B_0\|_{H^5} \leq C_5 \epsilon^4. \tag{5.14}$$

Likewise,

$$\|\bar{A}_0\|_{H^5} + \|\bar{B}_0\|_{H^5} \leq C_5, \tag{5.15}$$

$$\|\bar{A}_0\|_{H^6} + \|\bar{B}_0\|_{H^6} \leq C_5 \epsilon^{-1}, \tag{5.16}$$

$$\|\bar{A}_0\|_{H^7} + \|\bar{B}_0\|_{H^7} \leq C_5 \epsilon^{-2}, \tag{5.17}$$

and

$$\|\bar{\mathcal{A}}_0\|_{L^\infty} + \|\bar{\mathcal{B}}_0\|_{L^\infty} \le C_5.$$

Recall that $\bar{\mathcal{A}}_0(w) := \int_0^w \bar{A}_0(y)dy$ and $\bar{\mathcal{B}}_0(l) := \int_0^l \bar{B}_0(y)dy$. Note that these estimates imply

$$\|\bar{A}_0\|_{H^5} + \|\bar{B}_0\|_{H^5} + \|\bar{\mathcal{A}}_0\|_{L^\infty} + \|\bar{\mathcal{B}}_0\|_{L^\infty} \le C_5,$$

which, so to speak, says "$K_5 \le C_5$."

Let $\bar{A}(w,T)$ and $\bar{B}(l,T)$ be the unique global-in-time solutions of (3.39) initial conditions $\bar{A}_0$, $\bar{B}_0$ which are guaranteed by Theorem 4.5. These functions are in H^7 by virtue of (5.17). Now we return to our estimates of e and f. The triangle inequality gives

$$\begin{aligned} e \le& \sup_{|t|\le T_0\epsilon^{-3}} \left\| r_\epsilon(t) - \frac{1}{\kappa(\cdot)}\left[\epsilon^2 \bar{A}(\epsilon(\cdot - ct), \epsilon^3 t) + \epsilon^2 \bar{B}(\epsilon(\cdot + ct), \epsilon^3 t)\right] \right\|_{\ell^2} \\ &+ \sup_{|t|\le T_0\epsilon^{-3}} \left\| \frac{1}{\kappa(\cdot)}\epsilon^2 A(\epsilon(\cdot - ct), \epsilon^3 t) - \frac{1}{\kappa(\cdot)}\epsilon^2 \bar{A}(\epsilon(\cdot - ct), \epsilon^3 t) \right\|_{\ell^2} \\ &+ \sup_{|t|\le T_0\epsilon^{-3}} \left\| \frac{1}{\kappa(\cdot)}\epsilon^2 B(\epsilon(\cdot + ct), \epsilon^3 t) - \frac{1}{\kappa(\cdot)}\epsilon^2 \bar{B}(\epsilon(\cdot + ct), \epsilon^3 t) \right\|_{\ell^2}. \end{aligned}$$

Using (4.8) from Lemma 4.3 gives

$$\begin{aligned} &\sup_{|t|\le T_0\epsilon^{-3}} \left\| \frac{1}{\kappa(\cdot)}\epsilon^2 A(\epsilon(\cdot - ct), \epsilon^3 t) - \frac{1}{\kappa(\cdot)}\epsilon^2 \bar{A}(\epsilon(\cdot - ct), \epsilon^3 t) \right\|_{\ell^2} \\ \le& \sup_{|t|\le T_0\epsilon^{-3}} C\epsilon^{3/2} \|A(\cdot - ct, \epsilon^3 t) - \bar{A}(\cdot - ct, \epsilon^3 t)\|_{H^1}. \end{aligned}$$

The H^1 norm is shift invariant, and so

$$\begin{aligned} &\sup_{|t|\le T_0\epsilon^{-3}} \left\| \frac{1}{\kappa(\cdot)}\epsilon^2 A(\epsilon(\cdot - ct), \epsilon^3 t) - \frac{1}{\kappa(\cdot)}\epsilon^2 \bar{A}(\epsilon(\cdot - ct), \epsilon^3 t) \right\|_{\ell^2} \\ \le& \sup_{|T|\le T_0} C\epsilon^{3/2} \|A(\cdot, T) - \bar{A}(\cdot, T)\|_{H^1}. \end{aligned}$$

To control $A - \bar{A}$ we use the following.

Corollary 5.4 (KdV solution map is Lipschitz). *Suppose that U, $\bar{U}$ solve the KdV equation* (4.12) *with $a \ne 0$ and with $U(X,0), \bar{U}(X,0) \in H^3$. Then there exists $\mu > 0$ such that*

$$\|U(T) - \bar{U}(T)\|_{H^1} \le e^{\mu T} \|U(0) - \bar{U}(0)\|_{H^1}$$

for all $T > 0$. The constant μ depends on $\|U(0)\|_{H^3} + \|\bar{U}(0)\|_{H^3}$.

The proof follows from routine energy arguments similar to but easier than those used to prove the continuous dependence on initial conditions for KdV equations in [1]. We omit it. Using this, we see that

$$\sup_{|T|\le T_0} \|A(\cdot, T) - \bar{A}(\cdot, T)\|_{H^1} \le C_5 \|A_0 - \bar{A}_0\|_{H^1}.$$

Using (5.14) gives

$$\sup_{|T|\leq T_0} \|A(\cdot,T) - \bar{A}(\cdot,T)\|_{H^1} \leq C_5\epsilon^4.$$

The same reasoning applied to the "B" terms gives

$$e \leq \sup_{|t|\leq T_0\epsilon^{-3}} \left\| r_\epsilon(t) - \frac{1}{\kappa(\cdot)}\left[\epsilon^2\bar{A}(\epsilon(\cdot - ct),\epsilon^3 t) + \epsilon^2\bar{B}(\epsilon(\cdot + ct),\epsilon^3 t)\right]\right\|_{\ell^2} + C_5\epsilon^{11/2}$$

and

$$f \leq \sup_{|t|\leq T_0\epsilon^{-3}} \left\| p_\epsilon(t) - \sqrt{\overline{m}\breve{\kappa}}\left[\epsilon^2\bar{A}(\epsilon(\cdot - ct),\epsilon^3 t) - \epsilon^2\bar{B}(\epsilon(\cdot + ct),\epsilon^3 t)\right]\right\|_{\ell^2} + C_5\epsilon^{11/2},$$

and so we need to estimate

$$e_1 := \sup_{|t|\leq T_0\epsilon^{-3}} \left\| r_\epsilon(t) - \frac{1}{\kappa(\cdot)}\left[\epsilon^2\bar{A}(\epsilon(\cdot - ct),\epsilon^3 t) + \epsilon^2\bar{B}(\epsilon(\cdot + ct),\epsilon^3 t)\right]\right\|_{\ell^2}$$

and

$$f_1 := \sup_{|t|\leq T_0\epsilon^{-3}} \left\| p_\epsilon(t) - \sqrt{\overline{m}\breve{\kappa}}\left[\epsilon^2\bar{A}(\epsilon(\cdot - ct),\epsilon^3 t) - \epsilon^2\bar{B}(\epsilon(\cdot + ct),\epsilon^3 t)\right]\right\|_{\ell^2}.$$

5.3. Energy estimates. Now form $(\tilde{r}_\epsilon(t), \tilde{p}_\epsilon(t))$ from $\bar{A}(w,T)$ and $\bar{B}(l,T)$ as described in section 3. Addition, subtraction, and the triangle inequality give

$$\begin{aligned} e_1 \leq &\sup_{|t|\leq T_0\epsilon^{-3}} \|r_\epsilon(t) - \tilde{r}_\epsilon(t)\|_{\ell^2} \\ &+ \sup_{|t|\leq T_0\epsilon^{-3}} \left\| \tilde{r}_\epsilon(t) - \frac{1}{\kappa(\cdot)}\left[\epsilon^2\bar{A}(\epsilon(\cdot - ct),\epsilon^3 t) + \epsilon^2\bar{B}(\epsilon(\cdot + ct),\epsilon^3 t)\right]\right\|_{\ell^2} \end{aligned}$$

and

$$\begin{aligned} f_1 \leq &\sup_{|t|\leq T_0\epsilon^{-3}} \|p_\epsilon(t) - \tilde{p}_\epsilon(t)\|_{\ell^2} \\ &+ \sup_{|t|\leq T_0\epsilon^{-3}} \left\| \tilde{p}_\epsilon(t) - \sqrt{\overline{m}\breve{\kappa}}\left[\epsilon^2\bar{A}(\epsilon(\cdot - ct),\epsilon^3 t) - \epsilon^2\bar{B}(\epsilon(\cdot + ct),\epsilon^3 t)\right]\right\|_{\ell^2}. \end{aligned}$$

Using the definitions of $\breve{r}_\epsilon$, $\breve{p}_\epsilon$ together with (4.4) in Proposition 4.2, we see that

$$\begin{aligned} &\sup_{|t|\leq T_0\epsilon^{-3}} \left\| \tilde{r}_\epsilon(t) - \frac{1}{\kappa(\cdot)}\left[\epsilon^2\bar{A}(\epsilon(\cdot - ct),\epsilon^3 t) + \epsilon^2\bar{B}(\epsilon(\cdot + ct),\epsilon^3 t)\right]\right\|_{\ell^2} \\ &\qquad = \sup_{|t|\leq T_0\epsilon^{-3}} \|\breve{r}_\epsilon(t)\|_{\ell^2} \leq C_5\epsilon^{5/2} \end{aligned}$$

and

$$\begin{aligned} &\sup_{|t|\leq T_0\epsilon^{-3}} \left\| \tilde{p}_\epsilon(t) - \sqrt{\overline{m}\breve{\kappa}}\left[-\epsilon^2\bar{A}(\epsilon(\cdot - ct),\epsilon^3 t) + \epsilon^2\bar{B}(\epsilon(\cdot + ct),\epsilon^3 t)\right]\right\|_{\ell^2} \\ &\qquad = \sup_{|t|\leq T_0\epsilon^{-3}} \|\breve{p}_\epsilon(t)\|_{\ell^2} \leq C_5\epsilon^{5/2}. \end{aligned}$$

Thus we need to estimate

$$e_2 := \sup_{|t|\le T_0\epsilon^{-3}} \|r_\epsilon(t) - \tilde{r}_\epsilon(t)\|_{\ell^2} \quad \text{and} \quad f_2 := \sup_{|t|\le T_0\epsilon^{-3}} \|p_\epsilon(t) - \tilde{p}_\epsilon(t)\|_{\ell^2}.$$

The argument we use is based on a similar one developed in [21]. Define $\eta(j,t)$ and $\xi(j,t)$ via

$$r_\epsilon = \tilde{r}_\epsilon + \epsilon^{5/2}\eta \quad \text{and} \quad p_\epsilon = \tilde{p}_\epsilon + \epsilon^{5/2}\xi. \tag{5.18}$$

Clearly, then, $(\eta(t),\xi(t))$ exist for all $|t|\le T_0\epsilon^{-3}$. If we can show that they are $O(1)$ on that interval, we will be done.

A direct computation shows that η and ξ solve

$$\begin{aligned} \dot{\eta} &= \delta^+\xi + \epsilon^{-5/2}\mathrm{Res}_1(\tilde{r}_\epsilon,\tilde{p}_\epsilon),\\ \dot{\xi} &= \frac{1}{m}\delta^-\left[\tilde{\mathcal{V}}'(\eta;t)\right] + \epsilon^{-5/2}\mathrm{Res}_2(\tilde{r}_\epsilon,\tilde{p}_\epsilon), \end{aligned} \tag{5.19}$$

where $\tilde{\mathcal{V}}$ is defined as

$$\tilde{\mathcal{V}}_j(z;t) := \epsilon^{-5}\left[\mathcal{V}_j(\tilde{r}_\epsilon(j,t)+\epsilon^{5/2}z) - \mathcal{V}_j(\tilde{r}_\epsilon(j,t)) - \mathcal{V}_j'(\tilde{r}_\epsilon(j,t))\epsilon^{5/2}z\right].$$

The "primes" on $\tilde{\mathcal{V}}$ are derivatives with respect to z.

The $\epsilon^{-5/2}$ prefactors which multiply the residuals in (5.19) look troubling, but recall that (4.6) in Proposition 4.2 implies that the residuals are $O(\epsilon^{11/2})$. In particular if we combine that estimate with the smoothing estimates in (5.15), (5.16), and (5.17), we obtain

$$\sup_{|t|\le T_0\epsilon^{-3}} \|(\mathrm{Res}_1(\tilde{r}_\epsilon(t),\tilde{p}_\epsilon(t)), \mathrm{Res}_2(\tilde{r}_\epsilon(t),\tilde{p}_\epsilon(t)))\|_{\ell^2\times\ell^2} \le C_5\epsilon^{11/2}. \tag{5.20}$$

That is to say, these terms are, in fact, very small.

Since the residual terms are small, we can gain some insight into (5.19) by considering the system without them. In this case (5.19) looks formally very much like (2.1), which, given that $H(t)$ is constant for (2.1), leads us to the conclusion that

$$E(t) := \sum_{j\in\mathbf{Z}}\left[\frac{1}{2}m(j)\xi^2(j,t) + \tilde{\mathcal{V}}_j(\eta(j,t);t)\right]$$

might be "sort of conserved" for (5.19). We claim that $\sqrt{E(t)}$ is equivalent to the $\ell^2\times\ell^2$ norm of (η,ξ). We can use (5.7) to handle the "ξ" part of $E(t)$.

Handling the "$\tilde{\mathcal{V}}$" part of $E(t)$ is very similar to the estimate (5.6) for $H(t)$ above. Taylor's theorem gives

$$\tilde{\mathcal{V}}_j(\eta(j,t);t) = \frac{1}{2}\mathcal{V}_j''(b(j,t))\eta^2(j,t),$$

where $b(j,t)$ lies between $\tilde{r}_\epsilon(j,t)$ and $\tilde{r}_\epsilon(j,t)+\epsilon^{5/2}\eta(j,t) = r_\epsilon(j,t)$. Thus

$$|b(j,t)| \le \max\left\{\|\tilde{r}_\epsilon(t)\|_{\ell^\infty}, \|r_\epsilon(t)\|_{\ell^\infty}\right\}.$$

Using (5.8) and Proposition 4.2, this gives

$$\sup_{j\in\mathbf{Z}}\sup_{|t|\le T_0\epsilon^{-3}} |b(j,t)| \le C_5\epsilon^{3/2}.$$

Since the $b(j,t)$ are small, then we must have $\mathcal{V}_j''(b(j,t)) \sim \mathcal{V}_j''(0) = \kappa(j)$. More specifically, there exists $0 \le \epsilon_2 \le C_5$, with $\epsilon_2 \le \epsilon_1$ from above, such that $0 \le \epsilon \le \epsilon_2$ implies

$$\frac{1}{2}\min_{k\in\mathbf{Z}} \kappa(k) \le \sup_{j\in\mathbf{Z}}\sup_{t\in\mathbf{R}} \mathcal{V}_j''(b(j,t)) \le \frac{3}{2}\max_{k\in\mathbf{Z}} \kappa(k).$$

Thus we have

$$\frac{1}{4}\min_j \kappa(j)\|\eta(t)\|_{\ell^2}^2 \le \sum_{j\in\mathbf{Z}} \tilde{\mathcal{V}}_j(\eta(j,t);t) \le \frac{3}{4}\max_j \kappa(j)\|\eta(t)\|_{\ell^2}^2,$$

and so there exists $C_e > 1$ with

$$\frac{1}{C_e}\|(\eta(t),\xi(t))\|_{\ell^2\times\ell^2} \le \sqrt{E(t)} \le C_e\|(\eta(t),\xi(t))\|_{\ell^2\times\ell^2} \tag{5.21}$$

for all $|t| \le T_0\epsilon^{-3}$; this establishes the equivalence of $\sqrt{E(t)}$ to the $\ell^2\times\ell^2$ norm.

REMARK 4. *In* [21], *the authors develop an "alternate energy" to control the errors in an ad hoc way. Our $E(t)$ is (essentially) the same energy functional that they use; here we see that it arises naturally using the mechanical energy as a starting point.*

Moving on, we differentiate $E(t)$ to get

$$\dot{E}(t) = \sum_{j\in\mathbf{Z}} \left[m(j)\xi(j,t)\xi_t(j,t) + \tilde{\mathcal{V}}_j'(\eta(j,t);t)\eta_t(j,t) + \partial_t\tilde{\mathcal{V}}_j(\eta(j,t);t)\right].$$

Using (5.19) gives

$$\dot{E}(t) = \sum_{j\in\mathbf{Z}} \left[\xi\delta^-[\tilde{\mathcal{V}}'(\eta;t)] + \epsilon^{-5/2}m\xi\mathrm{Res}_2 + \tilde{\mathcal{V}}'(\eta;t)\delta^+\xi + \epsilon^{-5/2}\tilde{\mathcal{V}}'(\eta;t)\mathrm{Res}_1 + \partial_t\tilde{\mathcal{V}}(\eta;t)\right],$$

where we have hidden all explicit dependencies of the functions to make things more readable. Summing by parts kills a few terms, and we get

$$\dot{E}(t) = \sum_{j\in\mathbf{Z}} \left[\epsilon^{-5/2}m\xi\mathrm{Res}_2 + \epsilon^{-5/2}\tilde{\mathcal{V}}'(\eta;t)\mathrm{Res}_1 + \partial_t\tilde{\mathcal{V}}(\eta;t)\right]. \tag{5.22}$$

Cauchy–Schwarz gives

$$\sum_{j\in\mathbf{Z}} \epsilon^{-5/2}m\xi\mathrm{Res}_2 \le C_5\epsilon^{-5/2}\|\xi\|_{\ell^2}\|\mathrm{Res}_2\|_{\ell^2}.$$

Using (5.20) then implies

$$\sum_{j\in\mathbf{Z}} \epsilon^{-5/2}m\xi\mathrm{Res}_2 \le C_5\epsilon^3\|\xi\|_{\ell^2}.$$

Likewise, we have the following for the second term in (5.22):

$$\sum_{j\in\mathbf{Z}} \epsilon^{-5/2}\tilde{\mathcal{V}}'(\eta;t)\mathrm{Res}_1 \le C_5\epsilon^3\left\|\tilde{\mathcal{V}}'(\eta;t)\right\|_{\ell^2}.$$

Now, the definition of $\tilde{\mathcal{V}}$ tells us

$$\tilde{\mathcal{V}}_j'(\eta(j,t);t) = \epsilon^{-5/2}\left[\mathcal{V}_j'(\tilde{r}_\epsilon(j,t)+\epsilon^{5/2}\eta(j,t)) - \mathcal{V}_j'(\tilde{r}_\epsilon(j,t))\right].$$

The mean value theorem gives

$$\tilde{\mathcal{V}}_j'(\eta(j,t);t) = \mathcal{V}_j''(b(j,t))\eta(j,t),$$

where $b(j,t)$ is between $\tilde{r}_\epsilon(j,t)$ and $\tilde{r}_\epsilon(j,t)+\epsilon^{5/2}\eta(j,t)$. Using the same estimates and reasoning that led to (5.21) gives

$$\left|\mathcal{V}_j''(b(j,t))\right| \le \frac{3}{2}\max_{k\in\mathbf{Z}}\kappa(k).$$

Thus

$$\sum_{j\in\mathbf{Z}}\epsilon^{-5/2}\tilde{\mathcal{V}}'(\eta;t)\mathrm{Res}_1 \le C_5\epsilon^3\|\eta\|_{\ell^2}.$$

Finally, a direct calculation shows that

$$\begin{aligned}&\partial_t\tilde{\mathcal{V}}_j(\eta(j,t);t)\\&\quad=\epsilon^{-5}\partial_t\tilde{r}_\epsilon(t)\left[\mathcal{V}_j'(\tilde{r}_\epsilon(j,t)+\epsilon^{5/2}\eta(j,t)) - \mathcal{V}_j'(\tilde{r}_\epsilon(j,t)) - \mathcal{V}_j''(\tilde{r}_\epsilon(j,t))\epsilon^{5/2}\eta(j,t)\right].\end{aligned}$$

Taylor's theorem again gives

$$\mathcal{V}_j'(\tilde{r}_\epsilon(j,t)+\epsilon^{5/2}\eta(j,t)) - \mathcal{V}_j'(\tilde{r}_\epsilon(j,t)) - \mathcal{V}_j''(\tilde{r}_\epsilon(j,t))\epsilon^{5/2}\eta(j,t) = \frac{1}{2}V'''(b(j,t))\epsilon^5\eta^2(j,t)$$

for $b(j,t)$ between $\tilde{r}_\epsilon(j,t)$ and $\tilde{r}_\epsilon(j,t)+\epsilon^{5/2}\eta(j,t)$. As before, we know $|b(j,t)| \le C_5\epsilon^{3/2}$, and so we use the same ideas as above to get

$$\left|\partial_t\tilde{\mathcal{V}}_j(\eta(j,t);t)\right| \le C_5\left|\partial_t\tilde{r}_\epsilon(j,t)\right|\eta^2(j,t) \le C_5\|\partial_t\tilde{r}_\epsilon(t)\|_{\ell^\infty}\eta^2(j,t).$$

Then we have, using (4.5),

$$\left|\partial_t\tilde{\mathcal{V}}_j(\eta(j,t);t)\right| \le C_5\epsilon^3\eta^2(j,t),$$

and therefore

$$\sum_{j\in\mathbf{Z}}\partial_t\tilde{\mathcal{V}}_j(\eta,t) \le C_5\epsilon^3\|\eta(t)\|_{\ell^2}^2.$$

Putting the above together gives

$$\dot{E} \le C_5\epsilon^3\left(\|(\eta,\xi)\|_{\ell^2\times\ell^2} + \|(\eta,\xi)\|_{\ell^2\times\ell^2}^2\right).$$

Since $y \le 1+y^2$ for all $y\in\mathbf{R}$, this implies

$$\dot{E} \le C_5\epsilon^3\left(1+\|(\eta,\xi)\|_{\ell^2\times\ell^2}^2\right).$$

Using (5.21), this becomes

$$\dot{E} \le C_5\epsilon^3\left(1+C_e^2E\right) \le C_5\epsilon^3(1+E).$$

Gronwall's inequality then tells us that

$$E(t) \le e^{C_5 \epsilon^3 t}(1 + E(0)) - 1.$$

Thus,

$$\sup_{|t| \le T_0 \epsilon^{-3}} E(t) \le C_5(1 + E(0)).$$

Using (5.21),

$$\sup_{|t| \le T_0 \epsilon^{-3}} \|(\eta(t), \xi(t))\|^2_{\ell^2 \times \ell^2} \le C_5(1 + \|(\eta(0), \xi(0))\|^2_{\ell^2 \times \ell^2}). \tag{5.23}$$

A short calculation using the definitions of η and ξ reveals that

$$\eta(0) = \epsilon^{-5/2} \breve{r}_\epsilon(0) \quad \text{and} \quad \xi(0) = \epsilon^{-5/2} \breve{p}_\epsilon(0), \tag{5.24}$$

and so (4.4) indicates

$$\|(\eta(0), \xi(0))\|_{\ell^2 \times \ell^2} \le C_5.$$

Thus,

$$\sup_{|t| \le T_0 \epsilon^{-3}} \|(\eta(t), \xi(t))\|^2_{\ell^2 \times \ell^2} \le C_5.$$

With this estimate, we have proven Theorem 2.1.

Remark 5. *Due to the length of the argument, the origin of the discrepancy of a power of ϵ between our error estimates (which are $O(\epsilon^{5/2})$) and the results in* [3] *and* [21] *(which are $O(\epsilon^{7/2})$) is not immediately transparent. Let us first discuss the difference between our estimate and that of* [21] *(which treats the homogeneous problem). If one examines the formulas in* (3.24), (3.25), (3.31), *and* (3.33), *one sees that the χ_n are exactly zero when the material constants do not vary. This has the effect that (more or less) the expansion in* (3.6) *only has terms with even powers of ϵ. The lack of odd powers of ϵ in the expansion improves the estimate on the residuals in* (4.15) *from $O(\epsilon^{11/2})$ to $O(\epsilon^{13/2})$ "for free" because there* are *no terms of $O(\epsilon^{11/2})$ in them. Subsequently everything conspires together in a nice way to get the better error estimate.*

On the other hand, in [3], *the authors proceed using Bloch wave transforms instead of the homogenization approach we employ. This makes it somewhat difficult to precisely explain the difference. Roughly speaking, what they do is to include higher order terms in their definition of the KdV approximation. In our case, the KdV approximation given by* (2.7) *can be succinctly written as*

$$(r, p) \sim \epsilon^2 (R_0, P_0). \tag{5.25}$$

Steps analogous to those undertaken in [3] *would be to replace this with something akin to*

$$(r, p) \sim \epsilon^2 (R_0, P_0) + \epsilon^3 (R_1, P_1). \tag{5.26}$$

Ultimately this would improve the estimate on the residuals to $O(\epsilon^{13/2})$ and would result in the better overall error estimate.

So why do we not do this? Notice that since we made our choice (5.25)*, given the long wave data* (2.3) *for* (2.1)*, the choice of the initial conditions for the KdV equations is natural and can be done independently of ϵ. All we do is set the initial conditions of* (2.1) *equal to the approximation in* (2.7) *evaluated at $t = 0$ and solve for A_0 and B_0; all ϵ cancel.*

If we used (5.26)*, on the other hand, notice that the definitions of R_1 and P_1 in* (3.26) *involve χ_1 and χ_2, which in the nonhomogeneous case are nonzero. They also involve derivatives of A and B. It is no longer clear how one would go from* (2.3) *to choices for A_0 and B_0. It certainly cannot be done in a way independent of ϵ. Ultimately, the authors of* [3] *choose the initial conditions for the KdV equations first and then use their analogue of* (5.26) *to choose the initial conditions for* (2.1)*. This amounts to placing further restrictions on the initial conditions of* (2.1) *which are more complicated than ours are in* (2.3)*. The price we have paid for allowing a larger class of initial data is a less accurate estimate. We note that this is an occurrence of a common problem with the validity of homogenization approximations near the boundary* [4]*. Lastly we remark that the discrepancy is not connected in any way to the loosening of the restrictions on the regularity or decay of the initial data, but only to the process by which we "reconcile" the initial data of the lattice problem with that of its long wave limit.*

6. Simulations. To demonstrate our results, we carry our simulations of the lattice differential equation (2.1) and compare them with solutions of the KdV equations. We simulate (2.1) by truncating the lattice to include $M \gg 0$ sites, enforcing periodic boundary conditions at the ends, and then using a standard Runge–Kutta (RK4) algorithm to compute solutions of the resulting system of differential equations. As for the KdV equations, most of our simulations use the well-known explicit formulas for the soliton solutions (see [6]) and thus do not require simulation.

6.1. Solitary waves. First we specify ϵ, $m(j)$, $\kappa(j)$, and $\beta(j)$ in such a way so that $ab \neq 0$. We then set

$$\phi(X,0) = \frac{3}{b}\operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}X\right) \quad \text{and} \quad \psi(X,0) = \frac{3}{b\sqrt{\overline{m}\breve{\kappa}}}\operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}X\right), \tag{6.1}$$

which implies via (2.5) that

$$A(w,0) = \frac{3}{b}\operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}w\right) \quad \text{and} \quad B(l,0) = 0.$$

The initial condition for A is exactly that of the KdV solitary wave, and so

$$A(w,T) = \frac{3}{b}\operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}(w-cT)\right) \quad \text{and} \quad B(l,T) = 0$$

for all T. Given Theorem 2.1, we expect the solution of (2.1) to be approximately a solitary wave. We compute $(r(j,t), p(j,t))$ with our RK4 algorithm over the interval $t \in [0, \epsilon^{-3}]$. We repeat the process for a variety of $0 < \epsilon < 1$.

6.1.1. Mass dimer. In this case, we had

$$\kappa(j) = 1, \quad \beta = 1, \quad \text{and} \quad m(j) = 1.5 + (-1)^j 0.5. \tag{6.2}$$

Since $N = 2$ here, such a lattice is called a "dimer" [3]. Figure 1 contains snapshots of the solution of (2.1) together with the approximation at several times. For all ϵ,

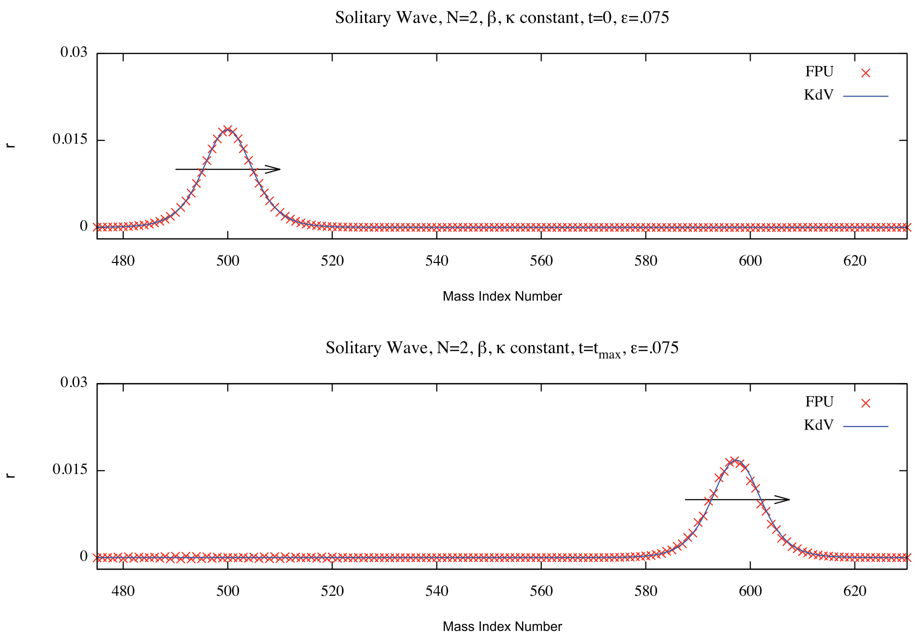

Fig. 1. *Solution (r-component only) of* (2.1) *with initial data given by* (6.1) *and coefficients* (6.2) *together with the KdV approximation (given by* (2.7)*). Here* $\epsilon = 0.075$.

the r-component of the solution is qualitatively a solitary wave which propagates to the right plus minor features which are much smaller than the amplitude of the wave. (These discrepancies are consistent with those observed in [23].)

In Figure 2 we plot the error between the numerically computed solution and the KdV approximation (namely, $e + f$ from (5.1) and (5.2) above) as a function of ϵ. The slope of the resulting line indicates the power of ϵ in the approximation. In this case we find the power is

$$\text{power} \sim 2.473,$$

which is in line with the power of ϵ stated in Theorem 2.1, i.e., 2.5.

6.1.2. General dimer. In this case, we had

$$\kappa(j) = 1.5 + (-1)^j 0.5, \quad \beta = 1.5 + (-1)^j 0.5, \quad \text{and} \quad m(j) = 1.5 + (-1)^j 0.5. \tag{6.3}$$

Again, $N = 2$. Figure 3 contains snapshots of the solution of (2.1) together with the approximation at several times. For all ϵ, the r-component of the solution is qualitatively a "spiky" solitary wave which propagates to the right. The irregular features are due to the prefactor of $1/\kappa(j)$ in (2.7). Note that in [15] and [14], the authors observe the same sort of solution for models of waves in layered elastic media. They call such solutions "stegotons," given their resemblance to the dinosaur stegosaurus; though we prefer "hedgehogons," we abide by their choice.

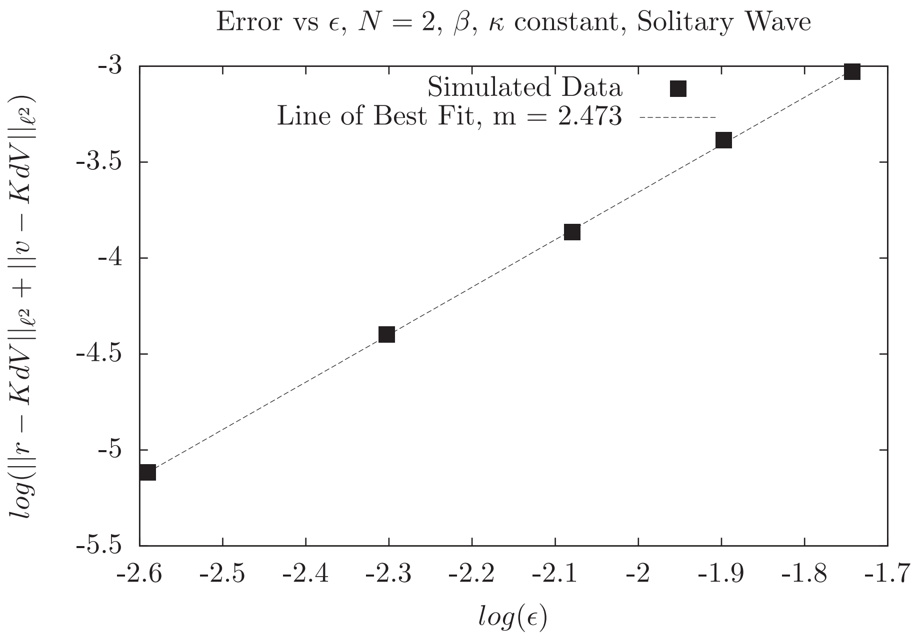

FIG. 2. *Error of the approximation of solutions of* (2.1) *by KdV equations with initial data given by* (6.1) *and coefficients* (6.2).

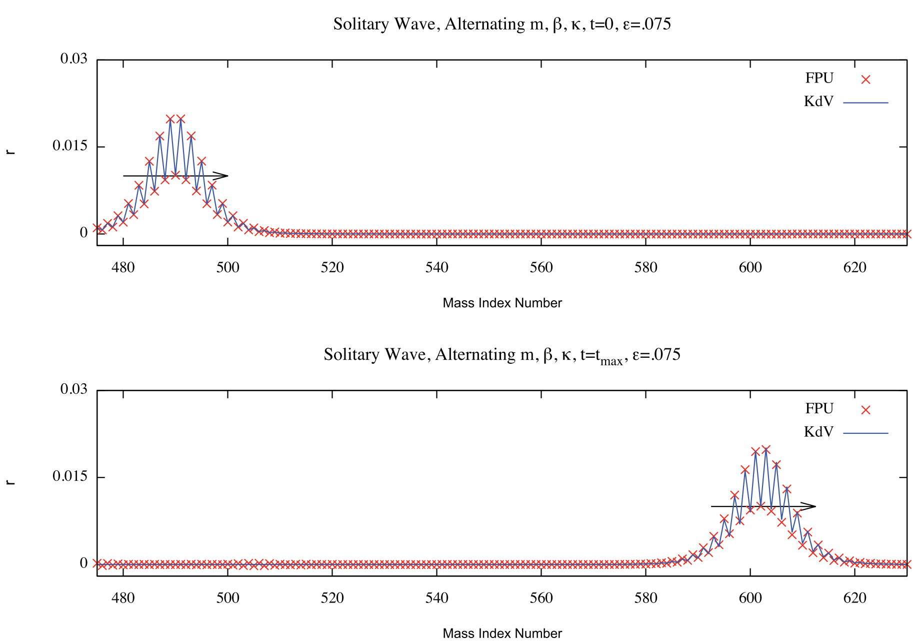

FIG. 3. *Solution (r-component only) of* (2.1) *with initial data given by* (6.1) *and coefficients* (6.3) *together with the KdV approximation (given by* (2.7)*). Here* $\epsilon = 0.075$.

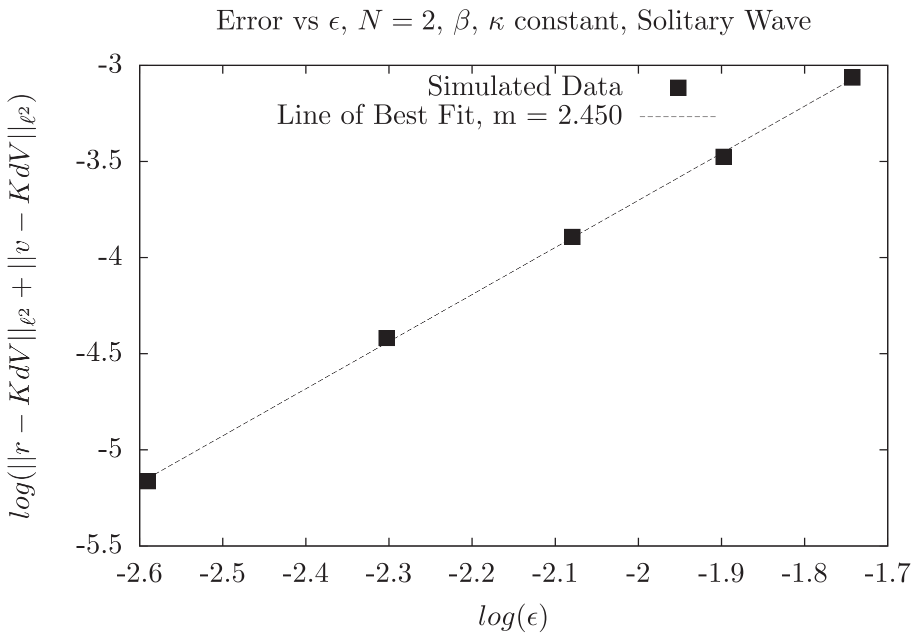

FIG. 4. *Error of the approximation of solutions of* (2.1) *by KdV equations with initial data given by* (6.1) *and coefficients* (6.3)*.*

In Figure 4 we plot the error against ϵ as above. In this case we find the numerically computed power of ϵ in the error is

$$\text{power} \sim 2.450,$$

which is in line with the power of ϵ stated in Theorem 2.1, i.e., 2.5.

6.1.3. Mass polymer, $N = 10$. In this case, we let $\kappa(j) = 1$ and $\beta(j) = 1$ and randomly selected ten positive numbers (taking values between 0.5 and 2.5) to be the masses $m(j)$. Figure 5 contains snapshots of the solution of (2.1) together with the approximation at several times. For all ϵ, the r-component of the solution is qualitatively a solitary wave which propagates to the right.

In Figure 6 we plot the error against ϵ as above. In this case we find the numerically computed power of ϵ in the error is

$$\text{power} \sim 2.682,$$

which is in line with the power of ϵ stated in Theorem 2.1, i.e., 2.5.

6.1.4. General polymer, $N = 100$. In this case, we let $\kappa(j)$, $\beta(j)$, and $m(j)$ each be one hundred randomly selected positive numbers (taking values between 0.5 and 2.5 for $\kappa(j)$ and $m(j)$ and between 0 and 1 for $\beta(j)$). Figure 7 contains a snapshot of the solution of (2.1) together with the approximation at several times. For all ϵ, the r-component of the solution is qualitatively a particularly spiky stegoton which propagates to the right.

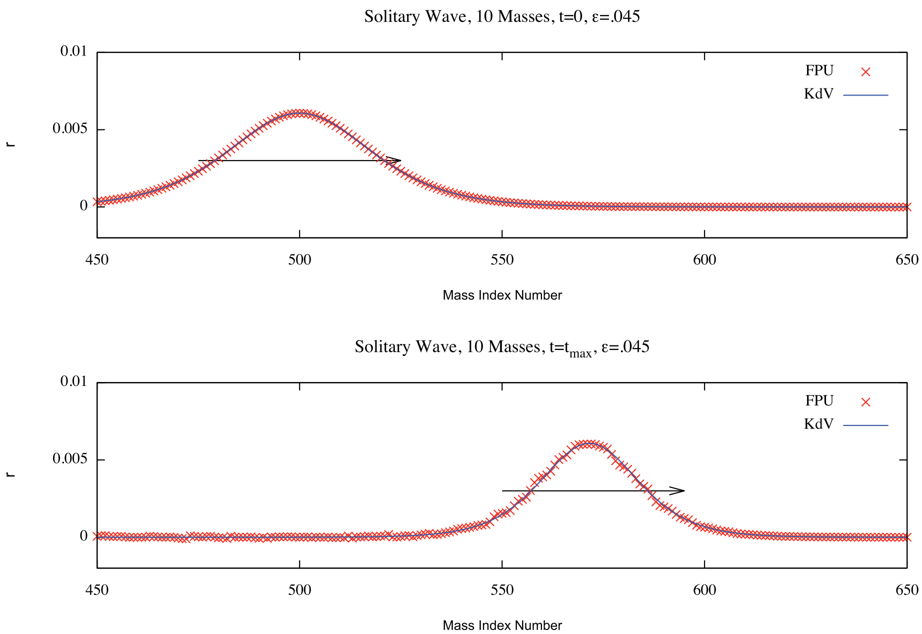

Fig. 5. *Solution (r-component only) of* (2.1) *with initial data given by* (6.1) *and coefficients where* κ *and* β *are constant and* $m(j)$ *varies with period* $N = 10$*, together with the KdV approximation (given by* (2.7)*).*

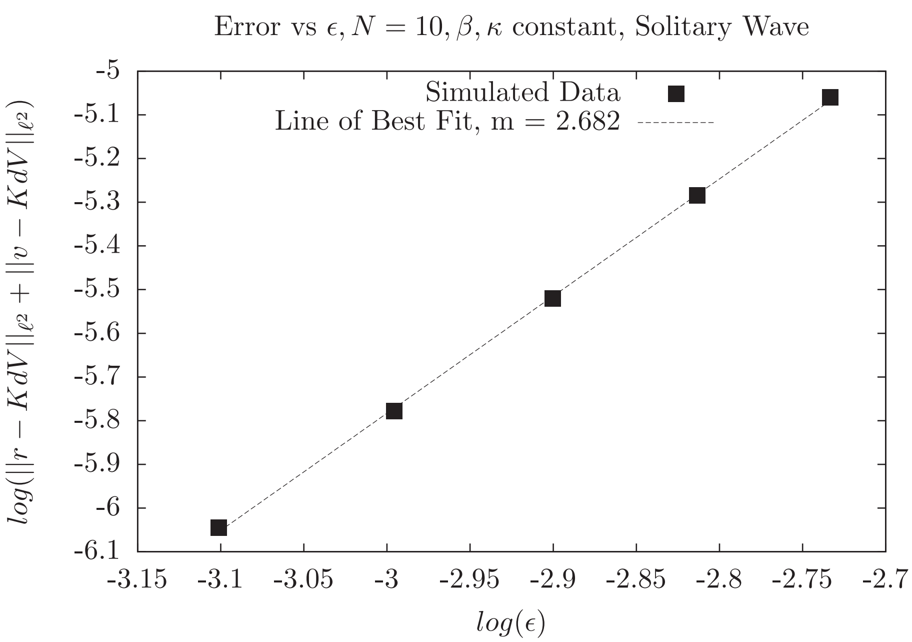

Fig. 6. *Error of the approximation of solutions of* (2.1) *by KdV equations when initial data given by* (6.1) *and coefficients where* κ *and* β *are constant and* $m(j)$ *varies with period* $N = 10$*.*

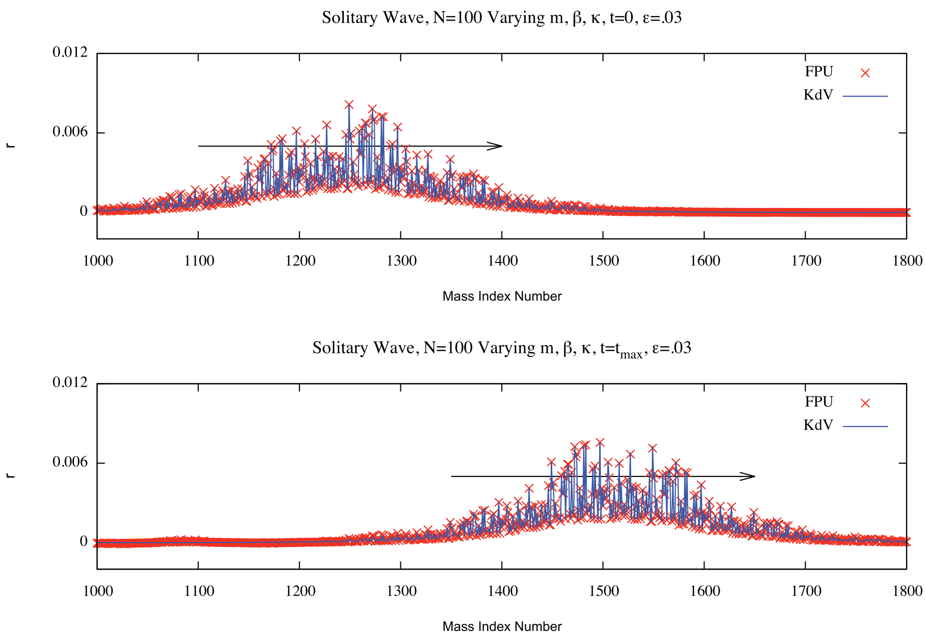

FIG. 7. *Solution (r-component only) of* (2.1) *with initial data given by* (6.1) *together with the KdV approximation (given by* (2.7)*). Here, the coefficients* $\kappa(j)$*,* $\beta(j)$*, and* $m(j)$ *vary with period* $N = 100$ *and* $\epsilon = 0.03$.

6.2. Head-on collision of equal amplitude solitary waves. First we specify ϵ, $m(j)$, $\kappa(j)$, and $\beta(j)$ in such a way so that $ab \neq 0$. We then set

$$\phi(X,0) = \frac{3}{b}\sum_{i=1}^{2} \operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}(X - x_i)\right)$$

(6.4) and

$$\psi(X,0) = \frac{3}{b\sqrt{\overline{m}\breve{\kappa}}}\sum_{i=1}^{2}(-1)^i \operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}(X - x_i)\right),$$

where x_1 and x_2 simply translate the peaks horizontally. For these simulations $\frac{x_1 - x_2}{\epsilon}$ was approximately 30. This implies via (2.5) that

$$A(w,0) = \frac{3}{b}\operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}(w - x_1)\right) \quad \text{and} \quad B(l,0) = \frac{3}{b}\operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}(l - x_2)\right).$$

These latter guarantee that

$$A(w,T) = \frac{3}{b}\operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}(w - x_1 - cT)\right) \text{ and } B(l,T) = \frac{3}{b}\operatorname{sech}^2\left(\frac{1}{2\sqrt{a}}(l - x_2 + cT)\right)$$

for all T. Given Theorem 2.1, we expect the solution of (2.1) to be the head-on collision of two equal amplitude solitary waves. We compute $(r(j,t), p(j,t))$ with our RK4 algorithm over the interval $t \in [0, \epsilon^{-3}]$. We repeat the process for a variety of $0 < \epsilon < 1$.

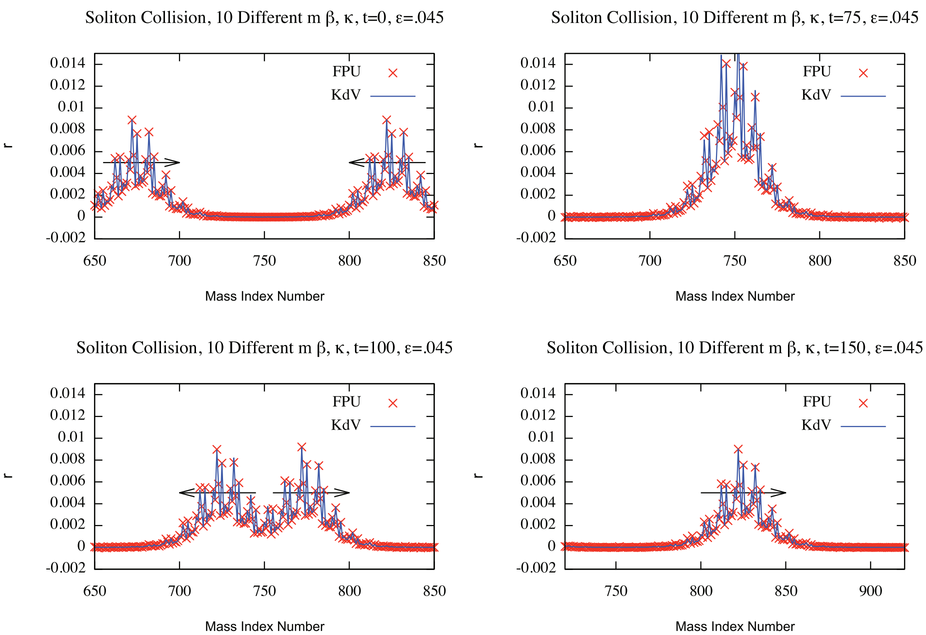

Fig. 8. *Solution (r-component only) of* (2.1) *with initial data given by* (6.4) *together with the KdV approximation (given by* (2.7)*). Here the coefficients* $\kappa(j)$*,* $\beta(j)$*, and* $m(j)$ *vary with period* $N = 10$ *and* $\epsilon = 0.045$*.*

6.2.1. General dimer, $N = 10$. In this case, we took $\kappa(j)$, $\beta(j)$, and $m(j)$ to be ten randomly selected positive numbers each (taking values between 0.5 and 2.5 for $\kappa(j)$ and $m(j)$ and between 0 and 1 for $\beta(j)$). Figure 8 contains snapshots of the solution of (2.1) together with the approximation at several times. In Figure 9 we plot the error against ϵ. In this case we find the numerically computed power of ϵ in the error is

$$\text{power} \sim 2.715,$$

which is in line with the power of ϵ stated in Theorem 2.1, i.e., 2.5.

6.3. Approximation by Airy's equation. It is possible that the coefficients $\kappa(j)$ and $\beta(j)$ are arranged in just such a way that $b = 0$ in (3.39). The simplest such case is when

$$m = 1, \quad \kappa = 1, \quad \text{and} \quad \beta(j) = (-1)^j. \tag{6.5}$$

We treat this situation here. In this case, the nonlinear problem (2.1) is approximated by two linear Airy's equations, which after appropriate rescaling, are of the form

$$U_T = U_{yyy}.$$

We compute solutions of this using the explicit formula $\widehat{U}(k,T) = e^{-ik^3T}\widehat{U}(k,T)$ and using standard techniques to approximate the Fourier transform with the FFT.

We take as initial conditions

$$\phi(X,0) = \operatorname{sech}^2(X) \quad \text{and} \quad \psi(X,0) = 0. \tag{6.6}$$

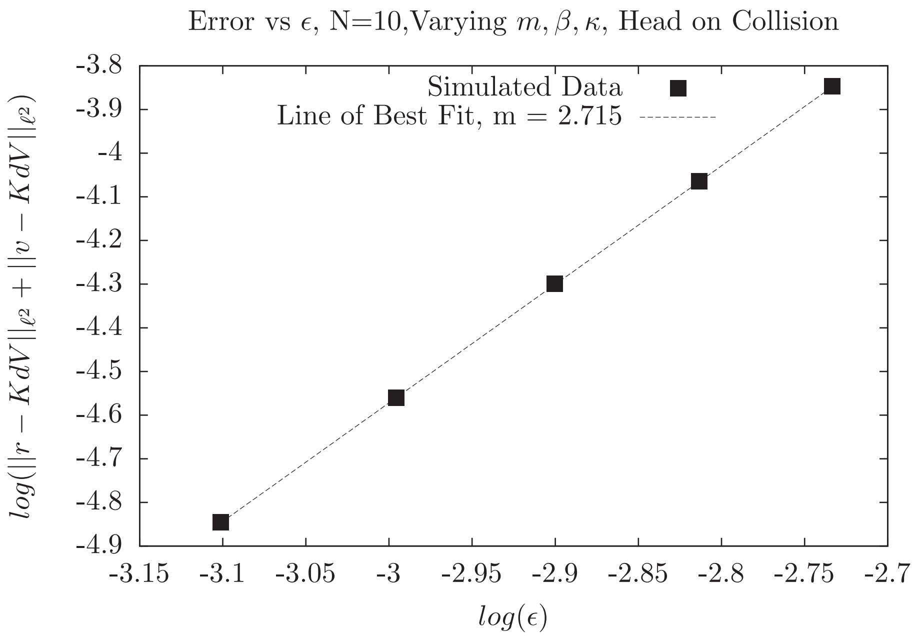

Fig. 9. *Error of the approximation of solutions of* (2.1) *by KdV equations when the initial data given by* (6.1) *and coefficients where* $\kappa(j)$, $\beta(j)$, *and* $m(j)$ *vary with period* $N = 10$.

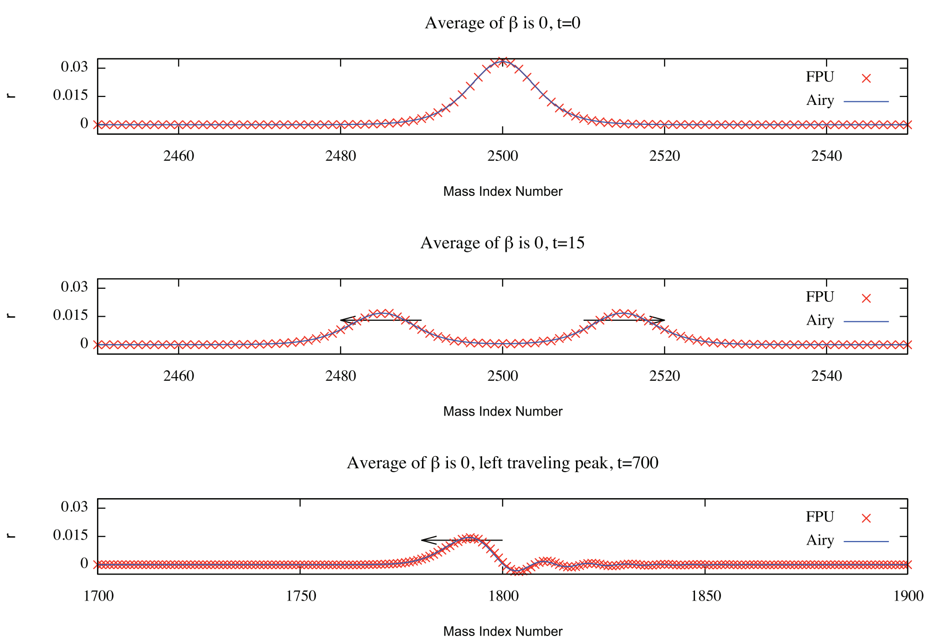

Fig. 10. *Solution (*r*-component only) of* (2.1) *with initial data given by* (6.6) *and coefficients given by* (6.5) *together with the Airy approximation (given by* (2.7)*). Here* $\epsilon = 0.2$*. Observe how the solutions break up into a piece which moves left and one which moves right. In the third graph, note the dispersive tail which forms behind the left-moving wave; this is behavior characteristic of Airy's equation solutions.*

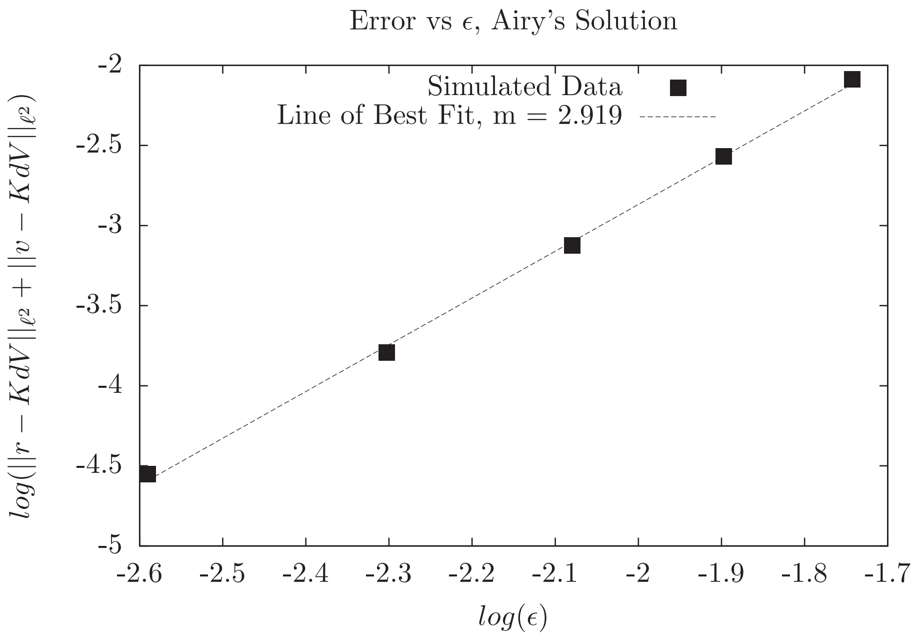

FIG. 11. *Error of the approximation of solutions of* (2.1) *by Airy equations with the initial data given by* (6.5) *and coefficients given by* (6.6).

Figure 10 contains snapshots of the solution of (2.1) together with the approximation at several times.

In Figure 11 we plot the error against ϵ. In this case we find the numerically computed power of ϵ in the error is

$$\text{power} \sim 2.919.$$

Note that this is quite a bit greater than the error expected. In this setting, since κ and m are constant, χ_1 and χ_2 are zero. Many of the terms in the approximation consequently vanish, and we expect a corresponding improvement in the error bound to 3.5.

Acknowledgment. J. Douglas Wright thanks Aaron Hoffman for his helpful comments.

REFERENCES

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

[2] J. F. Bukowski, *The Boussinesq limit of the Fermi-Pasta-Ulam equation*, Ph.D. Thesis, Brown University, ProQuest LLC, Ann Arbor, MI, 1997.

[3] M. Chirilus-Bruckner, C. Chong, O. Prill, and G. Schneider, *Rigorous description of macroscopic wave packets in infinite periodic chains of coupled oscillators by modulation equations*, Discrete Contin. Dyn. Syst. Ser. S, 5 (2012), pp. 879–901.

[4] D. Cioranescu and P. Donato, *An Introduction to Homogenization*, Oxford Lecture Ser. Math. Appl. 17, The Clarendon Press, Oxford University Press, New York, 1999.

[5] P. C. Dash, *Lattice solitons and nonlinear diatomic models*, in Solitons (Tiruchirapalli, 1987), Springer Ser. Nonlinear Dynam., Springer, Berlin, 1988, pp. 196–211.

[6] P. G. Drazin and R. S. Johnson, *Solitons: An Introduction*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, UK, 1989.

[7] E. Fermi, J. Pasta, and S. Ulam, *Collected Paper of Enrico Fermi, Vol.* 2, University of Chicago Press, Chicago, 1965.

[8] G. Friesecke and R. L. Pego, *Solitary waves on FPU lattices.* I. *Qualitative properties, renormalization and continuum limit*, Nonlinearity, 12 (1999), pp. 1601–1627.

[9] G. Friesecke and R. L. Pego, *Solitary waves on Fermi-Pasta-Ulam lattices.* IV. *Proof of stability at low energy*, Nonlinearity, 17 (2004), pp. 229–251.

[10] J. Giannoulis and A. Mielke, *The nonlinear Schrödinger equation as a macroscopic limit for an oscillator chain with cubic nonlinearities*, Nonlinearity, 17 (2004), pp. 551–565.

[11] M. Herrmann and J. D. M. Rademacher, *Heteroclinic travelling waves in convex FPU-type chains*, SIAM J. Math. Anal., 42 (2010), pp. 1483–1504.

[12] G. James and P. Noble, *Breathers on diatomic Fermi-Pasta-Ulam lattices*, Phys. D, 196 (2004), pp. 124–171.

[13] M. D. Kruskal, *The Korteweg-de Vries equation and related evolution equations*, in Nonlinear Wave Motion (Proc. AMS-SIAM Summer Sem., Clarkson Coll. Tech., Potsdam, NY, 1972), Lectures in Appl. Math. 15, AMS, Providence, RI, 1974, pp. 61–83.

[14] R. J. LeVeque and D. H. Yong, *Phase plane behavior of solitary waves in nonlinear layered media*, in Hyperbolic Problems: Theory, Numerics, Applications, Springer, Berlin, 2003, pp. 43–51.

[15] R. J. LeVeque and D. H. Yong, *Solitary waves in layered nonlinear media*, SIAM J. Appl. Math., 63 (2003), pp. 1539–1560.

[16] A. Mielke, *Macroscopic behavior of microscopic oscillations in harmonic lattices via Wigner-Husimi transforms*, Arch. Ration. Mech. Anal., 181 (2006), pp. 401–448.

[17] R. M. Miura, C. S. Gardner, and M. D. Kruskal, *Korteweg-de Vries equation and generalizations.* II. *Existence of conservation laws and constants of motion*, J. Math. Phys., 9 (1968), pp. 1204–1209.

[18] St. Pnevmatikos, N. Flytzanis, and M. Remoissenet, *Soliton dynamics of nonlinear diatomic lattices*, Phys. Rev. B, 33 (1986), pp. 2308–2321.

[19] M. A. Porter, C. Daraio, I. Szelengowicz, E. B. Herbold, and P. G. Kevrekidis, *Highly nonlinear solitary waves in heterogeneous periodic granular media*, Phys. D, 238 (2009), pp. 666–676.

[20] G. Schneider, *Bounds for the nonlinear Schrödinger approximation of the Fermi-Pasta-Ulam system*, Appl. Anal., 89 (2010), pp. 1523–1539.

[21] G. Schneider and C. E. Wayne, *Counter-propagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model*, in International Conference on Differential Equations, Vols. 1, 2 (Berlin, 1999), World Scientific, River Edge, NJ, 2000, pp. 390–404.

[22] A. Vainchtein and P. G. Kevrekidis, *Dynamics of phase transitions in a piecewise linear diatomic chain*, J. Nonlinear Sci., 22 (2012), pp. 107–134.

[23] C. E. Wayne and J. D. Wright, *Higher order modulation equations for a Boussinesq equation*, SIAM J. Appl. Dyn. Syst., 1 (2002), pp. 271–302.

[24] N. Zabusky and M. Kruskal, *Interaction of solitons in a collisionless plasma and the recurrence of initial states*, Phys. Rev. Lett., 15 (1965), pp. 240–243.